Lecture Notes in Mechanical Engineering

Halgamuge Saman
Yan Peng
Dingxuan Zhao
Yongming Bian *Editors*

The 9th International Conference on Advances in Construction Machinery and Vehicle Engineering

ICACMVE 2024

OPEN ACCESS

Springer

Lecture Notes in Mechanical Engineering

Series Editors

Fakher Chaari, National School of Engineers, University of Sfax, Sfax, Tunisia

Francesco Gherardini, Dipartimento di Ingegneria “Enzo Ferrari”, Università di Modena e Reggio Emilia, Modena, Italy

Vitalii Ivanov, Department of Manufacturing Engineering, Machines and Tools, Sumy State University, Sumy, Ukraine

Mohamed Haddar, National School of Engineers of Sfax (ENIS), Sfax, Tunisia

Editorial Board

Francisco Cavas-Martínez, Departamento de Estructuras, Construcción y Expresión Gráfica Universidad Politécnica de Cartagena, Cartagena, Spain

Francesca di Mare, Institute of Energy Technology, Ruhr-Universität Bochum, Bochum, Germany

Young W. Kwon, Department of Manufacturing Engineering and Aerospace Engineering, Graduate School of Engineering and Applied Science, Monterey, USA

Tullio A. M. Tolio, Department of Mechanical Engineering, Politecnico di Milano, Milano, Italy

Justyna Trojanowska, Poznan University of Technology, Poznan, Poland

Robert Schmitt, RWTH Aachen University, Aachen, Germany

Jinyang Xu, School of Mechanical Engineering, Shanghai Jiao Tong University, Shanghai, China

Lecture Notes in Mechanical Engineering (LNME) publishes the latest developments in Mechanical Engineering—quickly, informally and with high quality. Original research or contributions reported in proceedings and post-proceedings represents the core of LNME. Volumes published in LNME embrace all aspects, subfields and new challenges of mechanical engineering.

To submit a proposal or request further information, please contact the Springer Editor of your location:

Europe, USA, Africa: Leontina Di Cecco at Leontina.dicecco@springer.com
China: Ella Zhang at ella.zhang@cn.springernature.com
India, Rest of Asia, Australia, New Zealand: Swati Meherishi at swati.meherishi@springer.com

Topics in the series include:

- Engineering Design
- Machinery and Machine Elements
- Mechanical Structures and Stress Analysis
- Automotive Engineering
- Engine Technology
- Aerospace Technology and Astronautics
- Nanotechnology and Microengineering
- Control, Robotics, Mechatronics
- MEMS
- Theoretical and Applied Mechanics
- Dynamical Systems, Control
- Fluid Mechanics
- Engineering Thermodynamics, Heat and Mass Transfer
- Manufacturing Engineering and Smart Manufacturing
- Precision Engineering, Instrumentation, Measurement
- Materials Engineering
- Tribology and Surface Technology

Indexed by SCOPUS, EI Compendex, and INSPEC.

All books published in the series are evaluated by Web of Science for the Conference Proceedings Citation Index (CPCI).

To submit a proposal for a monograph, please check our Springer Tracts in Mechanical Engineering at https://link.springer.com/bookseries/11693.

Halgamuge Saman · Yan Peng · Dingxuan Zhao · Yongming Bian
Editors

The 9th International Conference on Advances in Construction Machinery and Vehicle Engineering

ICACMVE 2024

Springer

Editors
Halgamuge Saman
School of Electrical Mechanical
and Infrastructure Engineering
University of Melbourne
Melbourne, VIC, Australia

Dingxuan Zhao
Yanshan University
Qinhuangdao, Hebei, China

Yan Peng
School of Mechanical Engineering
Yanshan University
Qinhuangdao, Hebei, China

Yongming Bian
College of Mechanical Engineering
Tongji University
Shanghai, China



ISSN 2195-4356 ISSN 2195-4364 (electronic)
Lecture Notes in Mechanical Engineering
ISBN 978-981-96-5529-8 ISBN 978-981-96-5527-4 (eBook)
https://doi.org/10.1007/978-981-96-5527-4

This work was supported by Guangzhou KEO Info Technology Co., Ltd.

© The Editor(s) (if applicable) and The Author(s) 2026. This book is an open access publication.

Open Access This book is licensed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License (http://creativecommons.org/licenses/by-nc-nd/4.0/), which permits any noncommercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if you modified the licensed material. You do not have permission under this license to share adapted material derived from this book or parts of it.
The images or other third party material in this book are included in the book's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the book's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.
This work is subject to copyright. All commercial rights are reserved by the author(s), whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, reuse of illustrations, recitation, broadcasting, reproduction on microfilms or in any other physical way, and transmission or information storage and retrieval, electronic adaptation, computer software, or by similar or dissimilar methodology now known or hereafter developed. Regarding these commercial rights a non-exclusive license has been granted to the publisher.
The use of general descriptive names, registered names, trademarks, service marks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.
The publisher, the authors and the editors are safe to assume that the advice and information in this book are believed to be true and accurate at the date of publication. Neither the publisher nor the authors or the editors give a warranty, expressed or implied, with respect to the material contained herein or for any errors or omissions that may have been made. The publisher remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

This Springer imprint is published by the registered company Springer Nature Singapore Pte Ltd.
The registered company address is: 152 Beach Road, #21-01/04 Gateway East, Singapore 189721, Singapore

If disposing of this product, please recycle the paper.

Committee Members

Honorary Chairman

Shi Laide, China Construction Machinery Society, China

Conference Chairman

Zhao Dingxuan, China Construction Machinery Society, Yanshan University, China

Co-Chairmen

Bian Yongming, China Construction Machinery Society, Tongji University, China
Marcus Geimer, Professor, Karlsruhe Institute of Technology, Germany

Publication Chairmen

Bian Yongming, Tongji University, China
Zhao Dingxuan, Yanshan University, China
Yan Peng, Yanshan University, China
Saman Halgamuge, University of Melbourne, Australia

Academic Committee Chairmen

Zhong Zhihua, Chinese Academy of Engineering, China
Yang Huayong, Zhejiang University, China

Academic Committee Co-Chairmen

Chen Xuedong, Hefei General Machinery Research Institute, China
Zhu Hehua, Tongji University, China
Zheng Jianlong, Changsha University of Science and Technology, China
Yan Xinping, Wuhan University of Technology, China
Ge Shirong, China University of Mining and Technology-Beijing, China
Zongquan Deng, Harbin Institute of Technology, China
Zhong Jue, Central South University, China
Wen Bangchun, Northeastern University, China
Lin Baojun, University of China Academy Sciences
Wang Guobiao, Tianjin University, China
Feng Pei'en, Zhejiang University, China
Zheng Huiqiang, Tongji University, China
Jürgen Fleischer, German Academy of Engineering, Germany
Saman K. Halgamuge, The University of Melbourne, Australia
Michael Abramovici, Ruhr University Bochum, Germany

Academic Committee Members

Bao Weihua, Shanghai Automation Instrumentation Co., Ltd., China
Cai Hui, China Railway Engineering Equipment (Tianjin) Co., Ltd., China
Cao Ming, University of Shanghai for Science and Technology, China
Cao Donghui, SANY Heavy Industry Co., Ltd., China
Cao Gengxiang, CCCC Third Harbor Engineering Bureau Co., Ltd, China
Cao Yuanwen, Chongqing Jiaotong University, China
Chen Li, Tongji University, China
Cheng Haiying, Chang'an University, China
Cheng Wenming, Southwest Jiaotong University, China
Cheng Xiaoming, Ningbo University of Technology, China
Chi Feng, Shandong Lingong Construction Machinery, China
Cong Jinling, Shihezi University, China
Dong Xichen, Wuhan University of Technology, China
Duan Wenjun, China Railway Engineering Services Limited, China
Fu Maohua, Shanghai Zhenhua Heavy Industries Co., Ltd., China

Gao Rongzhi, Zoomlion Concrete Pumping Machinery Branch, China
Gao Shunde, Dalian University of Technology, China
Gao Yanbei, Chang'an University, China
Gong Mingde, Yanshan University, China
Guo Rui, Yanshan University, China
Guo Chuanxin, Beijing Institute of Construction Mechanization Co., Ltd., China
Guo Wenwu, Shijiazhuang Tiedao University, China
Guo Xiaohong, Chongqing Jiaotong University, China
Hao Rujiang, Shijiazhuang Tiedao University, China
He Qin, Shandong Jianzhu University, China
Hu Lin, Changsha University of Science and Technology, China
Hu Yongbiao, Chang'an University, China
Huang Hui, Huaqiao University, China
Huang Ming, Guangxi LiuGong Machinery Co., Ltd., China
Huang Yi, China Communications Construction Company, China
Huang Haibo, Ningbo University, China
Huang Heting, Xiamen XGMA International Trading Co., Ltd., China
Huang Qiufang, Fujian Longma Sanitation Equipment Co., Ltd., China
Huang Yunpeng, Shenji (Shanghai) Intelligent System R&D and Design Co., Ltd., China
Kou Ziming, Taiyuan University of Technology, China
Lan Peng, Xi'an University of Architecture and Technology, China
Li Hong, Yunnan Construction Machinery Industry Association, China
Li Zheng, Inner Mongolia University of Science and Technology, China
Li Zong, Xuzhou Construction Machinery Group Co., Ltd., China
Li Anhu, Tongji University, China
Li Suoyun, Xuzhou Construction Machinery Group Co., Ltd., China
Li Taijie, Da Gang Holding Group Co., Ltd., China
Li Yunhua, Beihang University, China
Li Zhonggang, Lanzhou Jiaotong University, China
Lin Deng, Zhejiang Yongan Construction Machinery Co., Ltd., China
Lin Mingzhi, Guangxi LiuGong Machinery Co., Ltd., China
Liu Guangjun, Tongji University, China
Liu Xiangdong, Inner Mongolia University of Technology, China
Liu Xinhui, Jilin University, China
Liu Yuxin, Zoomlion Construction Crane Branch, China
Liu Yuanhong, Beijing Institute of Construction Mechanization Co., Ltd., China
Liu Zhengfu, Kunshi Control Systems (Shanghai) Co., Ltd., China
Long Liang, Changsha Zoomlion Environmental Industry Co., Ltd., China
Lu Liang, Tongji University, China
Lu Nianli, Harbin Institute of Technology, China
Ma Biao, Shanghai Municipal Engineering Design Institute (Group) Co., Ltd., China
Meng Xianjun, China Railway Tunnel Bureau Group Company Limited Equipment Branch, China
Min Junying, Tongji University, China

Ni Tao, Yanshan University, China
Qin Lisheng, Shanghai Tongxin Electromechanical Control Technology Co., Ltd., China
Qiu Jiang, Shanghai Environmental Engineering Design and Research Institute Co., Ltd., China
Quan Long, Taiyuan University of Technology, China
Shang Haibo, Tianjin Construction Machinery Research Institute Co., Ltd., China
Shen Jijun, China Construction Machinery Co., Ltd., China
Sheng Jinliang, Tongji University, China
Shi Xianxin, XCMG Xuzhou Heavy Machinery Co., Ltd., China
Sun Hong, Shenyang Jianzhu University, China
Teng Ruming, Dalian University of Technology, China
Tian Zhicheng, CIMIC (Beijing) Vehicle Inspection Engineering Research Institute Co., Ltd., China
Wang Anlin, Tongji University, China
Wang Caiying, Huicong Construction Machinery Network, China
Wang Dajiang, Qinhuangdao Tianye Tolian Heavy Industry and Technology Co., Ltd., China
Wang Jiru, Hangzhou Huaxin Mechanical and Electrical Engineering Co., Ltd., China
Wang Jiangbo, Jiangsu Donghua Testing Technology Co., Ltd., China
Wang Zhiyong, Caofeidian Port Group Co., Ltd., China
Wang Zhongbin, China University of Mining and Technology, China
Wu Chong, Tongji University, China
Wu Binxing, Zoomlion Heavy Industry Co., Ltd., China
Wu Renzhi, Tongji University, China
Xie Haibo, Zhejiang University, China
Xin Ligang, Yantai Hyde Specialized Vehicle Co., Ltd., China
Xu Bing, Zhejiang University, China
Xu Huaiyu, Xuzhou Xugong Schwing Machinery Co., Ltd., China
Xun Zhaohong, Anhui Taiyuan Construction Machinery Co., Ltd., China
Yang bin, Phoenix (China) Investment Co., Ltd., China
Yang Wenbiao, Anhui Taiyuan Construction Machinery Co., Ltd., China
Yi Hong, Lanzhou University of Technology, China
Yin Chenbo, Nanjing Tech University, China
Yu Haoshan, Chinese Academy of Geological Sciences, China
Yuan Ye, Zoomlion Heavy Industry Co. Earthmoving Machinery Company, China
Yue Wenhui, Hunan University of Science and Technology, China
Zeng Qingliang, Shandong Normal University, China
Zhang Hao, Tongji University, China
Zhang Li, Anhui Taiyuan Construction Machinery Co., Ltd., China
Zhang Qing, Tongji University, China
Zhang Yuan, LIBO Heavy Industries Science and Technology Co., Ltd., China
Zhang Daqing, Shanhe Intelligent Equipment Co., Ltd., China
Zhang Guosheng, Tianjin Construction Machinery Research Institute, China
Zhang Junhui, Zhejiang University, China

Zhang Shejun, China Railway Construction Heavy Industry Corporation Limited, China
Zhang Shengjun, Beijing Institute of Construction Mechanization Co., Ltd., China
Zhang Weiguo, Shanghai Maritime University, China
Zhang Xuhui, Xi'an University of Science and Technology, China
Zhang Yanwei, Wuhan University of Technology, China
Zhang Zhonghai, Xugong Foundation Construction Machinery Co., Ltd., China
Zhang Zili, Jiangsu Tianzhou Testing and Inspection Ltd., China
Zhao Bin, Jiangsu Xugong Research Institute, China
Zhao Jingyi, Yanshan University, China
Zhao Keli, Jilin University, China
Zhao Rongyong, Tongji University, China
Zhao Shuen, Chongqing Jiaotong University, China
Zhao Youjun, Xi'an Coal Mining Machinery Co., Ltd., China
Zhao Jianping, Xinjiang University, China
Zhu Qiping, Shijiazhuang Tiedao University, China
Zhu Yutian, Tongji University, China
Zhuang Hongquan, Tsinghua University Press, China

Organizing Committee Chairman

Zhou Xianbiao, China Construction Machinery Society, China

Organizing Committee Co-Chairman

Chen Li, China Construction Machinery Society, China

Secretary

Chen Jianli, China Construction Machinery Society, China

Organizing Committee Members

Bu Wanghui, Tongji University, China
Cai Wei, Yanshan University, China
Chen Jianli, China Construction Machinery Association, China
Chen Maolin, Tongji University, China

Chen Qianpeng, Lanzhou University of Technology, China
Chen Qihuai, Huaqiao University, China
Chen Zhe, Tongji University, China
Cheng Min, Chongqing University, China
Christopher Ehrmann, Tongji University, China
Ding Ruqi, East China Jiaotong University
Fu Changhong, Tongji University, China
He Xiangyu, Changsha University of Science and Technology, China
Hou Zeran, Tongji University, China
Huang Jiahai, Taiyuan University of Technology, China
Lin Tianliang, Huaqiao University, China
Liu Changhui, Tongji University, China
Lu Liang, Tongji University, China
Min Junying, Tongji University, China
Morteza Mohammadzaheri, Birmingham City University, UK
Muhammad Saleem, Imam Abdulrahman Bin Faisal University, Saudi Arabia
Shi Youcheng, Lanzhou University of Technology, China
Sun Bo, Tongji University, China
Sun Yuantao, Tongji University, China
Tang Tang, Tongji University, China
Tang Yichao, Tongji University, China
Valerian Croitorescu, University Politehnica of Bucharest, Romania
Wang Jianwei, Yanshan University, China
Wang Kun, Tongji University, China
Wang Liang, Tongji University, China
Wu Yajing, Tongji University, China
Xiao Yao, Tongji University, China
Xu Dongdong, Tongji University, China
Yang Meng, Tongji University, China
Yu Chunlong, Tongji University, China
Yuan Zhicheng, Tongji University, China
Zhu Cuixiang, Tongji University, China

Preface

The 2024 9th International Conference on Advances in Mechanical and Vehicle Engineering (ICACMVE 2024) has successfully concluded, leaving behind a rich legacy of academic exchange and technological exploration in the fields of mechanical and vehicle engineering.

ICACMVE has a long and distinguished history. Since its inception, the conference has been a premier platform for presenting and discussing the latest advancements in mechanical and vehicle engineering. Over the years, it has witnessed significant developments in the field, from traditional mechanical systems to cutting-edge technologies such as intelligent and unmanned operations, electrification, and green development.

The theme of ICACMVE 2024 is particularly relevant to contemporary research. It focuses on topics such as modern design methods, transmission and drive control, intelligent and green technologies, and cloud applications in construction machinery and vehicle engineering. These topics are not only at the forefront of current research but also crucial for addressing the challenges faced by the industry in the era of digitalization and sustainability.

The conference features a series of keynote speeches delivered by renowned experts in the field. The keynote speeches, delivered by esteemed academics and industry leaders such as Prof. Hehua Zhu from Tongji University, China, Prof. Jürgen Fleischer of German Academy of Engineering, Germany, and many others, provided invaluable insights into the latest research developments and future trends in the field.

In addition to the keynote speeches, the conference also includes numerous oral and poster presentations. Researchers from various institutions shared their work on a wide range of topics, from fundamental theories to practical applications. These presentations facilitate in-depth discussions and collaborations among participants, fostering a vibrant academic atmosphere.

Moreover, the conference includes many papers, which were rigorously reviewed by a panel of experts. The accepted papers cover a wide range of topics, from new designs and structures of construction machinery and vehicles to intelligent and unmanned operations, electrification and greenification, fundamental theories and control methods, inspection and testing methodologies, human-machine interaction

and digital twins, as well as industry-academia-research cooperation models. These contributions reflect the diverse and dynamic nature of the field, highlighting its potential for further growth and innovation.

In a word, ICACMVE 2024 is a successful event that brings together experts, scholars, and industry representatives to discuss and explore the future of mechanical and vehicle engineering. The conference's focus on modern research themes, combined with its comprehensive agenda and strong emphasis on collaboration, makes it a significant contribution to the field. We are confident that the proceedings of this conference will serve as a valuable resource for researchers and practitioners alike, inspiring further innovation and progress in the years to come.

The Committee of ICACMVE 2024

Contents

Research and Application of Frontier Technology in Mechanical Engineering

Research on Object Pose Detection for Mobile Manipulators 3
Shutian Zhang, Ning Chen, Dengqiu Lin, Linhan Lin, Shen Wang, and Qiling Li

Analysis of the Causes of Abnormal Fracture of External Circular Spring Made of 60Si2CrVAT Steel 15
Luolan Zhu, Zhichao Hu, Yi Li, Yaokun Ding, Chengbin Zhao, Naixuan Xu, and Shaocheng Zhan

Dynamic Roll Stability Analysis of Independent Suspension of a New Type of Multi-Axis All-Road Crane 31
Rui Guo, Puxue Niu, Xifeng Zhang, Xin Wang, and Shuai Song

Research on Crane Motion State Monitoring System Based on Digital Twin 43
Xuyang Cao, Baokang Yan, Kai Lin, and Wanlong Shu

Research on Visual Image Processing Algorithm Based on Wind Turbine Hoisting 53
Xuyang Cao, Zhaoqing Zha, Jian Kang, and Yong Su

Analysis and Research on Split Front Wing and Vortex Generator of FSAE Racing Car 63
Chenguang Lai, Zirui Hu, Jie Song, and Lijuan Fu

Thermal Fatigue Life Prediction Under Emergency Braking of Disc Brake Based on FEA Simulation 73
Chuang Chen, Qiang Yang, Huayu Zhang, Xin Zhao, Cheng Gao, and Yuan Gao

Remote Troubleshooting APP Development for Excavators Based on Android System 85
Chao Yang, Chenbo Yin, Shoulei Ma, Donghui Cao, and Xiaozhong Liu

Research on the Influence of Cable Wind Ropes on Wind-Induced Response and Wind Vibration Coefficient of Tower Crane 99
Tong Qin, Yifan Lu, and Wenwu Liu

Fatigue Life Analysis and Improvement of Casting Crane Structures 109
Xin Wang, Cheng Gao, Rongpei Zheng, Tingting Zhang, and YiFan Zuo

Crawler Crane Lifting Path Planning Based on Improved RRT Algorithm 119
Zetao Li, Xiaochun Tie, Xin Wang, and Xuyang Cao

Multi-scale Assembly Feature-Driven Information Modeling for Diesel Engine Assembly Processes 131
Fei Tan, Yangang Zhang, Meiwen Cao, Yu Li, Guoxiang Liang, Shouqi Hu, and Jianjun Chen

Development and Test Study of a New Impact Barrier 143
Jialin Yang and Tengfei Bi

Simulation Research on the Three-Way Catalytic Converter of a Certain Type of Automobile Exhaust System 163
Kun Fang, Wei Ma, Wei Chang, and Pengjie Fan

Comparative Study on Single-Objective and Multi-objective Evolutionary Algorithms in the Parameter Optimization of Excavation Mechanism for Hydraulic Shovel 171
Gongyue Xu, Yuheng Zhou, Zhiwen Zhang, and Wei Hong

Numerical Study of the Characteristics of Spiral Teeth of Steam Turbine Shaft End Combined Seal 185
Zhaojun Meng, Keyu Wang, Shuyu Meng, Da Liu, Shilong Liu, and Lei Zhang

Refined Computation of Lifting Performance Considering Two Trolleys Distance on Shipbuilding Gantry Cranes 195
Yahui Fan, Zhenping Jiang, and Rumin Teng

Smoothness Analysis of Independent Suspension Based on Interconnection Mode for a Five-Axle All-Terrain Crane 209
Yunbo Ma, Ruming Teng, Xifeng Zhang, Shuai Song, and Xin Wang

Research on the Smoothness of the Suspension System of a Five Axis All Ground Crane 219
Rui Guo, Yuang He, Shuai Song, Rumin Teng, and Xifeng Zhang

Optimization of 3D Printing Process for Hydrogel-Based Thermal Management Patches in Engineering Machinery 235
Bowen Li, Zhen Wang, Chuanzhen Huang, Longhua Xu, Shuiquan Huang, Meina Qu, Zhengkai Xu, Dijia Zhang, Baosu Guo, and Chunhui Ji

Digital Twin-Based Structural Performance Prediction Method for Telescopic Boom Cranes 249
Xuyang Cao, Kai Lin, Baokang Yan, and Chengjie Zhang

Research on the Trajectory of Hydraulic Boom in Medium Sized Excavators Based on D-H Parameter Method 259
Tingting Luo, Chengqiang Liu, Xiliang Ma, and Nianning Luo

Weak Signal Detection of Construction Machines Based on Van Der Pol Oscillator 277
Ke-Feng Li and Qing Wang

Generating Adversarial Examples for Pedestrian Detectors in Dynamic Driving Scenarios 289
Zhuang Zhang, Lijun Zhang, Dejian Meng, Wei Tian, Ye Han, Kaikun Pei, and Zhou You

Research on Performance of Traction Winch Hydraulic System in Lifting Condition Based on AMESim 305
Zehao Chen, Junbo Ge, Jingying Xu, Jingqiao Liu, and Rui Guo

A Study on the Active Suspension System of a Wheel-Legged Excavator with Attitude Adjustment 317
Zeguang Hu, Hao Xiong, Dingxuan Zhao, and Haiwu Zheng

Construction and Application of Crawler Crane Digital Twin System 339
Juntao Liu, Bing Li, Yang Zhang, Yulan Wei, Tongxi Zhu, and Yuanqi Hu

Gearbox Fault Diagnosis Method Based on Model Transfer and Multi-sensor Self-learning Weighted Fusion 349
Yu Zhou, Long Chen, and Dongguang Zhang

Research on Automatic Control System of Sprinkler Truck Filling Robot Arm in Mining Area 363
Yandong Ma, Deli Liu, Kun Ma, Meng Zhang, and Zhong Wang

Temperature and Thermal Deformation Analysis of Key Components of Wheel Hub Electric Wheel Under Combined Braking Conditions 377
Kaikun Pei, Lijun Zhang, Dejian Meng, Wei Tian, and Zhuang Zhang

Gap Detection Scheme of Safety Mechanism of Three Gorges Ship Lift Based on Sensing Detection 391
Weibi Tang and Qingyun Liang

Mechanical Engineering and Dynamics Analysis Research

An Integrated Approach for Tolerance Synthesis of a Spiral Bevel Gearbox 407
Siyuan Hu, Yujie Huang, and Wenhui Zeng

Multi-body Dynamics Analysis of All-Terrain Crane Steering Member System 423
Ling Wang, Xin Wang, Shuai Song, Xifeng Zhang, and Chenghao Wang

Estimation of Coefficient of Road Adhesion Based on Stacked Generalisation and Bayesian Optimisation 435
Guan Zhou and Chenxi Gao

Bifurcation and Chaos Properties of Magnetic Gear Two-Speed Transmission System 445
Junliang Du, Dawei Liu, and Tingzhi Ren

Force Analysis of All-Terrain Crane Lowering Arm and Suspension Cylinder 457
Ling Wang, Xin Wang, Shuai Song, and Xifeng Zhang

Fluid–Structure Coupling Analysis of Corner Guide Vane in Large-Scale Low-Speed Wind Tunnel 467
Jiahang Li, Xuefei Wang, Rumin Teng, Linkai Zhou, and Zhenping Jiang

Research on Dynamic Envelope Calculation and Application of Metro Engineering Vehicles 477
Jinmei Xiao, Xuan Cai, Xiaobo Song, and Juntao Wang

Ground Point Cloud Segmentation Method in Open-Pit Mine Scenarios Based on Quadtree Voxel Partition 489
Ruoyao Li, Yafei Wang, Shi Sun, Yichen Zhang, Mingyu Wu, and Zexing Li

Research on Oil Film Pressure Distribution Simulation of Hydrodynamic Lubrication Sliding Bearing 497
Zhengxi Hu, Junbing Liao, Yingmiao Hu, Bin Peng, and Yongqiang Huang

Bucket Drop Point Prediction for Excavators Integrating Point Cloud Information 507
Yong Wang, Xinhui Liu, Changsheng Liu, Yang Li, Qiushi Bi, and Zongwei Yao

Research on Longitudinal Speed Estimation Methods for Multi-wheel Skidding of Multi-axle Distributed Electric Vehicles Based on Wheel Speed Prediction 525
Mingyang Tong, Junqiu Li, Yongxi Yang, Xiaohan Li, and Wuze Wang

Research on Visual SLAM Integrating Feature Method and Direct Method in Sparse Texture Environment 541
Yuxin Cui, Chunpeng Zhang, Tao Ni, Dingxuan Zhao, and Hongyan Zhang

Effect of Typhoon Load on the Attachment Height of a Rigid Attached Tower Crane 549
Xiangxiang Wang, Hongsheng Zhang, and Yue Wang

Dynamics Coupling Model and Stability Analysis of Special Vehicles Under Operating Conditions 559
Xiaoyu Sun, Jiandong Li, and Chengqian Su

Dynamic Analysis of Crawler Traveling Mechanism Steering Performance Analysis 575
Yongran Sui, Xin Wang, Zijun Li, and Junli Pi

Optimization of Formation Topology Based on the Leader–Follower Method 585
Linhan Lin, Ning Chen, Mingxing Fan, Shiyao Cheng, Hanteng Chen, and Yuhan Hou

Multi-body Rigid-Flexible Coupling System Dynamic Analysis of Spent Fuel Crane 597
Xuyang Cao, Wanlong Shu, Chengjie Zhang, and Kai Lin

Dynamics Modeling and Simulation of Dual-Crane Cooperative Lifting System 607
Xuyang Cao, Chengjie Zhang, Wanlong Shu, and Baokang Yan

Study on the Lean Analysis of Low Stretch for Automobile Sheet Metal 617
Jinwu Wang, Binhan Xu, Cancan Wang, and Kangkang Yan

A Ground Segmentation Algorithm for Slopes 627
Zeyu Liu, Dawei Li, and Dawei Jiang

Sparse Feature-Oriented 3D SLAM Mapping with Semantic Fusion 637
Zhangjie Ding, Tao Li, Linyuan Xie, and Yongming Bian

Research on the Dynamics Model of Specialized Heavy-Duty Multi-axle Electric Drive Vehicles Based on SIMPACK 649
Weize Wang, Xuelei Wu, and Xia Zuo

Research on Monocular Vision Indoor Construction Positioning Method Based on YOLOv5 663
Hang Yang, Bo Sun, Weize Lin, Taozheng Wu, Gangju Wang, and Guangjun Liu

Ergonomic Simulation Analysis of Forklift Drivers Based on JACK 677
Pingkai Wang, Dingxuan Zhao, Ruixuan Zhu, and Zeren Wang

Modeling and Analysis of Mechanical-Braking Thermomechanical Coupling for Direct-Drive Hub-Motor 693
Kaikun Pei, Lijun Zhang, Dejian Meng, Wei Tian, and Zhuang Zhang

Dynamics Modeling and Analysis of a Single Flexible Link-Flexible Joint Manipulator 705
Jianyuan Wang and Shanying Jiang

Design and Structural Optimisation of Vehicle and Its Equipment

ANSYS Contact Simulation Analysis Between Shed Tunnel Structure and Ultra-High Voltage Wire 717
Yue Wang, Xiangxiang Wang, and Hongsheng Zhang

Research on the Aerodynamic Performance of a Suspended Monorail Train Designed like a Whale 729
Wenhui Ma, Yazhuo Zou, Jiachun Zou, and Jingyi Wang

Experimental Study on Electro-Hydraulic Cooperative Braking by Wire Safety Performance of Mining Electric Vehicles 741
Zhi-jun Hao

Parameter Optimization Design of Air Springs for Subway Vehicles 753
Xianhui Su

Numerical Simulation Analysis of Vehicle Equipment Safety Fastening System Design Scheme 763
Jingqiao Liu, Xin Jiang, Guo Wang, Mengxiang Li, Kun Liu, and Shiye Liu

Mechanical Structure Design of a Continuously Rotating Flight Test Device 783
Shaozun Hong, Deng He, Daiyong Mao, and Xiaodong Zheng

Airworthiness Compliance Verification and Analysis of Turbofan Engine Bird Ingestion 795
Qing Zhang, Xigang Shen, Kun Niu, and Haiyang Zhang

Aerodynamic Optimization Design of Propellers for Multi-rotor Unmanned Aerial Vehicles Based on NSGA-II 815
Ming Gao, Wei Guo, Jiaqiang Sun, Yunlong Huang, and Yong Cai

A Quantitative Energy Efficiency Index for Heat Transfer and Energy Conservation of Plate Heat Exchanger in Commercial Vehicle 827
Haidong Wang, Yunjia Liang, Zhiao Wang, Jiaxin Liu, Lingqin Meng, and Weimin Liu

The Simulation Researching for Some Vehicle Based on Dragging to Tray Carriage Load Case 839
He Li and Yongfu Xiao

Research on Lightweight Design Methods of Cavity Structure 855
Ning Wang, Hongjun He, Baojun Yu, Yongfu Xiao, and Liwei Sun

U-ATTENTION-NET: A U-shaped Structure for No-Reference Low-Light Image Enhancement Method 867
Zinong Zhang and Lei Han

Research on Finite Element Modeling Method for Structural Strength Analysis of a C0-C0 Type Bogie Frame 879
Yuwen Huang, Xiaojuan Sun, Shijie Liu, Xinren Li, and Yu Wang

Medium-Sized Truck Brake Design 889
Wang Tiantian and Qichen Zhong

Power System Design and Performance Analysis of an M100 Methanol Range Extended Electric Dump Truck 901
Xiqing Zhang, Yanbo Dai, Yafeng Ren, Yaojie Wang, and Shiwei Zhao

Grayscale Image Classification Method for Auto Parts Based on MobileNet_V3_Small 917
Yujie Ye, Zhichao Hu, Long Diao, and Yunchao Wang

Response Characteristics of a Fourteen-Degree-Freedom Vehicle Under Four-Wheel Random Pavement Excitation 925
Lei Zou, Zhichao Hu, Haoyu Sun, and Yunchao Wang

Design, Modeling and Vibration Analysis of Piezoelectric Energy Harvesting Vehicle Suspension 933
Zewang Wu, Yunchao Wang, and Qi Hong

Structural Optimization Design of Cutter Plate of Grapevine Burying Machine Based on Response Surface Methodology 943
Beibei Zhang, Julaiti Maitirouzi, Hongcheng Deng, Haileiqguli Yasen, and Wenjing Bai

Multi-Element Coupling Model of Rail Grinding Vehicle 965
Zixuan Fan, Junjun Ding, Changquan Ding, Jianquan Zhang, Boxin Zhao, and Xiaoqian He

Research on the Influence of Rack Railway Irregularity on the Rack Railway Engineering Vehicle 977
Han Lv, Junjun Ding, Changquan Ding, Jianquan Zhang, Haoxiang Weng, and Yang Qin

Path Planning System Design for Autonomous Haulage in Semi-Structured Environments 987
Hang Xu, Dawei Li, and Dawei Jiang

Digital Twin Based Hydraulic System Design 1003
Yapeng Wen, Jiaqi Zhang, Yunzhong Deng, Rui Guo, and Yunfeng Gu

Research on Simulation and Optimization of Heat Sink for Motor Controller of Electric Forklift Truck 1015
Jiyuan Mao, Zhenyuan Zha, Hao Chen, Xun Hao, Yuntong Wang, and Huan Ma

Li-Ion Battery Life Prediction Based on RF-SVM 1027
Weijian Huang and Jixiang Zhou

Finite Element Analysis of Baja Racer Frame with Space Truss Structure 1039
Qiang Ding, Jiatong Qin, and Biao Jin

Influence of Hydro-pneumatic Suspension Structural Parameters on Off-Road Vehicle Ride Comfort 1049
Ziqi Huang, Liqiang Zhao, Hao Xiong, Liangpeng Li, and Dingxuan Zhao

Improved Particle Swarm Optimization for Rapid Tuning of Vehicle Suspension Damping 1069
Hao Xiong, Zeguang Hu, Ziqi Huang, Liqiang Zhao, Haiwu Zheng, and Dingxuan Zhao

Mobile Carrier Platform Design 1089
Hao Wang, Youqing Li, Rui Guo, Zebing Fan, Dianmin Chen, Pengfei Chen, Chenglong Yu, and Jingyi Zhao

Research on Vehicle Center of Mass Measurement Method Based on Active Suspension 1101
Jinming Chang, Dingxuan Zhao, Haoyu Jiang, Yunwang Ma, and Shiling Li

Design and Dynamic Performance Analysis of Rail Grinding on Metro Lines with Small Radius Curves 1115
Haifu Xing

Advanced Vehicle and Robot Control and Optimisation Technologies

Modal Analysis and Experimental Study of Wheeled Special Vehicle Body 1127
Wenshuo Li, Mingming Dong, and Weiwei Lv

Optimization of Burner Temperature Control Based on Genetic Algorithm 1137
Zhiming Zhang, Wengang Yan, Haiying Cheng, Tiejian Wang, and Fei Liu

Research on Anti-rocking Control of Wind Turbine Crane Based on Input Shaping Technology 1151
Xuyang Cao, Fengzhan Xu, Xiaochun Tie, and Dan Wang

Active Suspension Anti-rollover Control Strategy for Wind Turbine Blade Mountain Transport Vehicles 1161
Yiying Wang, Zhong Wang, Yuchen Wang, Mingyang Wang, and Yuhao Jiang

Automatic Calibration and Variable Scale Segmentation Algorithm of Driveless Dumping Truck Based on Mining Area 1171
Boxua Zhang, Zeyu Liu, Chenhui Li, DaWei Li, and Dawei Jiang

Effect of MHD Control on Turbulent Boundary Layer Separation Flow in Compression Corner 1183
Zhaoyu Wu, Mingsong Ding, and Weizhong Dong

Research on the Characteristics of Electro-Hydraulic System of Crane Pressure Flow Composite Control 1191
Pengfei Yang, Luhong Yue, and Jing Yang

Dynamic Path Planning Algorithm for Unmanned Inspection Vehicle Based on Improved DWA 1207
Zhiguo Zhao, Siyu Wang, Wen Xiao, Rong Zhou, Zhen Xu, Dong Xie, Xiaokang Wan, and Yeqin Wang

Simulation Analysis of Lifting Equipment for Underwater Robot 1219
Xin Wang, Xinyu Li, Tianchuan Pan, Hao Li, Sisi Zhu, and Zetao Li

Analysis of Steering Control of a Five-Axle All-Terrain Crane Considering Body Rolling 1227
Tingting Zhang, Xin Wang, Xifeng Zhang, and Shuai Song

Brief Introduction of Hybrid Electric Vehicle Research 1237
Shengqi Yao, Xiaoshan Yao, Yilu Liu, Chenfei Zheng, Weiqiang Zhan, Jie Ma, Ziliang Zhu, and Guang Yang

Research on Injury to Hybrid III 50th Percentile Dummy Based on FMVSS 208 Test 1249
Qian Zhang and Youjun Wang

Upper Limb Rehabilitation Robot Based on Planar Five-Bar Structure 1259
Mingjie Cui, Huixiang Yang, Dawei Jiang, Jianwei Sun, and Tiecheng Ji

Vehicle Trajectory Tracking and Optimization Based on Lateral and Longitudinal Control 1275
Jie Tang, Zhiwen Zhang, and Jiyuan Zhang

Full-Wheel Active Steering Control Method for Heavy-Duty Vehicles Based on GA-LQR 1283
Jining Fan, Jue Gong, Jian Zhao, Pengbo Liu, Jihong Zhou, and Xin Wang

Convolutional Neural Network-Based Autonomous Perception and Intelligent Decision System for Cabin Cleaning Machines 1297
Xinguo Yu, Yan Huang, Chunyang Liu, and Siyuan Zhou

Research on Mobile Grasping and Path Planning of Robotic Arms Based on Deep Learning 1315
Zhanhui Wang, Zhifei Ji, Ning Chen, Xiaochuan Lin, Kangyi Wu, and Xinrui Ma

Research on the Key Technology of Multi-sensor Fusion Environment Perception Based on Unmanned Patrol Vehicle 1329
Min Lin, Dongsheng Lin, Pengxin Yue, Wenyang Chi, Xinrui Ma, and Yingbing Huang

Research of Ergonomic Test on Human Whole-Body Vibration of Armored Vehicle Cabin 1345
Tushun Huang, Meng Cai, Wei Chang, and Weiwei Zhang

Workspace of Parallel-Legged Hexapod Walking Robot 1359
Xinxing Tang, Shuwei Bai, Huiqiu Mu, Dachuan Fan, and Fangyuan Han

Trajectory Tracking Control of Wind Turbine Blade Mountain Transporter Based on MPC with Disturbance Observer 1367
Yahui Zhang, Yuhao Jiang, Xiaohong Jiao, Ting Zhang, Chengye Wu, Mingyang Wang, Yiying Wang, and Yuchen Wang

Adaptive Control for Vehicle Active Suspension with Limited Working Space Based on Tire Stress Estimation 1377
Zhenjun Lin, Shuang Liu, Mingxuan Li, Dingxuan Zhao, Nayu Su, and Cong Zhang

A Study on Automobile Wind Resistance Prediction of E-vehicle Based on AOA-CatBoost Hybrid Modeling 1391
Chuanbo Liu, Yuhan Liu, Guangyuan Yang, and Zhu Bin

Yaw Rate Synchronisation Steering Control of Multi-axle Heavy Vehicle 1401
Mingyuan Han, Zhong Jin, Mingde Gong, Jingming Chang, Wei Zhang, and Dingxuan Zhao

A Self-adaptive PID Control Algorithm for Electromagnetic Bearings Based on RBF Adam 1417
Haiwu Zheng, Yinying Ren, Hao Xiong, Zeguang Hu, Dingxuan Zhao, and Ziqi Huang

Research on Active Suspension Preview Control of Vehicles Based on Real-Time Road Elevation Information 1435
Dingxuan Zhao, Liqiang Wang, Lirong Yang, Jinming Chang, Haoyu Jiang, Yang Li, and Rui Cheng

Defending Against Adversarial Examples with Adaptive Quantization Table-Guided Image Compression 1447
Zhuang Zhang, Lijun Zhang, Dejian Meng, Wei Tian, Ye Han, and Kaikun Pei

Hydraulic Characteristics Analysis of Working Device for Typical Working Conditions of Mini Excavators with Electric Drive Excavation 1463
Yifan Zhang, Li Chen, Zhou You, and Yongming Bian

Integral Backstepping Displacement Tracking Control for a Hydraulic Active Suspension Actuator 1473
Lizhe Wu, Dingxuan Zhao, Tao Ni, Tuo Jia, and Qinghe Guo

Impact of Friction Block Shape and Arrangement on High-Speed EMU Brake Noise Studied Using Complex Modal Analysis 1485
Bo Yu, Xingchao Li, and Han Zhou

Research and Application of Frontier Technology in Mechanical Engineering

Research on Object Pose Detection for Mobile Manipulators



Shutian Zhang, Ning Chen, Dengqiu Lin, Linhan Lin, Shen Wang, and Qiling Li

Abstract This study proposes an improved YOLO-GQ model and an optimized A* algorithm, which respectively enhance the efficiency and accuracy of object detection and trajectory planning for mobile grasping robots. By integrating autonomous navigation, detection, and trajectory planning algorithms, a physical prototype was successfully developed, and the grasping experiments were validated. The main contributions are as follows: (1) The YOLOv5 model was improved by incorporating an attention mechanism, MBConv module, and angle loss function. (2) The A* algorithm was optimized using a dynamic weight heuristic function, bidirectional search, and Bezier curve path smoothing, which improved the planning speed and path smoothness. (3) A physical prototype integrating a mobile chassis and robotic arm was built, verifying the feasibility of the pose detection and grasping algorithms, meeting the requirements for mobile grasping tasks.

Keywords Object detection · Path planning · Indoor localization · AGV

1 Introduction

1.1 A Subsection Sample

As urbanization and population aging intensify, traditional labor-intensive industries are facing recruitment challenges. Mobile manipulators, as a form of intelligent automation technology, have broad applications in reducing labor costs and alleviating labor intensity, especially in fields such as warehousing, logistics, and healthcare [1]. By combining the strengths of mobile robots and robotic arms, they can

S. Zhang · N. Chen (✉) · D. Lin · L. Lin · S. Wang · Q. Li
School of Marine Equipment and Mechanical Engineering, Jimei University, Xiamen, Fujian, China
e-mail: cn1972@jmu.edu.cn

© The Author(s) 2026

H. Saman et al. (eds.), *The 9th International Conference on Advances in Construction Machinery and Vehicle Engineering*, Lecture Notes in Mechanical Engineering,
https://doi.org/10.1007/978-981-96-5527-4_1

perform object grasping in complex environments, promoting the development of smart manufacturing and automation [2].

This paper explores the applications of mobile grasping robots in complex environments and proposes an optimized detection algorithm under low computational cost to improve target recognition and grasping efficiency. The effectiveness of the algorithm was validated, contributing to the wider adoption of mobile manipulators.

Significant progress has been made both domestically and internationally in pose estimation, trajectory planning, and SLAM (Simultaneous Localization and Mapping) [3] technology. Pose estimation, combined with deep learning, enhances the diversity of object recognition. Trajectory planning increasingly integrates Cartesian and joint space planning, with the introduction of artificial intelligence. SLAM technology has evolved from extended Kalman filtering to graph-based optimization, achieving sensor fusion and improving localization accuracy.

However, challenges such as dynamic obstacle handling, multimodal task coordination, and dynamic optimization remain. Future research should integrate machine learning and artificial intelligence to enhance the perception and adaptability of mobile manipulators, driving their intelligent and adaptive development in complex environments.

2 Improved Pose Detection Based on the YOLOv5 Network Model

Mobile grasping robots must accurately detect targets and ensure real-time performance in complex scenarios while providing planar pose information for robotic arm grasping. To achieve this, this chapter proposes the improved YOLO-GQ network, which reduces hardware requirements, enhances detection accuracy, and incorporates a rotational target detection mechanism to identify the planar pose of objects for robotic arm manipulation.

2.1 Introduction to the YOLOv5 Network

YOLOv5 was open-sourced in 2020 and includes four models differing in depth and width. To address issues such as insufficient frame rate and missed detections in mobile grasping robots, this study proposes an improved YOLOv5 model, YOLO-GQ, further optimizing its performance [4].

2.2 Proposed YOLO-GQ Network Structure

CBAM Attention Mechanism. CBAM (Convolutional Block Attention Module) is an attention mechanism that enhances the perception capabilities of convolutional neural networks [5]. In this study, it is integrated into the backbone and feature fusion layers of YOLOv5, utilizing channel attention and spatial attention to capture the correlations of key features, thereby improving the model's representation ability and robustness [6].

MBConv Module. The MBConv module, derived from MobileNetV2, improves efficiency through lightweight convolutions while maintaining accuracy [7]. In YOLO-GQ, MBConv replaces part of the CBS modules, forming the C3MB module. It utilizes channel expansion, depthwise separable convolutions, and linear fusion to reduce computational load and the number of parameters while preserving feature extraction capability, thereby enhancing multi-scale detection ability.

YOLOv5 Object Detection Network with Rotating Bounding Boxes. To improve accuracy, angle prediction is transformed into a classification problem, and Circular Smooth Labeling (CSL) is adopted to address discontinuity issues. The improved YOLO-GQ model, combining rotating bounding boxes with an attention mechanism, significantly enhances detection accuracy and robustness in complex outdoor environments.

Intersection Over Union (IOU) is used to measure the overlap between the detection box and the actual target, with higher values indicating more accurate detection, as shown in Eq. (1).

$$IOU = \frac{Overlap\ Between\ Detectedand\ Annotated\ Regions}{Total\ Area\ of\ Both\ \text{Regions}} \tag{1}$$

Therefore, the angle prediction approach can be modified into a classification problem. Circular Smooth Labeling (CSL) is designed to increase error tolerance between adjacent angles. The calculation formula for CSL is shown in Eq. (2).

$$CSL(x) = \begin{cases} g(x), \theta - r < x < \theta + r \\ 0, otherwise \end{cases} \tag{2}$$

In this context, g(x) represents a window function controlled by the radius r. Typically, the window function can be a pulse function, rectangular function, triangular function, or Gaussian function.

The confidence weight is calculated using the IOU value of the rotating box to enhance accuracy. In the bounding box loss calculation, the losses for the bounding box and angle are computed independently, using CIOU as the metric to optimizedetection performance. The structure of the improved network model, YOLO-GQ, is shown in Fig. 1.

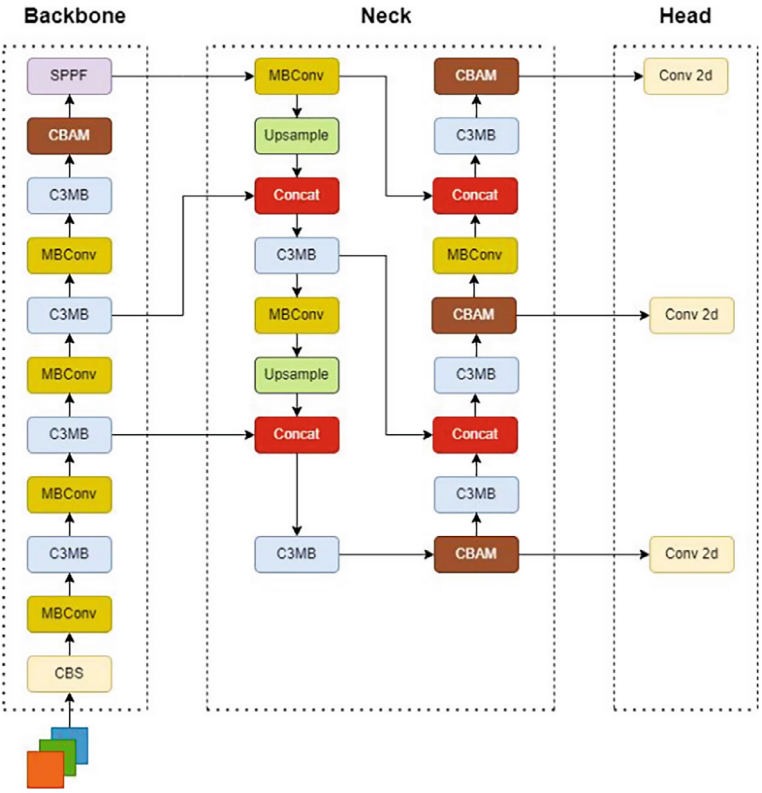


Fig. 1 Improved network architecture diagram

3 Autonomous Localization and Path Planning Algorithm Based on Improved a*

3.1 Global Path Planning

The A* algorithm is a path planning algorithm based on the evaluation function f(n) = g(n) + h(n), where g(n) represents the actual cost and h(n) is the heuristic function, estimating the cost from the current node to the target node. A reasonable heuristic function can improve search efficiency, with commonly used heuristics including Manhattan distance, diagonal distance, and Euclidean distance.

Improvement of the Heuristic Function. To improve path search efficiency, this paper replaces the Manhattan distance in the traditional A* algorithm with diagonal distance, which is more suitable for the omnidirectional movement of AGVs. The weight of the heuristic function is dynamically adjusted based on the distance between the AGV and the target point. When the AGV is far from the target, the weight of the

heuristic function is increased to accelerate the search; as it approaches the target, the weight is reduced, giving more importance to the actual cost to find a more optimal path. The improved heuristic function is shown in Eq. (3).

$$h_{improved} = wd(c) \tag{3}$$

$$w = [1 + \frac{d(c)}{d(e)}] \tag{4}$$

where w is as shown in Eq. (4), d(c) represents the diagonal distance from the current node to the target node, and d(e) represents the diagonal distance from the start node to the target node.

Bidirectional Alternating Search. As To reduce unnecessary node searches, this paper adopts a bidirectional alternating search method [], where searches are conducted simultaneously from both the start and end points, and the distance between the search nodes is calculated. This method narrows the search range by searching from both ends, improving search efficiency while ensuring the accuracy of path planning.

Path Smoothing. To make the planned path smoother and more continuous, this paper uses Bezier curves for path smoothing. Bezier curves generate smooth paths through a set of control points, ensuring that the AGV follows the planned path smoothly. An n-th order Bezier curve is generated using n + 1 control points, as shown in Eq. (5).

$$p_{(u)} = \sum_{i}^{n} P_i B_{i,n}(u), u \in [0, 1] \tag{5}$$

In conclusion, these improvements enhance the efficiency of the A* algorithm in path planning and increase the feasibility of the planned path, making it more suitable for the practical application requirements of AGVs.

3.2 Local Path Planning

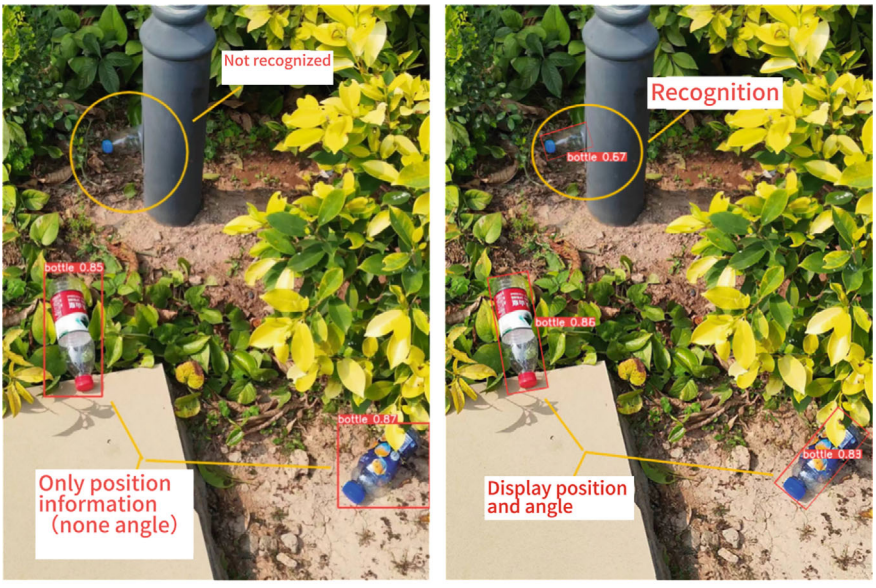
Local path planning for AGVs (Automated Guided Vehicles) is crucial to ensuring that they reach their target points without collisions, taking into account speed, kinematic constraints, and vehicle dimensions to ensure safe obstacle avoidance [8]. The core of the TEB (Timed Elastic Band) algorithm [9] is to optimize both the path and time simultaneously. The main constraints of the algorithm include: Distance constraints between path points and obstacles, Velocity and acceleration limits, Shortest-time path constraints and Differential drive robot constraints.

4 Experiment Result

4.1 Rotating Bounding Box Detection Experiment

To validate the effectiveness of the improved network model, this section captures animage with a complex outdoor background to verify the algorithm's performance, as shown in Fig. 2.

From the figure, it can be observed that even in an outdoor environment with ample lighting and occlusions, YOLO-GQ is still able to detect objects with high confidence and accuracy. In Fig. 2a, the water bottle partially obscured by the streetlight was not detected, and although the two plastic bottles below were accurately identified, only their positions were recognized, without pose information. In Fig. 2b, the improved YOLO-GQ successfully detected the plastic bottle obstructed by the streetlight with a confidence score of 0.67. Additionally, the poses of all plastic bottles in the image were correctly identified, proving the effectiveness of the improvements made to the YOLOv5 algorithm.



(a) (b)

Fig. 2 Model detection visualization

4.2 Pose Recognition Accuracy Evaluation Experiment

In this section, two types of objects (plastic bottles and plastic bags) were selected, with 10 images each taken in different environments and positions, for a total of 20 images used for recognition. The results show that all objects were accurately predicted. To evaluate positional accuracy, the error between manual annotations and network recognition results was compared, and Table 1 lists the coordinate values of both.

Table 1 shows the center point positions and rotation angles of plastic bottles obtained through manual annotation and network recognition. The center point coordinates are in pixels, and the rotation angles are in degrees. The center point is calculated as the average of the four corner points, with pixel values rounded to two decimal places to ensure accuracy in error calculation. The angles are also rounded to two decimal places for consistency. Table 2 presents the position and angle information for plastic bags.

The position error calculation for a single image is shown in Eq. (6).

$$e_{xy} = \sqrt{\left(c_x - c_x'\right)^2 + \left(c_y - c_y'\right)^2} \tag{6}$$

The angle error calculation for a single image is shown in Eq. (7).

$$e_\theta = \left|\theta' - \theta\right| \tag{7}$$

The average values of $\overline{e_{xy}}$ and $\overline{\theta}$ calculated from the 20 images are $\overline{e_{xy}} = 1.14px$ and $\overline{\theta} = 0.48°$.

Table 1 Plastic bottle position comparison

	Network identification			Manual annotation		
Bottle	$C_x(px)$	$C_y(px)$	$\theta(°)$	$C_x'(px)$	$C_y'(px)$	$\theta'(°)$
	472.5	145	−1.14	471.5	144	−1.29
	366.5	266.45	30.96	365	266.25	31.03
	878.5	662	−5.74	877.5	662.45	−5.54
	1081.25	779.5	16.82	1080	778.75	16.05
	810.2	345.5	8.07	811	345	8.12
	520	784	−9.07	519	784.35	−9.17
	476.45	986.75	1.32	475.5	986.25	1.52
	761.25	456	31.93	760.5	456.5	32.10
	177.35	345	−26.10	176.5	344	−26.60
	835.5	876.75	34.13	835	876.25	33.15

Table 2 Plastic bag position comparison

	Network identification			Manual annotation		
Bag	$C_x(px)$	$C_y(px)$	$\theta(^\circ)$	$C'_x(px)$	$C'_y(px)$	$\theta'(^\circ)$
	485	329.75	-4.98	485	329.75	-5.68
	1232.25	697.25	-5.52	1231.75	698.25	-5.92
	935.5	387.75	-3.59	936.75	386.75	-3.16
	531.75	327.75	1.78	531.5	327.25	1.75
	600.75	243.25	36.61	601.25	242	37.17
	799.25	714	28.02	799.5	714	28.07
	770.75	518.25	19.74	768.75	518.5	20.53
	686.25	716.5	28.18	688	716.25	-28.17
	479.5	614.5	3.80	477.75	615.25	3.10
	1103.5	430	19.91	1102	431.75	-19.06

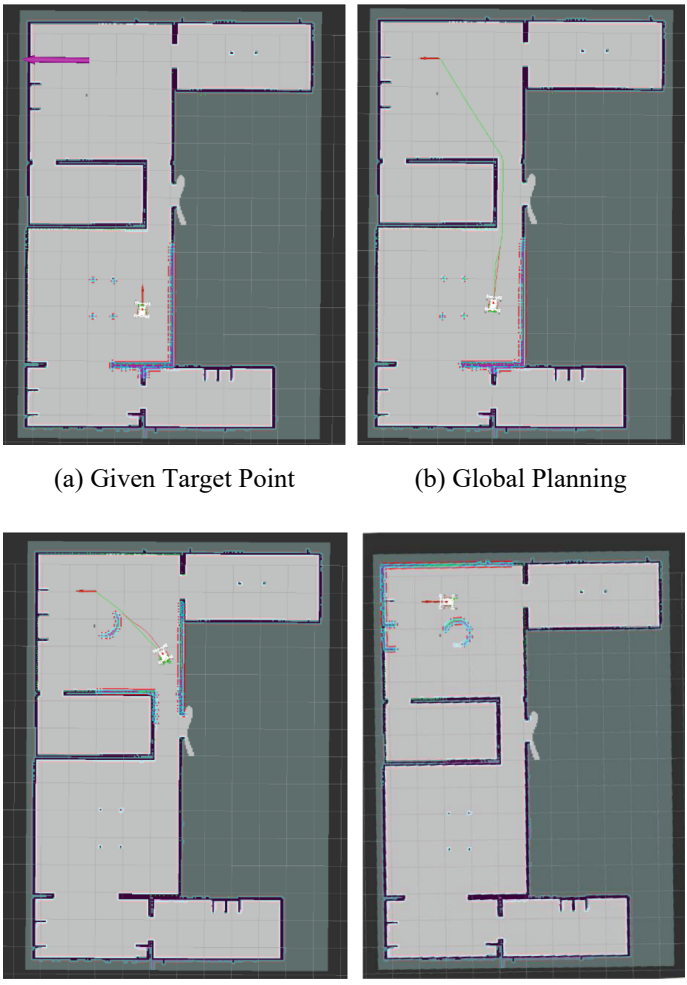
This indicates that the pixel error of the center point is around 1 pixel, and the angle error is within 1 degree, demonstrating that the proposed YOLO-GQ model can accurately obtain the pixel position and angle information of objects in different scenarios.

4.3 Simulation of Autonomous Localization and Path Planning Algorithm

After determining the global path planning algorithm and the local path planning algorithm, the effectiveness of the algorithms can be verified through a simulation platform. The simulation platform used is also the Gazebo simulation software. In this section, simulation tests will be conducted for the improved A* global path planning and TEB local path planning [10]. The simulation process is shown in Fig. 3.

The AGV first performed global path planning using the improved A* algorithm, and the generated green route initially passed through the obstacle. Subsequently, the TEB algorithm was used for local path planning, with the red route showing the AGV successfully avoiding the obstacle and reaching the target point. The green arrows indicate the AGV's pose probability distribution, while the purple and indigo areas represent the obstacle's inflated regions (with a radius of 0.1 m). The experiment demonstrated the AGV's ability to adjust its path in real-time in a dynamic environment, emphasizing the importance of updating map information and supporting its application in complex environments.

The improved A* algorithm in this paper may contact the corners of the wall after smoothing in some special environments, and the expansion of obstacles or dynamic processing for special terrain can be considered to avoid this situation.



(a) Given Target Point (b) Global Planning

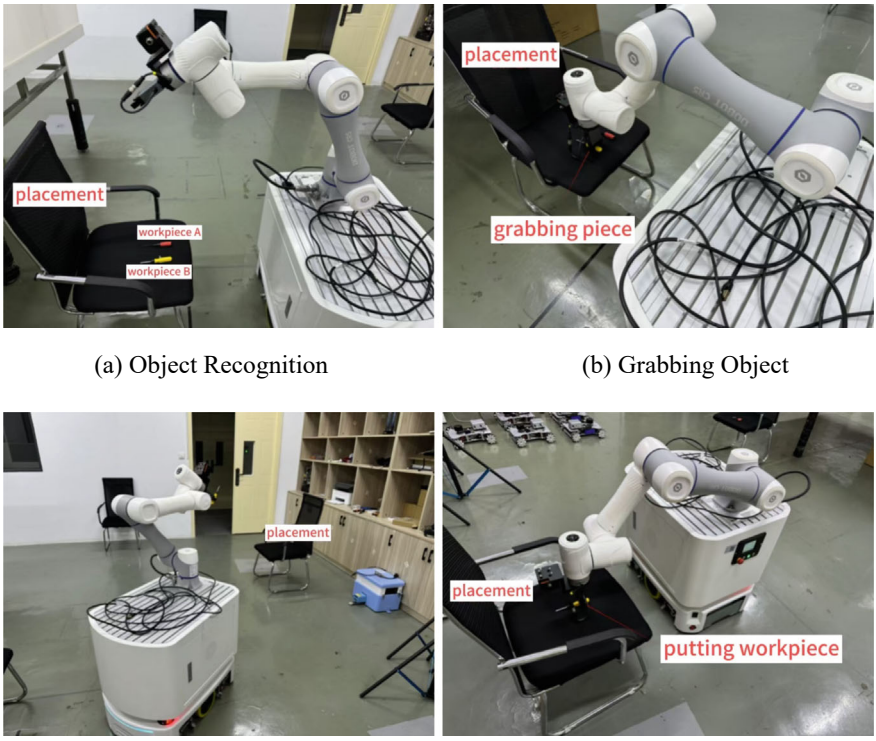
(c) Navigation Process (d) Navigation Target Point

Fig. 3 Navigation process Rviz visualization screenshot

4.4 Simulation of Autonomous Localization and Path Planning Algorithm

TIn the experiment, the mobile robotic arm is used for object transportation, as shown in Fig. 4.

First, the mobile base stays in a standby mode. When the robotic arm receives a visual command, it proceeds to grab the object. After successfully grasping the object, the arm moves to a safe position to ensure no collision occurs with the surroundings



(a) Object Recognition (b) Grabbing Object

(c) Navigation to Target (d) Object Placement

Fig. 4 Grasping and placing process

during the movement of the base. Once the base reaches the target location, the arm places the object, completing the transportation task.

The mobile robot successfully completed the task, and the experiment validated the feasibility of the grasping and movement process of the mobile manipulator, meeting the design requirements. In the process of grasping, the position of the object is deviated, and the gripper will not be able to complete the grasping process, and the accuracy of object pose recognition should be improved in the future.

5 Conclusion

This study addresses the challenges faced by mobile grasping robots in outdoor environments and proposes an improved YOLO-GQ model that enhances detection accuracy and processing speed by incorporating an attention mechanism and the MBConv module. The optimized A* algorithm, utilizing dynamic weighting and bidirectional search strategies, significantly improves path planning efficiency.

The final prototype of the mobile grasping robot successfully integrates autonomous navigation, object detection, and trajectory planning functions. Experimental validation confirms its effectiveness in real-world scenarios, laying a foundation for the development of mobile robot technology.

Acknowledgements The authors gratefully acknowledge the support of the Natural Science Foundation of Fujian Province, China (No. 2023J01792), Major Project of Science and Technology of Quanzhou (No. 2022GZ8), Xiamen Science and Technology Subsidy Project (No. 2023CXY0316), and the Young and Middle-aged Teacher Education Scientific Program of Fujian Province (Grant No. JAT220180), Fujian Provincial Department ofScience and Technology-University-industry cooperation project (No. 2024H6012).

References

1. Kotschy R, Bloom DE (2023) Population aging and economic growth: from demographic dividend to demographic drag. SSRN Electron J
2. Sun Y et al (2020) The transformation and upgrade of China's manufacturing industry in Industry 4.0 era. Syst Res Behav Sci 37
3. Kohlbrecher St et al (2011) A flexible and scalable SLAM system with full 3D motion estimation. In: IEEE international symposium on safety. IEEE
4. Sock J, Garcia-Hernando G, Kim TK (2020) Active 6D multi-object pose estimation in cluttered scenarios with deep reinforcement learning. In: IEEE/RSJ international conference on intelligent robots and Systems (IROS)
5. Lin TY et al (2017) feature pyramid networks for object detection. In: IEEE computer society
6. Woo, Sanghyun , et al.CBAM: Convolutional Block Attention Module.Springer, Chan, (2018)
7. Liu S et al (2018) Path aggregation network for instance segmentation. IEEE
8. Tan M, Le QV (2021) EfficientNetV2: smaller models and faster training
9. Fuentes-Pacheco J, Ruiz-Ascencio J, Rendón-Mancha JM (2015) Visual simultaneous localization and mapping: a survey. Artif Intell Rev 43(1)
10. Kummerle R et al (2011) G2o: A general framework for graph optimization. In: IEEE international conference on robotics & automation. IEEE

Open Access This chapter is licensed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License (http://creativecommons.org/licenses/by-nc-nd/4.0/), which permits any noncommercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if you modified the licensed material. You do not have permission under this license to share adapted material derived from this chapter or parts of it.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.



Analysis of the Causes of Abnormal Fracture of External Circular Spring Made of 60Si2CrVAT Steel



Luolan Zhu, Zhichao Hu, Yi Li, Yaokun Ding, Chengbin Zhao, Naixuan Xu, and Shaocheng Zhan

Abstract Multiple instances of abnormal fracture of 60Si2CrVAT steel outer circular springs have occurred in actual conditions of a certain open-top truck. To identify the causes of the fractures, we conducted tests on the fractured samples, including chemical composition analysis of the raw material, non-metallic inclusions and low-magnification structure inspection, geometric dimension checks, flaw detection and appearance inspection, hardness testing, optical metallographic examination, scanning electron microscopy observations of the fracture surface, and energy dispersive spectroscopy analysis. The results indicate that the raw materials and geometric dimensions of the fractured samples of 60Si2CrVAT spring steel meet industry standards and technical requirements. There are two main reasons for the abnormal fracture of the springs: firstly, during the grinding process, a newly introduced water-cooled wet grinding process was used instead of the dry grinding method due to environmental considerations, which effectively resulted in secondary quenching during grinding, causing thermal burn and creating a hardened layer of quenched martensite on the grinding surface of the support ring. This led to significant surface residual stress, initiating and propagating cracks at this location. Secondly, insufficient tempering of the springs resulted in an abnormal tempered sorbite matrix with a small amount of retained martensite, leading to excessive hardness, reduced toughness, and increased brittleness, which accelerated crack propagation. When the cracks reached a certain depth, the cross-sectional area of the spring could no longer withstand external forces, leading to rapid propagation and eventual fracture of the spring.

L. Zhu
College of Science, Jimei University, Xiamen, Fujian, China

Xiang'an Vocational and Technical School, Xiamen, Fujian, China

Z. Hu (✉) · Y. Li · Y. Ding · C. Zhao
College of Marine Equipment and Mechanical Engineering, Jimei University, Xiamen, Fujian, China
e-mail: hzc0204@163.com

N. Xu · S. Zhan
College of Marine Engineering, Jimei University, Xiamen, Fujian, China

© The Author(s) 2026

H. Saman et al. (eds.), *The 9th International Conference on Advances in Construction Machinery and Vehicle Engineering*, Lecture Notes in Mechanical Engineering,
https://doi.org/10.1007/978-981-96-5527-4_2

Keywords 60Si2CrVAT steel · Outer circular springs · Abnormal fracture · Secondary quenching · Tempering

1 Introduction

60Si2CrVAT steel is the most commonly used high-performance spring steel in the railway system, known for its high strength, high toughness, and excellent overall mechanical properties. It is highly susceptible to fatigue failure, necessitating stringent quality requirements regarding chemical composition, mechanical properties, grain size, and non-metallic inclusions. Additionally, the surface of the spring wire must be smooth and defect-free, with a product fatigue life not less than 3 million cycles [1, 2].

Springs are the main load-bearing components in open-top cars during operation. The two common failure modes of springs are fracture failure and elastic failure (such as relaxation or deformation), with fracture failure being the most prevalent and hazardous form of spring failure [3]. Improving spring quality can effectively extend their service life, prevent premature fracture failures, ensure the operational performance and quality requirements of open-top cars, and reduce the likelihood of major accidents [4].

This paper takes the case of multiple abnormal fracture failures of 60Si2CrVAT steel outer circular springs in a certain open-top car as a research object. Utilizing equipment such as the Shimadzu PDA5500II direct reading spectrometer, DMI5000M Leica metallographic microscope, and scanning electron microscope, we conducted tests on the fractured samples, including chemical composition analysis of the raw materials, examination of non-metallic inclusions and low-magnification structure, geometric dimension checks, flaw detection and appearance inspection, hardness testing, optical metallographic examination, and scanning electron microscopy observations of the fracture surfaces, along with energy dispersive spectroscopy analysis, to analyze the causes of the abnormal fractures.

2 Physical and Chemical Inspection

After a period of operation on the railway, multiple incidents of abnormal fracture failures occurred with the cushion springs of a certain type of open-top car. These cars utilized the mature and standardized K6 type bogie, and all fractured springs were outer circular springs. According to statistical data, over 90% of the broken springs were produced in January, February and June of a certain year. The raw material, 60Si2CrVAT steel, was specifically designed and customized for high-performance spiral springs used in bogies such as the K2 and K6 models, in accordance with the speed-up requirements set by the Ministry of Railways.



(a) (b)

Fig. 1 Actual broken springs

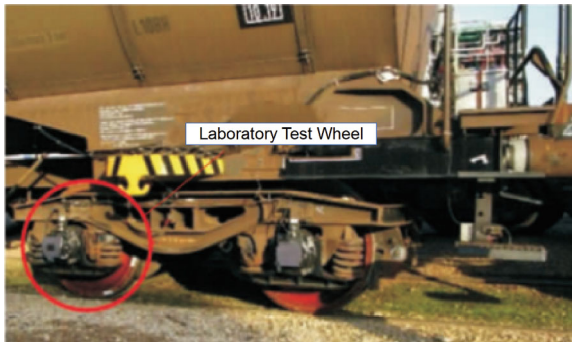


Fig. 2 Schematic diagram of the fracture position of the spring

Based on the condition of the broken springs (see Fig. 1) and macro inspection of the fracture surfaces (as shown in Fig. 2), it was observed that all fracture locations were situated on the spring support ring. The fracture originated from the ground surface of the support ring, extending inward and then rapidly expanding at a 45º angle relative to the axis of the spring coil until complete breakage occurred.

2.1 Raw Material Inspection

To investigate the causes of the abnormal fractures of the springs, 14 broken springs were selected. Samples were prepared from various locations, including the fractured support ring, the middle section of the spring, and the unbroken support ring (see Fig. 3) for testing and analysis, and assign numbers, for example, 1–1 for the fracture end sample of the first broken spring, 1–2 for the middle sample of the spring, and



Fig. 3 Broken spring samples

1–3 for the unbroken end sample. Among these, samples from the middle sections of springs numbered 1–2, 3–2, 6–2, 9–2, and 10–2 were selected for raw material testing. Figure 2 shows actual images of some of the samples.

Chemical Composition Testing. Using the Shimadzu PDA5500II direct-reading spectrometer, the chemical composition of the selected middle section samples of the 5 springs made from 60Si2CrVAT steel was tested in accordance with the requirements of TJ/CL 075-2004 and the specifications of GB/T 222. The test results are shown in Table 1 (taking sample 6–2 as an example). A comparison with the standard values reveals that the chemical composition of the raw materials fully meets the standard requirements.

Non-Metallic Inclusion Testing. To clearly observe the microstructure, samples were taken near the fracture surface of the failed spring steel. After embedding, sanding, and mechanical polishing, the samples were etched using a 4% nitric acid-alcohol solution to prepare metallographic specimens. The microstructure was then observed using a DMI5000M Leica metallographic microscope. The evaluation was conducted according to the ASTM standard grading chart in GB/T 10561, with test results rated as A0 or A0.5, B0, C0, and D1 (A, B, C, and D represent four types of inclusions: sulfides, alumina, silicates, and globular oxides, respectively. The numbers 0, 0.5 and 1 indicate the levels of the inclusions), indicating that the non-metallic inclusions in the raw materials fully meet the standard requirements.

Low Magnification Structure Testing. According to GB/T 226, each sample was subjected to etching treatment with a 1:1 hydrochloric acid aqueous solution. Subsequently, low magnification structure testing was performed using the DMI5000M Leica metallographic microscope. The test results indicate that all samples show no shrinkage cavities, cracks, delaminations, white spots, bubbles, scabs, or inclusions. In addition, the segregation or porosity level is generally classified as either general segregation at level 0.5 or central porosity at level 0.5, or both. So the low magnification structure of the raw materials fully meets the standard requirements.

According to the above inspection results, the raw material 60Si2CrVAT fully meets the requirements of the TJ/CL 075-2004 standard.

Table 1 Actual chemical composition testing results of samples (wt%)

Sample number	C	Mn	Si	P	S	Cu	Cr	Ni	V
Standard values	0.56 ~ 0.64	0.40 ~ 0.70	1.40 ~ 1.80	≤ 0.015	≤ 0.015	≤ 0.20	0.90 ~ 1.20	≤ 0.25	0.10 ~ 0.20
6–2	0.56	0.67	1.40	0.010	0.001	0.03	1.08	0.03	0.13



Fig. 4 Grinding burn marks

2.2 Spring Inspection

Geometric Dimension Check. In accordance with the requirements of the drawing numbered QCZ133-00-03 and the TJ/CL 075-2004 standard, inspections were conducted on 14 broken springs regarding their free height, inner diameter, outer diameter, round steel diameter, total number of coils, distance between the supporting ring end and the working ring, thickness of the supporting ring end, width of the supporting ring end, perpendicularity, direction of rotation, and flaw detection. The inspection results fully meet the standard requirements.

Flaw Detection and Appearance Inspection. To identify the source of the cracks, the coating on the support ring areas of the 14 broken springs was first removed. Upon examining the ground surface of the support rings, no visually detectable cracks were found. Fluorescent magnetic particle testing also revealed no cracks present. However, the appearance inspection noted that all support rings exhibited extensive blue grinding burn marks on the ground surfaces (see Fig. 4).

Hardness Inspection. Hardness is a measure of a metal's resistance to localized deformation [5].

Rockwell hardness of spring. According to the requirements of the TJ/CL 075-2004 standard, the hardness should be within the range of 42 to 48 HRC. Using the GRS-150S digital Rockwell hardness tester, the Rockwell hardness of the matrix of 42 samples was measured. The results indicated that the hardness of the broken spring samples mostly ranged from 49.5 to 54.5 HRC, indicating that almost all spring matrices exceeded the standard hardness.

Microhardness of burned surface and spring core during end face grinding. One sample was selected, and its surface was tested for microhardness using the HM-221 microhardness tester from the Japanese company MITUTOYO. The testing locations included the blue grinding burn area and the center of the sample. The microhardness test results are shown in Fig. 5. The microhardness of the blue grinding burn area was measured at 699 HV (equivalent to approximately 60.06 HRC), while the microhardness at the center was 480 HV (equivalent to approximately 47.7 HRC).

The hardness of the blue grinding burn area was significantly higher than that of the center and greatly exceeded the specified standards.

Optical Metallographic Inspection. Using the DMI5000M Leica metallographic microscope, metallographic inspections were performed on 42 samples. Except for a few samples with hardness at or below 48 HRC, which exhibited normal tempered bainite with dispersed carbide particles (see Fig. 6), most other samples displayed abnormal tempered bainite with a small amount of retained martensitic phase (see Fig. 7). Upon examining the fracture crack origins, it was found that the crack source was located at the grinding burn surface of the support ring. The grinding burn surface consisted of a layer of bright hardening, and the microstructure was identified as white acicular martensite (see Fig. 8).

Fracture Morphology Analysis. Using a Zeiss EVO-18 scanning electron microscope, two fracture areas were observed, revealing a distinct hardened layer (see

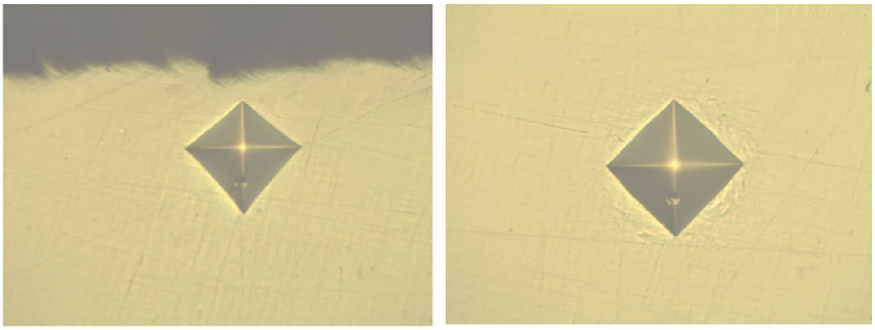


Fig. 5 Microhardness testing

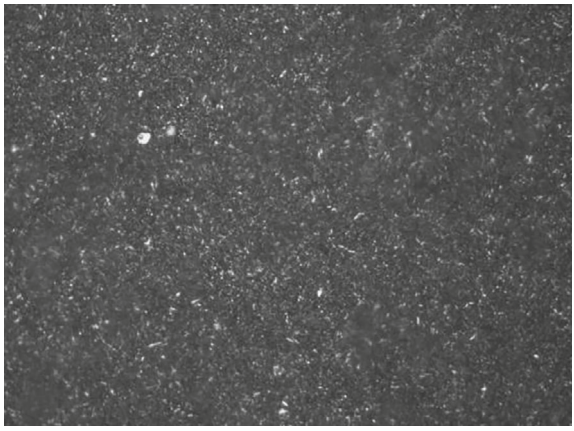


Fig. 6 Metallographic structure of 500X spring base material (48 HRC)

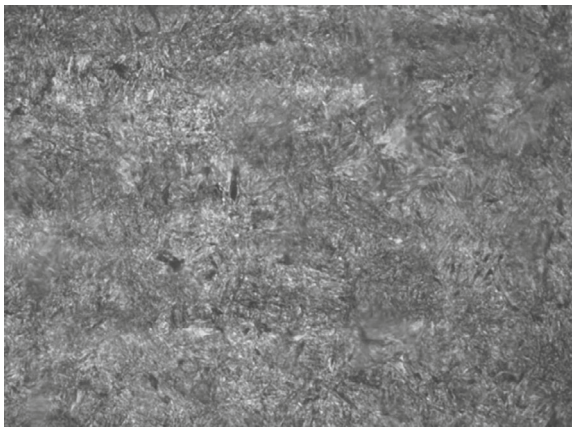


Fig. 7 Metallographic structure of 500X spring base material (52 HRC)

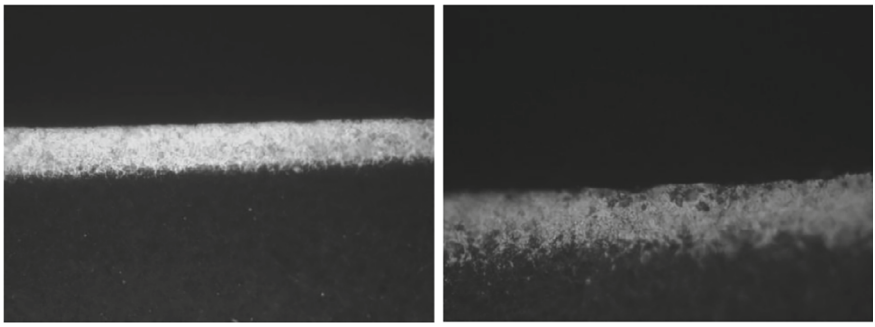
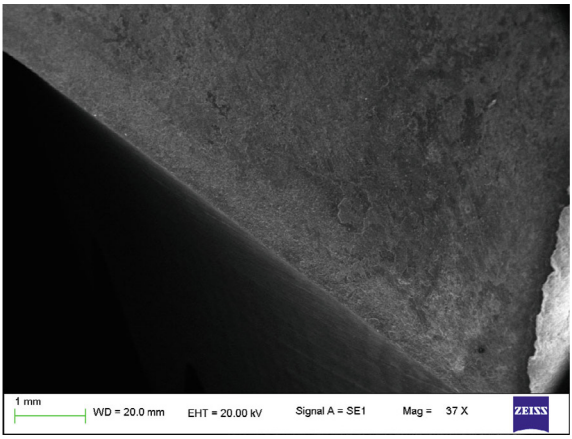


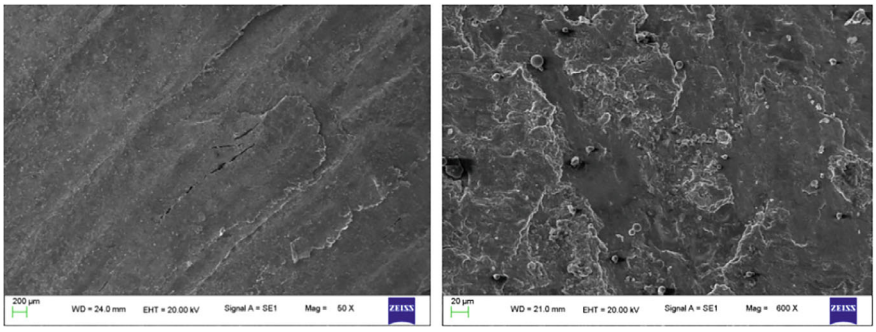
Fig. 8 Metallographic photos of the fracture crack source surface **a** 100X view of the fracture crack source surface; **b** 500X view of the fracture crack source surface

Fig. 9), with the thickness of the hardened layer measured at 0.14 mm. In Fig. 10a, noticeable extrusion marks can be seen along the crack propagation path, where the crack extends from the lower right corner to the upper left corner of the image. The enlarged view in Fig. 10b also shows indications of this crack propagation with groove patterns. Additionally, numerous microcracks were found at a depth of 0.09 mm to 0.10 mm below the surface of the hardened layer (see Fig. 11).

Fracture Energy Spectrum Analysis. Since the vehicle operated near the coast, energy spectrum analysis (EDS) was conducted on the fracture using a scanning electron microscope to eliminate the influence of other media on the fracture (see Figs. 12, 13 and Tables 2, 3). The energy spectrum shows that only a few elements present in the raw materials can be detected on the hardened layer and the substrate, with no components from other corrosive media. The presence of oxygen in the

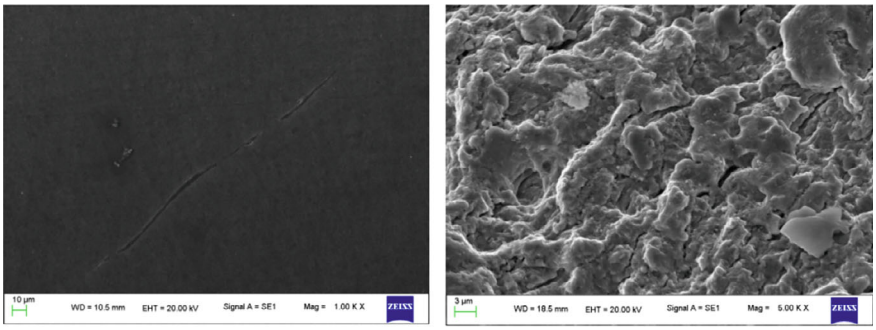
Fig. 9 Hardened layer





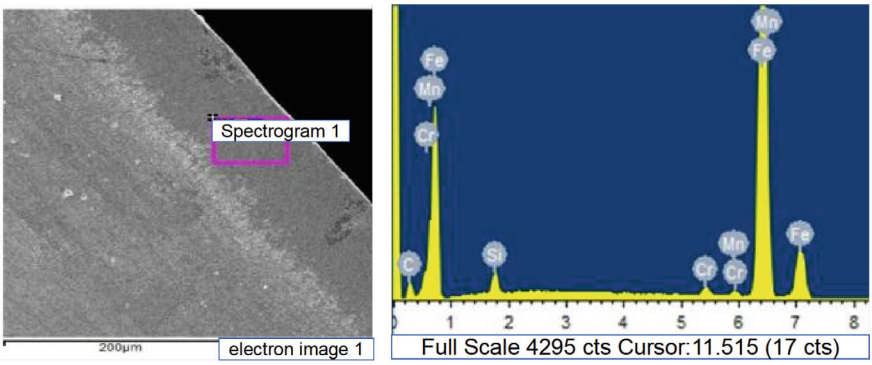
(a) (b)

Fig. 10 SEM image of crack propagation



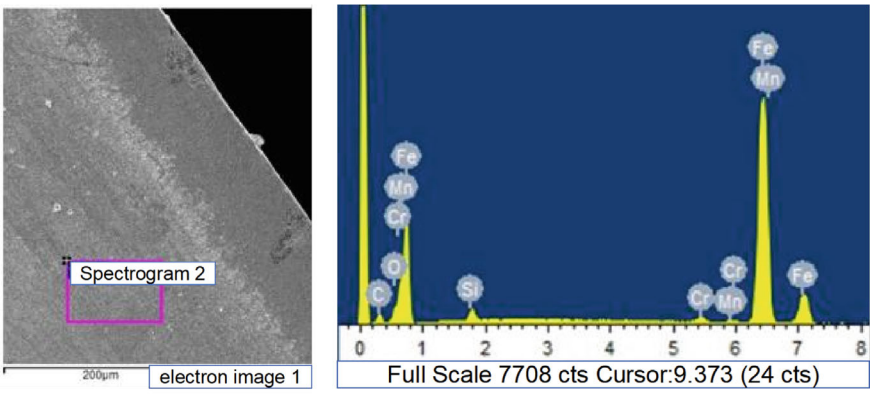
(a) (b)

Fig. 11 Microcracks



(a) (b)

Fig. 12 Energy spectrum of the fracture hardened layer



(a) (b)

Fig. 13 Energy spectrum of the fracture base material

substrate area is attributed to rusting at the fracture and does not affect the analysis results.

3 Analysis and Discussion

Based on the above inspections and tests as well as the on—site investigation of the production workshop, the preliminary analysis is as follows:

Table 2 Element percentage in the fracture hardened layer (%)

Element	Weight	Atomic
C	8.46	29.69
Si	1.52	2.28
Cr	1.19	0.97
Mn	0.81	0.62
Fe	88.01	66.44
Total amount	100.00	

Table 3 Element percentage in the fracture base material (%)

Element	Weight	Atomic
C	8.41	28.01
Si	1.27	1.81
Cr	1.09	0.84
Mn	0.63	0.46
Fe	85.57	61.29
O	3.04	7.59
Total amount	100.00	

- Through the detection of chemical composition, non—metallic inclusions, and macro—structure, all three indicators meet the requirements of the TJ/CL 075-2004 standard. The raw materials are qualified, and the abnormal fracture of the spring has nothing to do with the raw materials.
- The geometric dimensions of the spring meet the requirements of the drawing and the TJ/CL 075-2004 standard. The abnormal fracture of the spring has nothing to do with the geometric dimensions.
- There are blue grinding burn marks at the fracture part of the spring support ring. Through the investigation of the production site and process documents, it is found that the process document stipulates that spring grinding is a dry—grinding method, but in fact, the water—cooled wet—grinding process is used for spring grinding. It is understood that the new process was changed in January of that year. For the needs of dust reduction and environmental protection, it was changed to the current water—cooled wet—grinding process. The new process has not been verified. At the production site, it is observed that when the operator grinds the spring, the feed amount is large, the temperature of the grinding part rises rapidly, and a red—bright color appears (visually reaching 600 °C or above), and part of it exceeds the critical temperature of phase transformation. When the grinding is completed, it is suddenly cooled by water, which is equivalent to secondary quenching. Therefore, a white—bright quenched martensite hardened layer is formed on the grinding surface of the support ring. Quenched martensite is hard and brittle, and the local hardness reaches 60HRC, which greatly reduces the toughness and plasticity of the material. The reduction of area at fracture

and elongation are basically zero. This hardened layer is quite different from the matrix structure, and a large surface residual stress is generated during the structural change. During the vehicle operation, due to the combined action of large alternating loads, impact loads and other complex loads, cracks first initiate and propagate at the hard, high—stress, thin and brittle quenched martensite part. During the friction and wear process, stress concentration occurs at the notch. When the crack reaches a certain depth and the cross—sectional area of the spring cannot withstand the external force, the crack will propagate unstably, and finally lead to spring fracture.

- There are deviations in the hardness and metallographic test data of the spring matrix structure, indicating that there are problems in the spring heat treatment process. A company uses forging residual heat quenching + medium—temperature tempering process for the heat treatment of the outer—circle spring. The completed coiled spring immediately enters the annular soaking furnace at a high temperature for 16—minute soaking at 860 °C. After coming out of the furnace, it is immediately oil—quenched, and then enters the mesh—belt continuous tempering furnace for 120—minute tempering treatment at 540 °C. After years of production and application, this process is mature and applicable. It is understood that during January and February of that year, some individual operators did not place the springs in a single layer as required by the process during the spring tempering section, but stacked them in layers, resulting in insufficient spring tempering (low tempering temperature or short tempering time), causing the microstructure of the spring matrix to be an abnormal tempered troostite + a small amount of tempered martensite maintaining the martensite phase, and the hardness exceeds the specified 42–48HRC, reducing the toughness of the matrix, increasing the brittleness of the spring, which further accelerates the crack propagation and spring fracture [6].
- Since the helical spring is mainly subjected to torsion under the action of external force, the fracture surface presents the morphology of a torsional fracture surface. Due to the maximum and most complex stress on the variable—cross—section part of the spring support ring, as well as the combined action of the martensite—matrix interface stress generated by the grinding heat, most of the fractures occur at the variable—cross—section transition part of the spring support ring.
- During the investigation and testing process, except for the surface hardened layer generated by grinding causing crack initiation, propagation and fracture, no additional fatigue damage due to high hardness, and other forms of cracks and fractures were found in the springs with high overall hardness during the application cycle.

Through the above analysis, the process of abnormal spring fracture is described as follows: During the vehicle operation, cracks first initiate at the grinding hardened layer at the variable—cross—section of the support ring where the stress is the largest and most complex, and propagate under the action of alternating loads. When the cracks propagate to the relatively hard matrix structure, under the action of a large

impact load, they rapidly propagate unstably in the direction of 45° with the axis of the spring coil itself, and then the spring breaks instantaneously.

4 Conclusions and Prospect

In view of the phenomenon of multiple abnormal fractures of the outer—circle spring made of 60Si2CrVAT steel in the Zhuan K6 bogie of a certain gondola car, through a series of physical and chemical inspections, the following conclusions are drawn:

- Improper spring end—face spring—grinding operation causes grinding burns, and secondary quenching on the grinding surface of the support ring forms a quenched martensite structure, which is the main cause of this spring fracture.
- The spring tempering process is insufficient, and the matrix structure is an abnormal tempered troostite + a small amount of tempered martensite maintaining the martensite phase, which makes the spring hardness exceed the standard, increases the brittleness, and accelerates the crack propagation, which is another cause of this spring fracture.
- Other factors such as harsh operating conditions, large impact loads, and irregular wear of the wheel tread all accelerate the spring fracture.

In actual engineering, most metal parts work under variable—amplitude or random loads. Therefore, the damage caused by loads with different frequencies and amplitudes is continuously accumulated, and finally leads to failure and destruction. Due to the extreme complexity of load forms, structural details and service environments, ensuring and predicting the anti—fatigue fracture performance and service safety of bogie welded frames has become a core topic that urgently needs to be solved in the development of higher—speed railway vehicles. In recent years of research, many domestic and foreign experts have invested a great deal of energy in improving the design stress and strength of springs and extending the fatigue life, hoping for further achievements [7–11]. It is believed that in the near future, people will further improve the relevant content of this topic.

5 Process Improvement

Through the above analysis, it is found out that the spring grinding process is the main reason for the fracture of the spring support ring, and the original wet grinding process of the spring must be changed. The specific grinding process of the spring is to strictly control the feed of the grinding spring, control the temperature of the grinding spring, and adopt the method of first rough grinding and then fine grinding, especially the feed of the last fine grinding must be less than 1 mm. The final grinding surface of the spring shall be the natural color of metal, and no color burn mark is allowed.

The maintenance of heat treatment equipment shall be strengthened to eliminate the influence of equipment failure on product quality.

The spring with the new grinding technology has successfully passed 3 million times of fatigue tests, and the rationality of the new technology has been further verified through more than one year of operation.

Acknowledgements Sincere thanks to A. P. Zhichao Hu and A. P. Yi Li for their guidance on the paper.

References

1. Zhang C, Huang P, Tian Q (2006) Production practice of 60Si2CrVAT spring steel. China Metall 16(7):13–15
2. Eren A, Kenan G (2023) Failure analysis of automotive helical spring. Eng Fail Anal 153
3. Du P, Liang Y, Lei L et al (2016) Fracture failure analysis of HXD3B series springs for urban rail locomotives. Met Heat Treat 41(08):189–192
4. Infante V, Freitas M, Baptista R (2022) Failure analysis of a parabolic spring belonging to a railway wagon. Eng Fail Anal 140
5. Gao H, Shen Y, Guo S et al (2004) The influence of heat treatment process on the mechanical properties of spring steel 60Si2CrVAT. Liaoning Spec Steel 02:7–11
6. Lu B (2009) Research on the effect of grinding strengthening technology on the hardening effect of medium carbon steel. Mod Mach 03:82–84
7. Shi X, Gao H, Xu M (2011) Optimization design of automobile suspension springs based on BP. Hist Mech Technol Mech Des 42:82–85
8. Radhakrishna VD, Mayuri R, Samip K et al (2022) Failure analysis of leaf spring used in transport utility vehicles. J Fail Anal Prev 22(5):1844–1852
9. Sun S, Li Y, Sun M et al (2020) Effect of Ce on microstructure and properties of automotive valve spring steel 65Si2CrV. Trans Mater Heat Treat 41(10):51–59
10. Zhang Z, Zhao F, Tan W (2014) Influence of heat treatment process on microstructure and mechanical properties of 60Si2CrVA spring steel. Heat Treat Met 39(5):104–106
11. Jiang T, Wang K, Yu T et al (2019) Effect of heat treatment process on mechanical properties and microstructure of 55SiCrV spring steel. Heat Treat Met 44(10):96–99

Open Access This chapter is licensed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License (http://creativecommons.org/licenses/by-nc-nd/4.0/), which permits any noncommercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if you modified the licensed material. You do not have permission under this license to share adapted material derived from this chapter or parts of it.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.



Dynamic Roll Stability Analysis of Independent Suspension of a New Type of Multi-Axis All-Road Crane



Rui Guo, Puxue Niu, Xifeng Zhang, Xin Wang, and Shuai Song

Abstract In order to solve the problem of dynamic roll stability of multi-axis all-road crane, an independent suspension system with parallel adjustment of multi-accumulators is proposed. According to the mechanical model of steering dynamic force and roll moment of crane, the simulation model of multi-degree-of-freedom hydraulic suspension is established, and the effects of different speed, turning radius and road excitation on vehicle dynamic roll stability are studied. including output force characteristics and load transfer rate. The results show that under the steering condition of low speed 10 km/h and turning radius of 10.6, the load transfer rate is up to 0.023, while under the steering condition of high speed 80 km/h and turning radius of 39.27, the maximum load transfer rate is 0.292, and the vehicle driving is still stable. Even if driving on the dangerous road, the change range of load transfer rate is small, and the vehicle runs smoothly. The suspension system can adapt to multi-speed and different turning radius conditions, and has good robustness. This study can provide some reference for the suspension design of engineering vehicles.

Keywords Multi-axle all-road crane · Independent suspension · Dynamic roll stability · Load transfer rate

R. Guo (✉)
National Key Laboratory of Key Technologies of Lifting Appliances, Qinhuangdao, Hebei, China
e-mail: guorui@ysu.edu.cn

R. Guo · P. Niu
School of Mechanical Engineering, Yanshan University, Qinhuangdao, Hebei, China

X. Wang
School of Mechanical Engineering, Dalian University of Technology, Dalian, Liaoning, China

X. Zhang · S. Song
Taiyuan Heavy Industry Co. Ltd, Taiyuan, Shanxi, China

© The Author(s) 2026

H. Saman et al. (eds.), *The 9th International Conference on Advances in Construction Machinery and Vehicle Engineering*, Lecture Notes in Mechanical Engineering,
https://doi.org/10.1007/978-981-96-5527-4_3

1 Introduction

Oil–gas suspension is widely used in multi-axis all-road cranes at home and abroad, and the suspension system plays an important role in the stability of crane driving process. Scholars at home and abroad have done a lot of research on hydro-pneumatic suspension. Mu Xiaodong of Dalian University of Technology established a mathematical model of hydro-pneumatic suspension and studied the influence of hydro-pneumatic suspension on vehicle stability [1]. Banerjee. Sayan et al. established a mathematical model for the hydro-pneumatic suspension of special tracked vehicles and optimized the suspension structure according to its dynamic model, which improved the damping performance of the vehicle [2].

The change of body roll will cause changes in the load of left and right wheels and tire adhesion, which is directly related to the roll stability and handling stability of the vehicle [3]. LTR (Lateral Load Transfer Ratio) is a kind of effective Evaluation Indexes reflecting vehicle rollover Stability with a certain degree of versatility. It was first proposed by J. Preston-Thomas [4]. The load transfer rate is an important factor to determine the roll angle of the car body, and it is an important index to describe the lateral stability and maneuverability of the vehicle. In recent years, scholars at home and abroad have done a lot of research work on load transfer rate as the evaluation index of vehicle anti-rollover control and early warning control. Among them, CHEN and others carried on the early warning research of automobile rollover based on the simple model of vehicle rollover, and introduced the neural network method to improve the accuracy of the early warning algorithm. Zhu Tianjun and others combined with the classical Kalman state estimation technology to estimate the state of heavy vehicles real-time and accurately, and studied the heavy vehicle rollover early warning algorithm based on the improved rollover early warning tim [5].

In order to solve the problem of dynamic roll stability of multi-axis all-road crane, an independent suspension system with multiple accumulators and parallel adjusting joints is proposed in this paper. The mathematical model is built by using AMESim software, and the output force characteristics of hydro-pneumatic suspension are simulated and analyzed, and the effects of vehicle speed, turning radius, road excitation and other parameters on the output force characteristics and load transfer rate of hydro-pneumatic suspension are studied.

2 Establishment of Pavement Model and Vehicle Model

Dynamic modeling can not only be used in vehicle design, but also can monitor the status of vehicles in real time. However, the modeling is complex, the solution is difficult, and the parameters are too simplified. The use of computer-aided simulation modeling effectively improves the current situation of parameter setting, thus effectively improving the precision of the rollover model.

Firstly, the parameters of the whole vehicle and the characteristic parameters of each component need to be obtained, based on which the whole vehicle model is established, and finally the simulation driving condition is set. The influence factors of rollover stability can be analyzed.

2.1 Establishment of Pavement Modeling

It is mainly divided into random road model and pulse road model. Considering the main road conditions of the all-ground crane, this paper chooses the random road model.

According to the standard GB/T 7031-2005, the road work spectral density is used to describe the random pavement characteristics.

$$G_q(n) = G_q(n_0)(\frac{n}{n_0})^{-p} \tag{1}$$

In the form, p—the frequency index; n—Spatial frequency; n_0—Reference space frequency, Select $n_0 = 0.1\ \mathrm{m}^{-1}$; $G_q(n_0)$—Reference space frequency n_0.

According to different pavement grades, the root means square value of $G_q(n_0)$ has different values, and the pavement is divided into five grades.

According to different pavement grades, the root mean square value of Gq(n0) has different values, and the pavement is divided into A ~ F eight grades, as shown in Table 1.

When the frequency exponent pendant 2 is brought into the above formula and the spatial power density is converted into time frequency power density, the result can be obtained. The thematical modeling of road surface excitation is shown in Fig. 1.

Table 1 Schematic diagram of classification standards for road surface roughness

Pavement grade	$G_q(n_0)(n_0 = 0.1\ \mathrm{m}^{-1})(10^{-6}\mathrm{m}^3)$			Pavement description
	Lower limit	Geometric mean	Upper limit	
A	8	16	32	Highway, national highway, provincial highway and other paved roads
B	32	64	128	Flat dirt roads, gravel and other paved roads
C	128	256	512	Lots of dirt roads, gravel, continuous winding roads
D	512	1024	2048	There are many uneven and convex stones of about 10 cm
E	2048	4096	8192	20 cm of snow and long dirt roads
F	8192	16,384	32,768	There are 30 cm high rocks and 50 cm hard bottom wading

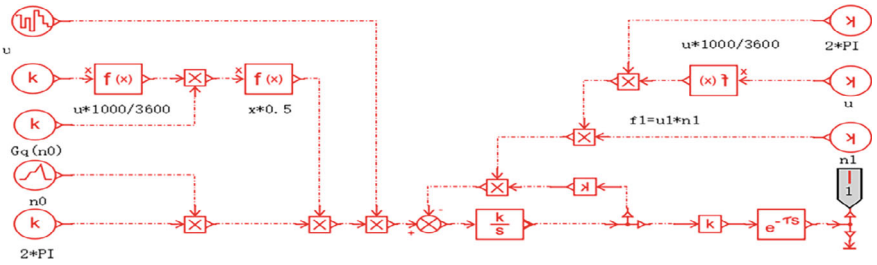


Fig. 1 a thematical modeling of road surface excitation

$$G_q(f) = (2\pi n)^2 G_q(f) = (2\pi n_0)^2 G_q(n_0) u \tag{2}$$

$$G_q(f) = (2\pi n)^2 G_q(f) = (2\pi n_0)^2 G_q(n_0) u \tag{3}$$

2.2 *Establishment of Vehicle Model*

The all-ground crane studied in this paper is a five-axle. In the whole vehicle system, the physical modeling is mainly aimed at its tire and car body, while in the hydro-pneumatic suspension system, the model of hydraulic part mainly includes suspension cylinder, accumulator, all kinds of control valves, damping hole and so on. The vehicle model is shown in Fig. 2.

According to the force analysis above, the force of the suspension cylinder is mathematically modeled, as shown in Fig. 3.

Aiming at the problem of vehicle tilt degree, it is generally described by load transfer rate, and the expression is as follows:

$$\eta = \frac{F_L - F_R}{F_L + F_R} \tag{4}$$

In the formula: η—Load transfer rate; F_l—Force on the left wheel; F_r—Force on the right wheel.

The value of load transfer rate should be between-1 and 1. When there is no tilt, the force on the left and right sides of the tire is the same, and the load transfer rate is 0; when the tilt occurs but no rollover occurs, the force on the left and right tire is not equal, for example, when the force on the left increases, the force on the right tire decreases until the contact force with the ground is zero, reaching the critical state of rollover, and the load transfer rate is between 0 and 1. When a rollover occurs, the force on one side of the tire becomes 0, and the load transfer rate is 1 or − 1[6].

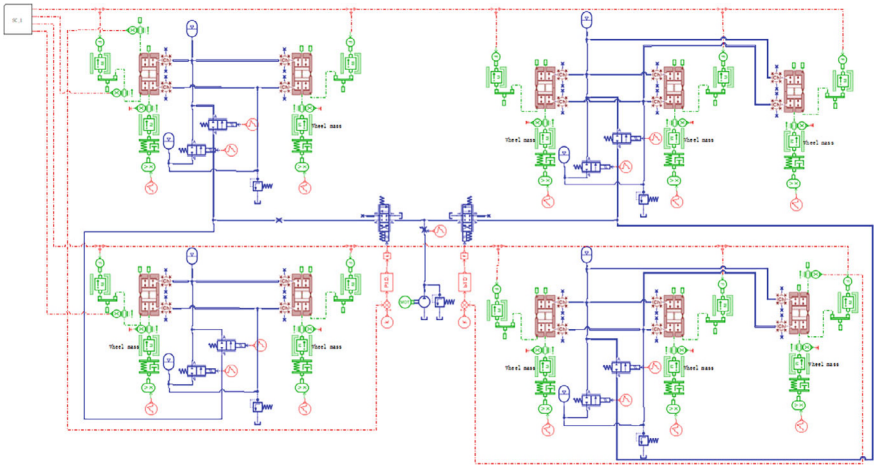


Fig. 2 Five-axle independent suspension vehicle model

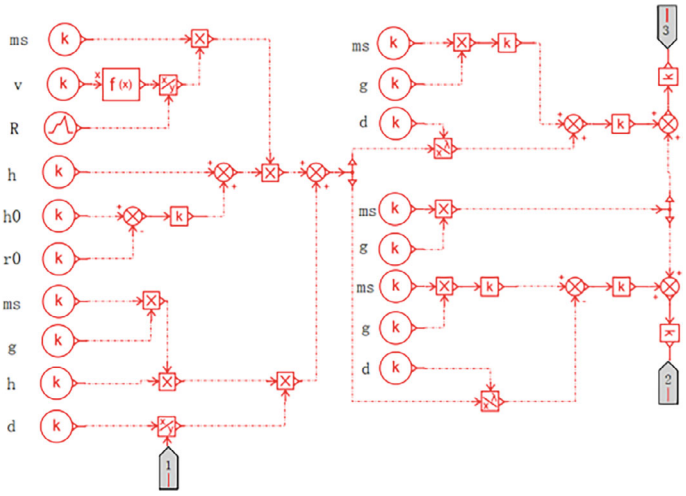


Fig. 3 Mathematical model of the force of the suspension cylinder

3 Dynamic Roll Stability Analysis of Whole Vehicle

3.1 Effect of Driving Speed on Vehicle Roll Stability

For the all-ground crane, the driving speed directly affects the safety of the vehicle in the turn· This section keeps the other parameters of the vehicle unchanged and studies the influence of the speed on the roll stability of the whole vehicle. Because

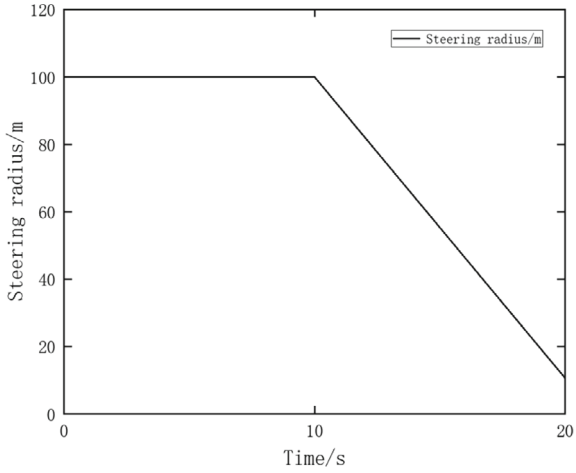


Fig. 4 Turn radius input

the steering mode and its minimum turning radius are different under low-speed steering condition and high-speed steering condition, the two steering conditions are discussed respectively.

Low speed steering condition. Under the condition of low-speed steering, the maximum static load of single suspension is 15 t. Set the road grade to B class, and select three speeds of 10, 20 and 30 km/h. The minimum turning radius is 10.6 m and the maximum steering angle is 40°. Keeping other parameters unchanged, the influence of turning radius on the roll stability of the whole vehicle is studied. The turning radius input is shown in Fig. 4.

As shown in Fig. 5, the speed load transfer rate varies with the turning radius of different vehicles under low-speed steering conditions. From the analysis results, it can be concluded that the load transfer rate increases with the decrease of the turning angle. Among them, the 10 km/h load transfer rate is 0.023 × 20 km/h, the load transfer rate is 0.072 × 30 km/h, the load transfer rate is 0.155. Therefore, a further study found that under the condition of turning at low speed, the load transfer rate increases with the increase of steering speed.

High-speed steering condition. Under the condition of high-speed steering, the maximum static load of single suspension is 6.5 t. The road grade is set to B class, and the speed is 60 km/h, 70 km/h, 70 km/h, 80 km/h. When driving at high speed, the steering angle is small, the input maximum steering angle is 10°, and the suspension force and load transfer rate at each speed are obtained by inputting different turning radius. When the steering angle is 10°, the corresponding turning radius is 39.2 7 m, that is, the minimum turning radius entered, as shown in Fig. 6.

As shown in Fig. 7, the vehicle load transfer rate varies with the turning radius at different speeds under high-speed steering conditions. From the analysis results, it can be concluded that the load transfer rate increases with the decrease of the turning angle. Among them, the load transfer rate of 60 km/h is 0.167 × 70 km / h, the load

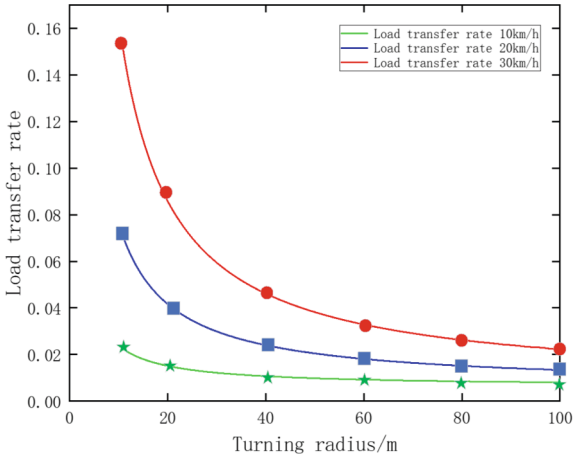


Fig. 5 Comparison of load transfer rates at low-speed steering

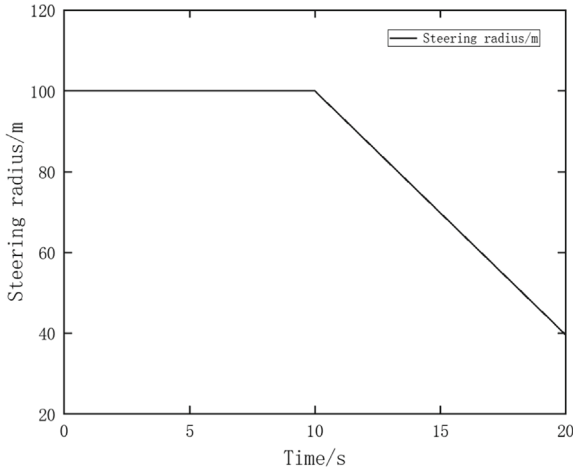


Fig. 6 Turn radius input

transfer rate is 0.225 × 80 km/h, the load transfer rate is 0.292. Therefore, further study found that under the condition of high-speed turning, the steering load rate increases with the increase of steering speed. Therefore, reducing the driving speed is beneficial to improve the roll stability of the vehicle.

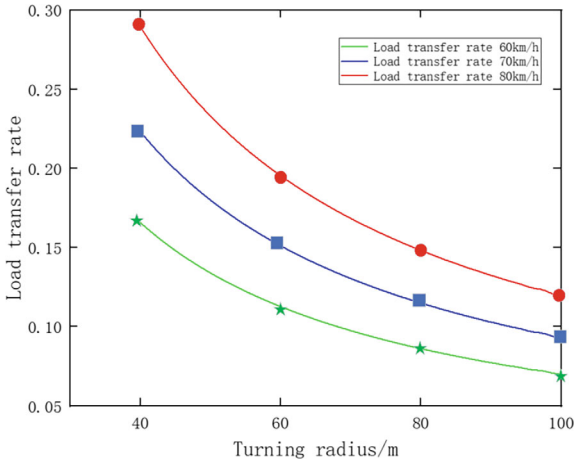


Fig. 7 Comparison of load transfer rates at high-speed steering

3.2 *Effect of Road Surface Excitation on Vehicle Roll Stability*

As most of the working sites of all-ground cranes are bad terrain, it is very necessary to analyze them in different road conditions. Figure 8 shows the road surface excitation input of different grades of pavement.

In order to study the influence of road excitation on vehicle roll stability, the speed of the simulation model is set to 10 km/h, the turning radius is set to 10.6 m, and the influence on vehicle roll stability is analyzed when the road grade spectrum is C, D and E respectively.

As shown in Fig. 9, the change law of load transfer rate is different from that under different speed, different working conditions and different turning radius above.

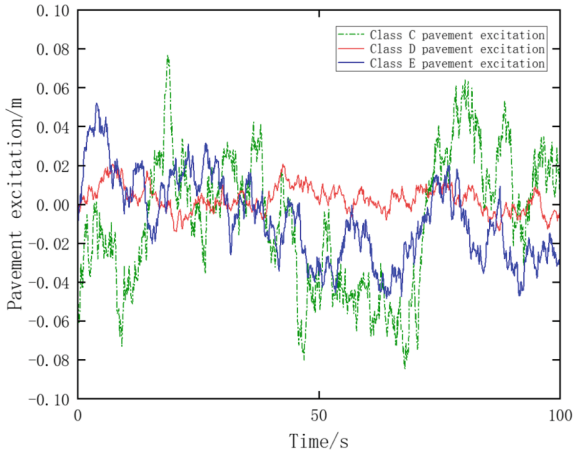


Fig. 8 Pavement excitation input

This result is due to the fact that vehicles experience different road surface shocks at different times under random road excitation. However, under the action of the independent suspension system, the load transfer rate is stable in the safe range. Therefore, in the face of complex and dangerous roads, the independent suspension system is beneficial to improve the roll stability of the whole vehicle.

Taking the change law of lateral load transfer rate with time as the evaluation index, this section focuses on the influence of driving speed, turning radius and road excitation on the dynamic rollover stability of the whole vehicle.

Perform Fourier transform on the acceleration vibration of the entire vehicle's cab to obtain the frequency domain diagram of the cab vibration, as shown in Figs. 10, 11, 12, and 13.

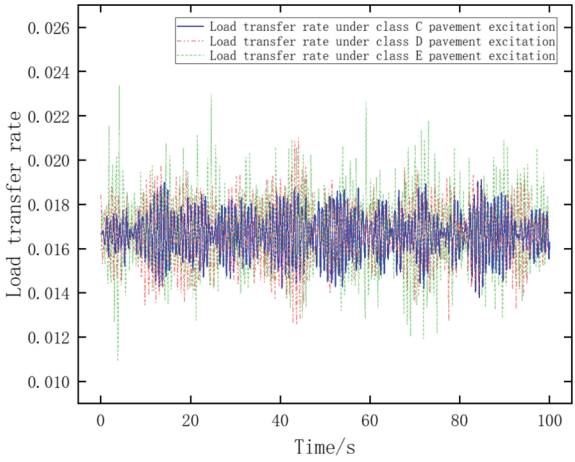


Fig. 9 Pavement incentives in relation to time

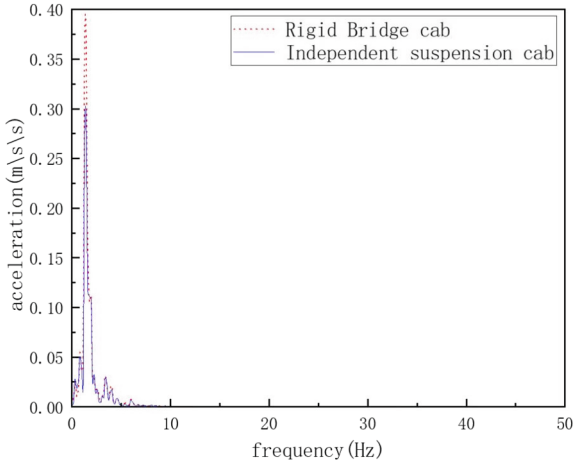


Fig. 10 Frequency domain diagram of 10 km/h E-class pavement vibration

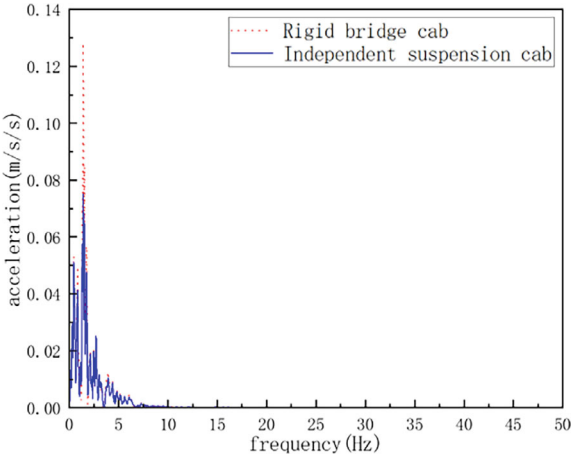


Fig. 11 Frequency domain diagram of 20 km/h D-class pavement vibration

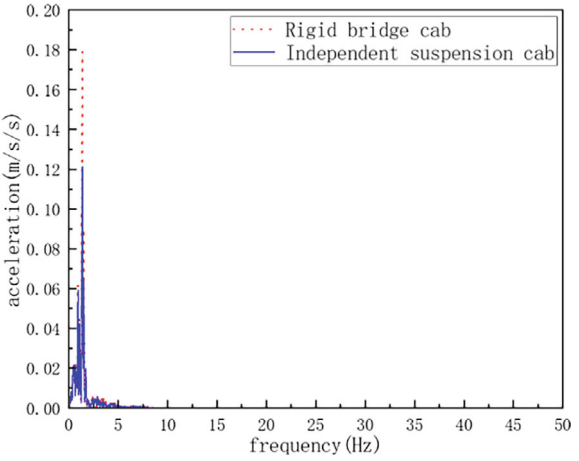


Fig. 12 Frequency domain diagram of 40 km/h C-class pavement vibration

As shown in Figs. 10, 11, 12, and 13, the vibration frequencies under different working conditions are mostly concentrated between 1–10 Hz, with low peak values.

4 Conclusion

(1) The results show that even under the condition of high-speed 80 km/h, the load transfer rate is 0.292 and the vehicle is still stable. The effects of vehicle speed and turning angle on vehicle roll stability are similar. Even in the case of driving on the dangerous road and ensuring the safe speed of the vehicle, the change range of the load transfer rate is small, and the vehicle runs smoothly.

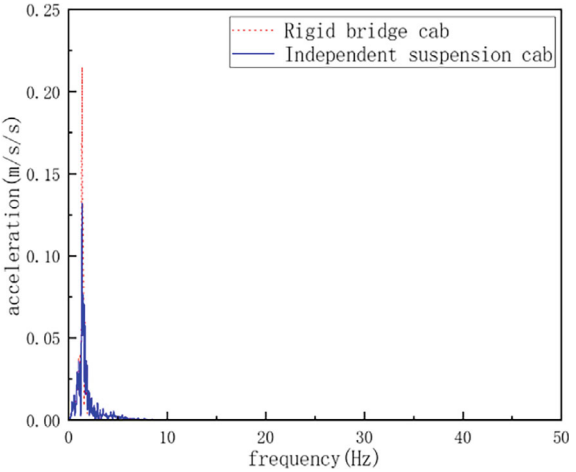


Fig. 13 Frequency domain diagram of 80 km/h B-class pavement vibration

(2) Through the simulation comparison and analysis of various working conditions, it is shown that the independent suspension system is effective and feasible to control the rollover of the whole vehicle. The suspension system improves the roll stability of the whole vehicle and has a certain application value.

Acknowledgements Supported by the National Natural Science Foundation of China (Grant No. 52075469, 12173054), This thesis is also funded by the Key Core Technology Research Project of Taiyuan City.

References

1. Mu X (2017) Design and analysis of hydropneumatics suspension system for 55 ton all ground crane. Dalian University of Technology
2. Banerjee S, Balamurugan V, Krishnakumar R (2014) Ride dynamics mathematical model for a single station representation of tracked vehicle. J Terrramech 53(1):47–58
3. Jianfu Y, Hongjie C, Zhihao L et al (2023) Optimization analysis of rollover index of multi-axle special vehicle. J Ordnance Equip Eng 44(09):134–142
4. Zhilin J, Zhenghua Y (2019) Research on automobile rollover early warning based on quadratically predicted lateral load transfer rate. China Mech Eng 30(15):1790–1795
5. Xia G, Li K, Tang X et al (2023) Research on anti-rollover control of counterbalanced forklift based on roll energy classification. J Mech Eng 59(10):275–289
6. Xin W, Sujie W, Panpan W et al (2022) Fault diagnosis and simulation of luffing hydraulic system of truck-mounted crane. Hydraul Pneum 46(07):173–181

Open Access This chapter is licensed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License (http://creativecommons.org/licenses/by-nc-nd/4.0/), which permits any noncommercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if you modified the licensed material. You do not have permission under this license to share adapted material derived from this chapter or parts of it.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.



Research on Crane Motion State Monitoring System Based on Digital Twin



Xuyang Cao, Baokang Yan, Kai Lin, and Wanlong Shu

Abstract In order to improve the efficiency of crane operation, ensure operational safety, and improve the intelligent level of crane operation, a crane motion state monitoring method based on digital twin is proposed. Firstly, according to the operating characteristics and real-time monitoring requirements of the existing experimental prototype-spider crane, a digital twin framework for telescopic cranes is proposed, and the functions of each part are elaborated in detail; secondly, a virtual model of crane digital twin is constructed, data acquisition, classification, transmission and parsing are carried out, and kinematics analysis of the crane is carried out to realize real-time state mapping of cranes, and it is verified that the crane digital twin motion state monitoring system can provide fine motion state monitoring, which provides a new idea for intelligent crane operation.

Keywords Digital twin · Crane · Motion status monitoring · Intelligence

1 Introduction

With the rapid advancement of global intelligent industrialization, in order to improve the efficiency of lifting operations and enhance safety and security, the need for real-time monitoring and remote control of the crane's operating status [1].

In recent years, digital twin technology has been widely used in production, manufacturing and other industries due to its advantages of dynamic consistency, virtual-reality interaction and feedback, and broad application prospects, etc. However, there are still many problems in its real-time information interaction, virtual-reality synchronization and mutual control, which need to be explored urgently. Zhang Wei et al. [2] established a tower safety monitoring system based on the Internet of Things, proposed a sensor arrangement scheme, and established a real-time monitoring system based on sensor data; Zhang Cheng et al. [3] proposed the selection

X. Cao · B. Yan (✉) · K. Lin · W. Shu
Dalian University of Technology, Dalian, China
e-mail: ybk1227@163.com

© The Author(s) 2026

H. Saman et al. (eds.), *The 9th International Conference on Advances in Construction Machinery and Vehicle Engineering*, Lecture Notes in Mechanical Engineering,
https://doi.org/10.1007/978-981-96-5527-4_4

of parameters for static and dynamic monitoring of structural capacity and the principle of measurement point arrangement, which provides a theoretical basis for the safety monitoring of the structure during the use of tower cranes; Gutiérrez et al. [4] proposed a vision-based 3D tower crane monitoring system, which utilizes image recognition technology to calculate the instantaneous displacements of each part of the tower crane in real time, and monitors the tower crane in real time according to the instantaneous displacements; Moi et al. [5] used an inverse method based on physical strain gauges to calculate force vector weights, thus realizing real-time condition monitoring of cranes; Lai et al. [6] introduced the implementation method of form-property-integrated digital twin technology, and illustrated how to design a digital twin method for structural analysis of complex equipment by taking a boom crane as an example. In summary, the existing crane monitoring system research function is relatively single, and the interactivity with the monitored physical system needs to be improved [7]. Based on this, this paper, with the help of digital twin technology, to carry out research in crane motion simulation and monitoring, to realize the faithful mapping of the physical world in the virtual environment, to truly reflect the state of the physical world, and experimental demonstration of the digital twin system on the accurate monitoring of the physical crane.

2 Digital Twin Monitoring System Framework

This paper takes a small telescopic boom spider crane as an example to construct a digital twin monitoring system, and the system framework is shown in Fig. 1. A virtual interaction mapping mechanism consisting of four components: physical entity-data connection-virtual model-application service is established to provide technical support for safe crane operation.

3 Real-Time Mapping Scheme

3.1 Twin Data Collection and Transmission

The essence of real-time mapping of physical entities to virtual environments is the real-time mapping of physical environment to virtual environment data [8]. The operational data is obtained by sensor acquisition and transmitted via CAN bus.

In this paper, we use Socket to establish two-way communication between the server-side and the client-side, the crane host controller as the Socket server-side, the digital twin monitoring system established by Unity3D as the client-side, and the server -side can transmit the twin data by completing the communication with the client-side. The TCP/IP protocol is a connection-oriented communication, i.e., both parties have to establish a reliable connection before communication can be carried

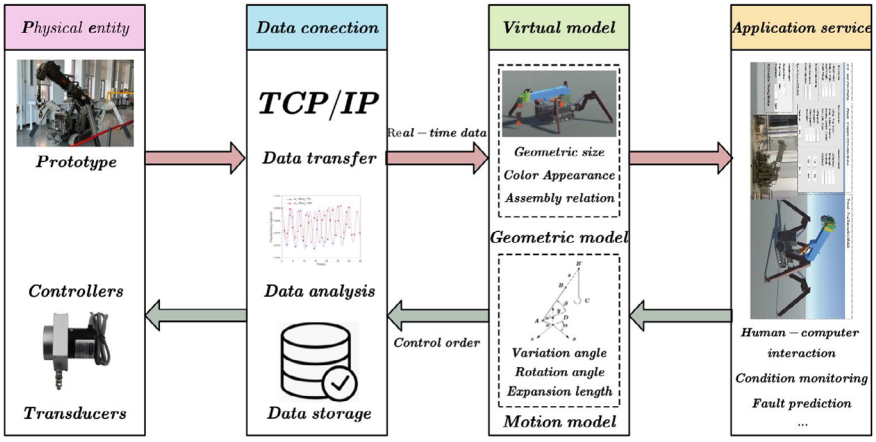


Fig. 1 Crane digital twin framework

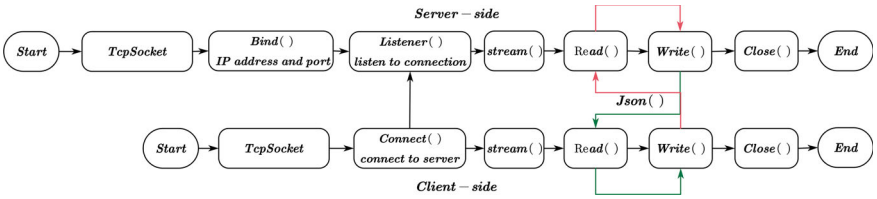


Fig. 2 Socket communication flowchart

out. To achieve stable and reliable transmission, to ensure that the transmitted data is not duplicated, not lost, and arrives in order, the communication flow is shown in Fig. 2.

3.2 Real-Time State Mapping

The research object spider crane and typical telescopic boom crane consistent, telescopic boom crane boom kinematics model shown in Fig. 3, o point for the crane slewing platform center, o' point for the A point on the crane slewing axis projection center, A point for the boom root and rotary table connection, B point for the head of the boom, C point for the hook point, D point for the luffing cylinder and rotary table connection. $oxyz$ For global coordinate system, AoD rotary table can rotate around y axis, which is indicated by α. The telescopic boom ABB' can be luffed in the plane of ABD, denoted by β, the luffing along the direction of the boom BB' is denoted by s, and $B'C$ is the wire rope. So far, three variables α, β, s can be used to describe the

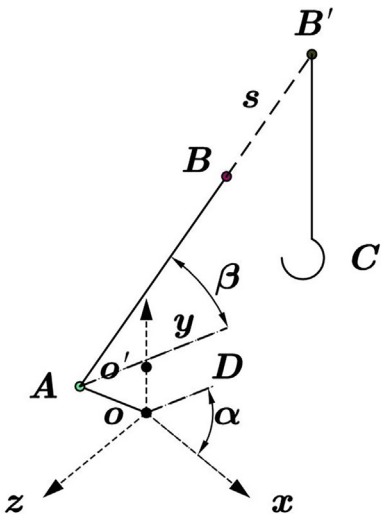


Fig. 3 Algorithm node expansion schematic

slewing kinematics, luffing kinematics, and weight lifting kinematics of the whole crane system.

(1) Slewing kinematics description

The actual slewing axis length of the boom is:

$$L = (L_{AB} + s)cos\beta - L_{Ao'} \tag{1}$$

Then the velocity at B' is:

$$\dot{x}_{B'} = -\dot{\alpha} L \sin\alpha \tag{2}$$

$$\dot{z}_{B'} = -\dot{\alpha} L \cos\alpha \tag{3}$$

The derivative with respect to time yields the acceleration respectively:

$$\ddot{x}_{B'} = -\ddot{\alpha} L \sin\alpha - \dot{\alpha}^2 L \cos\alpha \tag{4}$$

$$\ddot{z}_{B'} = -\ddot{\alpha} L \cos\alpha + \dot{\alpha}^2 L \sin\alpha \tag{5}$$

(2) Variable amplitude kinematic description

Boom luffing axle length:

$$L = (L_{AB} + s)\cos\beta \tag{6}$$

Variation induced velocities are:

$$\dot{y}_{B'} = \dot{\beta} L \cos\beta \tag{7}$$

$$\dot{x}_{B'} = -\dot{\beta} L \sin\beta \cos\alpha \tag{8}$$

$$\dot{z}_{B'} = \beta L \sin\beta \sin\alpha \tag{9}$$

The derivative with respect to time yields the acceleration due to the variable amplitude motion as:

$$\ddot{y}_{B'} = \ddot{\beta} L \cos\beta - \dot{\beta}^2 L \sin\beta \tag{10}$$

$$\ddot{x}_{B'} = -\ddot{\beta} L \sin\beta \cos\alpha - \dot{\beta}^2 L \cos\beta \cos\alpha + \dot{\alpha}\dot{\beta} L \sin\beta \sin\alpha \tag{11}$$

$$\ddot{z}_{B'} = \ddot{\beta} L \sin\beta \sin\alpha + \dot{\beta}^2 L \cos\beta \sin\alpha + \dot{\beta}^2 L \sin\beta \cos\alpha \tag{12}$$

(3) Stretching kinematic description

Arm head position:

$$x_{B'} = [(L_{AB} + s)\cos\beta - L_{Ao'}]\cos\alpha \tag{13}$$

$$y_{B'} = (L_{AB} + s)\sin\beta + L_{oo'} \tag{14}$$

$$z_{B'} = -[(L_{AB} + s)\cos\beta - L_{Ao'}]\sin\alpha \tag{15}$$

The position equation is derived for time:

$$\dot{x}_{B'} = (\dot{s}\cos\beta - (L_{AB} + s)\dot{\beta}\sin\beta)\cos\alpha - ((L_{AB} + s)\cos\beta - L_{AO'})\dot{\alpha}\sin\alpha \tag{16}$$

$$\dot{y}_{B'} = \dot{s}\sin\beta + (L_{AB} + s)\dot{\beta}\cos\beta \tag{17}$$

$$\dot{z}_{B'} = -(\dot{s}\cos\beta - (L_{AB} + s)\dot{\beta}\sin\beta)\sin\alpha - ((L_{AB} + s)\cos\beta - L_{AO'})\dot{\alpha}\cos\alpha \tag{18}$$

The velocity equation is derived for time:

$$\ddot{x}_{B'} = \ddot{s}\cos\beta\cos\alpha + \dot{s}(-\dot{\beta}\sin\beta\cos\alpha - \dot{\alpha}\cos\beta\sin\alpha)$$
$$-\dot{s}\beta\sin\beta\cos\alpha - (L_{AB}+s)(\ddot{\beta}\sin\beta\cos\alpha + \beta(\dot{\beta}\cos\beta\cos\alpha - \dot{\alpha}\sin\beta\sin\alpha))$$
$$-\dot{s}\dot{\alpha}\cos\beta\sin\alpha - (L_{AB}+s)(\ddot{\alpha}\cos\beta\sin\alpha - \dot{\alpha}(\dot{\beta}\sin\beta\sin\alpha - \dot{\alpha}\cos\beta\cos\alpha))$$
$$+L_{AO'}(\ddot{\alpha}\sin\alpha + \dot{\alpha}^2\cos\alpha) \tag{19}$$

$$\ddot{y}_{B'} = \ddot{s}\sin\beta + \dot{s}\dot{\beta}\cos\beta + \dot{s}\dot{\beta}\cos\beta + (L_{AB}+s)(\ddot{\beta}\cos\beta - \dot{\beta}^2\sin\beta) \tag{20}$$

$$\ddot{z}_{B'} = -\ddot{s}\cos\beta\sin\alpha - \dot{s}(-\dot{\beta}\sin\beta\sin\alpha + \dot{\alpha}\cos\beta\cos\alpha)$$
$$+\dot{s}\beta\sin\beta\sin\alpha + (L_{AB}+s)(\ddot{\beta}\sin\beta\sin\alpha + \dot{\beta}(\beta\cos\beta\sin\alpha + \dot{\alpha}\sin\beta\cos\alpha))$$
$$-\dot{s}\dot{\alpha}\cos\beta\cos\alpha - (L_{AB}+s)(\ddot{\alpha}\cos\beta\cos\alpha - \dot{\alpha}(\dot{\beta}\sin\beta\cos\alpha + \dot{\alpha}\cos\beta\sin\alpha))$$
$$+L_{AO'}(\ddot{\alpha}\cos\alpha - \dot{\alpha}^2\sin\alpha) \tag{21}$$

4 Crane Digital Twin Monitoring System

The spider crane, which is the subject of this study, is a telescopic boom crane, and its structure is shown in Fig. 4, where (1) is the off-loading part, (2) is the slewing mechanism, (3) is the telescopic boom, and (4) is the winch mechanism.

4.1 Functionality

According to the framework proposed in this paper, a digital twin monitoring platform for telescopic boom cranes is constructed with the help of Unity3D platform, which

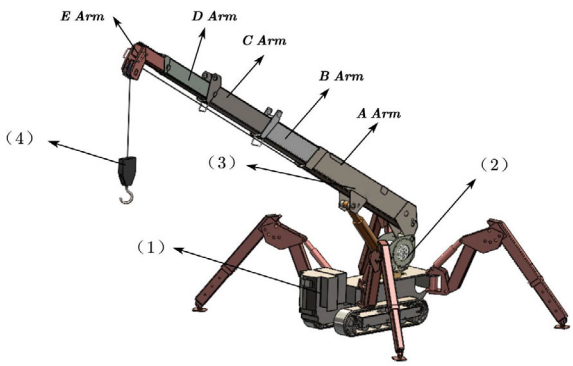


Fig. 4 Schematic diagram of telescopic boom crane structure

mainly consists of three major functional modules: crane motion state visualization, real-time data display and human–computer interaction.

(1) Crane motion state visualization

Crane motion state visualization is mainly real-time motion state visualization. Real-time motion state is mainly based on the physical entity and virtual model data real-time mapping mechanism, real-time data-driven virtual model movement, comprehensive monitoring and management of the crane operation status [9].

(2) Real-time data display

The data display module displays real-time information such as crane operation data on the display, which is convenient for users to fully grasp the operating status of the crane [10]. Specifically, it includes: luffing angle, slewing angle, telescopic boom length, hook height, running speed and other information when the crane is running.

(3) Human–computer interaction module

The human–computer interaction module is mainly through the further division of the specific functions of the digital twin monitoring system to help the user more convenient access to information [11]. Specifically, it includes crane reality interaction, multi-view roaming, historical motion playback and other functions.

4.2 Experimentation and Analysis

The experiment is to test the function of real-virtual interaction. The experimental process is that the system inputs the amplitude closed-loop angle and presses the motion command, sends the amplitude angle and the motion command to the industrial computer, realizes the amplitude motion, and stops the amplitude when it reaches the target value of amplitude, and the monitoring system monitors the whole motion process and records and stores the experimental data, and the monitoring process is shown in Fig. 5.

Through experiments, it is confirmed that the system in this paper can carry out crane motion state monitoring, and the data are shown in Fig. 6. After the monitoring is turned on, the system correctly sends control commands and amplitude target values to drive the crane to complete the amplitude action, and correctly receives and analyzes the real-time values of the sensors, so that the virtual model follows the physical crane to synchronize the movement, and maps the crane's movement status in real time. Experimental results show that the system has no significant delay and high real-time, real-time data acquisition, analysis, transmission without error, can be carried out on-site operation monitoring.

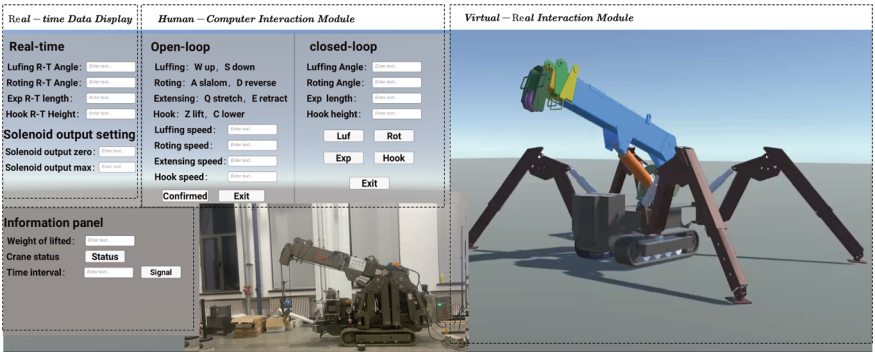


Fig. 5 System monitoring process

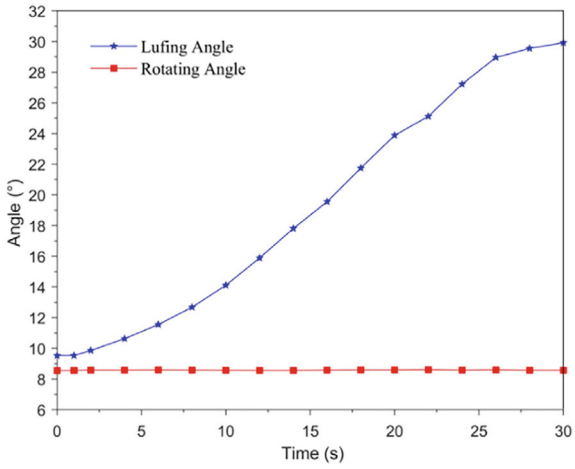


Fig. 6 Experimental angle change values

5 Conclusion

This paper starts from the establishment of a crane digital twin system framework, and introduces the key technologies of twin data acquisition, transmission and resolution, kinematic resolution, etc., and takes a telescopic boom spider crane as an example, and establishes a telescopic boom crane digital twin monitoring system according to the proposed framework and key technologies, realizing real-time monitoring and control of the experimental prototype. The results proved that the crane digital twin monitoring system operates stably and provides ideas for intelligent crane operation. Subsequently, data analysis and learning will be carried out on this basis, with a view to realizing the crane's fault prediction, life analysis and other functions.

Acknowledgements This research was supported by National Natural Science Foundation of China under Grant No.52275088 and the Fundamental Research Funds for the Central Universities under Grant (No.DUT22LAB506, No.DUT22LAB507).

References

1. Fan H (2023) Research on key technology of digital twin system for truss gantry crane. North Central University
2. Zhang W, Liao Y, Jiang L et al (2021) Safety monitoring system of tower crane based on internet of things. China J Saf Sci 31(2):55–62
3. Zhang C, Zhao T, Jiang L et al (2021) Selection of structural safety monitoring parameters and arrangement of measurement points for tower cranes. Chin J Saf Sci 31(8):112–118
4. Gutierrez R, Magallon M, Hernandez DC (2021) Vision-based system for 3D tower crane monitoring. IEEE Sens J 21(10):11935–11945
5. Moi T, Cibicik A, Rølvåg T (2020) Digital twin based condition monitoring of a knuckle boom crane: an experimental study. Eng Fail Anal 112:104517
6. Lai X, Wang S, Guo Z et al (2021) Designing a shape-performance integrated digital twin based on multiple models and dynamic data: a boom crane example. J Mech Des 143:1–15
7. Du Y, Luo Y, Peng Y et al (2023) Three-dimensional visualization and monitoring of industrial robots based on digital twin. Comput Integr Manuf Syst 29(6):2130–2138
8. Jia Z, Wang Xin, Gao S et al (2023) Online monitoring method of tower crane structural performance based on digital twin. Comput Integr Manuf Syst 1–14
9. Szpytko J, Duarte YS (2019) Digital twins model for cranes operating in container terminal. IFAC-PapersOnLine 52(10):25–30
10. Jiang W, Ding L, Zhou C (2022) Digital twin: stability analysis for tower crane hoisting safety with a scale model. Autom Constr 138:104257
11. Ritto TG, Rochinha FA (2021) Digital twin, physics-based model, and machine learning applied to damage detection in structures. Mech Syst Signal Process 155:107614

Open Access This chapter is licensed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License (http://creativecommons.org/licenses/by-nc-nd/4.0/), which permits any noncommercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if you modified the licensed material. You do not have permission under this license to share adapted material derived from this chapter or parts of it.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.



Research on Visual Image Processing Algorithm Based on Wind Turbine Hoisting



Xuyang Cao, Zhaoqing Zha, Jian Kang, and Yong Su

Abstract Wind turbine lifting is a crucial part of a wind energy generation project, and its installation has a decisive impact on the success of the whole project. Traditional wind turbine lifting relies on rich experience judgement, which is not only inefficient but also has high risk. In order to achieve real-time image shooting and processing, this study adopts visual sensors to capture images, extract image features, and realise real-time linkage between spreader operation and image analysis. By continuously feeding back the real-time positioning coordinates and correcting the positioning target, the precise control of the lifting process is achieved.

Keywords Wind turbine crane · Precision control · Vision sensor · Image features

1 Introduction

The current wind power installation process uses multiple cranes in a coordinated lifting method, i.e., a large crane is responsible for the main lifting operation, and another small crane is responsible for direction adjustment and balancing [1]. Due to the lifting operation staff line of sight is blocked and other factors, the staff can not obtain the position of the lifted components in a timely manner. This type of installation requires a large number of engineering machinery to cooperate with each other and a large number of staff to complete. In such a working mode, once a safety accident occurs, the life and property safety of the staff is easily threatened.

Researchers both domestically and internationally have built a solid foundation of studies in the field of machine vision, Zhang et al. [2] used image sequences and wheel inertial self-motion results to construct local maps describing the local environment, and then matched the local semantic maps with online map data to

X. Cao · Z. Zha (✉)
Dalian University of Technology, Dalian, China
e-mail: z2024489582@163.com

J. Kang · Y. Su
Taiyuan Heavy Industry Co., Ltd, Taiyuan, China

© The Author(s) 2026

H. Saman et al. (eds.), *The 9th International Conference on Advances in Construction Machinery and Vehicle Engineering*, Lecture Notes in Mechanical Engineering,
https://doi.org/10.1007/978-981-96-5527-4_5

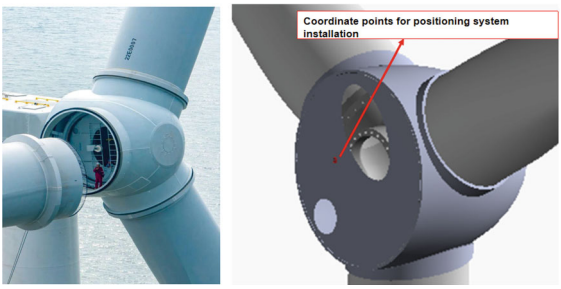


Fig. 1 Visual positioning system

a) Wind turbine hub b) Schematic diagram of the positioning system installation position

estimate the camera position, and experimentally verified that the average errors of this localisation method in the horizontal and vertical directions were 39 mm and 167 mm, respectively.

2 An Overhead Wheel Positioning Identification Method

The traditional position detection during the lifting of wind turbine hubs mainly relies on manual labour, which has many shortcomings, and the study of a high-altitude hub positioning recognition scheme is of great significance to promote the intelligent development of construction machinery. The new positioning identification method can accurately reflect the relative status of the hub and nacelle.

As shown in Fig. 1, it is a schematic diagram of the positioning system installation position, and the red points in the diagram are the positioning system installation coordinate points.

This method requires the use of a monocular camera and a range sensor, where the monocular camera captures image information of the mounting mating surfaces, and the range sensor detects the relative distance between the hub and the nacelle at the time when the corresponding image is taken.

3 Visual Localization Algorithms

3.1 Image Pre-Processing Algorithm

The three main common colour space models are RGB, HSV and Lab. The HSV colour space model is shown in Fig. 2. Compared to RGB colour space, HSV colour space is more suitable for image processing. In wheel mounting, light is the main factor that affects the image quality, and the HSV colour space is easier to track a

specific colour than the RGB colour space [3]. Therefore, in this paper for image colour extraction HSV colour space is used for colour extraction.

$$V = \max(R, G, B) \tag{1}$$

$$S = \begin{cases} \frac{[V - \min(R,G,B)]}{V}, & V > 0 \\ 0, & V = 0 \end{cases} \tag{2}$$

$$H = \begin{cases} \frac{[60*(G-B)]}{SV}, & V = R \\ 60 * \left[2 + \frac{(B-R)}{SV}\right], & V = G \\ 60 * \left[4 + \frac{(R-G)}{SV}\right], & V = B \\ 0, & V = 0 \\ H + 360, & H < 0 \end{cases} \tag{3}$$

where R, G, and B represent the magnitude of the image in the red, green, and blue components, respectively.

After converting the image to the colour space, the colour extraction was tested for blue colour. Checking the table shows that the mathematical expression for blue in HSV colour space is $125 \leq H < 150$, and the colour extraction effect is shown in Fig. 3.

As can be seen from the colour extraction effect diagram, the colour extraction based on HSV colour space can effectively extract the blue circular region in the image and get rid of the influence caused by environmental factors.

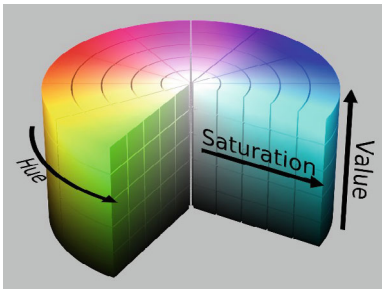


Fig. 2 HSV colour space model



Fig. 3 Colour extraction

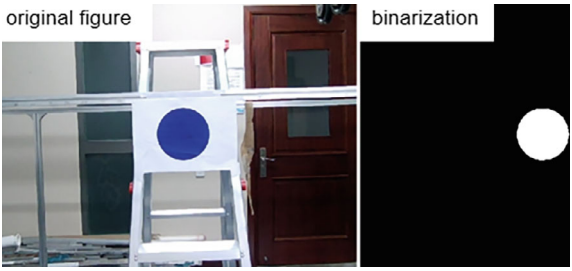


Fig. 4 Nacelle docking surface

3.2 *Image Segmentation*

Image segmentation is divided into semantic segmentation, instance segmentation, etc. Semantic segmentation is the classification of each pixel point in the image, to be specifically segmented to the edge of each instance; instance segmentation is the combination of target detection and semantic segmentation [4].

Image binarisation is the process of setting the grey value of a pixel point on an image to 0 or 255, giving the whole image a distinct black and white effect. There are various methods of image binarisation such as simple binarisation, OTSU method etc. Figure 4 shows the effect of image after binarisation.

The OTSU method is an improved maximum interclass variance method based on the difficulty of finding extreme points in the bimodal method [5].

3.3 *Camera Rotation Angle Algorithm Based on Perspective Transformation*

The circular markings on the nacelle docking surface take on an elliptical shape on the 2D image when the image sensor is moving with the turbine hub. A schematic diagram is shown in Fig. 5 for the case when there is only an inclination angle between the camera and the nacelle.

Assuming that there is a circle with radius r on the plane, when the camera tilt angle changes, the circle will form an ellipse with long axis 2*a and short axis 2*b in

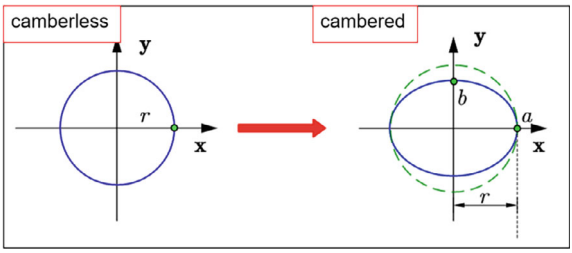


Fig. 5 Principle of circular perspective transformation

the camera imaging plane, the derivation can be obtained from the deflection angle calculation formula:

$$a = r \tag{4}$$

$$ang_y = \arccos(a/b) \tag{5}$$

In the formula ang_y is the camera tilt angle, a is the length of the circular radius in the camera imaging plane when there is no change in the tilt angle, and b is the length of the circular radius in the camera imaging plane due to the change in the tilt angle.

Firstly, the long and short axes of the ellipse are obtained by the ellipse detection algorithm, the radius of the circle in the image, assuming no deflection, is determined through calculations, and finally the calculation of the deflection angle is realised according to the principle of perspective.

In this paper, a least squares based ellipse detection method is used [6, 7], the algorithm works as shown in Fig. 6.

The concavity and convexity of a function can also be described as follows: when and only when the derivative of a unitary differentiable function is monotonically non-decreasing on a certain interval, this function is a convex function; conversely, it is a concave function. Let the function f(x) be on the interval I, take × 1, × 2,if for any point in the interval:

$$f(\lambda x_1 + (1-\lambda)x_2) \leq \lambda f(x_1) + (1-\lambda) f(x_2), \lambda \in (0, 1) \tag{6}$$

Then f(x) is a convex function and vice versa.

The line joining the midpoints of a set of parallel chords in an ellipse must pass through the centre of the ellipse, and by this property, the equation of the ellipse can be set as.

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1, (a > b > 1) \tag{7}$$

Let the equation of a chord on an ellipse that is not perpendicular to the x and y axes be:

$$y = kx + m \tag{8}$$

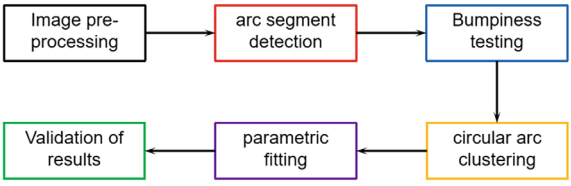


Fig. 6 Workflow of ellipse detection algorithm

This can be obtained by joining Eqs. (7) and (8):

$$x_0 = -\frac{kma^2}{b^2+a^2k^2} \tag{9}$$

$$y_0 = \frac{mb^2}{b^2+a^2k^2} \tag{10}$$

Since the ratio of y0 and × 0 is independent of m, the (×0, y0) trajectory is a straight line through the centre of the ellipse in the interior part of the ellipse. The three arcs of the ellipse are then used to estimate parameters such as the centre of the ellipse, the slope of the parallel chord and finally the long and short axes of the ellipse are calculated.

3.4 Coordinate Transformation Using DPI

The camera's internal and external parameters are obtained through calibration, and these parameters are used to convert the pixel coordinates of the target object into world coordinates [8]. However, this method does not meet the positioning accuracy requirements in all situations.

To improve the applicability of coordinate transformation, a new coordinate transformation technique is proposed, based on the relationship between the camera's DPI and the distance to the object [9]. The relationship between DPI and distance is obtained by varying the object-camera distance. A polynomial fitting method is used to represent the relationship between the DPI and the object distance. This process establishes a clear connection between DPI and distance, as shown in Fig. 7, with the distance expressed in millimeters.

Through the analysis of multiple data sets, a quadratic polynomial equation is established to describe the relationship between the camera's DPI and the distance transformation.

$$D = 1.38 * 10^{-3} * x + 6.13 * 10^{-8} * x^2 - 0.05 \tag{11}$$

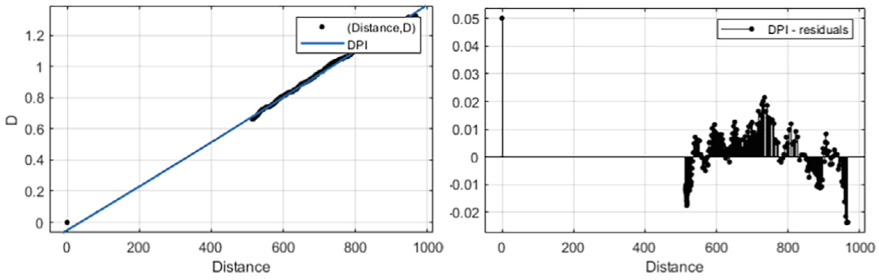


Fig. 7 Fitting the relationship between DPI and actual dimensions

In this formula, the distance between the principal point and the geometric center of the mounting surface is denoted by x, in millimeters. The reciprocal of the camera's DPI is represented by the variable D, with units of millimeters per pixel, indicating the correspondence between the number of pixels and the actual dimensions.

The accuracy of the derived equation is confirmed, with a squared error of 0.02 and an R-squared value of 0.998, indicating a high level of precision. Using this equation, the real-world distance L to the measured object can be calculated through image processing:

$$L = D * 1 \tag{12}$$

The geometric center of the nacelle mounting surface is set as the origin of the world coordinate system, and the object distance between the camera and the nacelle mounting center is calculated. The y-axis is defined along the direction perpendicular to the nacelle section. This allows the calculation of the hub's position matrix T, as shown in Fig. 8.

The rotation matrix R is obtained using the deflection angle data from the gyroscope. By combining it with the displacement matrix T, the pose matrix P of the hub can be calculated:

$$P = R * T \tag{13}$$

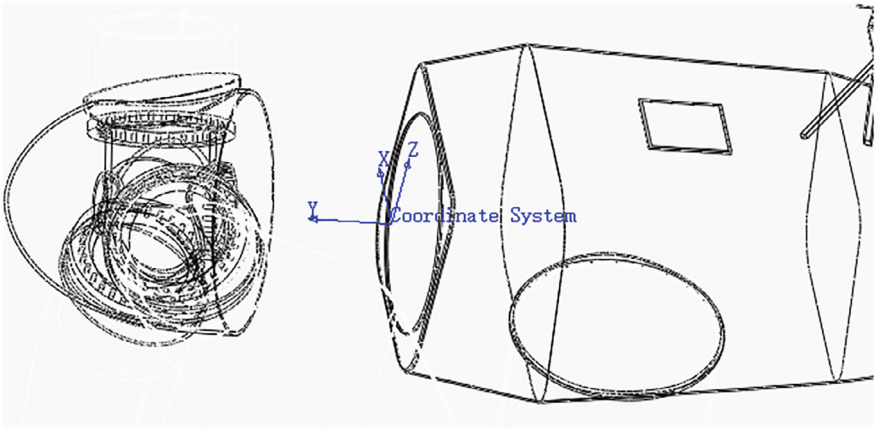


Fig. 8 Pose coordinate system

4 Conclusion

This paper combines visual object detection and various sensor data to develop a position detection system tailored to the characteristics of wind turbine hub lifting. The system processes images, adapts to the high-altitude working environment, and transmits the position information of the wind turbine hub to ground staff in real-time. It improves work efficiency, effectively protects the life and property safety of workers [10], and is of great significance for achieving efficient crane lifting automation.

Acknowledgements This work is supported by National Natural Science Foundation of China under Grant No.52275088 and the Fundamental Research Funds for the Central Universities under Grant (No.DUT22LAB506, No.DUT22LAB507).

References

1. Xu X, Chai Q, PGao P et al (2023) Research on construction technology of 6 MW onshore wind turbine impeller lifting. Hydropower New Energy 37(02):34–37
2. Zhang Z, Zhao J, Huang C et al (2023) Learning visual semantic map-matching for loosely multi-sensor fusion localization of autonomous vehicles. IEEE Trans Intell Veh 8:358–367
3. Liu Z, Liu F, Liu D et al (2024) Downhole image enhancement algorithm based on HSV colour space and improved two-dimensional gamma function. Coal Technol 43(01):203–206
4. Yu Y, Wang C, Fu Q et al (2023) Techniques and challenges of image segmentation: a review. Electronics 12(5)
5. Xu D (2024) Detection of surface defects of strain gauges based on image processing technology. Shandong University
6. Long C, Hu Q, Zhao M et al (2023) A triple-stage robust ellipse fitting algorithm based on outlier removal. Trans Instrum Meas 72:1–14
7. Michele F, Prati A, Cucchiara R (2014) A fast and effective ellipse detector for embedded vision applications. Pattern Recognit 47(11):3693–3708
8. Wang JZ (2020) Research on machine vision localization algorithm for stacking cartons. Huazhong University of Science and Technology
9. Zhao X, Yang HM, Qiang J et al (2020) A high-precision coherent laser ranging method based on Kalman filtering. J Opt 40(14):115–123
10. Ma J (2020) Design and implementation of workpiece recognition and localization system based on machine vision. University of Chinese Academy of Sciences. Shenyang Institute of Computing Technology, Chinese Academy of Sciences

Open Access This chapter is licensed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License (http://creativecommons.org/licenses/by-nc-nd/4.0/), which permits any noncommercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if you modified the licensed material. You do not have permission under this license to share adapted material derived from this chapter or parts of it.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.



Analysis and Research on Split Front Wing and Vortex Generator of FSAE Racing Car



Chenguang Lai, Zirui Hu, Jie Song, and Lijuan Fu

Abstract Aerodynamic design is one of the key elements in FSAE (Formula SAE) racing car design, and it plays a decisive role in the car's handling performance. This paper uses CFD (Computation Fluid Dynamics) method to conduct a simulation analysis of the split front wing and vortex generator design application, and the results show that the split front wing and vortex generator design can significantly improve the overall aerodynamic performance of the car and effectively improve the position of the pressure center. This study provides a new reference for the design of Formula SAE car front wings.

Keywords FSAE racing car · CFD · Split frond wing · Vortex generator

1 Introduction

The aerodynamic performance of FSAE racing cars plays a crucial role in the overall handling limits of the car. Excellent race car aerodynamics requires an adequate lift-to-drag ratio as well as a reasonable wind pressure center position, which is determined by the bodywork and the design of each aerodynamic package [1–5], Sneh and his team [6] conducted a numerical simulation study of an FSAE race car equipped with a rear-mounted engine. They adopted the SST k-ε turbulence model and succeeded in reducing the drag of the race car while seeking to increase the downforce and speed of the vehicle. It is worth noting that the effect of aerodynamic torque on the driving process of the race car was not fully considered in the study, and in order to improve the accuracy of the simulation calculations, Lu and his team [7] established a fine CAD model to study the effect of front and rear wings on the flow field of an open-wheeled racing car, and made suggestions to optimize the aerodynamic characteristics through quantitative calculations of pressure distributions, Rijns et al. [8] studied the aerodynamic characteristics of a high-performance

C. Lai · Z. Hu · J. Song (✉) · L. Fu
School of Vehicle Engineering, Chongqing University of Technology, Chongqing, China
e-mail: songjie@cqut.edu.cn

© The Author(s) 2026

H. Saman et al. (eds.), *The 9th International Conference on Advances in Construction Machinery and Vehicle Engineering*, Lecture Notes in Mechanical Engineering,
https://doi.org/10.1007/978-981-96-5527-4_6

vehicle, explored the interaction with the rear wing under straight-line driving and cornering conditions, and provided an important foundation for future research on static and active aerodynamic design. Design studies, providing an important basis for future static and active aerodynamic design studies.

The front wing, as the aerodynamic package located at the very front of the car, has the easiest aerodynamic performance to develop. Since it is directly facing the "clean" incoming flow from the front, most of the teams can develop a downforce level of nearly 200 N. However, the outflow field of the car is a relative whole, so how to make a good relationship with the other aerodynamic packages or the bodywork is very important. However, the outflow field of a race car is a relative whole, and it is worth noting how well it works with other aerodynamic packages or the bodywork [9].

The vortex and upwash generated by the high negative lift level of the front wing may cause downforce loss to the other aerodynamic packages located behind it, which may result in the center of the wind pressure being in front of the car due to poor aerodynamic distribution.

To sum up, whether the design of the front wing is reasonable or not can greatly determine the overall aerodynamic performance of an FSAE racing car, therefore, in order to improve the phenomenon of wind pressure center forward and enhance the overall aerodynamic performance of the racing car, this paper proposes a new design idea based on the split front wing and the addition of the vortex generator [10] method.

2 Numerical Theory Methods

2.1 Fluid Mechanics Control Equations

In incompressible flow, the fluid motion satisfies the three conservation laws. Most of the steady state numerical simulation studies of automotive aerodynamic external flow fields are based on the Reynolds time-averaged method (RANS) approach. The solution process requires a temporal and spatial discretization of the three major conservations where the continuity equation, also known as the mass conservation equation, has the differential form:

$$\frac{\partial \rho}{\partial t} + \mathrm{div}(\rho u) = 0 \tag{1}$$

In the formula, ρ is the density; t is the time; u is the velocity vector.

The law of conservation of energy is the basic law that must be satisfied by flow systems containing heat exchange, and is given by the following equation:

$$\frac{\partial (\rho T)}{\partial t} + \mathrm{div}(\rho u T) = \mathrm{div}\left(\frac{\kappa}{Cp}\mathrm{grad}T\right) + ST \tag{2}$$

In the formula, Cp is the specific heat capacity; T is the temperature; κ is the heat transfer coefficient of the flow; ST is an internal heat source for the fluid.

Momentum equation, also known as the Navier–Stokes equation, derived from Newton's second theorem, The expressions are as follows:

$$\frac{\partial}{\partial \chi j}(\rho \mathrm{uiuj}) = \frac{\partial}{\partial \chi \mathrm{j}}\left[\mu\left(\frac{\partial \mathrm{ui}}{\partial \chi \mathrm{j}} + \frac{\partial}{\partial \chi \mathrm{i}}\right) - \frac{2}{3}\mu\frac{\partial \mathrm{ui}}{\partial \chi \mathrm{k}}\right] - \frac{\partial \mathrm{p}}{\partial \chi \mathrm{i}} \tag{3}$$

In the formula, p is the pressure on the fluid micromeres; ρ is the density; μ is the dynamic viscosity.

2.2 Definition and Calculation of Racing Related Parameters

Wind resistance and lift coefficient are the basic parameters of a racing car's aerodynamic performance. A good aerodynamic design of a racing car should have a lower wind resistance and a higher negative lift at the same time.

The coefficient of drag of a car can be expressed by the following equation:

$$\mathrm{C_D} = \frac{2\mathrm{F_D}}{\left(\rho V^2 \mathrm{A}\right)} \tag{4}$$

And the following equation can be used to express the lift coefficient:

$$\mathrm{C_L} = \frac{2\mathrm{F_L}}{\left(\rho V^2 \mathrm{A}\right)} \tag{5}$$

In the formula, $\mathrm{C_D}$ is the coefficient of resistance, $\mathrm{C_L}$ for lift, $\mathrm{F_D}$ is the resistance, $\mathrm{F_L}$ for lift, ρ is the density of air, V is the speed of the vehicle, A is the windward area (positive projection area).

The position of the wind pressure center is also an important design parameter of the racing car, which represents the equivalent point of action of the negative lift force on the car, and its position will directly affect the stability of the car's handling, the use of the wind pressure center can be determined by the size of the aerodynamic downforce distributed to the front and rear axles, which is generally designed to overlap with the center of mass or a slightly backward position.

The downforce acting on the rear axle and the proportional distribution of the downforce can be expressed by the following equation:

$$\mathrm{F_{Rl}} = \frac{\mathrm{a}}{\mathrm{a}+\mathrm{b}}\mathrm{F_L} \tag{6}$$

$$\mathrm{k_R} = \frac{\mathrm{a}}{\mathrm{a}+\mathrm{b}} * 100\% \tag{7}$$

In the formula, F_{Rl} is the downforce distributed to the rear axle, a and b are the horizontal distance between the center point of the front and rear axes and the center point of wind pressure l, k_R is the proportion of downforce allocated to the rear axle.

3 Simulation Analysis

Racing car aerodynamics is a connected whole, in the research of FSAE aerodynamic performance, a large number of studies are based on the enhancement of the aerodynamic performance of a component alone, while ignoring the cooperation between aerodynamic packages. For the development of aerodynamic performance of FSAE front wing, on the one hand, it is required to have a high level of negative lift, and on the other hand, it is necessary to reduce the impact on the aerodynamic performance of the side wing and rear wing, so as not to cause the center of the wind pressure to be in an unreasonable position [11]. Therefore, this paper is based on the 2024 season of the FSAE race car model (Fig. 1), for the split front wing and front wing vortex generator design and application for in-depth discussion.

3.1 Split Front Wing CFD Simulation and Analysis

At present, for most of the participating teams, the front wing monolithic main wing design is still the mainstream. Due to the higher angle of attack of the airfoils, the monolithic main wing design has the advantage of high negative lift, but the smaller airflow inlet formed by it and the lower side of the nose is prone to airflow blockage, which will result in the consequences of increased drag and insufficient underfloor intake, and thus affect the level of negative lift on the lower surface of the bodywork and the side wings.

The split front wing design in this study is to replace the middle section of the main wing with a symmetrical wing with 0° angle of attack and longer chord length, and at the same time move it downward to increase the area of the airflow channel

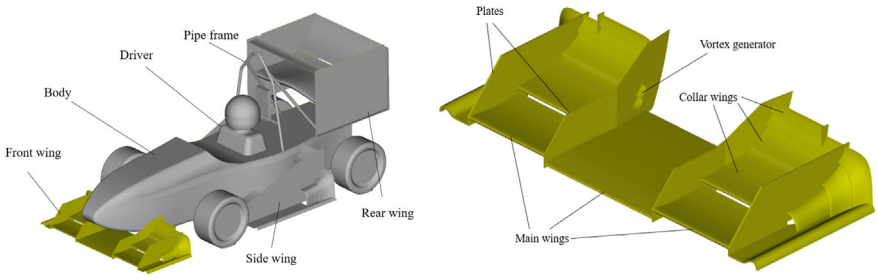


Fig. 1 Schematic model of the FSAE car for the 2024 season of the Shiji formula racing team

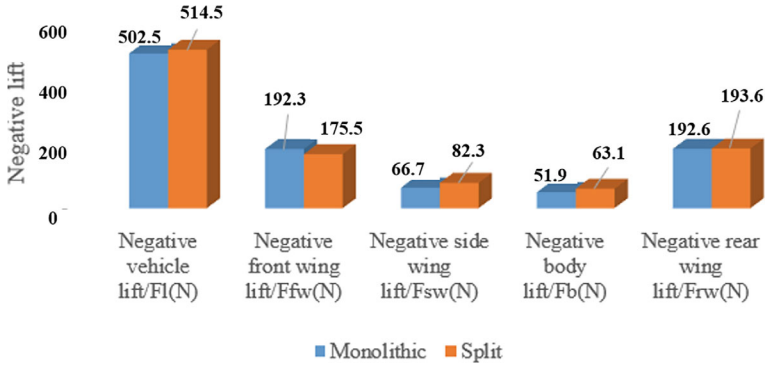


Fig. 2 Negative lift of monolithic main wing and split main wing

with the bottom of the nose of the car, and the closer distance between the airfoils and the ground is also more likely to produce a stronger ground effect. Figure 2 and Table 1 show the CFD simulation results of the whole car with a monolithic and split front wing at a speed of 15 m/s.

From Fig. 2 and Table 1, it can be seen that the aerodynamic performance of the split front wing is obviously more advantageous than that of the one-piece front wing, with the overall negative lift and lift-to-drag ratios improved while the rear axle aerodynamic force percentage is restored by 5.1%.

Figures 3 and 4 show the flow state and surface pressure distribution of both airflows near the front wing and the underside of the nose at 15 m/s vehicle speed. From Fig. 3, it can be seen that the one-piece main wing has a smaller airflow channel due to the curvature of the airfoil with a larger angle of attack and the lower part of the nose, which causes a certain airflow congestion and additional drag, which is also confirmed by the positive pressure in the ② region of Fig. 4. At the same time, due to the existence of positive pressure in the ② region, the upward wash of the airflow after passing through the lower surface of the airfoil collides with the airflow inside the channel, which causes the stalling of the lower surface of the wing and aggravates the positive pressure level of ② region, which undoubtedly increases the pressure distribution. At the same time, it exacerbates the positive pressure level in region ②, which is undoubtedly disastrous for the drag and negative lift level of the body.

The split front wing is a better way to avoid the above problems. Firstly, the airflow congestion problem is better solved due to the enlarged inlet area of the channel, and

Table 1 lift-Drag ratio and rear axle aerodynamics ratio of monolithic main wing and split main wing

	Lift-drag ratio/L/D	Rear axle aerodynamics ratio/ k_R (%)
Monolithic	−2.52	48.2
Split	−2.63	53.3

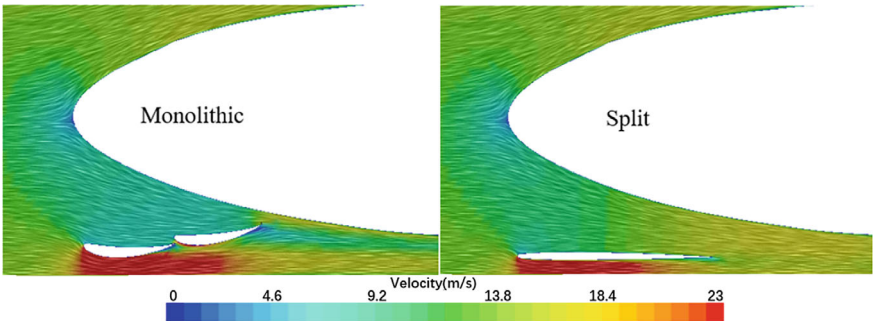


Fig. 3 Velocity vectors in the mid-section of monolithic and split main wing

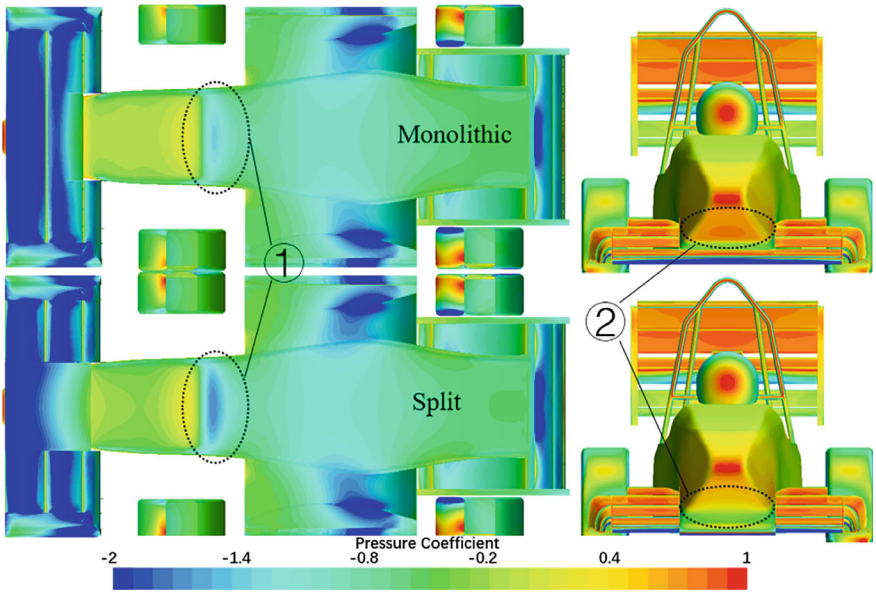


Fig. 4 Surface pressure coefficient cloud of Monolithic and split main wing

the symmetrical wing shifted downward forms a stronger Venn diagram with the ground, which effectively utilizes the ground effect without causing additional drag, and further increases the airflow into the underbody. The difference in air flow can be seen in the negative pressure level in area ① in Fig. 4. The faster air flow rate also increases the level of negative lift on the body and side wings. It is worth noting that as the negative lift performance area is moved back, the characteristic of the center of air pressure being forward is also improved.

3.2 Vortex Generator CFD Simulation and Analysis

Vortex generators have a high research value in the aerodynamic design of racing cars, and the utilization of vortexes in the outer flow field of racing cars is crucial. The classic Y250 vortex generator was designed to create a vortex located on the outside of the body to isolate the outside air so as to protect the negative pressure level of the inner bottom plate, which was a widely used vortex generator in F1 races, but it is not suitable for the short wheelbase and short wheelbase of the FSAE race car. Therefore, this study proposes a vortex generator design idea located on the inside of the front wing end plate, and the aerodynamic performance after CFD simulation is shown in Table 2 below:

Table 2 shows that with the addition of the inboard vortex generator in the front wing, the negative lift levels of the side wing and rear wing gained 6.5 N and 7.9 N, respectively. Figure 5 shows the design and installation of the vortex generator on the inside of the front wing end plate. Due to the downward position of the split main wing and the faster air flow rate in the channel to facilitate turning, the air passes through the vortex generator and enters the side wing, and then the air with higher Reynolds number is less prone to stalling on the underside surface of the airfoil, which makes the side wing can be designed for a higher angle of attack.

Figure 6 captures the vortex created by the vortex generator, which can be seen more clearly from its creation to its dissipation along the body to the side wings. The dissipation of the vortex increases the local Reynolds number, allowing more air to attach to the lower surface of the side wings as shown in Fig. 7, which increases the area of the negative pressure region on the lower surface of the airfoils and enhances the level of negative lift on the side wings.

Table 2 With vortex generators and without vortex generators pneumatic performance parameters

	Negative side wings lift /F_{sw}(N)	Negative rear wings lift /F_{rw}(N)
Vortex-free generator	75.8	185.7
With eddy current generator	82.3	193.6

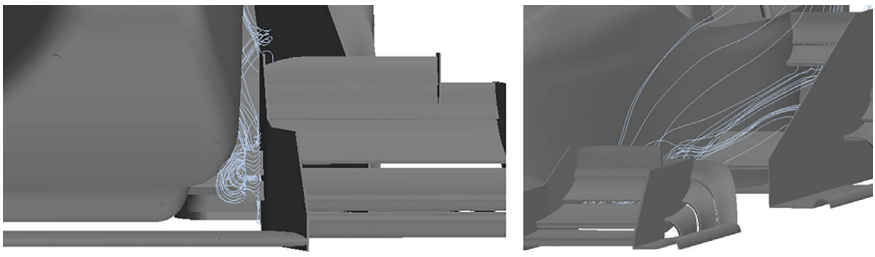


Fig. 5 Vortex generator installation location and flow diagram

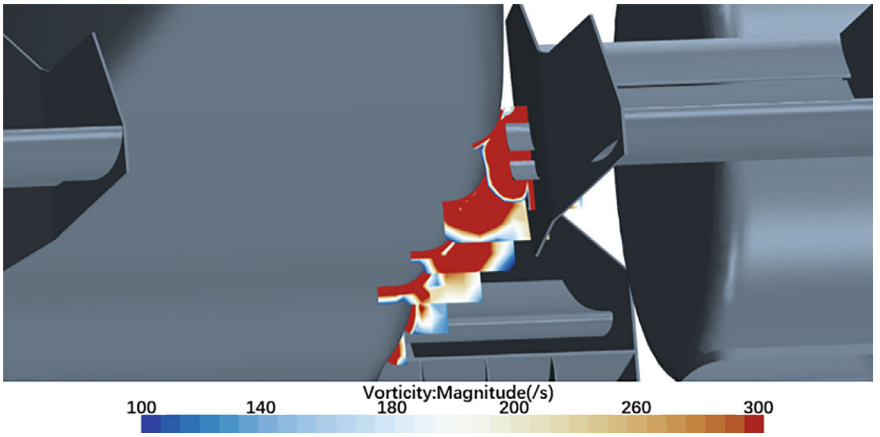


Fig. 6 Vortex generator vortex cloud

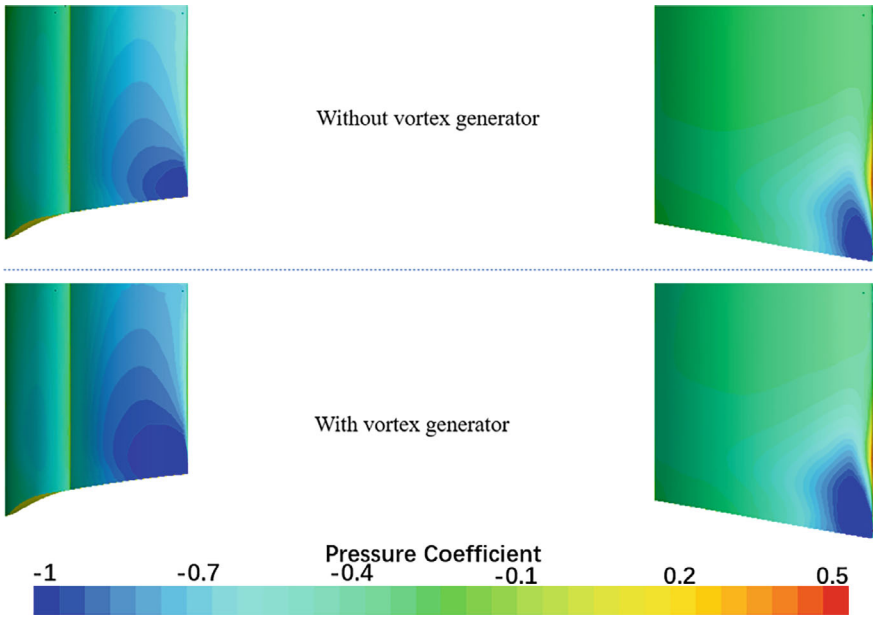


Fig. 7 Pressure clouds on the undersurface of the side wings with and without vortex generators

4 Conclusions

(1) Based on the current aerodynamic design status of the front wing of the FSAE racing car, this paper proposes a split main wing and inboard short-plate vortex generator design method, which puts forward a new design idea for improving

the aerodynamic performance of the whole car and balancing the aerodynamic center.

(2) The adoption of split main wing for the front wing can effectively utilize the ground effect, on the one hand, reduce the drag without sacrificing too much negative lift of the front wing and improve the lift-to-drag ratio, on the other hand, it promotes the transfer of aerodynamic loads and avoids the dangerous aerodynamic characteristics that the wind pressure center is too far forward.

(3) The vortex generator can effectively promote air turning, increase the local Reynolds number so that the air flow through the lower surface of the wing is less likely to produce flow separation, which can effectively improve the negative lift level of the wing.

References

1. Basso M, Cravero C, Marsano D (2021) Aerodynamic effect of the gurney flap on the front wing of a F1 car and flow interactions with car components. Energies 14(8)
2. Qi X, Ou Y, Zhang H et al (2022) Efficiency enhancement design approach in the side wing of a FSAE car utilizing a shutter-like fairing structure. Appl Sci 12(13)
3. Zhang GQ, Wang D, Zhou FQ et al (2024) Aerodynamic design and performance research of racing cars with adaptive control of attitude. Int J Automot Technol 25(4):901–912
4. Katz J (2006) Aerodynamics of race cars. Annu Rev Fluid Mech 38(27–63)
5. Katz J (2021) Aerodynamics in motorsports. Proc Inst Mech Eng Part P-J Sport Eng Technol 235(4):324–338
6. Hetawal S, Gophane M, Ajay BK et al (2014) Aerodynamic study of formula SAE car. Procedia Eng 97(1198–207)
7. Lu X (2015) Research on the Flow Field around a Formula SAE Car. SAE International
8. Rijns S, Teschner TR, Blackburn K et al (2024) Effects of cornering conditions on the aerodynamic characteristics of a high-performance vehicle and its rear wing . Phys Fluids 36(4)
9. Martins D, Correia J, Silva A (2021) The influence of front wing pressure distribution on wheel wake aerodynamics of a F1 car. Energies 14(15)
10. Koike M, Nagayoshi T, Hamamoto N (2004) Research on aerodynamic drag reduction by vortex generators. Mitsubishi Mot Tech Rev 16(2)
11. Lai CG, Zhu SJ, Feng S et al (2023) Flow characteristics and wake topology of two-seat convertibles. Phys Fluids 35(1)

Open Access This chapter is licensed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License (http://creativecommons.org/licenses/by-nc-nd/4.0/), which permits any noncommercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if you modified the licensed material. You do not have permission under this license to share adapted material derived from this chapter or parts of it.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.



Thermal Fatigue Life Prediction Under Emergency Braking of Disc Brake Based on FEA Simulation



Chuang Chen, Qiang Yang, Huayu Zhang, Xin Zhao, Cheng Gao, and Yuan Gao

Abstract The accumulation of emergency braking at high speed and heavy load may result in the initiation and expansion of thermal fatigue cracks on the surface of brake, then cause the potential risk such as brake wheel crack or even burst. Aiming at the emergency braking condition of the SBD80-A disc brake used in a company's port crane, a finite element analysis (FEA) simulation model was established for the thermal–mechanical performance analysis applying ABAQUS software. Firstly, FEA simulation was employed to obtain the variation process of thermal stress, and the stress cycle load of a single emergency brake was analysed by using the rain flow counting method simultaneously. And the stress–strain curve of the brake material was used to convert stress amplitude into the strain amplitude. Then, the strain life of one-time emergency brake was calculated under different strain amplitudes by utilizing the Manson-Coffin strain-life formula corresponding to the actual temperature. Finally, the Miner's rule was adopted to propose thermal fatigue life prediction

C. Chen · Q. Yang · H. Zhang · X. Zhao (✉) · Y. Gao
School of Mechanical Engineering, Northeastern University, Shenyang, China
e-mail: zxvir@163.com

C. Chen
e-mail: 2370090@stu.neu.edu.cn

Q. Yang
e-mail: qiangyang@mail.neu.edu.cn

H. Zhang
e-mail: 1401950094@qq.com

Y. Gao
e-mail: 2370122@stu.neu.edu.cn

C. Chen · Q. Yang · X. Zhao · C. Gao
Key Laboratory of Lifting Equipment's Safety Technology, State Administration for Market Regulation, Shenyang, China
e-mail: dutgao@sina.com

X. Zhao · C. Gao
Liaoning Inspection, Examination and Certification Centre, Liaoning Academy of Safety Science, Shenyang, China

© The Author(s) 2026

H. Saman et al. (eds.), *The 9th International Conference on Advances in Construction Machinery and Vehicle Engineering*, Lecture Notes in Mechanical Engineering,
https://doi.org/10.1007/978-981-96-5527-4_7

method under emergency braking of disc brake considering the influence of high temperature. The research provides a theoretical basis for failure mechanism analysis and life prediction of thermal fatigue for the mechanical products with similar structures, offers a design guidance for determining the inspection and maintenance cycle of brakes under the condition of emergency braking.

Keywords Thermal fatigue · Life prediction · Emergency braking · Simulation · Failure mechanism analysis

1 Introduction

Emergency braking is prevalent in most industrial applications such as metallurgy and mining, plays an important role in ensuring safe production and preventing vicious safety accidents. However, emergency braking conditions is prone to trigger the initiation and expansion of thermal fatigue cracks under extreme loads, and subsequently lead to the collapse or even cracking. Therefore, carrying out thermal fatigue failure mechanism analysis and life prediction method research on brakes contains important theoretical value and practical engineering significance for ensuring the safe operation of lifting machinery and maintenance cycle of brakes under extreme loads.

Bi, et al. conducted simulations analysis on the brake based on the finite element software ANSYS, established a life prediction model to calculate the lifespan of the dangerous parts, and explored the impact of different working conditions on fatigue life [1]. Considering the geometric and kinematic attributes of the brake, Liu, et al. proposed the displacement gradient cycle method to calculate the friction heat flow, thereafter, the heat flow values were utilized as inputs for simulating the brake discs, and the resulting data were compared with the findings of the research to validate the results [2]. Wang, et al. converted the obtained stress curve into a strain curve using the 500 °C R-O mode, subsequently, the prediction of the thermal fatigue life of the brake drum was accomplished by utilizing the rain-flow counting and integrating it into the Manson-Coffin-Basquin strain life model [3]. Yevtushenko, et al. developed a mathematical model of the effect of friction on the stresses in the brake disc and completed numerical calculations, the analysis demonstrated that radial thermal microcracks on the friction surface of the brake disc are a consequence of circumferential stresses [4]. Moghanlou, et al. conducted a three-dimensional heat-stress coupled simulation analysis and used the numerical results to predict the fatigue life based on the Smith–Watson–Topper model, and ultimately investigated the impact of braking hotspots on the reduction of brake lifespan [5]. Yin, et al. conducted a joint simulation with FE-SAFE and ABAQUS to analyse the thermal stress and thermal fatigue life of a commercial vehicle brake, and the thermal fatigue life of the brake disc was calculated by treating HT250 as a linearly elastic material and applying the M-C low circumference fatigue formula [6].

This paper focuses on the SBD80-A disc brake under emergency braking condition as the subject of investigation and establishes a simulation model of thermal-machine coupling emergency braking performance based on finite element software. The damage degree of brake disc thermal fatigue life is calculated by a single emergency braking. Finally, a thermal fatigue life prediction method is proposed based on the Miner linear damage accumulation theory.

2 FEA Simulation Modelling of Disc Brakes

2.1 Properties of Brake Material

The SBD80-A disc brake employed in the main lifting mechanism of a container is constituted of HT250 disc material, featuring a material density of 7220 kg/m^3 and a Poisson ratio of 0.3. The temperature-dependent material parameters of the brake are presented in Table 1. The friction sheet material is asbestos-free, with basic properties such as the density of 1788 kg/m^3, the Poisson's ratio of 0.3, the Young's modulus of 2.2 GPa, and the thermal expansion coefficient of 11 $10^{-6}{\cdot}K^{-1}$. Other properties such as specific heat capacity, the thermal conductivity exhibit considerable variation with temperature, and the thermophysical parameters are shown in Table 2.

Table 1 Thermophysical material parameters of brake discs under emergency braking

Temperature (°C)	Young's modulus (GPa)	Thermal conductivity (W/(K·m))	Thermal expansion coefficient ($10^{-6}{\cdot}K^{-1}$)	Specific heat (J/(K·kg))
20	100	42.38	4.386	503
100	~	43.06	11.653	530
200	99	44.23	12.836	563
300	~	43.55	13.580	611
400	96	40.67	13.578	641

Table 2 Thermophysical parameters of friction pads

Temperature (°C)	Specific heat (J/(K·kg))	Thermal conductivity (W/(K·m))
20	998	1.701
120	1178	1.595
210	1295	1.489
300	1340	1.410

Table 3 Parameters of container crane braking conditions

Parameters	Values	Parameters	Values
Lifting capacity LL (N)	637,000	Quality of spreader and rack LS (N)	176,400
Lifting speed of the weight V (m/min)	90	Reel wire rope line speed V_0 (m/min)	180
Reduction ratio i	21.389	Reel diameter D_1 (m)	1.365
Inertia moment of motor shaft J (kg·m^2)	70	Rotation speed of the motor n_0 (r/min)	898
Clamping force F_n (KN)	80	Lag time Δt (s)	0.3
Inertia moment of the other parts to the motor shaft J_1 (kg·m^2)	64.96		

2.2 Conditions of Loads and Boundary

The emergency braking working condition parameters for a certain type of container loading and unloading task are presented in Table 3.

According to the braking conditions in Table 3, it can be calculated that two SBD80-A brakes with clamping force of 80 KN will stop after 0.88 s under the condition of initial angular velocity of 119.84 rad/s and friction coefficient of 0.3.

The empirical formula [7] can be used to obtain the convective heat transfer coefficient h of disc brake when disc braking, as shown in Eq. (1):

$$h = \begin{cases} 0.70R_e^{0.55}k_a/D, \ R_e \le R_{e_c} = 2.4 \times 10^5 \\ 0.04R_e^{0.8}k_a/D, \ R_e > R_{e_c} = 2.4 \times 10^5 \end{cases} \tag{1}$$

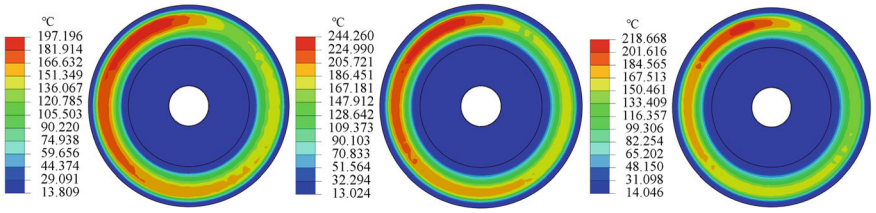
where: h is the thermal conductivity of air, W/(m·°C); D is the outer diameter of the brake disk, m; R_e is the Reynolds number.

According to the literature [7], it can be found that the convective heat transfer coefficient relationship of the brake disk, as shown in Eq. (2):

$$h = \begin{cases} 5.37 \times (119.54 - 135.84t)^{0.8}, \ t < 0.76 \\ 3.37 \times (119.54 - 135.84t)^{0.55}, \ t \ge 0.76 \end{cases} \tag{2}$$

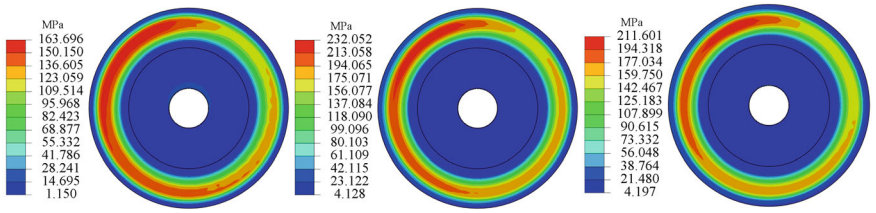
2.3 Analysis of Simulation Results

In the FEA software ABAQUS, the material properties of the brake disk and friction pads are set according to Tables 1 and 2. Interaction property was utilized to set the convective heat transfer coefficients of the brake disc, as shown in Eq. (2). Finally,



(a) $t = 0.300$ s (b) $t = 0.660$ s (c) $t = 0.880$ s

Fig. 1 Temperature contours of brake disc under emergency braking conditions



(a) $t = 0.300$ s (b) $t = 0.720$ s (c) $t = 0.880$ s

Fig. 2 Thermal stress contours of brake disc under emergency braking conditions

the temperature and thermal stress contours under emergency braking condition, as shown in Figs. 1 and 2.

As can be seen from Fig. 1: in the initial stage of braking, the friction area on the surface of the brake disc exhibits a high temperature, and the temperature gradually decreases with the rotation of the brake disc. When the speed of heat generation is greater than the speed of heat dissipation, the high temperature area of the brake disc will increase, the temperature in the radial direction is a considerable gradient. From Fig. 2, it can be found that the thermal stress in the main friction area is the largest on the surface of the brake disc, and decreases gradually to the radial sides. In the initial period of braking, due to the combined effects of high temperatures and mechanical stresses, the brake disc generates great thermal stresses. In the later stage of braking, the speed of the brake disc decreases significantly, heat conduction and convection heat dissipation are greater than the heat generated, thermal stress is reduced.

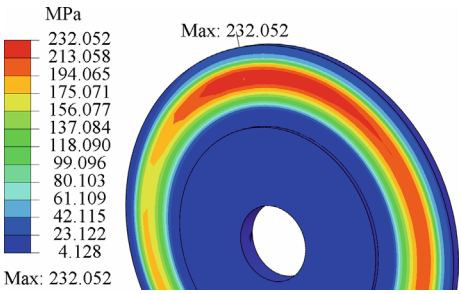


Fig. 3 Equivalent stress of the brake disc at $t = 0.72$ s

3 Calculation of Thermal Fatigue Life

3.1 Form and Dangerous Parts Confirmation of Fatigue Failure

The brake disc and the friction disc generate heat due to intense friction, and when the thermal deformation of the brake disc is subject to constraints, there will be far greater than the mechanical stress of the thermal stress, and the appearance of thermal stress will lead to local "hot spots" on the surface of the brake disc. When the local "hot spot" occurs frequently, it will trigger the emergence of thermal fatigue cracks and lead to low circumference thermal fatigue failure. Figure 3 shows the contours of equivalent force distribution of the brake disk at $t = 0.72$ s. Meanwhile, the maximum equivalent force value of 232.1 MPa appears on the surface of the brake disc, which is located at the centre of the friction region.

To investigate the mechanical mechanism of thermal fatigue cracking, a three-way stress analysis of the maximum stress point is performed. The curves of stresses in three directions with time at the point of maximum equivalent stress are shown in Fig. 4. Obviously, the effect of circumferential stress σ_{11} on the thermal fatigue life of the friction region is of the utmost significance. In summary, the brake disc initially generates localized micro-cracks under the effect of circumferential stress, and then the micro-cracks expand along the radial direction under the action of radial stress.

3.2 Thermal Fatigue Life Prediction Method Considering the Influence of High Temperature

The maximum circumferential stress during emergency braking was firstly split using the rain flow counting method, and the stress cycles with higher degree of influence on thermal fatigue life were retained. The stress amplitude was then taken into the stress–strain curve at the corresponding temperature to find out the corresponding strain amplitude, and the strain amplitude was brought into the Masson-Coffin life

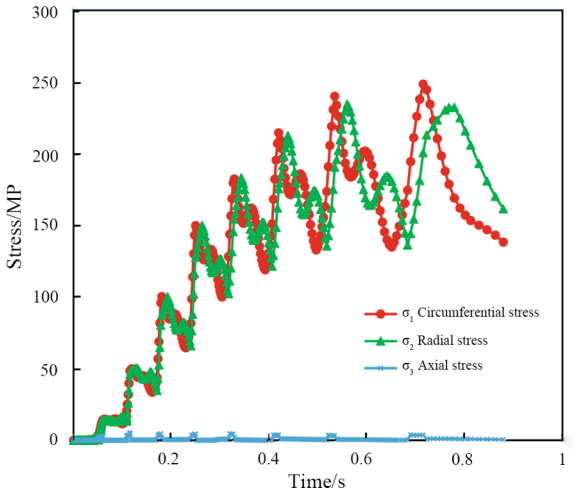


Fig. 4 Three-dimensional stress curve

prediction formula at the corresponding temperature. Finally, the thermal fatigue life was calculated using the Miner's rule. The maximum circumferential stress cycle of emergency braking was shown in Fig. 5, the stress fluctuations after applying the rainfall counting method for splitting was shown in Fig. 6.

The stress–strain curves of HT250 material at different temperatures can be obtained by using linear interpolation as shown in Fig. 7. Due to the surface temperature of the brake disc reaches approximately 200 °C when the stress cycle occurs, it has a significant impact on the thermal fatigue life. Accordingly, based on the four stress fluctuations depicted in Fig. 6 and the stress–strain curve of HT250 at 200°C presented in Fig. 7, four strain fluctuations can be derived, as shown in Fig. 8, and the strain amplitude is provided in Table 4.

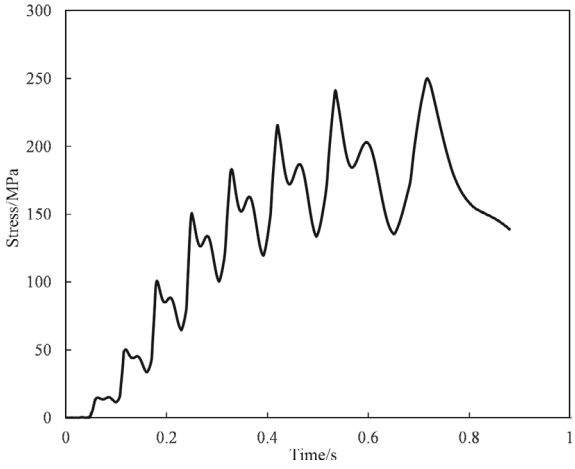


Fig. 5 Curve of circumferential stress

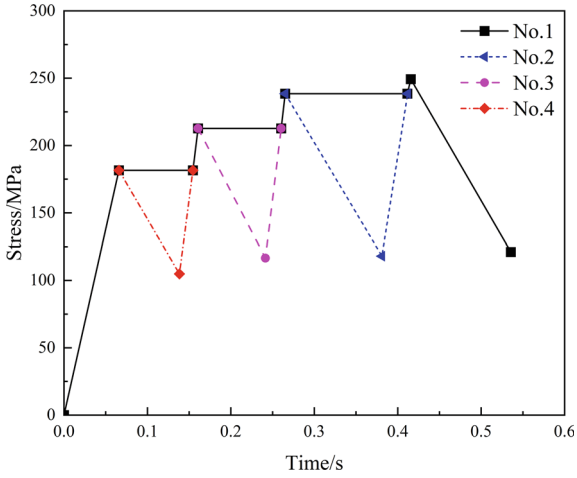


Fig. 6 Stress rain-flow counting method

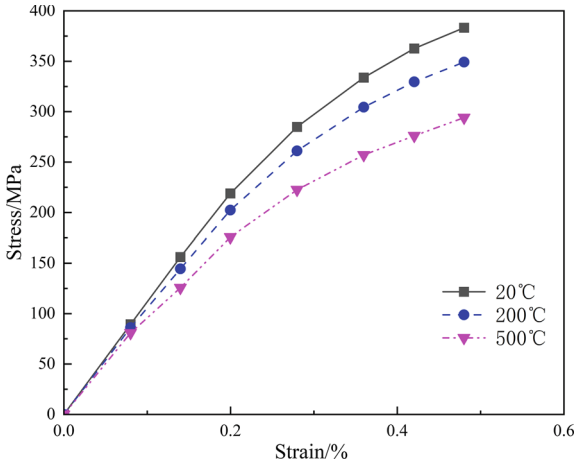


Fig. 7 Stress–strain curves of HT250 material

The Manson-Coffin formulation at 200 °C is applied to low-cycle fatigue with Eq. (3). The parameters and meanings of the Manson-Coffin formula are shown in Table 5 [8]. The four strain amplitudes presented in Table 4 are utilized in Eq. (3), The subsequent four stress load damages are calculated in accordance with the specifications outlined in Table 6.

$$\frac{\Delta\varepsilon}{2} = \frac{\Delta\varepsilon_e}{2} + \frac{\Delta\varepsilon_p}{2} = \frac{\sigma_f'}{E}(2N_f)^b + \varepsilon_f'(2N_f)^c \tag{3}$$

Combining the damage amount of each stress load as presented in Table 6 and the Miner's rule, the cumulative number is 856. In conclusion, when the disc brake

Fig. 8 Strain cycle curves

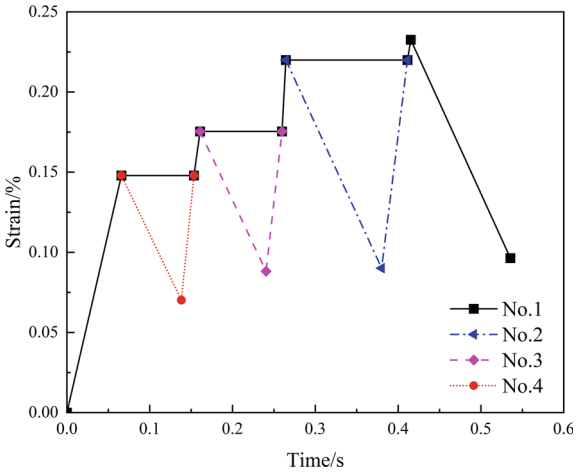


Table 4 Strain amplitude by rain flow counting method

Strain number	No.1	No.2	No.3	No.4
Strain amplitude	0.000758	0.000914	0.001208	0.002423

Table 5 Manson-Coffin coefficient value at 200 °C

Temperature (°C)	Fatigue strength coefficient $\sigma_f{}'$	Fatigue ductility coefficient $\varepsilon_f{}'$	Fatigue strength exponent b	Fatigue ductility exponent c
200	514	0.0451	−0.1316	−0.5924

Table 6 The amount of damage during an emergency braking

Strain number	No. 1	No. 2	No. 3	No. 4
Impairment D_i	$\frac{1}{976000}$	$\frac{1}{252000}$	$\frac{1}{37600}$	$\frac{1}{880}$

is in the initial state with an initial speed of $\omega = 119.84$ rad/s and a clamping force of $F = 80$KN, the service life is 856 times.

4 Conclusion

A thermal–mechanical coupled FEA simulation model of a disk brake under emergency braking conditions is established. Simulation results show that: at the braking time $t = 0.66$ s, the maximum surface temperature of the disc brake is 244.3 °C, braking time $t = 0.72$ s, the maximum stress of the disc brake is 232.1 MPa, the highest temperature and the maximum stress both appeared in the entrance to the

friction plate which located on the braking radius. A novel method was put forward for predicting the thermal fatigue life of discs brake based on the Manson-Coffin formula considering the influence of temperature. The stress cycle of one-time emergency braking is divided into multiple stress loads by utilizing the rain-flow counting method and converted into corresponding strain amplitudes according to the stress–strain curve of the brake material. The Manson-Coffin formula at the appropriate temperatures were then adopted to calculate the strain life corresponding to different strain amplitudes. Finally, considering the temperature effect, the thermal fatigue life prediction of disc brake was completed based on the Miner's rule. For the emergency braking condition of the SBD80-A brake used for a harbour crane, the thermal fatigue life is 856 times.

Acknowledgements Authors wishing to thank to the Applied Basic Research Program of Science and Technology Projects in Liaoning Province (Grant No. 2023JH2/101300137), Open Project of Key Laboratory of Lifting Equipment's Safety Technology for State Administration for Market Regulation (Grant No. 2023KF002) and the Fundamental Research Funds for the Central Universities (Grant No. N2203010).

References

1. Bi X, Hu H, Lv W, Sun M (2017) Research on fatigue life of industrial disc brake based on Manson-Coffin. J Min Process Equip 45(06):59–62
2. Liu N, Jin X, Shi J, Wang B, Shi Y, Feng M (2024) Study on the thermal-structural coupling method for disc brake of high-speed train. J Railw Sci Eng 21(02):456–464
3. Wang X, Fan Z, Bian J, Gui L (2021) Drum brake fatigue life predictions. J Tsinghua Sci Technol 61(1):21–22
4. Yevtushenko A, Kuciej M, Topczewska K (2018) Analytical model to investigate distributions of the thermal stresses in the pad and disk for different temporal profiles of friction power. J Adv Mech Eng 10(10):23–34
5. Moghanlou RM, Googarchin SH (2020) Three-dimensional coupled thermo-mechanical analysis for fatigue failure of a heavy vehicle brake disk: Simulation of braking and cooling phases. Proc Inst Mech Eng Part D J Automob Eng 234(13):3145–3163
6. Yin X, Zhang S, Guo F, Pang Z, Rong Y, Zhang B (2023) Thermal stress simulation and fatigue life of commercial vehicle disk brakes under emergency braking conditions. J AIP Adv 13:125220
7. Limpert R (2011) Brake design and safety. SAE international, Warrendale, p 150
8. Han MJ, Lee CH, Park TW, Lee SP (2018) Low and high cycle fatigue of automotive brake discs using coupled thermo-mechanical finite element analysis under thermal loading. J Mech Sci Technol 32:5777–5784

Open Access This chapter is licensed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License (http://creativecommons.org/licenses/by-nc-nd/4.0/), which permits any noncommercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if you modified the licensed material. You do not have permission under this license to share adapted material derived from this chapter or parts of it.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.



Remote Troubleshooting APP Development for Excavators Based on Android System



Chao Yang, Chenbo Yin, Shoulei Ma, Donghui Cao, and Xiaozhong Liu

Abstract Fault diagnosis technology is an important means to maintain the normal operation of the machine. Currently, excavator faults occur frequently, which not only affects the progress of the project but also brings potential safety hazards, and the service engineers need to visit the site for inspection and troubleshooting, which has high time cost and condition limitations. Based on this, a remote troubleshooting tool is developed to improve the troubleshooting efficiency. Firstly, the fault diagnostic device is connected to the excavator to realize the mutual conversion of WIFI signal and CAN signal, and the APP developed by Android system is used to send commands, based on the signal comparison method and the case inference method, the one within a certain threshold error is regarded as no fault, and the other way around is based on the source case to guide the fault case, so as to realize the remote one-button inspection. The APP has been tested by the dSPACE simulation system in the laboratory and the on-board test of SY980, with excellent diagnostic effect, and has been put into use.

Keywords Remote diagnosis · Android system · Signal comparison method · Case-based reasoning method

1 Introduction

In the field of excavator operation, the harsh working environment, diverse working conditions and frequent failure problems not only challenge the operator's technology and resilience, but also bring considerable trouble to the construction project. In order to minimize the loss of engineering projects, fault diagnosis technology has become an essential and important means [1]. However, service engineers usually need to visit

C. Yang · C. Yin (✉) · S. Ma
Institute of Vehicles and Construction Machinery, Nanjing Tech University, Nanjing, China
e-mail: 1197933989@qq.com

D. Cao · X. Liu
Sany Heavy Machinery, Kunshan, China

© The Author(s) 2026

H. Saman et al. (eds.), *The 9th International Conference on Advances in Construction Machinery and Vehicle Engineering*, Lecture Notes in Mechanical Engineering,
https://doi.org/10.1007/978-981-96-5527-4_8

the site one by one after receiving the news of fault problems in order to determine the source of faults, which is a complicated and inefficient process. Therefore, this paper aims to integrate wireless transmission, CAN bus and Android system with multiple elements to develop an excavator remote troubleshooting APP, which realizes real-time information interaction and precise judgment of fault sources, and effectively overcomes the problem of low troubleshooting efficiency.

With the complexity of engineering projects and the expansion of production needs, the fault diagnosis technology has been updated from the traditional experience-based to a new technology based on integrated and intelligent diagnostic methods. Whether at home or abroad, excavator fault diagnosis technology is taking a qualitative leap. Mo and Wong [2] from Xi'an University of Architecture and Technology, for the problem of hydraulic system failure, put forward a diagnostic network based on ACA-BP algorithm, and test it by using Komatsu excavator model, which has a high fault recognition rate and good performance. He et al. [3] from Changsha University of Science and Technology proposed a dynamic generalized regression neural network (GRNN) diagnosis scheme for hydraulic system, which has strong nonlinear convergence ability and accurately identifies more than 90% of the faults. Feng [4] from Shijiazhuang Railway University, carried out non-stop online listening and fault diagnosis on various parameters of excavators, designed a virtual instrument and carried out fault diagnosis through WiFi at the construction site. Wang et al. [5] from Sichuan University chose to use wavelet packet analysis to calculate the energy entropy of the fault signal and constructed feature vectors to input into the vector machine, and then determine the type of faults, and its accuracy reached 87.3%. Bagyalakshmi et al. [6] designed a fuzzy logic controller (FLC) for the fault problem of radiators, and utilized the FLC through the method of fixing thresholds to carry out fault classification with significant results.

Based on the above analysis of the status quo, considering the limited arithmetic power of the excavator controller and the timeliness of processing after the occurrence of faults, we develop a remote fault diagnosis App based on the Android system, which can carry out one-key inspection and generate a graphic and textual fault repair instruction sheet, solving the problem of having to determine the source of faults on site and the problem of low efficiency of troubleshooting. The App has been tested by Dspace, a real-time simulation system in excavator physics laboratory, and on-board tests of domestic excavator SY980 and other models, and has good diagnostic performance and has been put into use.

2 Design of the Remote Troubleshooting System

The remote troubleshooting APP based on Android system is developed with the help of two tools, Visual Studio Code (hereinafter referred to as VS Code) and HBuilderX, and the functional design of each module is carried out on the VS Code side using the JavaScript language and compiled by HBuilderX, and finally installed and utilized in the Android system cell phone terminal installation and use.

After preliminary market research and with the study of automotive diagnostic system, the diagnostic system developed in this paper has the following main functions: fault diagnosis, port inspection, auxiliary maintenance and firmware upgrade four major functions [7]. Fault diagnosis function can realize the operation of reading fault code, parameter monitoring, generating fault repair guide page, etc., which is convenient for fault detection and maintenance of the excavator; with remote data interaction function, carry out one-key inspection, solving the problem of having to physically come to the scene to check the system one by one, reducing the consumption of resources and costs; provide the function of pushing the APP version to the cloud for real-time upgrading, which is convenient for the update of the fault library and the solution of the user software problems. The user's software problems are solved.

The overall functional framework of the mobile APP is shown in Fig. 1. The service objects are customers, service engineers and drivers, who use smart phones to carry out remote one-key inspection of the faulty excavator, and reflect the current equipment fault problems and occurrence trends through the fault list and figurative image curves, and ultimately feed back to the users and maintenance personnel, and carry out maintenance based on the fault repair guidance page. Diagnostic engineers are mainly responsible for the daily maintenance of APP and cloud system, and remotely assist users in diagnosis when necessary.

The system logic diagram of remote fault diagnosis is shown in Fig. 2, including fault diagnosis, smart phone terminal, cloud service center and other modules. The remote diagnosis system is connected to the excavator through the CAN line on the fault diagnosis device to carry out reliable data acquisition, and by setting the communication frequency that meets the requirements, the corresponding working data is extracted based on TCP protocol. The communication terminal module uses the mutual conversion function of the fault diagnosis CAN signal and WIFI signal to implement the mutual transmission of data signals. As the actuator of remote fault diagnosis, smart phone terminal also serves as the carrier of most functions. It uses the perfect APP platform to realize data storage and fault diagnosis and other business operations. In addition, the application side and the cloud can also exchange

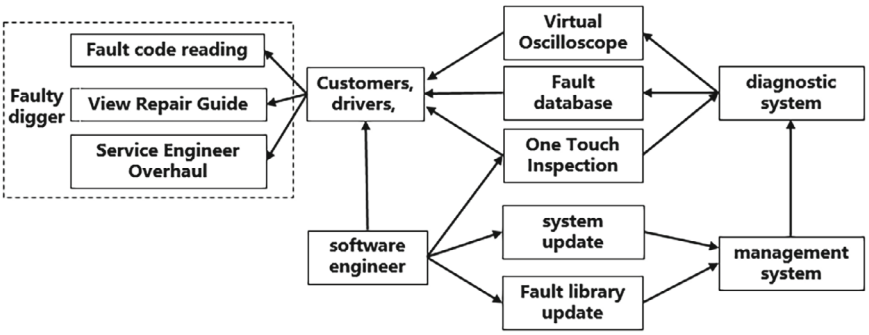


Fig. 1 System frame diagram

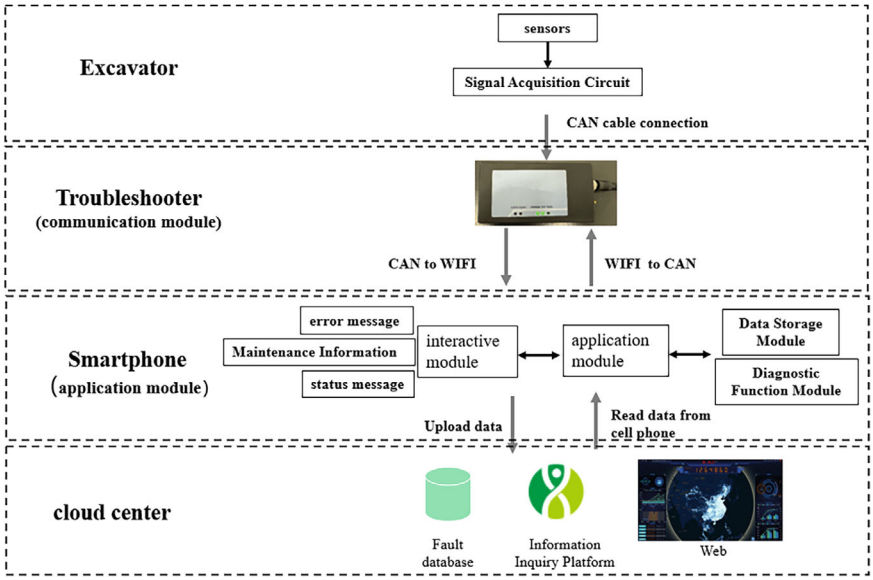


Fig. 2 Logical structure diagram of remote diagnosis system

information, and the relevant information of the fault maintenance database is stored in the cloud. After the successful connection of the APP, the relevant model fault database information is downloaded to the user's mobile phone in the form of html, and the fault information in the database can be invoked and displayed on the page for analysis.

3 Realization of the Smartphone Remote Diagnosis System

3.1 Establishment of Communication Methods

The remote troubleshooting APP can be used through the troubleshooting box, CAN cable and signal antenna as the external hardware connection, and the Android OS as the software base.

Android operating system as the software base, as shown in Fig. 3, after successful debugging to use the software.

APP to realize online monitoring and remote diagnosis and other functions, shall build data transmission platform to complete the function business. The specific communication process is shown in Fig. 4. Data transmission platform. It mainly relies on the mutual conversion of CAN signal and WIFI signal of the fault diagnostic device to realize. In the process of the underlying data interaction, in order to solve the synchronization problem between the two sides of the communication, the flow



Fig. 3 Fault diagnostics

control frame and N_Bs and N_Cr timer are used to coordinate the communication and monitoring role, the specific process is shown in the figure below, which is mainly divided into the data parsing and processing module, the encapsulation module, the unencapsulation module and the storage module. Bs in the figure is the size of the data block that can be received by the receiver within a certain period of time.

In order to ensure the integrity and reliability of data transmission and to simplify the data processing work of the application layer, the TCP protocol is selected as the transport layer protocol. To prevent the sticky packet phenomenon caused by the characteristics of the protocol itself and the information processing capability of the system when the device receives the packet, it is necessary to customize the packet structure and carry out caching processing to ensure the smooth transmission.

3.2 Remote Inspection Principle

The structure of excavator is complex, the operation requires the cooperation of various institutions, and the system is closely connected. Ordinary single fault diagnosis method can not meet the current needs of enterprises. Based on this, the combination of fault diagnosis and intelligent network technology can effectively improve the diagnosis accuracy and efficiency. Signal comparison method and case reasoning method are commonly used fault diagnosis methods, when the equipment fails, its signal characteristics will change significantly, based on this principle by comparing the difference between the actual system output signal or feedback and the expected output signal (or expected feedback), so as to identify and locate the fault case, and the corresponding reasoning from the source case guides the solution of the fault case. The learning of source cases requires a large amount of equipment failure data analysis, which is not easy to collect and obtain in the actual study, and can be collected and integrated through the network and diagnostic technology collection of the system, and stored in the cloud system, which is easy to analyze and solve the failure cases.

In excavator fault problems, a source case is a complete solution to a fault phenomenon, including reading fault codes, clearing fault codes, fault time,

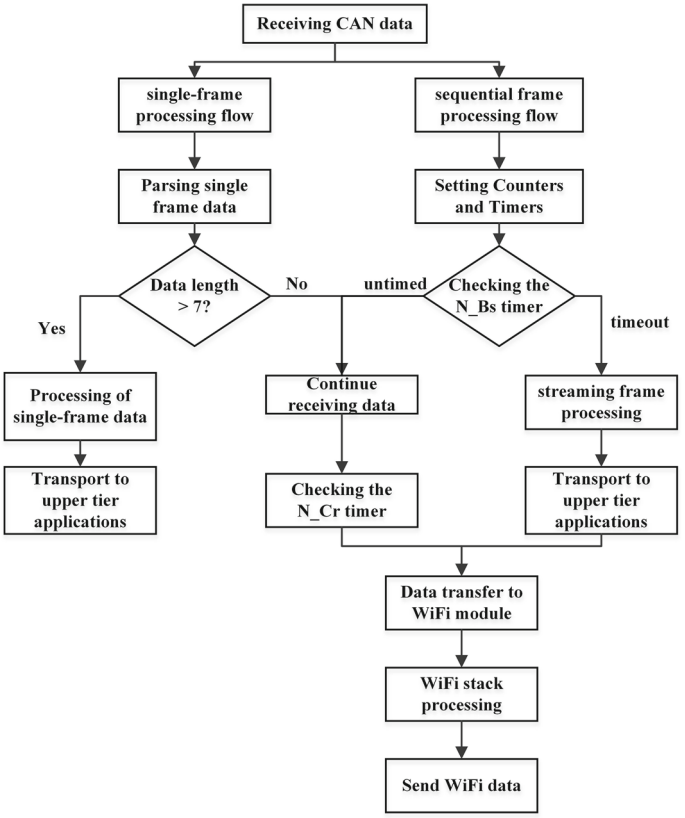


Fig. 4 General process of conversion communication between CAN and WIFI

fault phenomena, causes, schematic diagrams, and guiding maintenance steps, etc., where the set of fault phenomena and causes can be expressed as $M_p = \{M_{p1}, M_{p2}, \cdots, M_{pm}\}$; $M_c = \{M_{c1}, M_{c2}, \cdots, M_{cn}\}$.

Failure phenomenon and cause may be many-to-many relationship, let its matrix is

$$R = \begin{bmatrix} r_{11}\ r_{12} \cdots r_{1n} \\ r_{21}\ r_{22} \cdots r_{2n} \\ \cdots \\ r_{m1}\ r_{m2} \cdots r_{mn} \end{bmatrix} \tag{1}$$

where r_{ij} is the influence possibility of the jth failure cause M_{cj} on the ith failure phenomenon M_{pi}, 0 means no influence, and vice versa is $1.r_{ij} = \sum a_{ij} / \sum a_i \in [0, 1]$, where $\sum a_{ij}$ indicates the sum of the occurrence times of the i corresponding fault phenomenon caused by the j fault cause, and $\sum a_i$ is the sum of the number of

times the ith fault phenomenon occurs. However, in the actual operation period, the fault is not a single occurrence, may be caused or caused by the relationship, so the correlation matrix between the fault phenomena is as follows:

$$L = \begin{bmatrix} l_{11}\, l_{12} \cdots l_{1m} \\ l_{21}\, l_{22} \cdots l_{2m} \\ \cdots \\ l_{m1}\, l_{m2} \cdots l_{mm} \end{bmatrix} \tag{2}$$

where l_{ij} represents the relationship of the ith fault phenomenon to the jth phenomenon, 0 means irrelevant, and vice versa is 1. $l_{ij} = \sum b_{ij} / \sum a_i \in [0, 1]$, where $\sum b_{ij}$ is the sum of the number of times that the jth fault phenomenon occurs at the same time as the ith fault phenomenon also occurs. This results in a comprehensive set of relationships between phenomena and phenomena and between phenomena and causes is

$$Q = LR \tag{3}$$

When the user starts the inspection program, all port options are traversed with one click and based on the signal comparison method $\left|D_\mu\right| > \theta_\mu$, fault points with differences greater than the threshold range are identified and displayed in the APP fault diagnosis list. The fault vector $\mathrm{X} = \{x_1, x_2, \cdots, x_m\}$ is inputted into the inference program, and the program searches for the corresponding fault cases in the source cases, and finally gives the corresponding maintenance guidance in the form of html.

4 Main Functions

4.1 One-Touch Inspection Function

One-click inspection is an automatic and intelligent inspection function, which aims to simplify the inspection process and improve efficiency, and is also the main function of APP. Users can start one-click inspection through the button on the port check function page of the APP, and directly iterate through all port options of the controller (including DO port and PWM port), and use the signal sent by the mobile phone, after the signal conversion of the fault diagnoser, compare and judge with the received feedback signal. Based on the signal comparison method, If the difference is within a certain threshold range, the port status is considered normal. If the difference is abnormal, the operation process is as shown in Fig. 5. The PWM port is judged by the difference between the initially set threshold and the actual feedback value, and

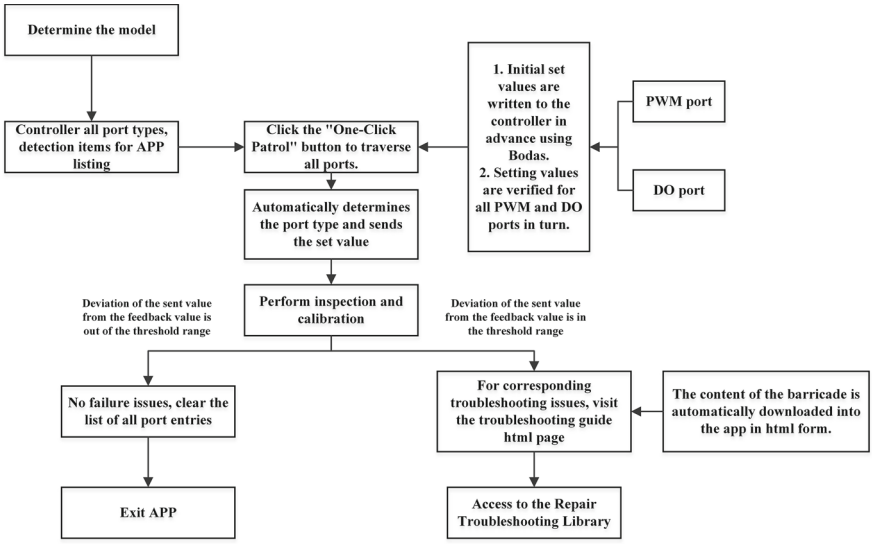


Fig. 5 Flow of one-click inspection

the DO port is judged by the feedback of the digital signal 1 using the switch button. If the signal value sent and received are both 1, it is considered a success.

After the inspection is completed, if a fault problem is found, the corresponding fault name and code will be automatically generated in the port detection list. Users can click on the corresponding fault option, according to the source case reasoning guidance to the html web page, to view the specific fault repair guidance, which includes the fault code, fault description, fault description, fault schematic diagram and fault repair steps, as shown in Fig. 6.

4.2 Failure Data Analysis System

APP implements fault analysis function, as shown in Fig. 7. Based on Flask framework, the front and back end are connected, that is, the cloud SQL fault database is shared with the mobile phone data, the data transmitted in the cloud is updated in real time, and the fault data is visualized on the mobile phone, which is easy to query the data at any time and analyze the system fault trend. As shown in the following figure, according to the logic flow chart to realize the cycle data comparison and concrete data two parts of the function, the former can compare the number of failures in this cycle and the previous cycle, the latter in the form of an image, in the number of days as a unit to generate a curve to analyze the trend of failure.

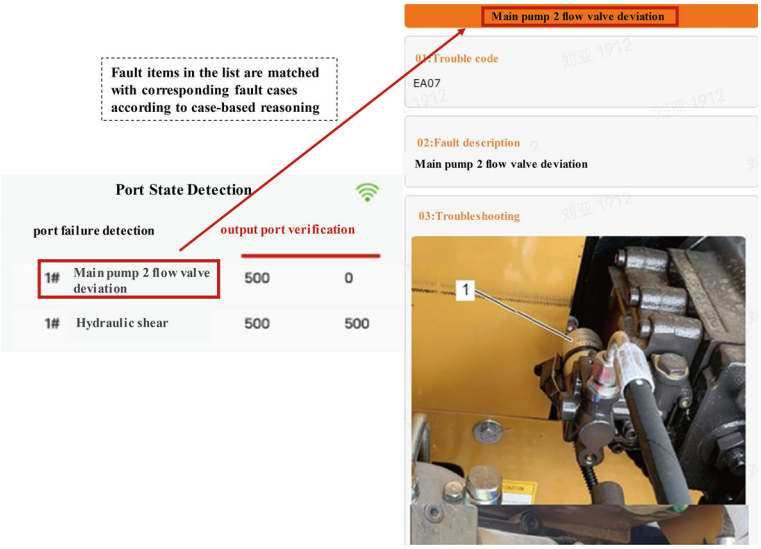
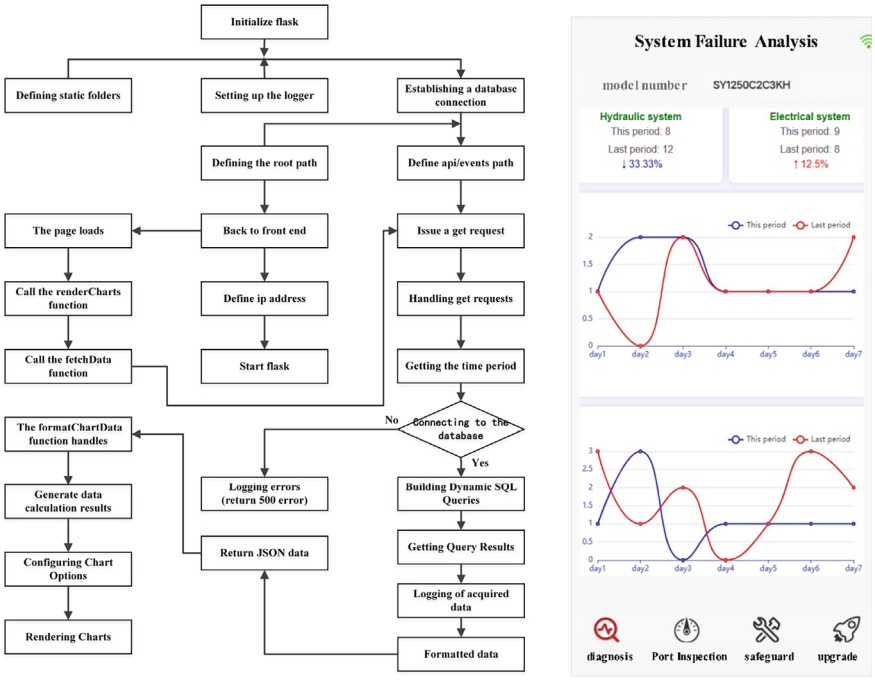


Fig. 6 One-click inspection running result page



(a) (b)

Fig. 7 **a** Fault analysis system logic diagram **b** fault analysis function page

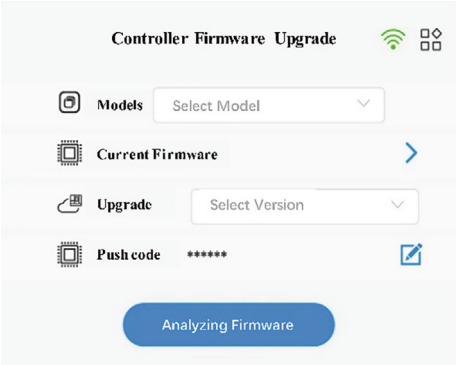


Fig. 8 Firmware update function page

4.3 Auxiliary Maintenance and Firmware Upgrade Functions

Fault libraries and problems in the use of APP need to be constantly tracked and updated, so the firmware update function is deployed, as shown in Fig. 8, using the cloud network to push the new version to solve the user's use of the problem.

5 Diagnostic Tests

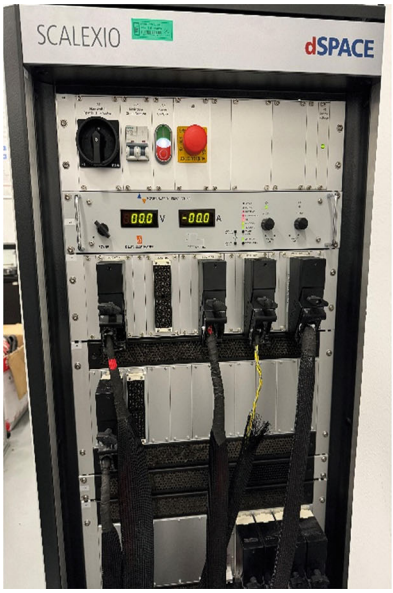
In order to test the effectiveness of the diagnostic function of the APP on the smartphone terminal, this paper verifies the function by both laboratory validity testing and on-board testing.

5.1 Laboratory dSPACE Tests

The effectiveness test is carried out in the semi-physical digital simulation laboratory, based on the dSPACE equipment to fully simulate the operating characteristics of the excavator, the use of Bodas software to import the relevant programs into the controller and with the fault diagnostic instrument and other hardware as a means of connecting the diagnostic APP developed by the use of Android system for testing.

Based on the Dspace equipment, the active setup system malfunction, as shown in the Fig. 9 for the functional effectiveness of the test. Active unplugging of the main pump 4 flow valve, linear travel and hydraulic shear three ports, test the diagnostic function is no feedback signal due to the port, and output 0 value.

According to the diagnostic principle proposed above, the diagnosis is carried out by using the transmission of the device signal and the feedback from the port, and the results are shown in Fig. The figure clearly shows that when the port is unable to



Port State Detection

port failure detection		output port verification	
1#	Hydraulic shear	500	498
1#	Main Pump 4 Flow Valve	500	500
1#	linear walking	500	501

Before port unplugging

1#	Hydraulic shear	500	0
1#	Main Pump 4 Flow Valve	500	0
1#	linear walking	500	0

After port unplugging

Fig. 9 Dspace test equipment and effectiveness test results

feed back the signal value, the number of faulty ports increases by 3 and the value of the feedback column in the port status detection list becomes 0, which satisfies the desired goal.

5.2 On-Board Testing

The experiment takes a domestic 98-ton excavator as the test object, connects the CAN cable of the fault diagnostic device with the CAN port behind the seat of the excavator, and connects the cell phone network and the diagnostic device network to complete the preparation before the test.

After powering up the whole machine, start the one-key inspection function in the function page, and send inspection commands through the set values written in advance in the controller, traversing the 27 PWM ports and 5 DO ports of the controller. The result is shown in Fig. 10, when part of the solenoid valve receives the output signal, it will produce a slight action accordingly, and the display will synchronously show the fault data code of the cell phone terminal to meet the APP function.

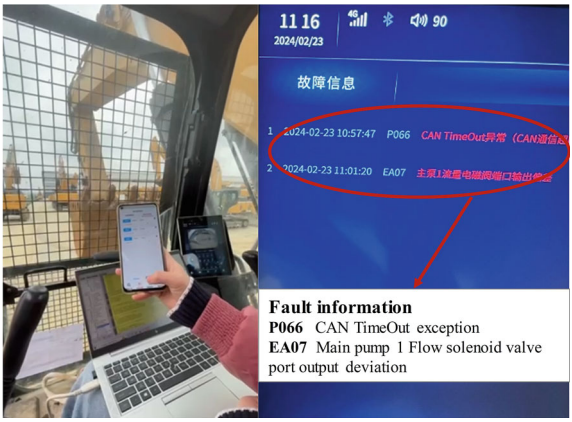


Fig. 10 Test diagram of 98-ton machine type

6 Conclusion

This paper takes the remote fault diagnosis APP developed based on Android system as the research object, and introduces the main functions of APP and the steps and principles of remote fault diagnosis. Finally, the effectiveness and functionality of the remote diagnosis function is verified through the laboratory Dspace equipment simulation of excavator faults and SY980 on-board testing, and the following conclusions are obtained:

(1) In the dSPACE simulated excavator fault experiment, using the comparison of the number of ports and the state, it is proved that the remote diagnostic function meets the expected goal when the APP is run under the simulated excavator working condition.
(2) Using the 98-ton machine for on-board testing, the diagnostic and excavator are connected successfully, and all port options are traversed with one click. After different solenoid valves receive the signal, different mechanisms of the excavator will have a slight action response of 1-2 s. Observing the list of port status on the cell phone, the feedback value floats up and down within a certain range, the functionality is intact, which meets the expected goal, and the display in the cab also displays the name of the fault and the fault code at the same time.
(3) When auxiliary diagnostic APPs are used for complex fault problems that intersect multiple systems or components, there are diagnostic inaccuracies and the updating and maintenance of the APPs may not be able to keep up with the emergence of new fault problems and solutions, resulting in lagging diagnostic capabilities.
(4) In the future, diagnostic logic and algorithms will be improved to enhance diagnostic accuracy through big data and algorithms and to realise the function of using data to achieve predictive maintenance, so as to avoid downtime and costly repairs due to malfunctions; and flexible updating mechanisms and a

comprehensive malfunction database will be designed to realise full coverage of a wide range of aircraft models.

Acknowledgements Thanks for all the people in School of Mechanical and Power Engineering United Institute of Excavator Key Technology and the Institute of Automobile and Construction Machinery Nanjing Tech University, the help from Sany Heavy Machinery Company Ltd., and the support of National Key R&D Program of China (2021YFB2011904).

References

1. Hong Z (2014) Vehicle remote fault diagnosis based on fuzzy theory. Automot Eng 05:50–52
2. Mo M, Wong Y (2020) Research on fault diagnosis of excavator hydraulic system based on ACA-BP neural network. Eng Constr Des 13:137–140
3. He Xiangyu, He Q, Jiang P et al (2010) Fault detection of excavator hydraulic system based on dynamic GRNN model. Chin J Constr Mach 8(03):335–339
4. Feng X (2018) Research and development of parameter monitoring and fault diagnosis system of construction machinery based on WiFi and virtual instrument. Shijiazhuang Railway University
5. Wang H, Fu B, Tian P et al (2024) Fault diagnosis of excavator hydraulic system based on energy entropy feature. Hydraul Pneum Seals 40(10):84–89
6. Bagyalakshmi N, Subramanian SS, Kumar VD (2024) Model based fault detection diagnosis and classification in shell and tube heat exchanger using intelligent techniques. Desalination Water Treat 319
7. Wang C, Xu L, Peng W (2007) Conceptual design of remote monitoring and fault diagnosis systems. Inf Syst (7):32

Open Access This chapter is licensed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License (http://creativecommons.org/licenses/by-nc-nd/4.0/), which permits any noncommercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if you modified the licensed material. You do not have permission under this license to share adapted material derived from this chapter or parts of it.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.



Research on the Influence of Cable Wind Ropes on Wind-Induced Response and Wind Vibration Coefficient of Tower Crane



Tong Qin, Yifan Lu, and Wenwu Liu

Abstract In order to reveal the influence of cable wind ropes on the dynamic response and vibration coefficient of tower cranes under the excitation of wind, the theoretical equation for the vibration coefficient of tower cranes under the combined action of wind excitation and cable wind rope was firstly established. Secondly, simulate the pulsating wind speed by using the Davenport wind speed spectrum, and a finite element dynamic model of QTZ25 tower crane under the action of cable wind rope was constructed. The harmonic synthesis method was considered to obtain the wind excitation speed. Finally, the time history response of displacement at different nodes and the power spectral density curve of displacement at corresponding nodes of QTZ25 tower crane were obtained. Furthermore, the wind vibration coefficients at different nodes of the crane that vary with the height of the tower structure and the windward angle under wind excitation are obtained. It found that when the angle between the cable wind rope and the tower body is 45°, and the projected height of the cable wind rope on the tower body accounts for 50% of the total height of the tower body, the coefficient value of wind vibration decreases when the height is same. When the windward angle is 0, the maximum wind vibration coefficient value at different nodes of the tower body drops by 20.33%, and which drops by 19.80% when the windward angle is 45°. The results indicate that the cable wind rope can effectively reduce the dynamic characteristics, and the reasonable placement of cable wind ropes can achieve significant vibration reduction effects.

Keywords The tower crane · Cable wind rope · Dynamic response of wind · Vibration coefficient of wind · Speed spectra of wind

T. Qin
Guangxi Construction Engineering Group Construction Machinery Manufacturing Co., Ltd, Nanning, P. R. China

Y. Lu
School of Mechatronics Engineering, Harbin Institute of Technology, Harbin, P. R. China

W. Liu (✉)
College of Engineering, Northeast Agricultural University, Harbin, P. R. China
e-mail: liuwenwu963@163.com

© The Author(s) 2026

H. Saman et al. (eds.), *The 9th International Conference on Advances in Construction Machinery and Vehicle Engineering*, Lecture Notes in Mechanical Engineering,
https://doi.org/10.1007/978-981-96-5527-4_9

1 Introduction

Wind load is a common load in engineering structures [1, 2]. Cable wind ropes are often used in engineering to improve the wind and vibration resistance of tower cranes, reduce vibration fatigue of local components and overall structures, and increase reliability of use. One end of the cable wind rope is connected to the tower body, and the other end is fixedly connected to the ground, which can effectively reduce the vibration of the tower crane under wind load. Therefore, it has important theoretical significance to study the influence of cable wind ropes on the dynamic response of tower cranes induced by wind. Existing research has shown that steel wire ropes can act as dampers.

Lai et al. [3, 4] established an optimization method to determine that the diameter of the steel wire rope and the ring diameter have a significant impact on the damping performance of the steel wire rope damper. Luo et al. [5] studied the vortex induced vibration suppression effect of cable wind ropes on deepwater pipeline racks and found that the placement of cable wind ropes has a significant impact on the vibration reduction effect. Guo et al. [6] proposed a set of calculation methods for offshore steel structure vertical section cable wind ropes by drawing on existing cable wind rope design methods. In addition to steel wire ropes as cable wind ropes, nylon materials have nonlinear average tension and amplitude stiffness, and are widely used as vibration reduction measures for offshore wind turbines [7]. In addition to serving as a vibration reduction measure for heavy structures such as cranes, cable wind ropes can also be used for full model wind tunnel flutter tests, providing support for simulating free flight of aircraft [8]. Rytömaa et al. [9] made a research that the nonlinear hysteresis behavior of steel wires has advantages compared to linear isolators and the results indicate that the effects of preloading and frequency are not significant, while the influence of amplitude is more significant.

In summary, current research mainly focuses on the influence of wind loads on structures such as lifting systems, offshore steel structures, and steel frame bridges. However, there is relatively little research on the dynamic response of tower cranes under wind loads using cable wind ropes and their different arrangements. This paper aims to explore the dynamic response of tower cranes under the action of cable wind ropes. It will also reveal the influence of cable wind rope's distribution on the dynamic characteristics of tower cranes under wind loads.

2 Wind Response Theoretical Analysis Under the Action of Cable Wind Rope

Under the influence of cable wind ropes and wind load, the dynamic formulation of the tower crane can be established as:

$$\mathrm{M}\ddot{\mathrm{x}}(\mathrm{t}) + \mathrm{K}\mathrm{x}(\mathrm{t}) + \mathrm{C}\dot{\mathrm{x}}(\mathrm{t}) = P(\mathrm{t}) + F(\mathrm{t}) \tag{1}$$

where, M, K, and C are the mass, stiffness, and damping matrix of the tower crane structure, respectively, $\ddot{x}(t)$, $\dot{x}(t)$ and $x(t)$ are the different node's acceleration, velocity, and displacement, respectively, $P(t)$ is the wind load, $F(t)$ is the force vector of cable wind rope.

Under the influence of cable wind rope and wind load, it can be found that, the all displacement of the beam structure is expressed as:

$$x_i = \beta_z \overline{x}_i \tag{2}$$

In which

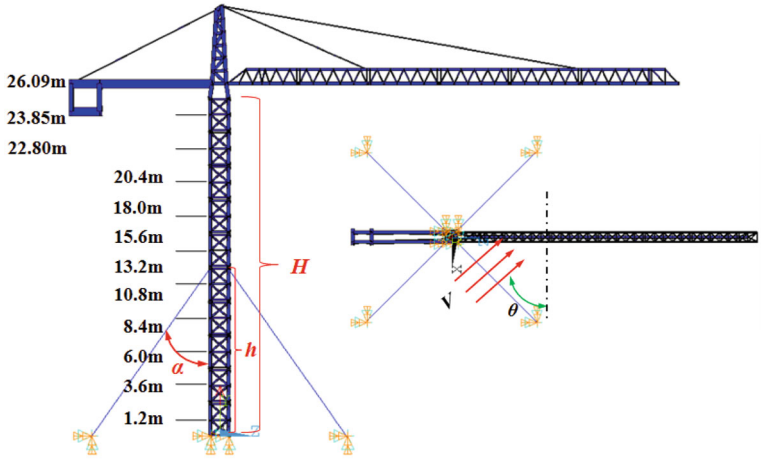
$$\beta_z = 1 + \frac{\mu \sigma_{xi}}{\overline{x}_i}$$

where, x_i is the all displacement response, $\overline{x}_i$ is the average of static displacement, μ is the factor of peak under probability conditions, β_z is the vibration coefficient by wind, σ_x is the root variance of the output power spectral density (PSD) fluctuation response.

3 A Finite Element Dynamic Model Under the Action of Cable Wind Ropes

Under the action of cable wind ropes, the finite element dynamic analysis model of the QTZ25 tower crane was established as shown in Fig. 1. The structural parameters of QTZ25 tower crane are as follows: the total height is 30470 mm, and the body of tower is made of angle steel, which parameter is 1300 mm × 1300 mm. The boom has an isosceles triangular cross-section with a side length of 1000 mm. The cross-section of the balance arm is 300 mm × 100 mm of channel steel, and the tower height H is 25000 mm. The length of the balance arm is 10000 mm. Four cable wind ropes are used to connect the tower body and the ground, symmetrically.

The QTZ25 is layered from low to high and which is divided into 12 sections. The value of the wind speed at different height is taken as the middle position of each section, as shown in Fig. 1a, and α represents the angle between the cable wind rope and the tower body, and h represents the projected height of the cable wind rope on the tower body. The windward angle θ is assumed as a calculate variable, and there are two working conditions in which θ equals 0 and 45°, V represents the wind speed, as shown in Fig. 1b, in which, the shape factor μ_s is1.628 and 2.04, respectively.



(a) Schematic diagram (b) Schematic diagram of different wind angles

Fig. 1 The finite element model of QTZ25 under action of four cable wind ropes

4 Dynamic Response and Vibration Coefficient Under the Action of Cable Wind Rope

4.1 Theoretical Analysis of Wind Induced Load Simulation

This article uses spectrum of Davenport to simulate the effect of wind, which can be represented as follows:

$$\frac{nS_v(n)}{v_*^2} = \frac{4x^2}{(1+x^2)^{4/3}}, x = 1200\frac{n}{\overline{v}_{10}} \tag{3}$$

where

v_* is the load shear velocity of wind: $v_*^2 = K(\overline{v}_{10})^2$,
$S_v(n)$ is the load pulsating of wind speed PSD,
n is the load frequency of wind,
K is the speed spectrum correlation coefficient of wind,
$\overline{v}_{10}$ is the average speed of wind at a height of 10 m above the ground.

Multiply the average pressure of wind ϖ by the size factor of crane body μ_s as well as the height factor μ_z, then the wind pressure p can be got as Eq. (4) shows:

$$p = \mu_s \mu_z \varpi \tag{4}$$

Assuming the terrain is Class B, $\overline{v}_{10} = 10.0$ m/s, and the average pressure of wind in the simulation condition is 150N/m^2. The calculation formula for wind load at

different nodes along the tower body from low to high is as follows:

$$P(z,t) = 0.5\mu_s \rho A \left[V(z,t)\right]^2 \tag{5}$$

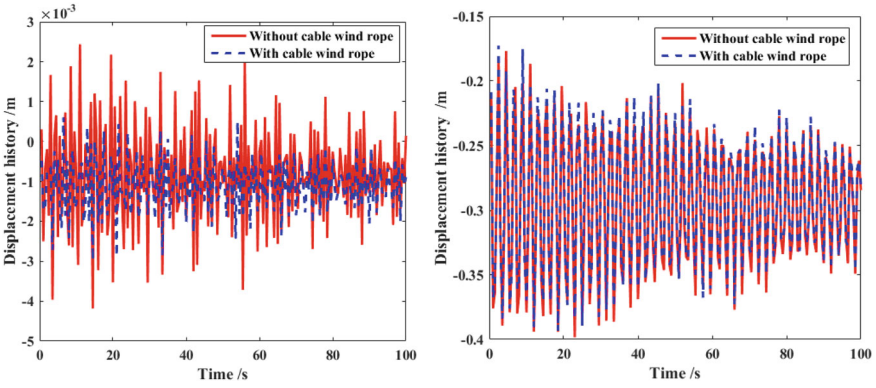
where

μ_s is the coefficient of different windward side shapes of the simulated structure,
ρ is the air density under operating conditions, $\rho = 1.25\text{kg/m}^3$,
A is the projection area of different windward side members of the crane,
$V(z,t)$ is the speed of wind at height z as the load of wind changes over time.

4.2 Response Induced by Wind Under the Action of Cable Wind Rope

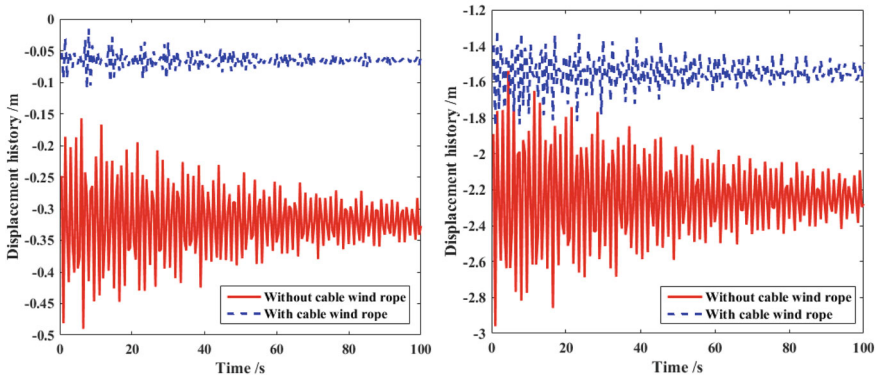
Under the action of cable wind ropes, the excitation load of wind is applied along the height of the tower to the finite element dynamic analysis model shown in Fig. 1, and history of displacement at various windward angles and different heights under wind excitation are obtained as shown in Figs. 2 and 3 when $\alpha = 45°$, $h/H = 0.5$.

From the analysis of Figs. 2 and 3, it can be seen that when the windward angle equals 0 and 45°, the response amplitude of displacement increases from low to high along the tower body. Under the action of cable wind ropes, when the windward angle is 0, the maximum amplitude of the displacement of the structure at heights of 10.8 m and 23.85 m decreases by 75.67% and 10.58% respectively and which decreases by 87.50% and 52.91% when the windward angle is 45°. The results indicate that the cable wind rope has a significant impact on reducing the response of displacement of tower cranes under various windward angles.



(a) 10.80m height (b) 23.85m height

Fig. 2 Time history of displacement when windward angle is 0



(a) 10.80m height (b) 23.85m height

Fig. 3 Time history of displacement when windward angle is 45°

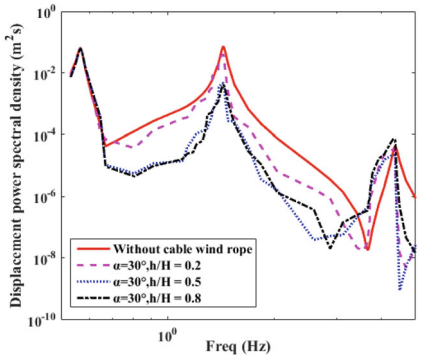
4.3 Vibration Coefficient of Wind Under the Action of Cable Wind Rope

At windward angles of 0 and 45°, as well as at different tower heights, the PSD of displacement of the QTZ25 Tower crane under the action of cable wind rope are shown in Figs. 4 and 5.

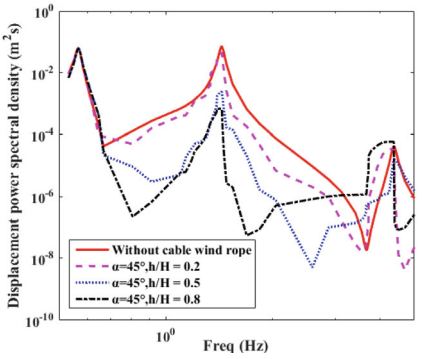
It can be seen from Figs. 4 and 5 that the PSD of displacement at different heights changes significantly with the wind excitation frequency increase. The PSD of displacement value decreases as the ratio of the projected height of the cable wind rope on the tower body to the total height of the tower body increases when the height is same, while the PSD of displacement value increases As the angle between the cable wind rope and the tower body increases. It can also be shown from Figs. 4 and 5 that under the action of cable wind rope, when the angle between the cable wind rope and the tower body is 45°, and the projected height of the cable wind rope on the tower body accounts for 50% of the total height of the tower body, the PSD of displacement value of QTZ25 tower crane is the smallest.

Analyze and consider certain crossing probability conditions, the peak factor μ is 3.24 [1]. The comparison of the vibration coefficients induced by wind of the crane structure with and without cable wind ropes along the height variation of the tower is shown in Fig. 6 when $\alpha = 45°$, $h/H = 0.5$.

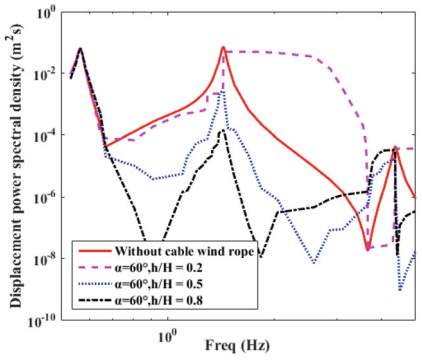
Figure 6 indicates that the wind-induced vibration coefficient gradually increases from low to high along the tower body. The wind vibration coefficient decreases significantly under the action of cable wind ropes when the height is same. Maximum wind vibration coefficient decreases directly 20.33% when the windward angle is 0 and which 19.80% when the windward angle is 45°. The reason is that the cable wind rope has a vibration absorbing effect on the tower cranes dynamic displacement under



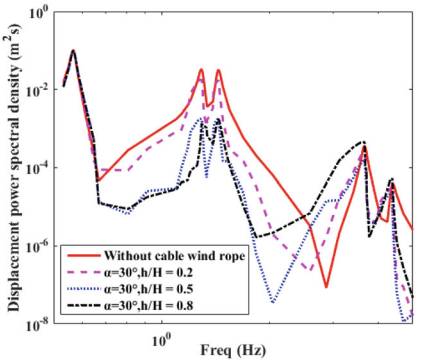
(a) θ=0, α=30°



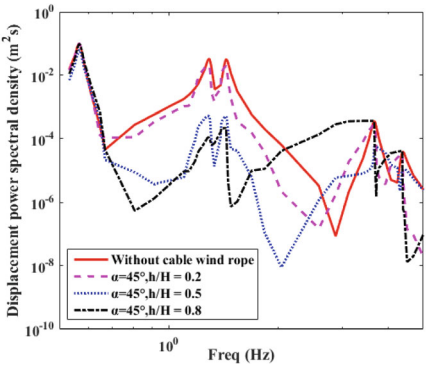
(b) θ=0, α=45°



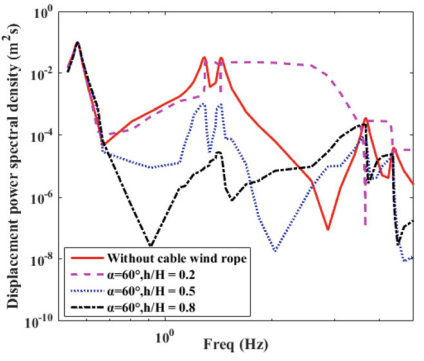
(c) θ=0, α=60°



(d) θ=45°, α=30°

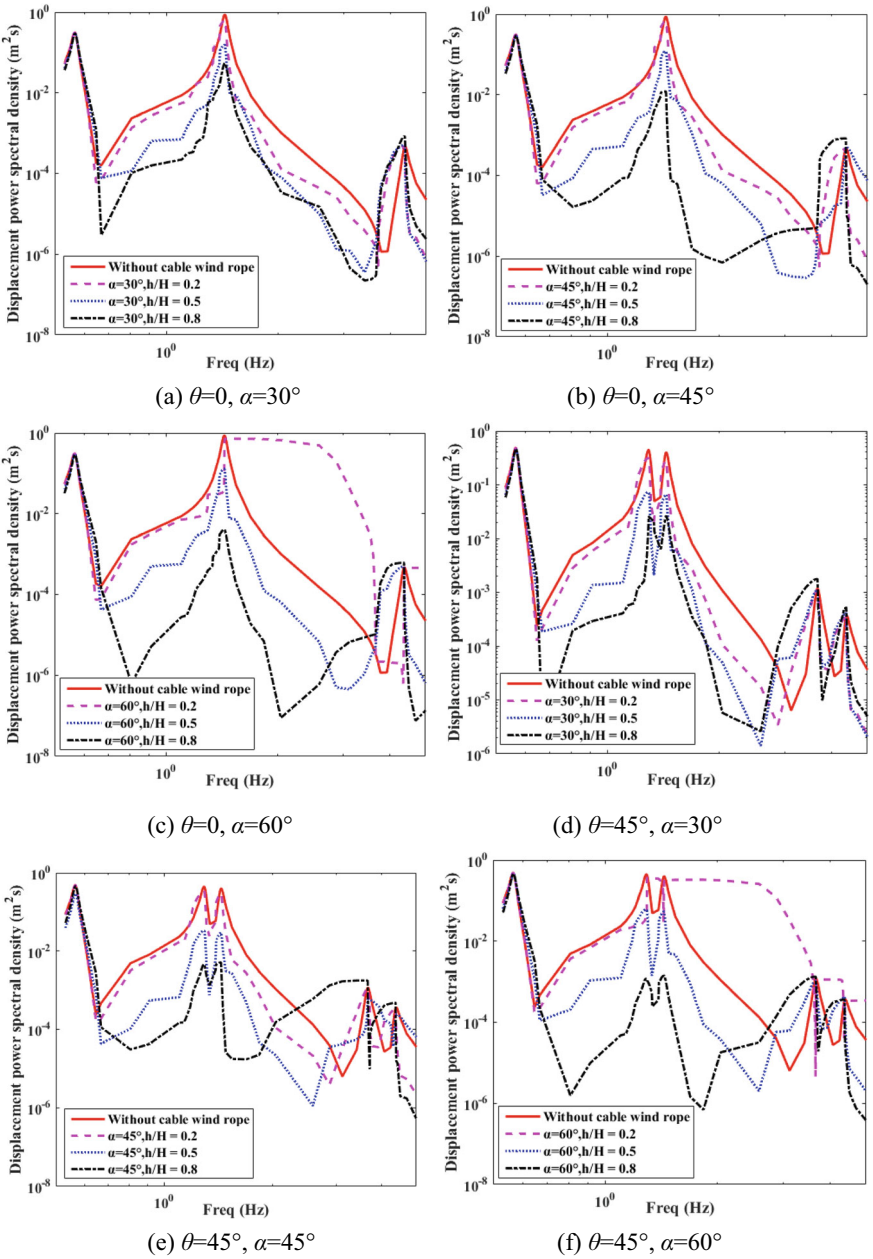


(e) θ=45°, α=45°



(f) θ=45°, α=60°

Fig. 4 PSD of Displacement at a height of 10.80 m



(a) θ=0, α=30°

(b) θ=0, α=45°

(c) θ=0, α=60°

(d) θ=45°, α=30°

(e) θ=45°, α=45°

(f) θ=45°, α=60°

Fig. 5 PSD of displacement at a height of 23.85 m

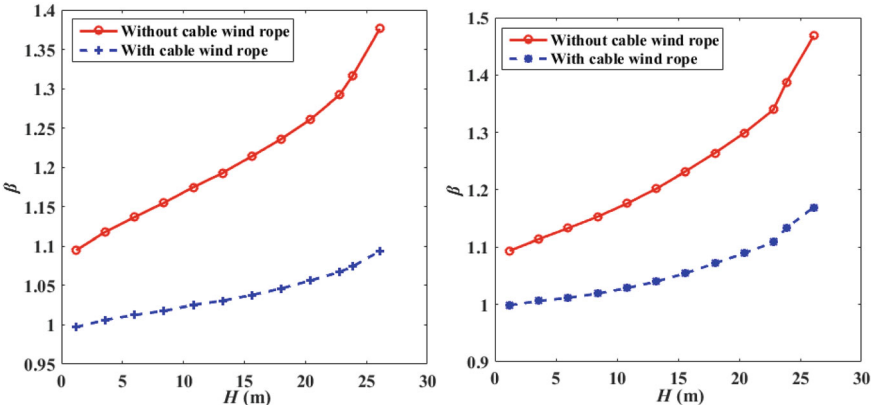


Fig. 6 Vibration coefficient of wind at different heights

wind load excitation. The higher the height and the greater the wind speed, the more obvious the vibration reduction effect of the cable wind rope.

5 Conclusion

The tower crane structure's theoretical equation of the vibration coefficient of wind under the action of cable wind ropes is derived. This paper also establishes the finite element model of tower crane for dynamic analysis. Wind load on finite element dynamic analysis model is simulated. Displacement response analysis and spectral analysis are performed under wind load excitation. Then the wind-induced vibration coefficients at various node heights and windward angles under the action of cable wind ropes are obtained. The specific conclusions which can give more information are as follows:

(1) The maximum displacement at heights of 10.80 m and 23.85 m decreases by 75.67% and 10.58% when the windward angle is 0, respectively, and which decreases by 87.50% and 52.91% when the windward angle is 45°, respectively.
(2) The PSD of displacement changes significantly with the increase of wind excitation frequency. The PSD of displacement value decreases as the ratio of the projected height of the cable wind rope on the tower body to the total height of the tower body increases when the height is same, while the PSD of displacement value increases as the angle between the cable wind rope and the tower body increases.

(3) The wind vibration coefficient decreases significantly under the action of cable wind ropes at the height is same. The maximum reduction in wind vibration coefficient is 20.33% and 19.80% when the windward angle is 0 and 45°, respectively.

Acknowledgements This paper is supported by the Major Science and Technology Projects of Nanning City (Grant No. 20221242).

References

1. Xu J, Geng S, Xiao F (2016) Research on wind vibration response of inner suspension method and outer cable holding pole. J Hefei Univ Technol (Nat Sci) 39(11):1531–1535
2. Hamada A, King JPC, EI Damatty AA (2017) The response of a guyed transmission line system to boundary layer wind. Eng Struct 139:135–152
3. Lai K, Fan W, Chen Z et al (2024) Optimal design of transverse dampers incorporating inherent stiffness effects for stay cable vibration control: a case study with wire rope dampers. Eng Struct 308(1):118015–118020
4. Lai K, Fan W, Chen Z et al (2023) Performance of wire rope damper in vibration reduction of stay cable. Eng Struct 278(1):115527–115530
5. Luo C, Hu W, Li W (2022) Suppression of vortex induced vibration of pipe of deep-water jacket by wind rope. Ship Sci Technol 44(4):87–90
6. Guo N, Yang F, Shan J, Niu J, Li G (2019) Design of offshore steel structure vertical cable wind rope. China ShipRepair 32(2):54–56
7. Thuilliez H, Davies P, Cartraud P et al (2023) Characterization and modelling of the dynamic stiffness of nylon mooring rope for floating wind turbines. Ocean Eng 287(1):115866–115870
8. Zhao Z, Zhao Z, Yan Y et al (2023) Dynamics, stability, and control of a four-cable mount system for wind tunnel test. Chin J Aeronaut 36(2):58–75
9. Rytömaa S, Malmi O, Laine S (2024) Wire rope isolator identification and dynamic modeling for small amplitude vibrations. Eng Struct 318(1):118721–118726

Open Access This chapter is licensed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License (http://creativecommons.org/licenses/by-nc-nd/4.0/), which permits any noncommercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if you modified the licensed material. You do not have permission under this license to share adapted material derived from this chapter or parts of it.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.



Fatigue Life Analysis and Improvement of Casting Crane Structures



Xin Wang, Cheng Gao, Rongpei Zheng, Tingting Zhang, and YiFan Zuo

Abstract To analyze the reason of crack caused at the variable section of the main girder, the finite element model of 260 t casting crane is built under 4 cases of a full load working cycle. The results show that the stress concentration at the crack (node 1) is much higher than in the mid-span structure where is usually high stress zone (node 2 and node 3 are selected to compare with node 1). Based on the stress amplitude method, these three modes' fatigue lives and remaining lives are estimated. The results show that node 1 has the lowest life. For the girder to keep using, 3 improvement schemes are proposed for the stress concentration of the main girder, according to the structural characters of the main girder and the existing experience. One of them is to increase the web near to the main web, and two of them is to open holes in the stress concentration, and the last one is to replace the nearby main web plate with a thick plate. The reinforcement effect of each scheme is compared after stress simulation, and the last one is selected as the best one. It is provided favorable reference for structure to obtain high fatigue life.

Keywords Casting cranes · Fatigue life · Finite element analysis · Structural improvement

X. Wang (✉) · R. Zheng · T. Zhang
School of Mechanical Engineering, Dalian University of Technology, Dalian, China
e-mail: wangxbd21@163.com

C. Gao · Y. Zuo
Key Laboratory of Lifting Equipment's Safety Technology for State Market Regulation, Shenyang, China

Liaoning Academy of Safety Science, Shenyang, China

© The Author(s) 2026

H. Saman et al. (eds.), *The 9th International Conference on Advances in Construction Machinery and Vehicle Engineering*, Lecture Notes in Mechanical Engineering,
https://doi.org/10.1007/978-981-96-5527-4_10

1 Introduction

A casting crane is one of the key operating equipment in metallurgical industry for material handling. Due to the high temperature, dusty operating environment, full load operation and high frequency, it makes the casting crane structure dominated by fatigue damage, which is prone to cracks and affects the safety of operation [1].

Scholars at home and abroad have carried out extensive research on the strength and fatigue life of casting crane structures. Frankowski [2] retrofitted the main girder webs of casting cranes that were already damaged, and performed complex strength and fatigue life analysis of the critical structures before and after the refurbishment. Euler Mathias [3] conducted an experimental study on the fatigue behavior of crane runway girder web joints and proposed transverse reinforcement in the runway girder wheel load region. Liu Lijuan [4] carried out theoretical calculations and finite element analysis of cracked bridge, and combined with the theoretical and finite element analysis results to derive the reasons for the cracking of the bridge, based on the original parameters of 165 t casting crane. Yu Wancheng [5] conducted stress tests on the main stress plates of the crane main girder and the key parts of the main girder span, and analyzed and estimated the fatigue life of the main girder based on the stress spectrum obtained from the stress tests, and made a safety assessment of the equipment. Ren Xiaoyu [6] used inductive statistical methods to derive the typical number of lifts per month for tire-type container gantry cranes, and obtained the load spectrum through the rainfall counting method. Combined with fatigue-related theories and finite element analysis, the fatigue life of tire-type container gantry crane metal structure is analyzed.

It can be seen that most of the research literature in the load spectrum, fatigue life analysis methods and other programs have achieved better research results, which provides a reference for this paper to carry out the fatigue life analysis of casting crane structure. In this paper, 260 t casting crane structure is as the research object. Due to the long-term full-load work, cracks appeared at the variable section of the main girder, and a finite element model is built to analyze the construction and the cause of cracks here. According to the stress amplitude method provided by ISO standard, the remaining fatigue lives of typical structural nodes are analyzed. In order to ensure working safety of the structure in the stress concentration zone, three improvement schemes are proposed. Through the comparison of calculation results and further analysis of the reason for stress concentration, a more effective improvement scheme is selected.

2 Structural Finite Element Modeling

Due to long-term full-load work, inspection found that the 260 t casting crane had cracks at the variable section of the main girder, as shown in Fig. 1. From the corresponding drawing information, as shown in Fig. 2, it can be seen that the structure here

is composed of multiple plates spliced and welded, so maybe cause stress concentration and reduce its life expectancy. It is necessary to build the finite element model to analyze the structural stress.

The structure of 260 t casting crane mainly consists of double main girders, double sub girders and end girders, with main trolley installed on the main girder and sub trolley installed on the sub girder. The main girder, sub girder and end girder are box-type plate structure, so the shell element is selected to build the finite element model. And only the main stress components are modeled, and other non-main stress components such as railings, platforms, escalators, distribution pipes, etc. are not modeled. Other structural details are ignored when modeling, such as welding process edges, small process holes, small rounded corners and chamfers. The trolleys are also not modeled, and as mass points to apply to the contacted nodes between each trolley wheel and the main girder. At the variable section, the model is



Fig. 1 Crack condition

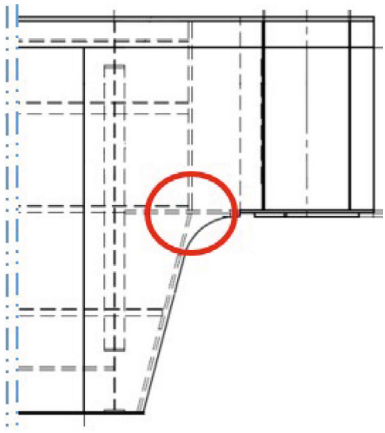


Fig. 2 Drawing information at cracks

built in detail according to its construction drawings. The structural model is meshed using quads and triangles with a mesh size of 50 mm.

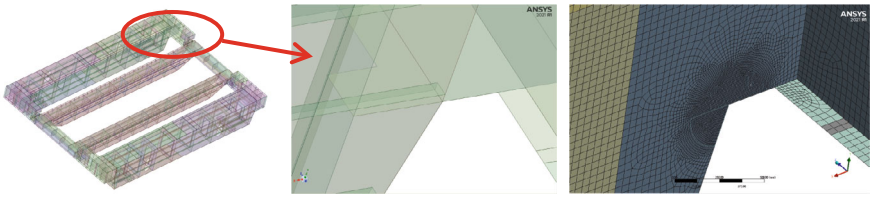
For boundary constraints, MPC elements are used to connect the wheel center of the large vehicle with the bottom flange plate of end girder. And then full displacement constraints are applied at the wheel center.

For loads, due to the relatively fixed trajectory of the casting crane, the working cycle is basically that the main and sub trolleys lift the ladle, and move from one end of the main girder near the center of the span to the other end. When approaching the target position, the ladle is overturned and dumped, and the trolleys return. The ladle is usually full load. Therefore the analysis is based on a single full load working cycle. In this working cycle, there exists loaded and unloaded, also exists acceleration and deceleration movement, so the actual loads and load positions are various, and the structural stress are also various. This variation in stress is one of the main factors to generate fatigue cracks. For a comprehensive comparison, 4 load cases are selected in one working cycle, including no-load and full load in the span of main girder, no-load and full load at the end main girder, where can obtain the up and down limits of stress in one working cycle. At the same time, according to the provisions of the GB/T 3811–2008 standard, the dynamic loading coefficients are selected to consider the possible impact load under unfavorable conditions. The specific load cases are shown in Table 1.

The finite element model is shown in Fig. 3 with the total number of nodes 347,742 and the total number of elements 349,916.

Table 1 Loading cases

Case	Loading conditions
Case 1 Full load at the end span of main girder	Lifting capacity is 260 t, main trolley weight is 115 t Dynamic load coefficient is 1.25, gravitational acceleration is 1.1 g Load direction is vertical The trolley is at the end of main girder The loads are applied at the positions where the wheels are contacted with the runway
Case 2 Full load at the mid span of main girder	Loads are same as Case 1 The trolleys is at the mid span of main girder
Case 3 No load at the end span of main girder	the main trolley weight is 115 t, gravitational acceleration is 1 g Load direction is vertical The trolley is at the end span of main girder The trolley weight is applied at the point where the wheels are contacted with the runway
Case 4 No load at the mid span of main girder	Loads are same as Case 3 The trolley is at the mid span of the main girder



a) Geometric model b) Local enlargement c)Local meshed model

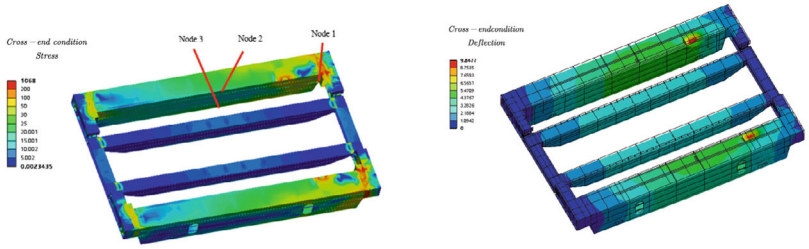
Fig. 3 Finite element model

3 Static Analysis

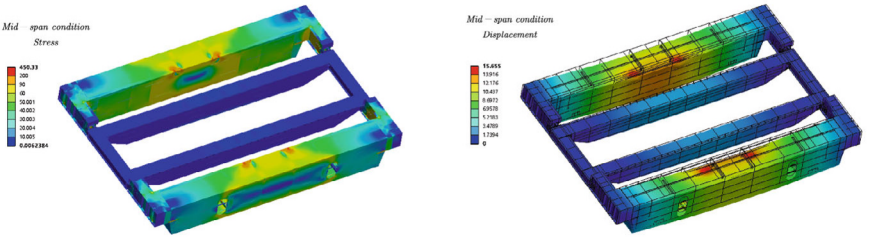
According to the above boundary constraints and loads, the stress and displacement cloud diagrams are calculated and shown in Fig. 4 as examples of case 1 and case 2.

It can be seen that the highest stress of both Case 1 and Case 2 under full load are near to the cracked structure. The next higher stress is at the upper and lower flange plates of the mid-span girder, so the nodes at these three locations are selected to compare their stresses, and the specific locations of the nodes are shown in Table 2.

The stresses of 3 nodes may be tensile or compressed, so the principal stresses are extracted with " − " for compression and " + " for tension, as shown in Table 3.



a) Stress distribution of case 1 b) Displacement distribution of case 1



a) Stress distribution of case 2 b) Displacement distribution of case 2

Fig. 4 Stress and displacement distribution

Table 2 Node Location

Node number	Location
Node 1	Intersection point between bottom flange plate and web at the end span of main girder
Node 2	Intersection point between upper flange plate and web at the mid span of main girder
Node 3	Intersection point between bottom flange plate and web at the mid span of main girder

Table 3 Stresses of cases (in MPa)

Load case	Node 1	Node 2	Node 3
Case 1	132.0	−19.0	23.6
Case 2	109.0	−71.5	98.6
Case 3	61.3	−7.5	8.5
Case 4	34.0	−21.3	28.0

It can be seen that the stress at node 1 is much higher than the other two, and also higher than the stresses of nearby nodes. Due to high stress concentration, the cracks are easily generated.

4 Fatigue Life Analysis

As one of the most common methods, the stress amplitude method is used to assess the fatigue life of a structure based on the changes of stress and the fatigue properties of the material. GB/T 30024-2020 standard [7] provides the characteristic fatigue strength and the corresponding cycles under different structural details, according to S–N curve of fatigue life, namely $\triangle\sigma_c\, N = C$. For single stress range, the corresponding total cycles can be obtained by Eq. (1), and the difference from the used cycles is the remaining cycles. Based on the cycles of per year, it can be converted to the remaining fatigue life expressed as time.

$$\mathrm{N_i} = \frac{(\Delta\sigma_c/\gamma_{\mathrm{mf}})^m}{\Delta\sigma_i^m} N_{ref} \tag{1}$$

where $\triangle\sigma_c$ and N_{ref} denote the characteristic fatigue strength and the corresponding cycles (2e6). $\Delta\sigma_i$ and Ni denote the i^{th} stress range and the corresponding cycles. m is the exponential coefficient, and takes as 3 if welded structures. γ_{mf} is the fatigue strength specific resistance coefficient, and take as1.25 if unsafe after failure members. The stress range here is the difference between the upper and lower stress limits in a working cycle.

Table 4 Stress ranges and cycles of 3 nodes

Item	Node 1	Node 2	Node 3
Fatigue characteristic strength/MPa	71	71	71
Corresponding number of cycles	2e6	2e6	2e6
Stress range/MPa	98.0	64.0	90.1
Corresponding number of cycles	389,400	1431,376	511,217
Remaining number of cycles	9240	1,051,216	131,057
Number of cycles per year	23,760	23,760	23,760
Remaining fatigue life/year	0.39	44.24	5.52

$$\Delta\sigma_i = |\sigma_{\max} - \sigma_{\min}| \tag{2}$$

where σ_{max} is the upper stress limit under one working cycle. σ_{min} is the corresponding lower stress limit. $\triangle\sigma_i$ is the stress range under the i^{th} operating cycle.

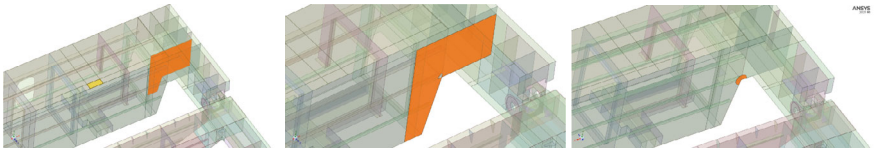
Characteristic fatigue strength can be selected based on the structural details, weld quality conditions, etc. provided in the standard.

Based on the stress results, the stress range at full load duty cycle can be calculated for 3 nodes. The work plan of the casting crane is relatively fixed, so the fatigue life is calculated according to the stress rang obtained from full load case. Based on the current work plan, it can be estimated that the work cycles per year are about 23,760. Thus, based on the aforementioned stress amplitude method, the remaining fatigue life of 3 nodes are calculated, as shown in Table 4. it can be seen that the remaining fatigue life of node 1 is the lowest, less than half a year, because of its high stress range. Nodes 2 and nodes 3 have high remaining fatigue lives, and the other nodes will be close to or above these two node lives, indicating that most of the structure has possibly keep using. In order to ensure node 1 to keep to use, the improvement should be considered.

5 Improvement Schemes

As shown in Fig. 5, according to the structural details at node 1, there is a sudden change in the cross section of the main girder, the existence of splices at the bottom flange plate and the spacer plate inside, so the welds here are concentrated and also the stresses are concentrated. At the end span, the stresses are almost caused by supporting loads. Therefore, the local structure can be improved by sharing the load or removing the stress concentration, thus 3 improvement schemes are proposed, as shown in Table 5.

The remaining fatigue life of node 1 for the improvement schemes are shown in Table 6. Compared the three schemes, the improvement S1 expected to add a new web to help share the original web load, has the effect to improve the life, but the load sharing is limited. Further analysis shows that the stress concentration at the mid span



a) Improvement Program 1 b) Improvement Program 2 c) Improvement Program 3

Fig. 5 Shape and dimensions of the reinforcement plate

Table 5 Improvement schemes

Scheme	Improvements	Figure
S1	Adding a web near the main web to share the load	Figure 5a
S2	Opening holes in stress concentrations to release stresses	Figure 5b
S3	Based on S2, the nearby main web is replaced with a thick plate	Figure 5c

Table 6 Results of 3 improvement schemes at node 1

Item	Improvement S 1	Improvement S 2	Improvement S 3
Stress range/MPa	85.0	103.0	62.0
Corresponding number of cycles	556,459.2	335,491.2	1,537,747.2
Remaining number of cycles	176,299.2	0	1,157,587.2
Number of cycles per year	23,760	23,760	23,760
Remaining fatigue life/year	7.42	0	48.72

of main girder is caused by the welding of multiple plates, and the web is relatively thin to cause high stress range. Therefore, thickening the web is one of the effective ways. Compared improvement S2 and S3, the latter increases the thickness of the web, so the remaining fatigue life increase is higher. Improvement S2 opens a hole in this position to try to reduce stress, but the calculated stress increases on the contrary. It shows that the stress concentration is not caused by the structural rigidity, but by the weak structural strength. Therefore, improvement S3 removes the web and replaces it with a thicker plate. It can be seen that improvement S3 has the smallest area to be modified, then the improvement cost is also low and maintainability is convenience. Compared with other two improvements, improvemen S3 is more effective.

6 Conclusion

Through finite element modeling and analysis, the crack of he girder are caused by stress concentration. Compared with the stresses and remaining lives of the mid-span girder, it is necessary to improve the cracked structure to keep using. So,

three improvement schemes are proposed. According to comparing their stresses and remaining lives by the stress amplitude method, and considering the cost and maintainability, improvement S3 is finally selected. The structure is now in the improvement stage, and will be load tested to verify the availability of improvement S3 in the future. After that, a maintenance plan will be proposed and implemented.

Acknowledgements This research was supported by Open-ended Fund for Key Laboratory of Lifting Equipment's Safety Technology for State Market Regulation.

References

1. Xu G-N, Zhang Y-C, Zhang WW (2020) Analysis of fatigue crack extension in metal structures of cast cranes. Mech Des Manuf (03):13–17
2. Frankovsky P et al (2020) Experimental assessment of time-limited operation and rectification of a bridge crane. Materials 13(12)
3. Euler M, Kuhlmann U (2019) Fatigue of transverse stiffeners in light crane runway beams. Stahlbau 88(12):1176–1183 2019
4. Liu L (2020) Analytical study on the cracking of a 165 t casting crane bridge. Shanxi Metall 43(01):31–33
5. Yu W (2024) Research on stress test and fatigue life of main beam of casting crane. Shanxi Metall 47(01):4–7
6. Ren X (2012) Fatigue life study of tire-type container gantry crane structure based on MSC. Fatigue. Wuhan University of Technology, Wuhan
7. International Organization for Standardization (2016) Cranes—proof of competence of steel structures. ISO 20332-2016

Open Access This chapter is licensed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License (http://creativecommons.org/licenses/by-nc-nd/4.0/), which permits any noncommercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if you modified the licensed material. You do not have permission under this license to share adapted material derived from this chapter or parts of it.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.



Crawler Crane Lifting Path Planning Based on Improved RRT Algorithm



Zetao Li, Xiaochun Tie, Xin Wang, and Xuyang Cao

Abstract For simulating crawler crane lifting, a kinematic model is established by an improved DH (Denavit-Hartenberg) method. The regular geometry enveloping crane boom, the lifted object and the obstacles are modeled in the lifting task. In order to improve the lifting efficiency and safety, a crane lifting path planning algorithm is proposed, based on the bi-directional RRT (Rapidly Random Exploring Tree) algorithm where the principle of artificial potential field method is introduced. At the same time, through applying the B spline curve, the strategy of deleting redundant nodes of the path is generated to reduce the frequent action switching of the crane. Then the sequence of crawler crane lifting actions is solved by inverse kinematics. The simulation experimental results for the traditional RRT algorithm, the bi-directional RRT algorithm and the improved algorithm show that the improved algorithm proposed in this paper can effectively reduce the path search time and path length, and achieve the smooth crane lifting action sequence.

Keywords Crawler crane · Path planning · Kinematic model · Collision detection · Artificial potential field method · Improved bi-directional RRT algorithm

1 Introduction

With the rapid development of the global economy industrialization, the traditional crane industry is growing up towards the direction of intelligence and information technology. Simultaneously, crane lifting projects are facing more and more objects, which involve various fields. In addition, enterprises also gradually improve the requirements of lifting, and pay more attention to the crane's efficiency, functional

Z. Li · X. Wang (✉) · X. Cao
Dalian University of Technology, Dalian, China
e-mail: wangxbd21@163.com

X. Tie
Taiyuan Heavy Industry Co., LTD, Taiyuan, China

© The Author(s) 2026

H. Saman et al. (eds.), *The 9th International Conference on Advances in Construction Machinery and Vehicle Engineering*, Lecture Notes in Mechanical Engineering,
https://doi.org/10.1007/978-981-96-5527-4_11

diversity and safety [1]. Therefore, through application of path planning technology, accurate positioning and real-time monitoring combined with digital twin, it is great significance for the realization of the crane's intelligence, informationization and unmanned operation [2].

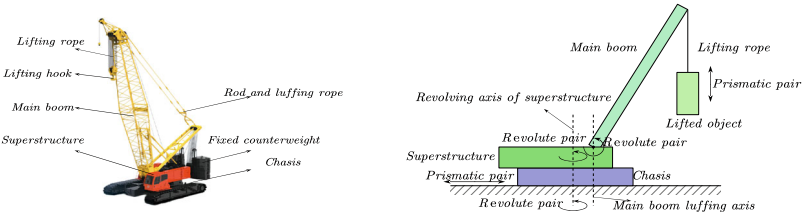
More and more algorithms on crane lifting path planning are generally studied at home and abroad, such as genetic algorithm [3], A* algorithm [4], Rapidly Random Exploring Tree [5], ant colony algorithm [6] and artificial potential field method [7]. However, current researches, to a certain extent, still exist with low path planning efficiency, lack of goal orientation, low path smoothness and other problems [8–10]. In this paper, the rigid body model and kinematics model are firstly established for a crawler crane. And the traditional RRT algorithm is improved by combining the artificial potential field method. Then, the collision detection algorithm of crane lifting path search is optimized, where the redundant nodes are deleted to smooth the lifting path by B-spline curve. Finally, the sequence of crane lifting actions is obtained through the inverse kinematics solution. Based on the simulation and comparison of MATLAB algorithm, the rationality of the proposed algorithm is verified.

2 Multi-degree-of-Freedom Mathematical Model of Crawler Crane

2.1 Crawler Crane Rigid Body Model

As the research object in this paper, TGZC180 crawler crane with main boom is shown in Fig. 1a, which mainly includes chassis, superstructure, and main boom. The mechanisms include lifting, luffing, slewing and traveling mechanism. From these mechanisms, the object can be lifted, lowered, rotated or travelled from starting position to target position.

In a crane rigid model, according to geometrical shapes of its parts, the chassis and superstructure can be abstracted into two cuboid bodies, and the boom can be abstracted into a cylindrical block, and the lifting ropes are abstracted into a linear



a) TGZC180 crawler crane with main boom b) Rigid body model

Fig. 1 Crawler crane and its rigid body model

rigid body that can be moved up and down, as shown in Fig. 1b. The lifted object is set as a cylinder, so that the rotational degree of freedom of the lifting hook can be canceled out, at this time, the rigid body model has 5 degrees of freedom, including the lifting rope up and down, boom luffing, superstructure slewing, chassis travelling longitudinal, transverse and turning.

2.2 Kinematic Modeling

According to the crane load capacity with its radius up and lower limit, the lifting capacity ring ranges both at starting position and target position can be sketched (top view), as shown in Fig. 2. In the intersecting area of the two ring ranges, a suitable standing position of crane is selected. At this standing position, the whole crane can be simplified to be a three-degree-of-freedom robotic arm if no travelling, and the kinematic model is built by the improved DH (Denavit-Hartenberg) method, as shown in Fig. 3.

In Fig. 3, d_1 is the total height of the crawler crane superstructure and chassis, a_2 is the distance between the slewing axis of the superstructure and the boom foot

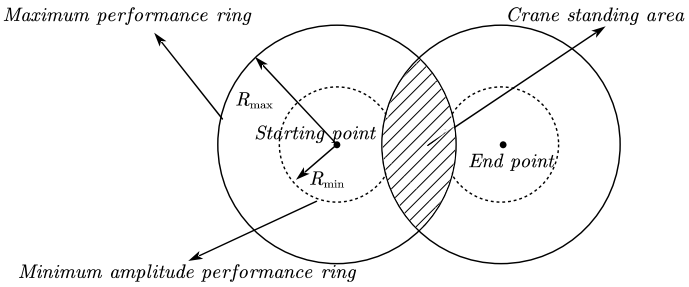


Fig. 2 Schematic diagram of crane station area

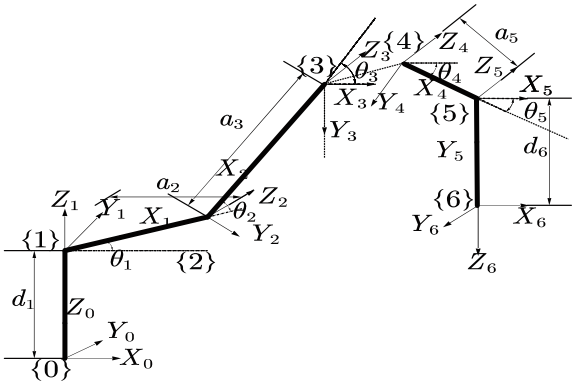


Fig. 3 Kinematic model of crane lifting system

point, a_3 is the length of the main boom, a_5 is the offset distance of the pulley block, d_6 is the lifting rope height, θ_1 is the slewing angle of the superstructure, and θ_2 is the boom angle. In addition, the lifting rope is always vertically, so the relationship among the boom angle θ_2 and the three angles of the boom head , θ_3, θ_4 and θ_5 can be set as follows,

$$\theta_3 = -\theta_2 \; \theta_4 = 90 - \theta_3 \; \theta_5 = -\theta_4 \tag{1}$$

According to the kinematic model, the position at the end of the lifted object can be expressed as:

$$T_0^6 = T_0^1 T_1^2 T_2^3 T_3^4 T_4^5 T_5^6 \tag{2}$$

From Eq. (1), θ_3, θ_4 and θ_5 can be expressed by θ_2, then,

$$T_0^6 = \begin{bmatrix} \cos(\theta_1) & \sin(\theta_1) & 0 & \cos(\theta_1) * (a_2 + a_3\cos(\theta_2) + a_5\sin(\theta_2)) \\ \sin(\theta_1) & -\cos(\theta_1) & 0 & \sin(\theta_1) * (a_2 + a_3\cos(\theta_2) + a_5\sin(\theta_2)) \\ 0 & 0 & -1 & d_1 - d_6 - a_5\cos(\theta_2) + a_3\sin(\theta_2) \\ 0 & 0 & 0 & 1 \end{bmatrix} \tag{3}$$

With known coordinates of the end of the lifted object (P_x, P_y, P_z), it is possible to invert the data for each joint angle variable. Therefore, it is possible to set the lifted object position transformation matrix $T = T_0^6$ as follows,

$$T = \begin{bmatrix} n_x & o_x & a_x & P_x \\ n_y & o_y & a_y & P_y \\ n_z & o_z & a_z & P_z \\ 0 & 0 & 0 & 1 \end{bmatrix} \tag{4}$$

It can be obtained from Eqs. (3) and (4) as follows,

$$\begin{cases} \theta_1 = \text{atan2}(P_y, P_x) + \text{atan2}(0, \pm 1) \\ \theta_2 = \text{atan2}(a_3, a_5) - \text{atan2}\left(\dfrac{\sqrt{P_x^2 + P_y^2} - a_2}{\sqrt{a_3^2 + a_5^2}}, \pm\sqrt{1 - \left(\dfrac{\sqrt{P_x^2 + P_y^2} - a_2}{\sqrt{a_3^2 + a_5^2}} \right)^2} \right) \\ d_6 = d_1 - P_z + a_3\sin(\theta_2) - a_5\cos(\theta_2) \end{cases} \tag{5}$$

3 Lifting Path Planning Algorithm

3.1 Collision Detection

From the rigid body model of the crane, it can be seen that the main boom and the lifted object are enveloped by a cylinder. Obstacles in the lifting environment can be modeled by regular rectangular, cylindrical, and spherical bounding boxes. In this way, the collision detection during lifting process can be simplified to the collision detection between any two of the regular geometries, such as the main boom bounding box, obstacles of various bounding boxes, and the lifted object bounding box.

In the case of collision detection between main boom and lifted object, when the lifting rope rises or boom luffing, the top of the lifted object and the main boom may have possibility of collision, as shown in Fig. 4. Therefore, the non-collision criteria between should satisfy as follows,

$$L > L_{\min} \tag{6}$$

In the case of collision detection between main boom and obstacle, the main boom of the crane may collide with the obstacles while working, as shown in Fig. 5. Therefore, the non-collision criterion should satisfy as follows,

$$\begin{cases} \theta_2 > \alpha | L_2 > L_{2\min} | (\theta_2 \le \alpha \& L_2 \le L_{2\min} \& L_1 > L_{1\min}) \ (\textit{If the obstacle bounding box as sphere}) \\ \theta_2 > \alpha | L_2 > L_{2\min} | (\theta_2 \le \alpha \& L_2 \le L_{2\min} \& L_3 > L_{3\min}) \ (\textit{If the obstacle bounding box as cylinder}) \\ \theta_2 > \alpha | P | \theta_2 \le \alpha \& \overline{P} \& L_{5_i} > R_1 (i = 1, 2, 3) \ (\textit{If the obstacle bounding box as cuboid}) \\ P = ((x_o < x_{\min}) | (x_o > x_{\max})) \& ((y_o < y_{\min}) | (y_o > y_{\max})) \end{cases} \tag{7}$$

In the case of collision detection between suspended objects and obstacles, the lifted object may collide with the obstacle, as shown in Fig. 6. Therefore, the non-collision criteria should satisfy the as follows.

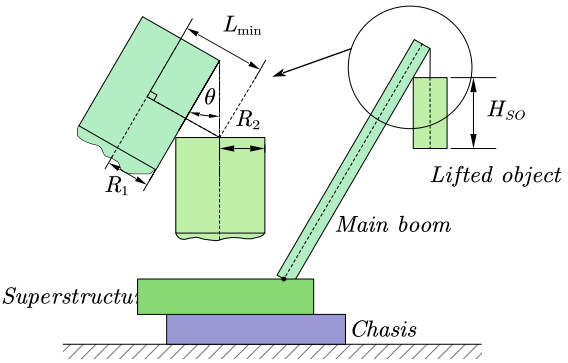
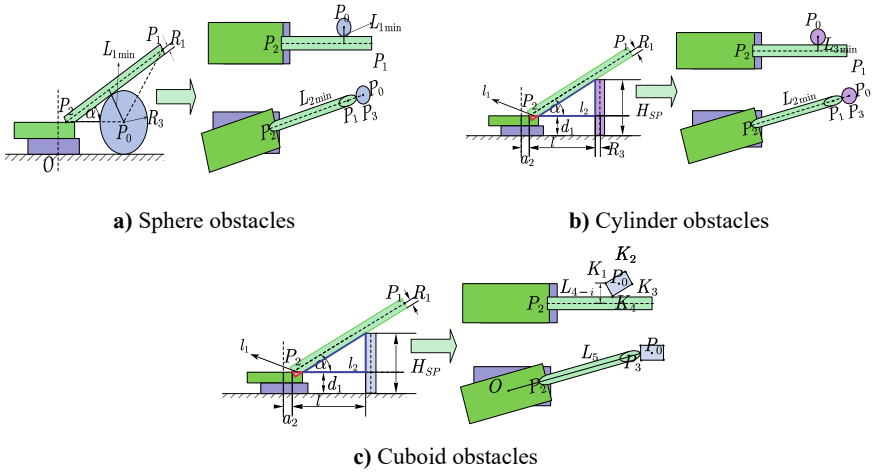


Fig. 4 Schematic diagram of boom and suspended object collision test

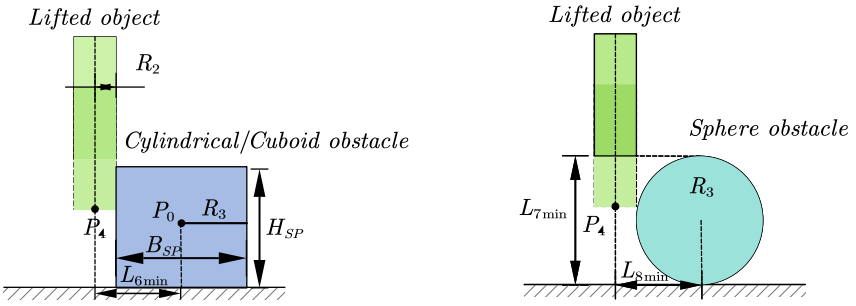


a) Sphere obstacles **b)** Cylinder obstacles

c) Cuboid obstacles

Fig. 5 Schematic diagram of boom and obstacle collision test

$$\begin{cases} z_4 > H_{SP} | (((x_4 < x_{1_\min}) | (x_4 > x_{1_\max})) \& ((y_4 < y_{1_\min}) | (y_4 < y_{1_\max}))) \\ (\textit{If the obstacle bounding box as sphere}) \\ z_4 > H_{SP} | ((z_4 \le H_{SP}) \& (L_6 > L_{6\min})) \\ (\textit{If the obstacle bounding box as cylinder}) \\ z_4 > L_{7\min} | ((z_4 \le L_{5\min}) \& (L_8 > L_{8\min})) \\ (\textit{If the obstacle bounding box as cuboid}) \end{cases} \tag{8}$$



a) Cylindrical/rectangular obstacle **b)** Spherical obstacle

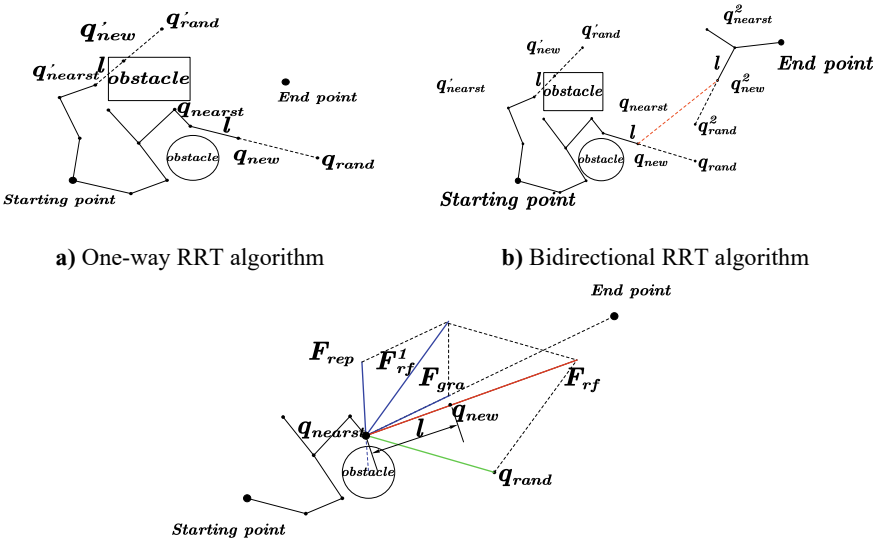
Fig. 6 Collision diagram between suspended object and obstacle (side view)

3.2 Improved Two-Way RRT Algorithm

The basic principle of RRT (Rapidly Random Exploring Tree) algorithm is to randomly sample node information in the path planning environment and construct a path search tree, which is gradually expanded from the starting position to the target position according to a certain search step. Eventually, the feasible nodes in the search tree are connected as collision-free paths. Bidirectional RRT algorithm is the introduction of a second search tree in the basic RRT algorithm, i.e., the path from the end point to the beginning of the expansion, when the two search trees converge, the path search is completed (as shown in Fig. 7b).

In this paper, to make its search path more oriented, the path expansion principle of the artificial potential field method is applied in the two-way RRT algorithm. The artificial potential field method is to add the concepts of gravitational force and repulsive force in mechanics into the path expansion method, i.e., the target point is attractive to the newly expanded point, and the obstacle has repulsive effect on the newly expanded point. The gravitational force equation F_{gra} and the repulsive force equation F_{rep} are shown in Eqs. (9) and (10):

$$F_{gra} = \varepsilon\rho_{(q_{goat}-q_{neast})} \tag{9}$$

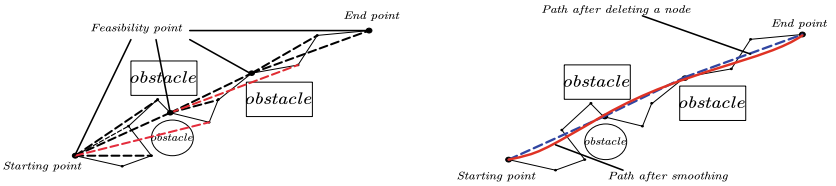


a) One-way RRT algorithm

b) Bidirectional RRT algorithm

c) Improved RRT algorithm

Fig. 7 Algorithm node expansion schematic



a) Flowchart for screening of redundant nodes **b)** Path after screening

Fig. 8 Schematic diagram for removing redundant nodes

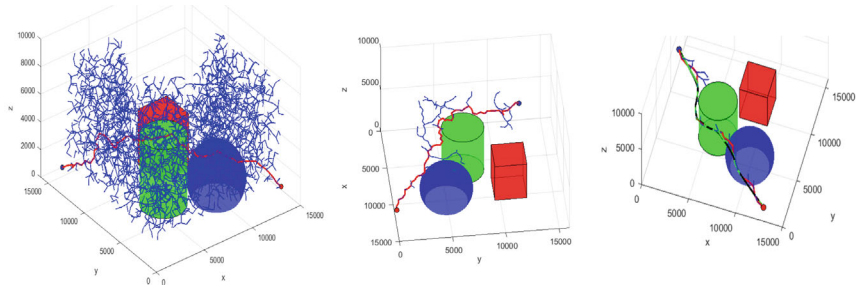
$$F_{rep} = \begin{cases} \sigma\left(\frac{1}{\rho_{(q_{nearst},q_{obj})}} - \frac{1}{\rho_0}\right)\frac{1}{\rho^2_{(q_{nearst},q_{obj})}} & \rho_{(q_{nearst},q_{obj})} < \rho_0 \\ 0 & \rho_{(q_{nearst},q_{obj})} > \rho_0 \end{cases} \tag{10}$$

Combined with the artificial potential field method, the direction of the path extension is changed into the vector direction both the combined force F^1_{rf} of the potential field and the force F_{rf} of the random point, as shown in Fig. 7c. Such an extended path increases the goal orientation of the algorithm search due to the gravitational force, and improves the algorithm search efficiency, and also allows the search path to be smoother.

Valid nodes can be filtered and redundant nodes removed from the planned path to further simplify the path. The principle is shown in Fig. 8, where the searched path nodes are connected sequentially and collision detection is performed to find all feasible nodes and connect them into a new path. The trimmed path is smoothed using the cubic B-spline curve to get the final reasonable and effective lifting path.

4 Algorithm Simulation Case and Analysis

In order to verify the practicality of the improved algorithm on crane lifting work, the algorithm comparison case is carried out through Matlab software, and the number of tests is 30 times. In Fig. 9, the simulation results are shown that the black path of the improved algorithm is more goal-oriented than other two blue (RRT) and red paths (Bi-directional RRT), and the comparison between the red path (Bi-directional RRT) and the black path (the improved algorithm) can be clearly observed that the last one is smoother. The relevant average data of the path results are shown in Table 1 and Fig. 10, which can be observed that the improved algorithm takes less time to plan, and is more efficient, also the number of planned path nodes is smaller, and the length of the path is shorter.

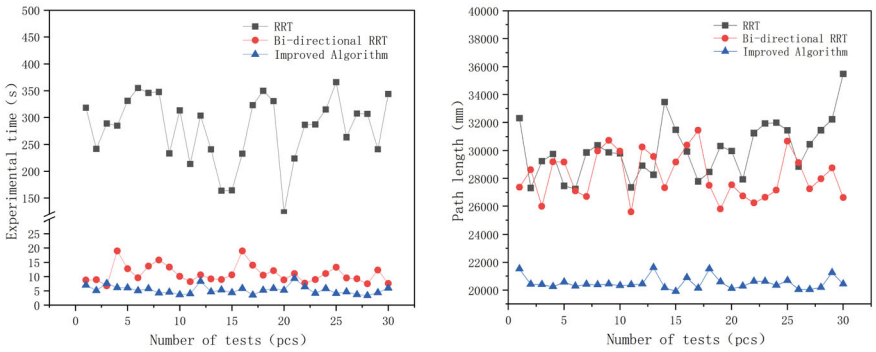


a) RRT **b)** Bi-directional RRT **c)** Improved algorithm

Fig. 9 Algorithm simulation comparison experiment

Table 1 Simulation results (average values)

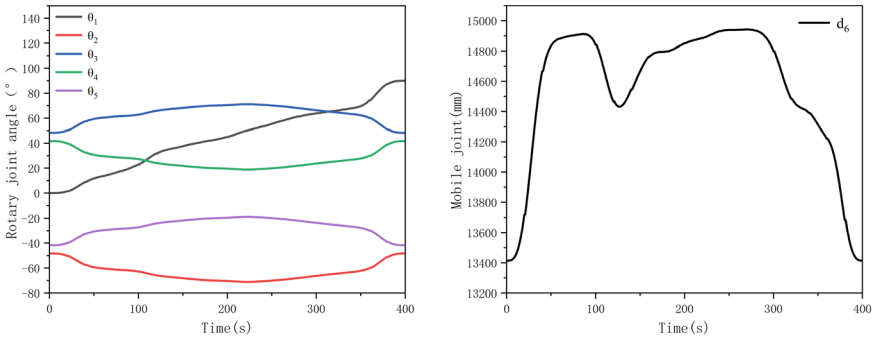
Algorithm type	Planning time (s)	Path length (mm)	Number of nodes (pcs)
RRT	282	30,076	62
Bi-directional RRT	11	28,134	72
Improved Algorithm	5	20,363	15



a) Comparison chart of test time **b)** Comparison of test path lengths

Fig. 10 Comparison experimental graph of parameters related to algorithm simulation

Figure 11 shows the sequence diagram of the crane action planned by the improved algorithm. Figure 11a is the rotating joint angle change diagram, and Fig. 11b is the lifting rope length diagram (showing the change of the height of the lifted object). It can be found that the crane planned by the improved algorithm has a smoother curve of the joints, which makes the crane lifting movement smoother and improves the efficiency and safety of the crane lifting.



a) Rotary joints **b)** Moving joints

Fig. 11 Improved algorithm to plan the sequence of crane movements

5 Conclusion

In this paper, for the crawler crane lifting path planning task, the crawler crane is simplified into a three-degree-of-freedom robotic arm, and the kinematic model of the crawler crane is established using the improved DH method. Meanwhile, in order to solve the problems of strong randomness of search path, more path inflection points and low planning efficiency of traditional algorithms in crane lifting path planning, an improved path planning algorithm is proposed. The cases of different algorithms are simulated. The results show that the improved algorithm not only effectively improves the efficiency and stability of crane lifting path planning, but also shortens the length of the path and improves the smoothness of the path. Dynamic obstacles should be taken into account in the lifting environment in future studies.

Acknowledgements This research was supported by the Fundamental Research Funds for the Central Universities (No. DUT22LAB506, DUT22LAB507).

References

1. Xiao Z (2023) Research on intelligent path planning of tower crane. Taiyuan University of Science and Technology
2. Geng MW (2023) Research on intelligent crane path planning. Taiyuan University of Science and Technology
3. Anwaar A, Ashraf A, Bangyal WHK et al (2022) Genetic algorithms: brief review on Genetic Algorithms for Global Optimization Problems. IEEE Global Assoc Acad Supervis; Shanghai Univ; London S Bank Univ;Xian Jiaotong Liverpool Univ; Zhejiang Univ Sci & Technol, pp 80–85.
4. Miyombo ME, Liu Y, Mulenga CM et al (2024) Optimal path planning in a real-world radioactive environment: a comparative study of A-star and Dijkstra algorithms. Nucl Eng Des 420

5. Kuffner JJ, La Valle SM (2000) RRT-connect: an efficient approach to single-query path planning. In: IEEE international conference on robotics and automation, pp 995–1001. San Francisco, CA, USA
6. Khanesar MA, Yan M, Kendal P et al (2023) Intelligent static calibration of industrial robots using artificial bee colony algorithm. Inst Elect& Elect Engineers. IEEE Ind Applicat Soc
7. Hwang YK, Ahuja N (1992) A potential field approach to path planning. Robot Autom, IEEE Trans 8(1):23–32
8. Kang Y-W, Chen P-Y, Wang X et al (2023) A kinematic modelling method for crawler cranes oriented to lifting path planning. Mechatron Eng 40(2):307–316
9. Qian S, Qian P, Wang C et al (2022) Dynamics and path planning of multi-crane cooperative lifting robot. J Mech Eng 58(7):20–31
10. You Y, Quanyin C, Peng G et al (2023) Construction technology of onshore wind turbine crawler crane assembly and disassembly. Hydropower New Energy 37(7):72–75

Open Access This chapter is licensed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License (http://creativecommons.org/licenses/by-nc-nd/4.0/), which permits any noncommercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if you modified the licensed material. You do not have permission under this license to share adapted material derived from this chapter or parts of it.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.



Multi-scale Assembly Feature-Driven Information Modeling for Diesel Engine Assembly Processes



Fei Tan, Yangang Zhang, Meiwen Cao, Yu Li, Guoxiang Liang, Shouqi Hu, and Jianjun Chen

Abstract A high-quality assembly process information model forms the key foundation for the digital assembly of mechanical products. This paper addresses the challenges inherent in traditional diesel engine assembly process information modeling, including the large number of components, complex structures, vast amounts of process data, and the high precision required for local assembly. It proposes a multi-scale, assembly feature-driven information modeling method tailored to the diesel engine assembly process. This method is designed to meet the actual needs of the entire diesel engine assembly process and involves reconstructing the diesel engine assembly process BOM through a comprehensive analysis of all relevant elements. The process is divided into three scales: assembly units, assembly elements, and assembly components. At each scale, different methods are employed to define assembly process information: skeleton models for units, ontology models for elements, and interactive feature pairs for components. This approach not only enhances the structuring and integration of assembly process information but also reinforces hierarchical relationships within the assembly, laying a solid foundation for the subsequent design of three-dimensional process procedures and visual guidance in the diesel engine assembly workshop.

Keywords Diesel engine · Assembly process information model · Digital definition · Multi-scale · Assembly features

F. Tan · Y. Zhang (✉)
North University of China, Taiyuan, China
e-mail: zyg31124@163.com

M. Cao · Y. Li · G. Liang · S. Hu · J. Chen
Northern General Power Group Co., Ltd, Datong Shanxi, China

© The Author(s) 2026

H. Saman et al. (eds.), *The 9th International Conference on Advances in Construction Machinery and Vehicle Engineering*, Lecture Notes in Mechanical Engineering,
https://doi.org/10.1007/978-981-96-5527-4_12

1 Introduction

Assembly processes play a crucial role in the design and manufacturing of modern mechanical products, serving as one of the core factors influencing product performance, quality, development cycles, and costs. The final performance of a product is often determined by the quality of its assembly process [1]. As China's manufacturing industry transitions towards digitalization and intelligent manufacturing, a significant gap has emerged between the rapid digital development of parts manufacturing processes and the slower digitalization of assembly processes. Diesel engine products, with their large number of parts, complex structures, extensive process data, and stringent local assembly requirements, impose even stricter demands on the digital design and implementation of assembly processes. A high-quality assembly process information model is essential for the digital definition of assembly processes, consistent expression of assembly quality, and optimization of assembly workflows [2].

Research on diesel engine assembly process information modeling, both domestically and internationally, primarily focuses on geometric modeling, process simulation, and information integration, and has achieved certain results. Huang [3] and Bian [4] researched methods for defining assembly information within MBD (Model-Based Definition), using digital 3D models to comprehensively express design and manufacturing data during the assembly process, thus advancing the digitalization of diesel engine manufacturing. Wang [5] studied the application of BOM (Bill of Materials) technology in diesel engine manufacturing, focusing on full lifecycle data management and resource optimization, thereby enhancing overall manufacturing efficiency and data management capabilities. Sun [6] investigated the design and key technologies of assembly tooling for marine diesel engines, addressing accuracy and efficiency issues in the assembly process, which laid the foundation for the digital manufacturing of diesel engines. Yang [7] developed an advanced engine manufacturing model based on digital twin technology, improving manufacturing efficiency and accuracy through simulation and model lightweighting. Oleksiy et al. [8] created a digital twin model for diesel engines to predict the dynamic behavior of propulsion systems. Bao et al. [9] proposed a part-oriented digital twin modeling method for assembly, using ontology theory to define and identify machining and key assembly features. Hiraoka et al. [10] studied a lifecycle simulation-oriented assembly model for mechanical products, which includes information about assembly structure, behavior, aging, and deviations of parts and assemblies during use. Ding et al. [11] researched a multi-stage rotation optimization method for aircraft engine assembly, based on a unified Jacobian-Torsor model. Currently, the digital assembly of diesel engines is in a stage of both development and exploration. Although progress has been made in digital technology, information management, and simulation optimization, there are still significant challenges in achieving widespread application and deep integration.

Despite advances in digitalization, the diesel engine assembly process still faces challenges such as complex processes, data consistency issues, and the need for

standardization and systematization of assembly process information. Different data formats and protocols can lead to information loss or distortion during system transfers, and semantic differences between data sets may cause misunderstandings or operational errors. To address these challenges, a multi-scale assembly feature-driven information modeling method was developed for assembly processes. By dividing the entire diesel engine into three scales—assembly units, assembly elements, and assembly components—different information modeling methods are applied at each scale, which are then integrated into a 3D model to create an MBD assembly process information model. This multi-scale approach effectively tackles the challenges posed by the numerous parts and complex processes of diesel engines. MBD technology, which uses the 3D model as a carrier, seamlessly combines geometric information, assembly sequences, and process specifications, integrating the information from various scales into the 3D model. This resolves data consistency issues and achieves standardized, integrated information management, enhancing data sharing and reusability. Consequently, the multi-scale assembly feature-driven MBD assembly process information modeling is a crucial step in achieving the digital transformation of diesel engine manufacturing and serves as an effective means to improve assembly process efficiency and quality.

2 Multi-scale Diesel Engine MBD Assembly Information Modeling

The MBD assembly process information modeling for diesel engines aims to achieve digitalization, standardization, and visualization of assembly process information by integrating geometric data, assembly sequences, process parameters, and other essential elements. As illustrated in Fig. 1, this step involves defining the scope and classification of the content that the information model needs to express, laying the groundwork for subsequent information definition. Next, the focus shifts to studying methods for expressing assembly process information. By dividing the entire diesel engine into multiple scales—assembly units, assembly elements, and assembly components—and analyzing assembly features, different methods are employed to define information at each scale. Assembly components are defined using the interactive feature pair method [12], assembly elements are defined based on ontology methods, and assembly units are defined using a skeleton line approach. Finally, these information definition methods are used to model the fully defined assembly process information, resulting in a comprehensive diesel engine assembly process information model. By integrating this model with the diesel engine's 3D model, the diesel engine MBD assembly information model is created.

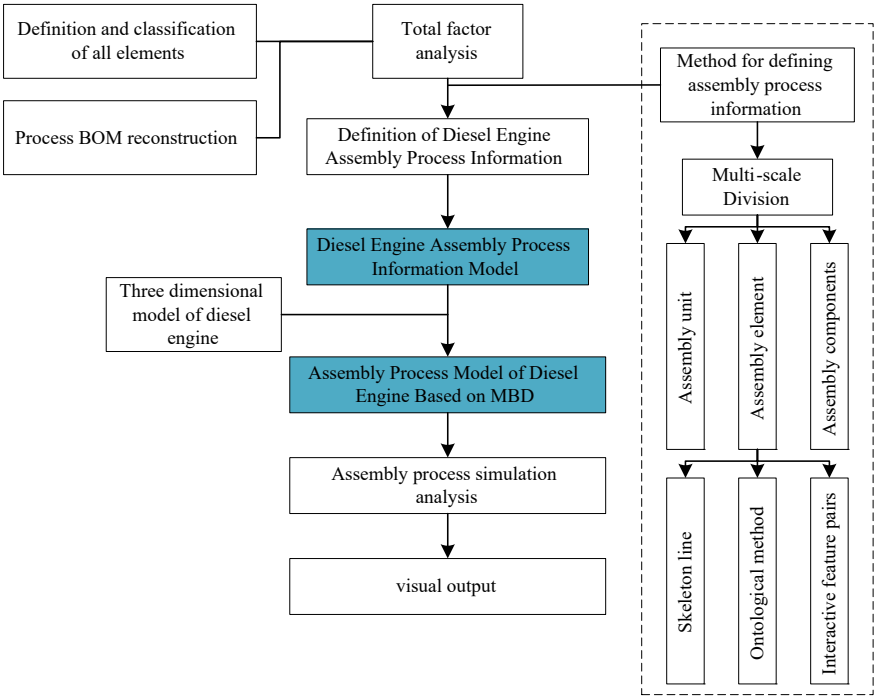


Fig. 1 Diesel engine assembly process information modeling workflow

3 Comprehensive Analysis of Diesel Engine Assembly Process Elements

In diesel engine assembly process modeling, a comprehensive analysis of the relevant datasets is required as an initial step. This includes the transformation from the design BOM to the process BOM and the definition and classification of all involved elements. The design BOM primarily focuses on parts information during the product design phase, while the process BOM is based on the assembly process workflow and includes information such as assembly sequences, process requirements, and assembly relationships. Therefore, reconstructing the process BOM is fundamental to building an assembly process information model. The information related to diesel engine assembly processes encompasses various types, which can be broadly categorized into geometric and non-geometric information. These cover all the critical elements involved in the assembly process. After thoroughly analyzing all elements, the design BOM is reconstructed into the process BOM, based on the distinctions between design information and assembly information.

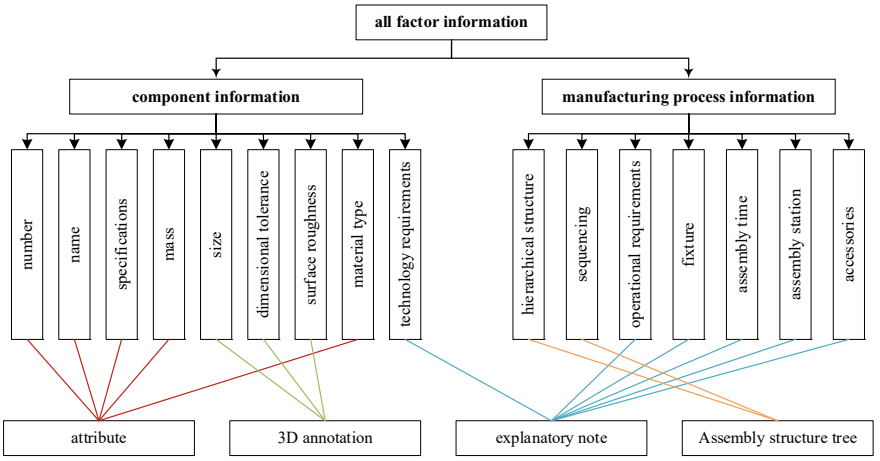


Fig. 2 Classification and annotation methods of all elements information

3.1 Definition and Classification of All Elements in Assembly Processes

Under the framework based on MBD technology, the definition of comprehensive information for diesel engines includes both geometric and non-geometric data, which are precisely described and managed using different methods. As shown in Fig. 2, all elements are categorized into part information and assembly information, each with corresponding annotation methods.

3.2 Process BOM Reconstruction Method

Diesel engines are characterized by complex structures, diverse components, and strict assembly sequence requirements. It is essential to consider specific manufacturing processes, assembly order, required tools and equipment, and the dependencies between different parts during assembly. In this context, a well-established rule-based method for process BOM reconstruction has been selected. This method allows for the pre-definition of transformation rules, ensuring that the generated process BOM accurately reflects these complex process requirements and assembly logic. The rule-based approach involves creating assembly coding rules, as illustrated in Fig. 3. Each level of nodes, components, or parts in the design BOM is assigned corresponding assembly codes as basic attributes. This enables the structural reconstruction from the design BOM to the process BOM. The process starts by inheriting the basic information of the parts, followed by assigning attributes related to the assembly process to each part based on a comprehensive analysis of all elements.

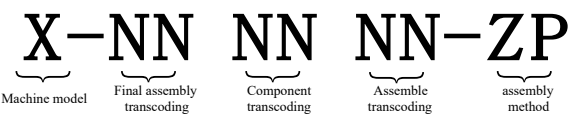


Fig. 3 Assembly coding rules

4 Method for Defining Diesel Engine Assembly Process Information

This paper constructs a diesel engine assembly process information modeling framework based on the assembly semantic knowledge information model, as shown in Fig. 4. Based on this assembly semantic process knowledge information model, a multi-scale assembly feature-driven diesel engine assembly process information modeling method is proposed. This method divides the entire diesel engine into multiple scales, including assembly units, assembly elements, and assembly components. Assembly units are positioned and ordered, forming a diesel engine assembly skeleton based on these units. Ontological information is defined between assembly elements, and interactive feature pairs are used for information modeling of assembly components.

4.1 Multi-scale Division

When dividing the entire diesel engine into assembly units, assembly elements, and assembly components, the division rules should follow these four principles:

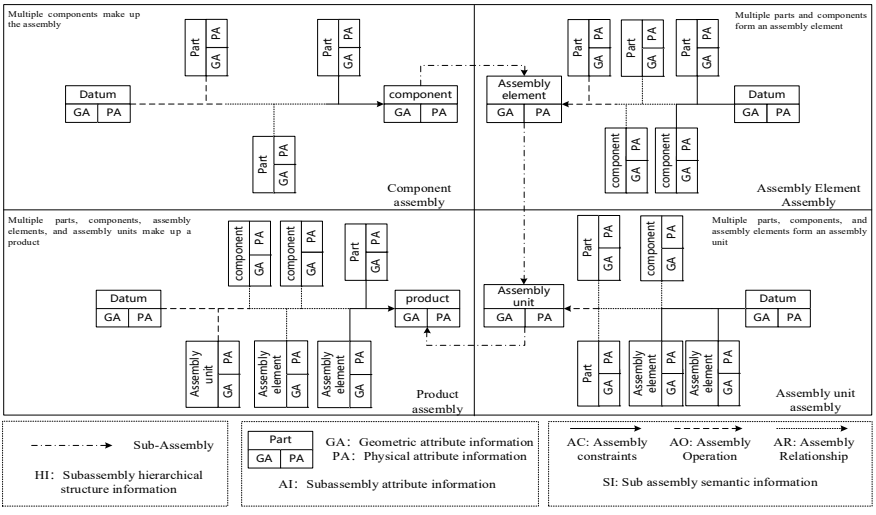


Fig. 4 Assembly semantic process knowledge information model

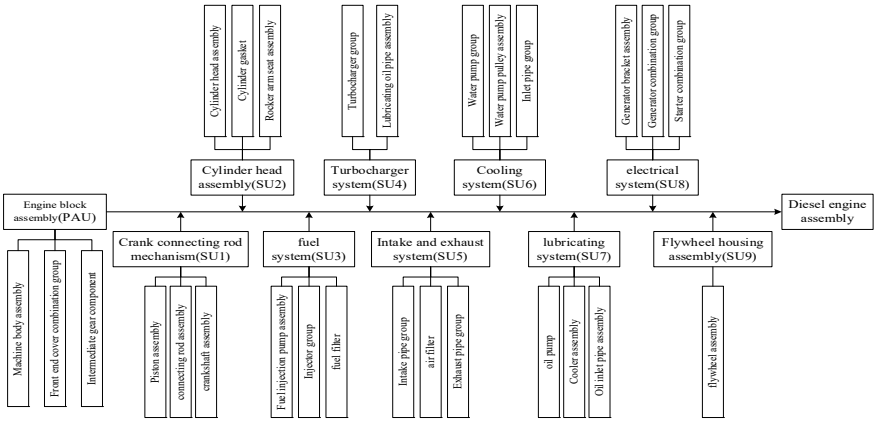


Fig. 5 Diesel engine assembly skeleton model based on assembly units

Functional independence and integrity: Each assembly component should possess a certain degree of functional independence and completeness.

Complexity of assembly relationships: The division of assembly elements should consider the complex relationships among parts and components during the assembly process.

Spatial position and geometric features: The division of assembly units should be based on the spatial positions and geometric features of parts and assembly elements.

Rationality of process and assembly sequence: The division should take into account the actual assembly processes and sequences, ensuring that each assembly unit, assembly element, and assembly component can be assembled in the planned order, avoiding conflicts or unnecessary operations.

Following the aforementioned rules, the entire diesel engine is divided into various assembly units, and the assembly sequence is arranged to form the diesel engine assembly skeleton, as shown in Fig. 5. Based on the assembly skeleton line, the assembly units can be categorized into primary units and subordinate units. The engine block assembly is defined as the primary assembly unit (PAU), while other subordinate units are sequentially installed onto the primary unit according to their assembly positioning and order. The subordinate units are sequentially labeled as SU1, SU2, and so on, based on their assembly order relative to the primary assembly unit.

4.2 Definition Method

The process information definition between assembly elements is carried out using the ontology method. This method systematically describes the entire assembly

process by establishing logical connections, assembly features, and sequence relationships between assembly units. The process information definition for assembly components is carried out using the interactive feature pair method. This method describes the assembly relationships by defining interactive feature pairs between parts.

4.3 Analysis of a Crank and Connecting Rod Mechanism Example

Through structural and assembly process analysis of the diesel engine, a diesel engine assembly semantic process knowledge information model was constructed using the piston → piston connecting rod assembly → crank and connecting rod mechanism → diesel engine as an example, as shown in Fig. 6. In the diesel engine assembly process, the crank and connecting rod mechanism is a critical assembly unit.

During the assembly process of the crank and connecting rod mechanism, assembly modeling is carried out sequentially according to the assembly hierarchy shown in Fig. 7. These parts are identified, and their positions and functions within the mechanism are determined. These parts, with clearly defined interactive feature pairs, are then combined into assembly components to create an assembly component

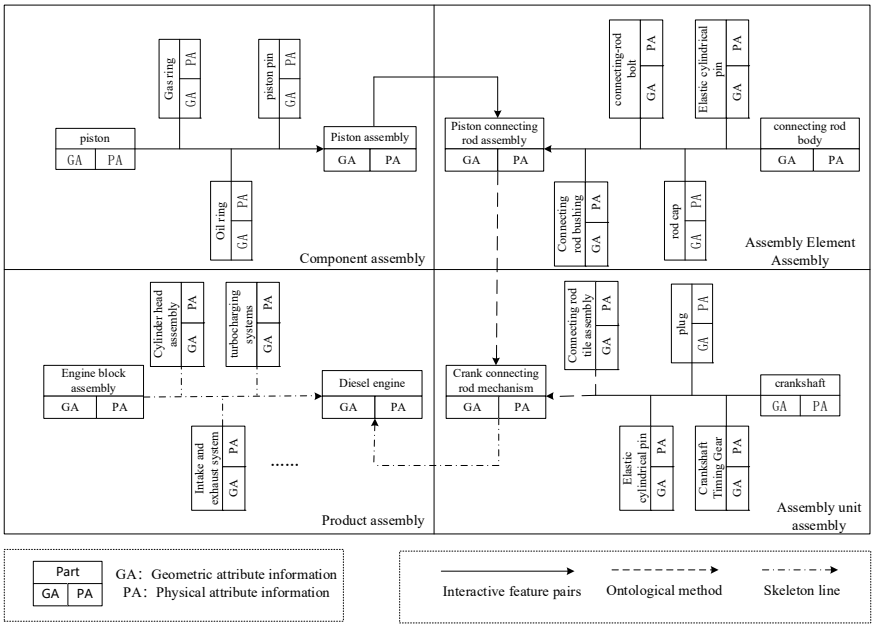


Fig. 6 Assembly semantic process knowledge information model of the diesel engine crank and connecting rod mechanism

information model. In the ontology, the attributes and assembly relationships of the assembly elements are defined. Based on the assembly relationships defined in the ontology model, specific process steps between assembly elements are broken down and embedded into the ontology model to create the assembly element information model. As shown in Fig. 7, these represent the MBD assembly process information models for the piston assembly and the piston connecting rod group, respectively.

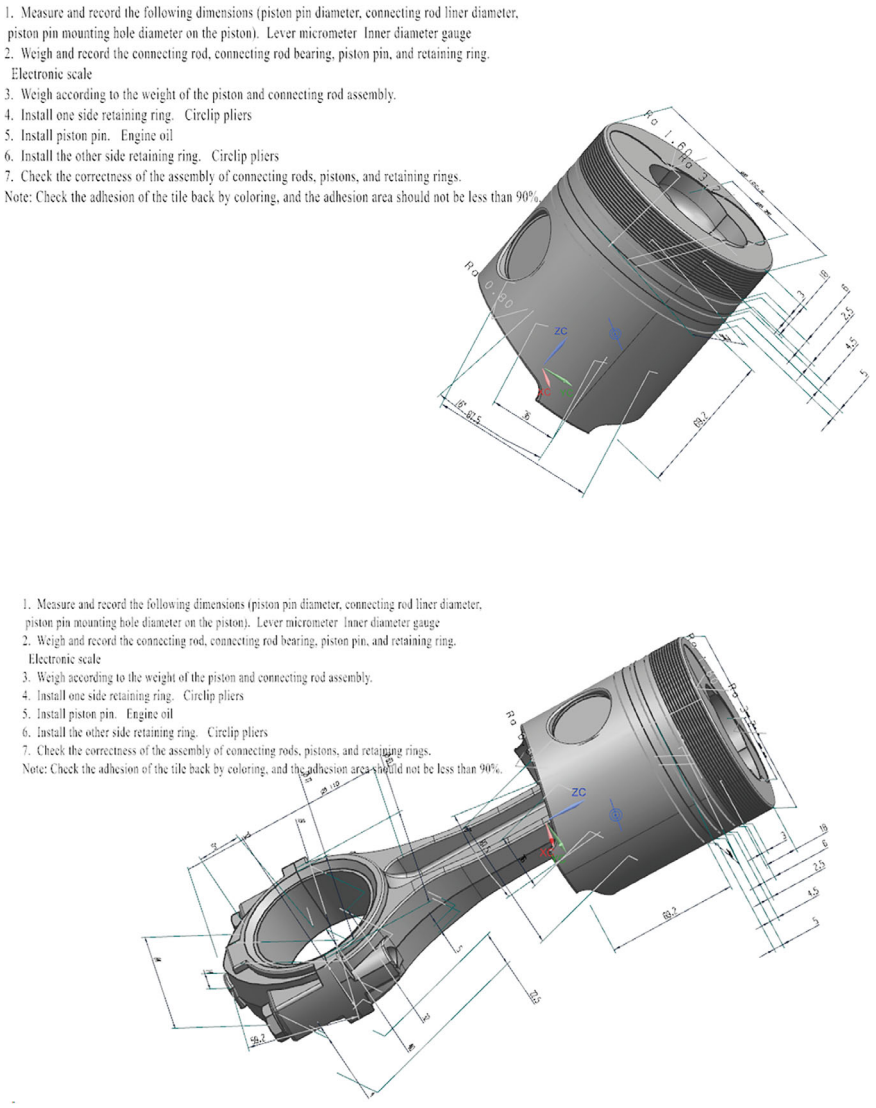


Fig. 7 Piston assembly MBD assembly information model, piston connecting rod group MBD assembly information model

5 Conclusion

The construction of a comprehensive, hierarchically clear, and semantically explicit assembly process information model is crucial for improving the digital assembly quality and efficiency of mechanical products. This article addresses the shortcomings in the current digital definition of diesel engine assembly processes by proposing a multi-scale, assembly feature-driven information modeling method. This method overcomes the large data volume limitations of traditional assembly process ontology modeling by focusing on assembly features at three scales: assembly units, assembly components, and assembly elements. It uses skeleton models, ontology models, and interactive feature pairs to define the full range of process information at different scales, enhancing the structuring and integration of assembly process information, and reinforcing the assembly hierarchy and constraints at various levels. The research outcomes provide methodological support for the future digital design and intelligent assembly of diesel engine products.

Acknowledgements This work is supported by Basic Product Innovation project of China and Basic Research Programs of Shanxi Province in China (Grant No. 202403021211075).

References

1. Liu JH, Xia HX, Gong H et al (2023) Connotation, technical system, and development trends of precision assembly. J Mech Eng 59(20):436–450
2. Li JY, Zhang PF, Guo YC et al (2024) Assembly analysis of mating interfaces in aero engines based on digital twin. Acta Aeronautica et Astronautica Sinica 1–19
3. Huang S (2016) Research on 3D assembly and manufacturing technology of marine diesel engines based on MBD. Harbin Engineering University
4. Wang YJ (2015) Research on full lifecycle data management and resource allocation of diesel engines based on MBOM. East China University of Science and Technology
5. Bian HJ (2017) Research on assembly information definition technology for marine diesel engines based on MBD. Jiangsu University of Science and Technology
6. Sun HR (2022) Research on design and key technologies of assembly tools for marine diesel engines. Jiangsu University of Science and Technology
7. Yang ZP (2023) Construction and application of digital twin models for advanced engine manufacturing processes. Shenyang University of Technology
8. Oleksiy B, Tetsugo F (2020) Development of a diesel engine's digital twin for predicting propulsion system dynamics. Energy 196:117–126
9. Bao Q, Zhao G, Yu Y et al (2022) Ontology-based modeling of part digital twin oriented to assembly. Proc Inst Mech Eng, Part B: J Eng Manuf 236(1–2):16–28
10. Hiraoka H, Tanaka T, Sugahara T et al (2021) Assembly model of mechanical products for their life cycle simulation. Procedia CIRP 98(2):553–558
11. Ding SY, Jin S, Li Z et al (2019) Multistage rotational optimization using unified Jacobian-Torsor model in aero-engine assembly. Proc Inst Mech Eng, Part B: J Eng Manuf 233(1):251–266
12. Xu ZJ, Wang QH, Li JR (2018) Design intent modeling method based on assembly feature Pairs. J Mech Eng 54(01):214–222

Open Access This chapter is licensed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License (http://creativecommons.org/licenses/by-nc-nd/4.0/), which permits any noncommercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if you modified the licensed material. You do not have permission under this license to share adapted material derived from this chapter or parts of it.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.



Development and Test Study of a New Impact Barrier



Jialin Yang and Tengfei Bi

Abstract The European side impact barrier which was developed more than 20 years ago have become difficult to meet the stiffness and structural characteristics of vehicles in China vehicle crash tests. It cannot represent the typical vehicle model. In this paper, the new side impact barrier is designed and developed independently based on the traffic accident data and typical vehicle data. The appearance dimension, mechanical performance and other specifications of the new barrier are introduced. With the typical target vehicles, the car to car side impact tests are conducted and the performance is verified. And the difference comparison indexes are established to improve and optimize the side impact barrier. So that the barrier test results are in accord with the car to car crash test results. The new impact barrier that meets the requirement is finally launched.

Keywords Side impact test · New barrier · Test study

1 Introduction

For a long time, the deformable barriers (aluminum honeycomb) used in domestic crash tests of China have been the same as those in European crash standards [1, 2]. With the increasing development of the domestic automotive market in China and the continuous rise in vehicle ownership, there are significant differences in road vehicle characteristics, forms, and types compared to Europe [3]. The current AE-MDB side impact barrier used in China vehicle crash tests is based on European vehicle models from around 2000, which is more than 20 years ago [4]. According to statistical analysis of test evaluation results, the proportion of vehicle models that received full scores is 45.10%, and the average side impact score rate is 96.94% [5].

J. Yang (✉) · T. Bi
CATARC Automotive Test Center (Tianjin) Co., Ltd., CATARC, Tianjin, China
e-mail: yangjialin@catarc.ac.cn

T. Bi
e-mail: bitengfei@catarc.ac.cn

© The Author(s) 2026

H. Saman et al. (eds.), *The 9th International Conference on Advances in Construction Machinery and Vehicle Engineering*, Lecture Notes in Mechanical Engineering,
https://doi.org/10.1007/978-981-96-5527-4_13

These results indicate that the performance of vehicles in side impact tests is continuously improving. The currently used AE-MDB side impact barrier does not clearly distinguish the side impact safety performance of vehicles and cannot adequately demonstrate the differences in side impact performance. Therefore, a new side impact barrier is urgently needed to be developed which based on actual traffic accident situations and vehicle characteristics in China. This paper aims to develop a new side impact barrier and improve the barrier through real-vehicle crash test, ultimately introducing a new side impact barrier that meets evaluation requirements.

2 Preliminary Research on the Side Impact Barrier

2.1 Traffic Accident Data Analysis

According to China In-Depth Accident Study (CIDAS) traffic accident data from 2011 to 2020, side impact accidents account for 29.5% of severe injury accidents and 17.1% of fatal accidents. It indicates that side impact accidents are still one of the significant forms causing severe injuries or deaths to occupants. Further analysis of 103 typical side impact accident cases reveals that SUV and larger passenger vehicle models account for 54% of fatal accident cases, making them the target vehicle types for side impact project research. The statistics on the speeds of offending vehicles are shown in Fig. 1. A collision speed of 60 km/h covers 86.67% of minor injury cases, 40% of severe injury cases, and 50% of fatal cases for occupants of the impacted vehicles [6]. Therefore, 60 km/h is chosen as the test development collision speed. The statistics on side impact angles are shown in Fig. 2. A 90-degree impact angle covers 50% of accident cases, so 90 degrees is selected as the test impact angle.

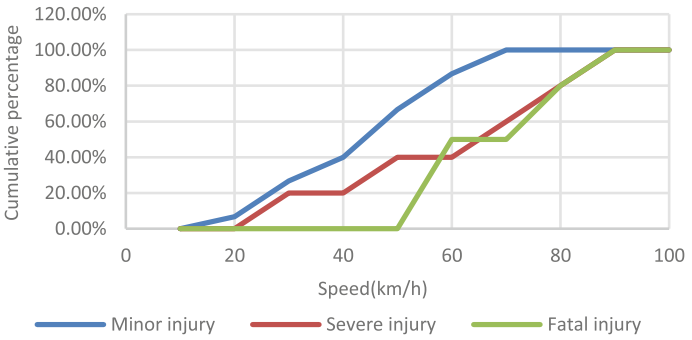


Fig. 1 Statistical analysis of offending vehicle speeds

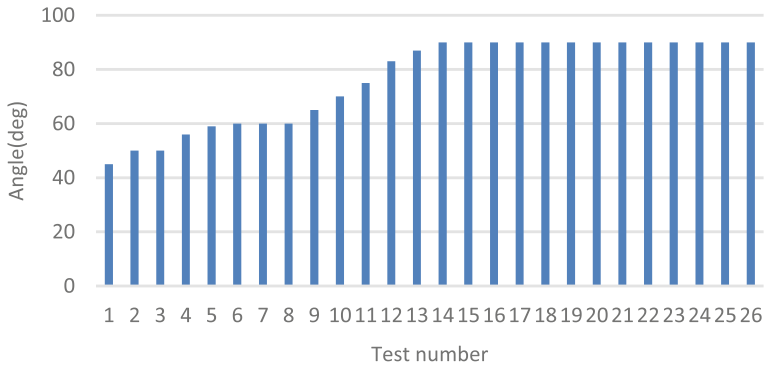


Fig. 2 Statistical analysis of side impact angles

2.2 Preliminary Development on Side Impact Barrier

Based on sales data of 206 SUV models from 2017 to present, the average curb weight based on sales is calculated to be 1622.23 kg, with an average front and rear axle load ratio of 1.3/1. Statistics indicate that the 50th percentile weight for Chinese males is 75 kg, and for Chinese females, it is 59.8 kg [7], the male-to-female driver ratio is 1.97/1. Consequently, the total weight of the side impact barrier should be 1692.11 kg, rounded to 1700 kg.

Combining traffic accident data, the project initially collected appearance size data from 46 SUV models and front-end collision stiffness data from 25 SUV models. Using sales-weighted averaging, a theoretical stiffness curve for the test barrier was derived. After multiple rounds of collision force wall tests and iterations, the first version of the aluminum honeycomb's dimensions, internal structure, and mechanical performance curve parameters were developed. The aluminum honeycomb has a width of 1800 mm, a depth of 600 mm, and a height of 500 mm, with a front bevel angle of 45 degrees. The aluminum honeycomb barrier is divided into six blocks, which is consistent with AE-MDB parameters [8]. Blocks 1/3 are 650 mm wide and 300 mm high, blocks 4/6 are 650 mm wide and 200 mm high, block 2 is 500 mm wide and 300 mm high, and block 5 is 500 mm wide and 200 mm high. Regions 1/3 and 4/6 have graded stiffness, while regions 2 and 5 have uniform stiffness. The specific details are shown in Fig. 3. Combining the aluminum honeycomb specifications, corresponding test samples were produced in collaboration with an aluminum honeycomb manufacturer, with collision stiffness curves aligning with the theoretical stiffness curves, as shown in Figs. 4, 5, 6, 7, 8.

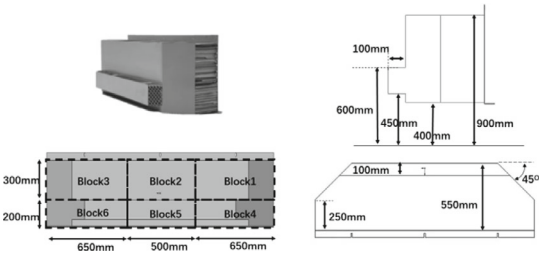


Fig. 3 MDB V1.0 dimensions

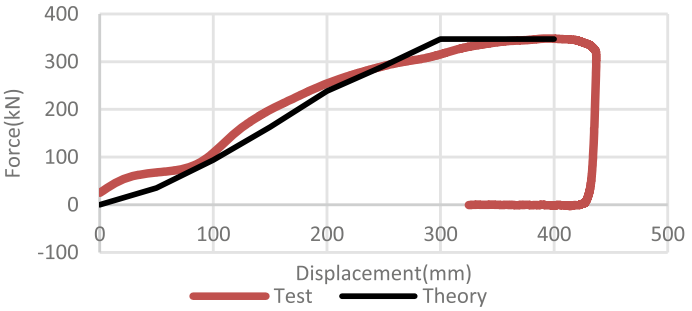


Fig. 4 Overall test curve and theoretical curve

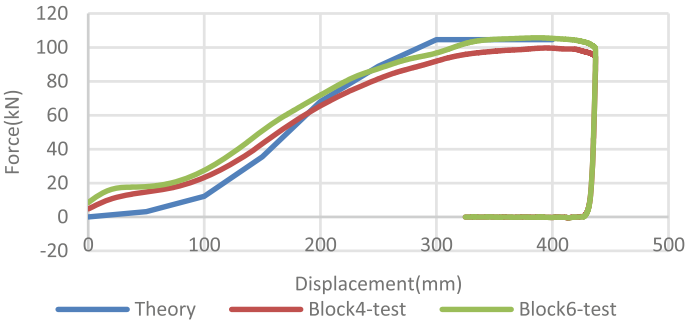


Fig. 5 Test curve and theoretical curve for block 4/6

3 Car-to-Car Side Impact Test

3.1 Car-to-Car Side Impact Test Method

Based on Sect. 2.1, statistical data from the preliminary research indicates that the collision speed for the car-to-car side impact test is 60 km/h, the impact angle is 90 degrees. The impact location is chosen to be 250 mm rearward from the R-point

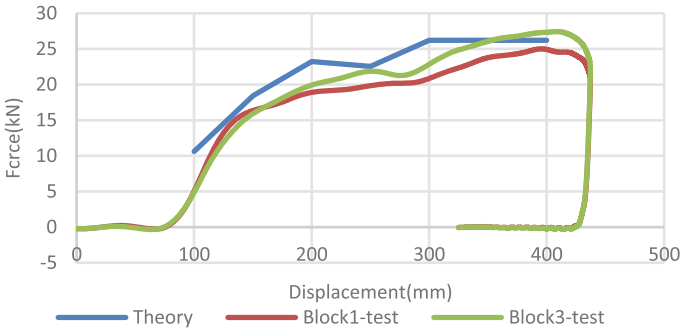


Fig. 6 Test curve and theoretical curve for block 1/3

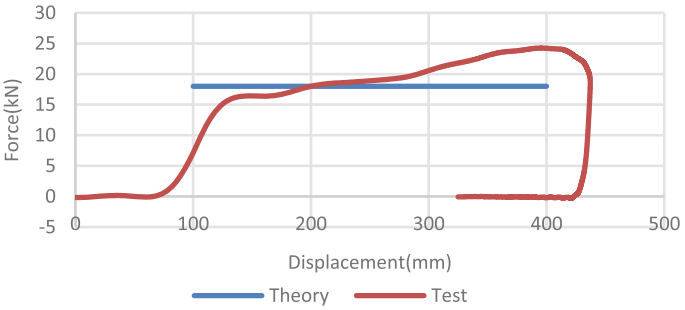


Fig. 7 Test curve and theoretical curve for block 2

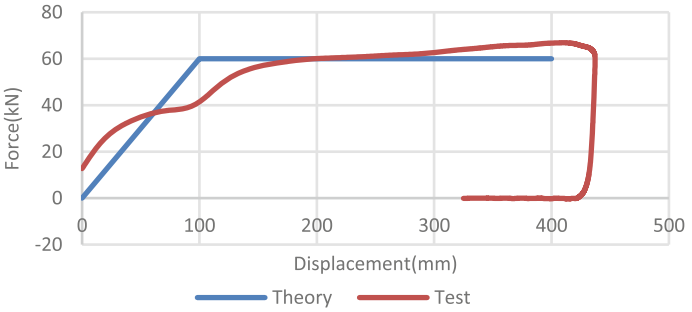


Fig. 8 Test curve and theoretical curve for block 5

of the impacted vehicle, which is used in Euro NCAP side impact protocol [9]. Statistical data from Sect. 2.2 indicates that the trolley mass is set to 1700 kg, with 960 kg on the front axle and 740 kg on the rear axle, and its surface is equipped with the new deformable barrier.

Typical SUV vehicles are selected for car-to-car impact tests as a benchmark for side impact barrier development. Vehicle model 1 represents China brand SUVs,

vehicle model 2 represents Japan brand SUVs, and vehicle model 3 represents German brand SUVs. The specific selection reasons and front-end structures are shown in Table 1. The target vehicle used in the impact test is a typical B-class sedan with a wheelbase of 2830 mm and a curb weight of 1480 kg.

The deformation of the impacted vehicle's side body are the primary data for comparing side impact tests and are the main factors causing side impact injuries to vehicle occupants [10]. Therefore, the following sections describe the arrangement of feature point positions and body speed acquisition locations on the impacted vehicle's side body.

Mark the feature points on the side of the impacted vehicle using point R as a reference. Mark vertically every 125 mm downwards to the threshold and upwards to the top of the door panel, and horizontally every 100 mm. Define six evaluation areas on the impacted vehicle model. The upper edge of areas 1, 2, and 3 is 900 mm from the ground, and the lower edge is 600 mm from the ground. The upper edge of areas 4, 5, and 6 is 600 mm from the ground, and the lower edge is 350 mm from the ground. The entire evaluation area is centered 250 mm backward from the vehicle's R-point. Area 3 is symmetrical to area 1, and area 4 is symmetrical to area 6. Areas 2 and 5 have a width of 500 mm, while areas 1, 3, 4, and 6 have a width of 650 mm. Mark the body feature points horizontally from top to bottom in each area for measuring the deformation before and after the test, as shown in rows A, B, C, D, E, F in Fig. 9. The coordinates are measured of the body feature points before and after the collision test to obtain the corresponding deformation of these feature points.

Table 1 Bullet vehicle information

Model	Vehicle structure	Representative
Model 1	Front longitudinal beam and subframe, with the subframe extending to the front	Representative of China brand SUVs
Model 2	Front longitudinal beam and subframe, with the subframe extending to the front	Representative of Japan brand SUVs
Model 3	Front longitudinal beam and subframe, with the subframe not extending to the front	Representative of German brand SUVs

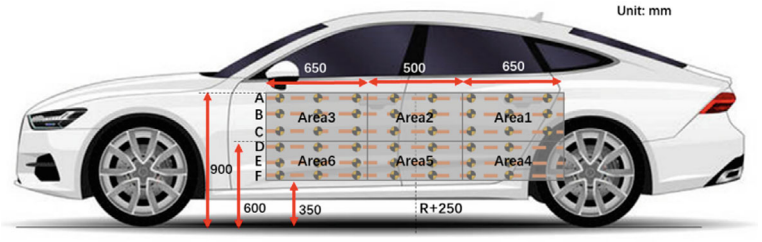


Fig. 9 Measurement Features of the Impacted Vehicle

3.2 *Results of Vehicle-to-Vehicle Side Impact Tests*

Following the vehicle-to-vehicle side impact test method described in Sect. 3.1, collision tests were conducted. The side deformation of the impacted vehicle is shown in Fig. 10. The corresponding collision deformation contours are shown in Figs. 11, 12, 13, 14, 15, 16.

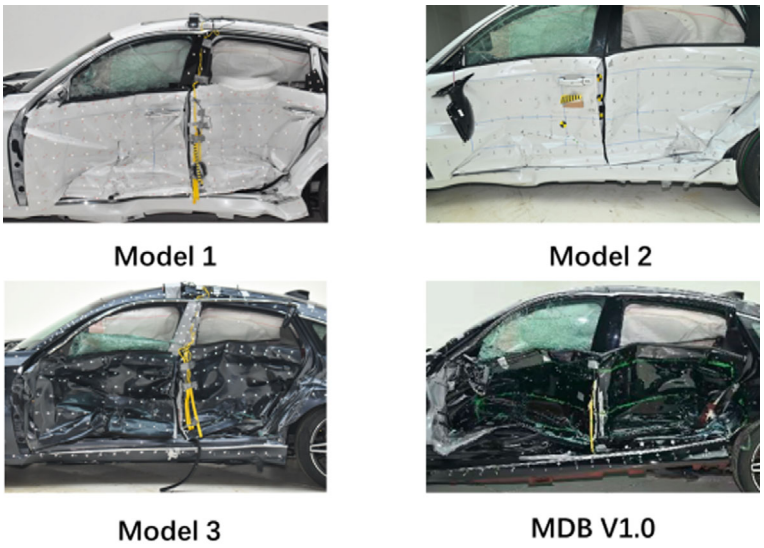
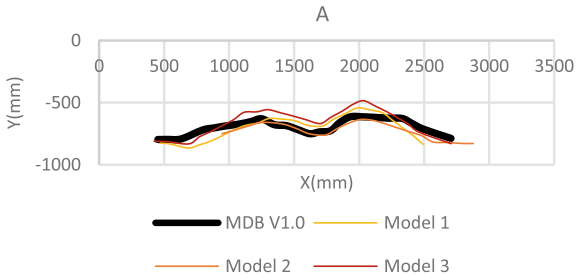


Fig. 10 Comparison of vehicle body deformation

Fig. 11 Deformation contour at row A position



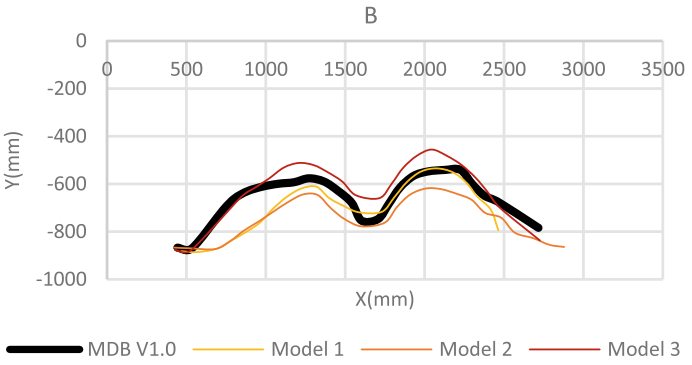


Fig. 12 Deformation contour at row B position

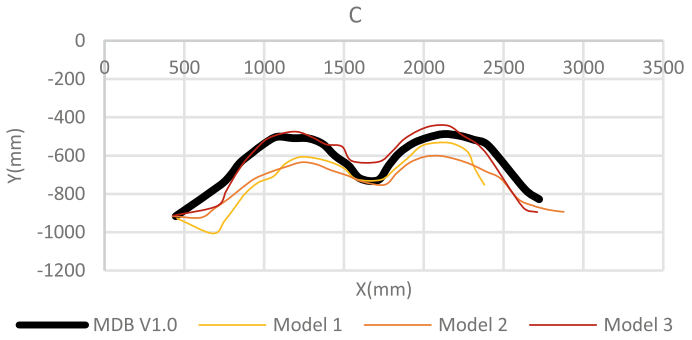


Fig. 13 Deformation contour at row C position

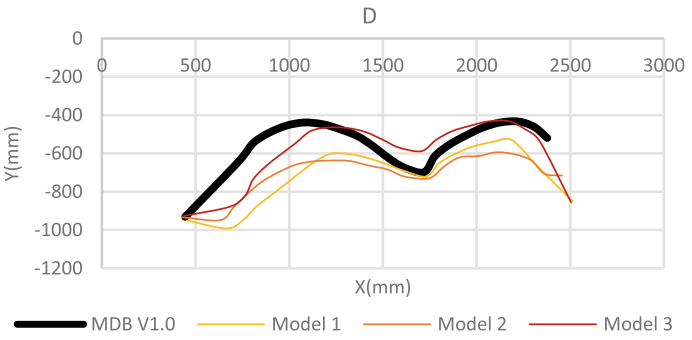


Fig. 14 Deformation contour at row D position

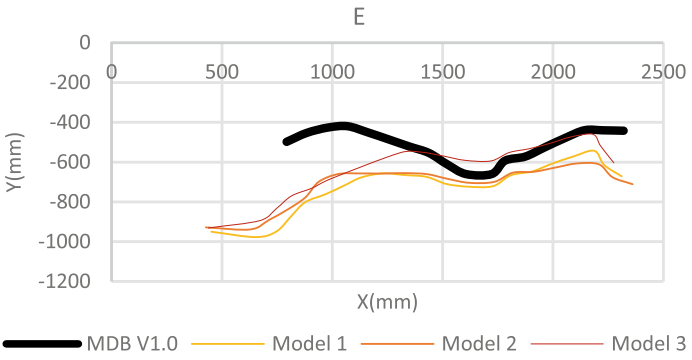


Fig. 15 Deformation contour at row E position

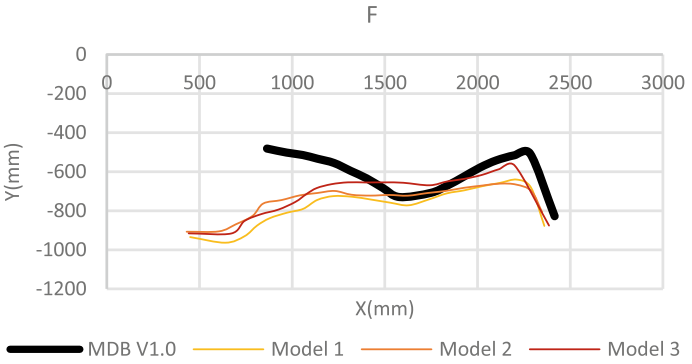


Fig. 16 Deformation contour at row F position

3.3 Calculation and Analysis of Evaluation Indexes for the Side Impact Barrier

From the comparison of the collision deformation contours, it can be observed that the body deformation trends at row A, B, and C positions are relatively similar between the side impact barrier test and the vehicle-to-vehicle collision test. However, there are significant differences in deformation at the corresponding positions of rows D, E, and F, with the side impact barrier causing greater impact deformation in the front door area. The trends in body deformation velocity are generally consistent between the side impact barrier test and the vehicle-to-vehicle collision test. However, the deformation velocities are slightly lower at positions such as the lower part of the B-pillar, the abdomen of the front door, and the pelvis of the front door. To quantitatively compare the similarity between the side impact barrier test and the actual vehicle-to-vehicle test, the following three indicators are proposed for comparative analysis. The difference between the collision tests is evaluated through these indicator values.

Body Structure Deformation Difference Index (BSI). For a single evaluation area, calculate the Body Structure Deformation Difference Index (BSI) according to the following steps:

Referring to the setup of body deformation feature points in Sect. 2.1, calculate the absolute difference (S_{ij}) between the body deformation caused by the crash barrier (S_{Tij}) and the body deformation caused by the car-to-car test (S_{Cij}) for each feature point in the area.

$$S_{ij} = \left| S_{Tij} - S_{Cij} \right| \tag{1}$$

Calculate σ_i for S_{ij} in the i-th row of this sub-evaluation area.

$$\sigma_i = \frac{\sqrt{\sum_{j=1}^{m} S_{ij}^2}}{m} \tag{2}$$

For this sub-evaluation area, compare σ_i for each row and take the maximum value as BSI_x for this sub-evaluation area.

$$BSI_x = \mathrm{Max}(\sigma_i) \tag{3}$$

Perform weighted calculations for each sub-evaluation area to obtain the Body Structure Deformation Difference Index (BSI) for the entire evaluation area.

$$BSI = \frac{\sum_{x=1}^{6} BSI_x}{6} \tag{4}$$

A smaller BSI indicator indicates that the crash barrier is closer to the reference vehicle model in overall body structure deformation.

Structure Height Index (SHI). Besides consistent deformation distribution with the reference vehicle model, the overall structure height of the crash barrier should also align with the reference vehicle model. Calculate the Structure Height Index (SHI) using body deformation and feature point heights with the following steps:

Within the evaluation area, use the corresponding feature point deformation (S_{Tij}) and height (h_i) of the crash barrier to derive the structure height (h_{Tx}) of this evaluation area using the provided formula.

$$h_{Tx} = \frac{\sum_{j=1}^{m} \sum_{i=1}^{n} S_{Tij} * h_i}{\sum_{j=1}^{m} \sum_{i=1}^{n} S_{Tij}} \tag{5}$$

Within the evaluation area, use the corresponding feature point deformation (S_{Cij}) and height (h_i) of the actual vehicle to derive the structure height (h_{Cx}) of this evaluation area using the provided formula.

$$h_{Cx} = \frac{\sum_{j=1}^{m} \sum_{i=1}^{n} S_{Cij} * h_i}{\sum_{j=1}^{m} \sum_{i=1}^{n} S_{Cij}} \tag{6}$$

Calculate the absolute difference (Δh_x) between h_{Tx} and h_{Cx}, denoted as. *SHI*.

$$SHI = \Delta h_x = |h_{Tx} - h_{Cx}| \tag{7}$$

A smaller *SHI* indicator indicates that the crash barrier is closer to the reference vehicle in overall impact height.

Analysis of Evaluation Indicators. Analyze the Body Structure Deformation Difference Index for Vehicle Models 1, 2, 3, and the side crash barrier as shown in Table 2. The comparison reveals that the side crash barrier is closest to Vehicle Model 3 in overall structural deformation, followed by Vehicle Model 1, and has the largest gap with Vehicle Model 2. For Vehicle Models 1 and 2, the side crash barrier shows significant body deformation differences in Areas 1, 3, 4, and 6, while the body deformations are closer in Areas 2 and 5. For Vehicle Model 3, the side crash barrier exhibits larger body deformation differences in Areas 1, 2, and 6, and relatively closer body deformations in other areas. Notably, the BSI6 values for all three vehicle models are relatively high, indicating significant differences between the side crash barrier and the three vehicle models in Area 6, which requires further optimization and adjustment.

Analyze the Structure Height Index for Vehicle Models 1, 2, 3, and the side crash barrier test as shown in Table 3. The comparison reveals that the side crash barrier is closest to Vehicle Model 3 in overall structure height, followed by Vehicle Model 2, and has the largest gap with Vehicle Model 1. The structure height of the side crash barrier is 610.99 mm, which is relatively higher than the actual vehicle structure height, and further adjustments are needed to reduce the overall structure height of the side crash barrier in conjunction with body deformation.

Table 2 Comparison of BSI

Vehicle model	BSI_1	BSI_2	BSI_3	BSI_4	BSI_5	BSI_6	BSI
Model 1	71.91	14.78	82.02	66.04	33.70	94.53	60.49
Model 2	73.53	24.19	73.67	85.73	26.95	83.14	61.20
Model 3	58.43	41.87	38.58	36.42	23.78	70.95	45.01

Table 3 Comparison of SHI

Bullet vehicle	Structure height (mm)	SHI
Model 1	597.27	13.72
Model 2	603.58	7.41
Model 3	607.04	3.95
MDB V1.0	610.99	–

4 Improvement of the Side Impact Barrier

Based on the calculation and analysis of the side crash barrier evaluation indicators in Sect. 3.3, it is evident that the side crash barrier differs significantly from the three vehicle models in Area 6 (lower part of the front door) in terms of body deformation contour. In terms of structure height, the side crash barrier is generally higher than the actual vehicle structure height, necessitating adjustments to reduce the overall structure height in conjunction with body deformation. In terms of body deformation velocity, the largest differences between the side crash barrier and the actual vehicle are concentrated in the lower part of the front door and the lower part of the B-pillar. Through the analysis of these three differences, it is found that the reasons for the significant differences between the side crash barrier and the car-to-car test are concentrated in the lower part of the side crash barrier. Therefore, the following analysis of the side crash barrier is conducted in conjunction with the actual vehicle structure.

In the initial design of the side impact barrier, zones 4 and 6 correspond to the longitudinal beam areas of the vehicle, which are areas with higher stiffness in the aluminum honeycomb structure. The height range is set to 400 mm ~ 600 mm, referencing the actual distribution of the vehicle's longitudinal beams. Upon further examination of the front-end structures of Vehicle Types 1, 2, and 3, it is observed that in collision tests, apart from the front longitudinal beams and energy-absorbing boxes, structures such as the subframe and radiator support at the lower end also participate in collision energy absorption. The dispersed energy-absorbing structure at the front of the vehicle allows for a larger contact area with the target vehicle during a collision, avoiding excessive localized impact loading, which could lead to severe deformation of the doors. Therefore, it is necessary to optimize the design of the side impact barrier in consideration of the lower front-end structure of the vehicle.

For the lower ground clearance H of the side impact barrier, the combined effects of vehicles with and without subframe should be considered. The following formula can be used for calculation:

$$H = \alpha \cdot J + \beta \cdot K \tag{8}$$

where J represents the lower ground clearance of the front longitudinal beam, K represents the upper ground clearance of the subframe, and α and β are scaling

factors determined by the proportion of vehicles with subframes. After statistical analysis of the front-end structures of 46 typical SUVs, the ratio of vehicles with and without subframe is approximately 1:1. The average lower ground clearance of the front longitudinal beams is 462.87 mm, and the average upper ground clearance of the subframes is 238.11 mm. Through weighted calculations, the lower ground clearance of the side impact barrier should be 350.49 mm, rounded to 350 mm. Therefore, while maintaining the overall stiffness and regional stiffness levels of the side impact barrier, the barrier is extended downward by 50 mm. To avoid puncturing the aluminum honeycomb, the thickness is increased by 50 mm. The updated side impact barrier sample and its dimensions are shown in Fig. 17.

The mechanical performance curve of the improved side impact barrier, obtained from the force wall collision test, is compared with the theoretical curve in Fig. 18. The overall stiffness curve of the aluminum honeycomb meets the requirements of the theoretical curve.

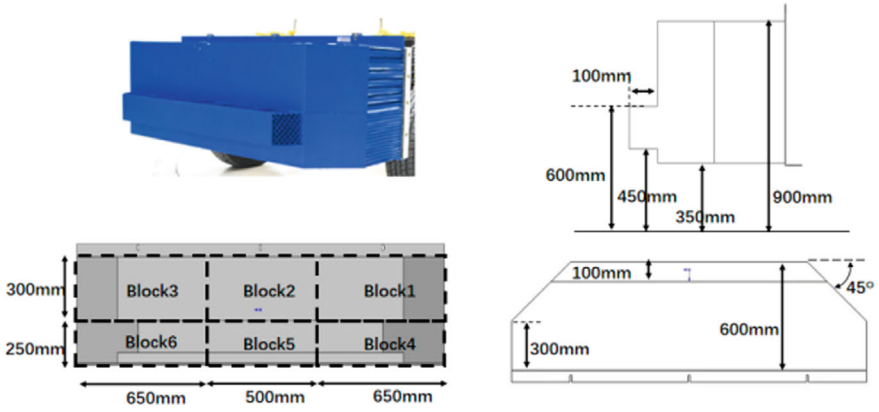


Fig. 17 MDB V2.0 sample and dimensions

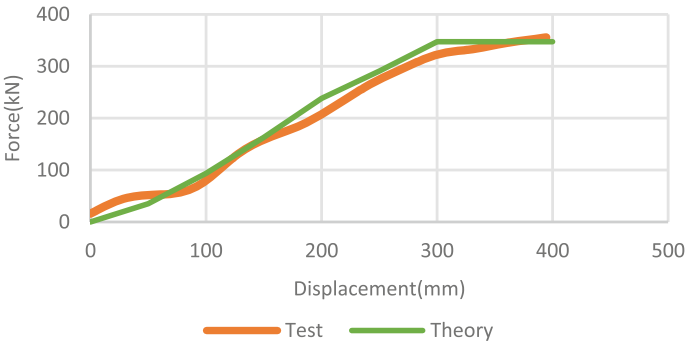


Fig. 18 MDB V2.0 test and theoretical curves

Real vehicle collision tests were conducted using the Side Impact Barrier V2.0 (MDB V2.0). A comparison of the side body deformation profiles of the impacted vehicles is shown in Fig. 19. The corresponding deformation profiles at different positions are shown in Figs. 20, 21, 22, 23, 24, 25.

Analyzing the body deformation profiles and deformation speeds of the MDB V2.0 based on the test curves, it is observed that the deformation profiles at positions A to F align well with the results of vehicle-to-vehicle collision tests, showing consistent trends. The deformation speed curves of the MDB V2.0 also closely resemble those of vehicle-to-vehicle tests in terms of trend and peak values. The MDB V2.0 is extended downward by 50 mm, increasing the contact area with the impacted vehicle body

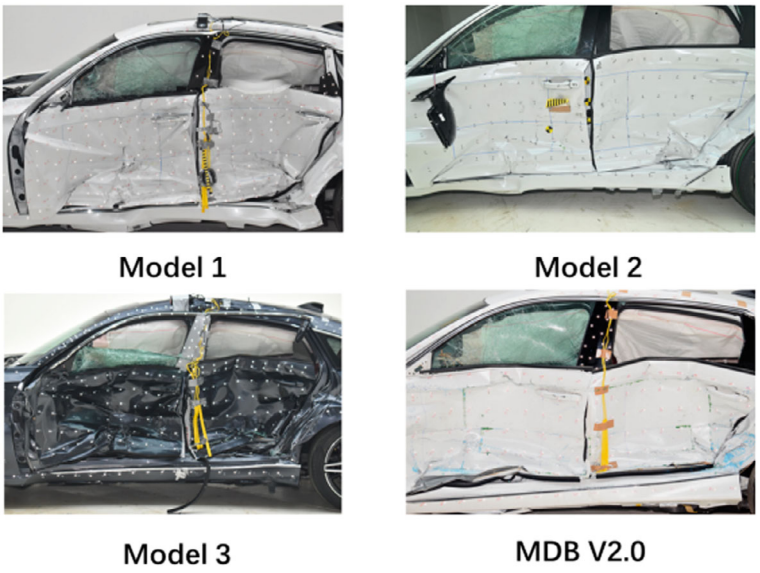


Fig. 19 Comparison of vehicle body deformation

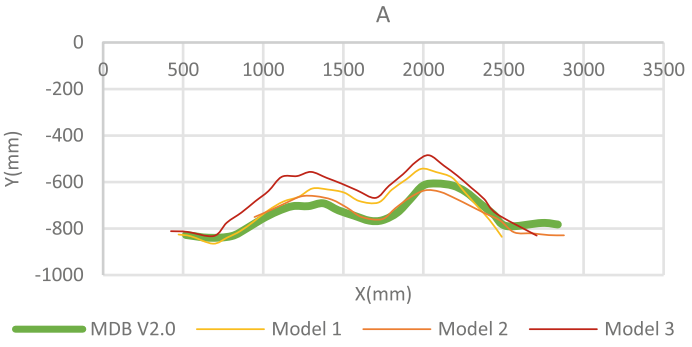


Fig. 20 Deformation contour at row A pPosition

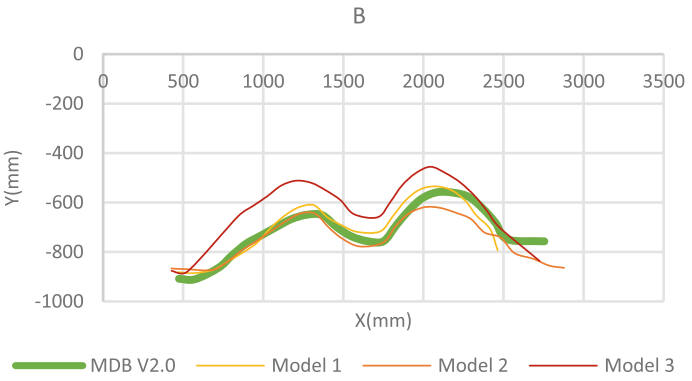


Fig. 21 Deformation contour at row B position

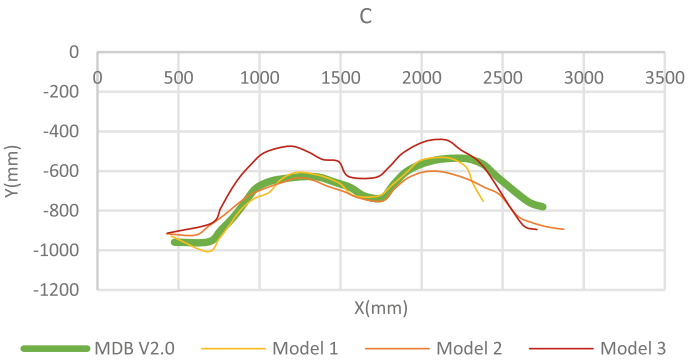


Fig. 22 Deformation contour at row C position

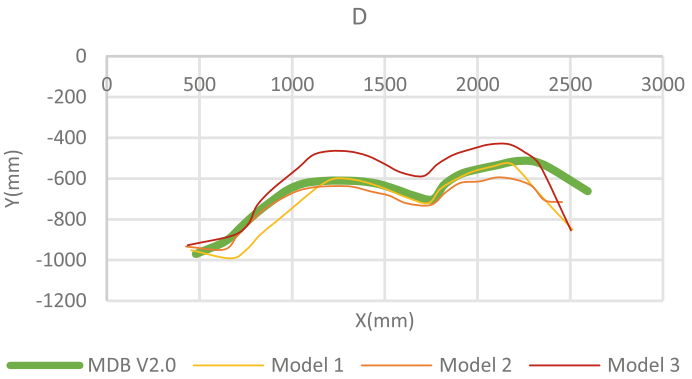


Fig. 23 Deformation contour at row D position

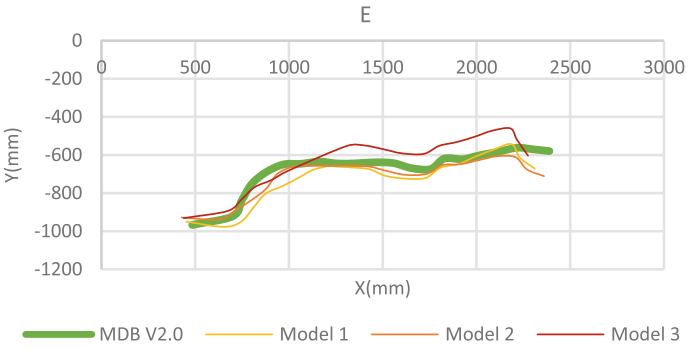


Fig. 24 Deformation contour at row E position

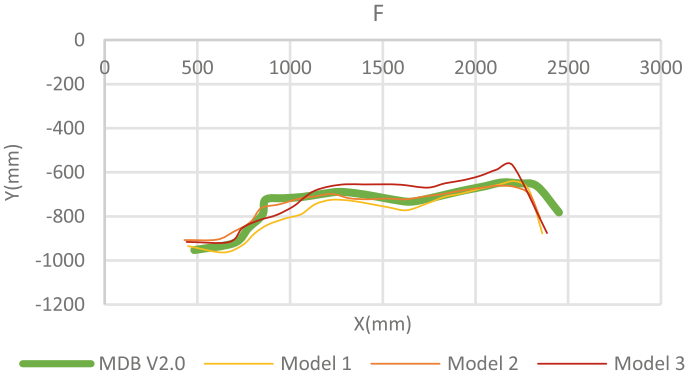


Fig. 25 Deformation contour at row F position

during a collision. Due to the maintained overall and regional stiffness, the collision intensity per unit area of the aluminum honeycomb in zones 4, 5, and 6 is reduced, mitigating severe deformation at the lower part of the front door and making the body deformation pattern more consistent with vehicle-to-vehicle collision test results.

Further analysis of body structure deformation differences between Model 1, 2, 3, and the MDB V2.0, as shown in Table 4, reveals that compared to V1.0, the MDB V2.0 exhibits a decreasing trend in body structure deformation difference index values for all three vehicle types, with reduction percentages of 58.83% (Model 1), 63.07% (Model 2), and 12.46% (Model 3). This indicates that the MDB V2.0 more closely resembles the results of vehicle-to-vehicle collision tests in terms of body deformation patterns compared to V1.0.

Analyzing the structural height index of Model 1, 2, 3, and as shown in Table 5, it is found that the overall structural height of the MDB V2.0 is 602.25 mm, which is 8.74 mm lower than that of the MDB V1.0. The structural height index (SHI) decreases by 8.74 for Model 1 and 6.08 for Model 2, while there is a slight increase

Table 4 Comparison of BSI

Model	BSI_1	BSI_2	BSI_3	BSI_4	BSI_5	BSI_6	BSI
Model 1	38.3	17.9	23.9	20.8	22.4	26.1	24.9
Model 2	40.4	16.1	13.1	35.8	14.7	15.3	22.6
Model 3	43.0	49.8	50.3	37.6	34.8	20.7	39.4

Table 5 Comparison of SHI

Vehicle model	Structure height (mm)	SHI (V2.0)	SHI (V1.0)	Difference
Model 1	597.27	4.98	13.72	8.74
Model 2	603.58	1.33	7.41	6.08
Model 3	607.04	4.79	3.95	− 0.84

of 0.84 in the SHI for Model 3. This indicates that the MDB V2.0 is structurally closer to actual vehicles in terms of height, meeting the expected design requirements.

Based on the analysis results of the above two evaluation indicators, the MDB V2.0 is closer to the car-to-car test results in terms of body deformation contour and structural height. It can represent the collision characteristics of the benchmark vehicle models and meets the design expectations.

5 Conclusion

Combined with traffic accident data and typical vehicle dimensions and collision stiffness data, a new side impact barrier has been autonomously designed and developed. By selecting typical benchmark vehicle models and conducting real-vehicle collision tests for development and performance verification, improvements and optimizations have been made to the side impact barrier, making it consistent with car-to-car side collision test results and capable of truly reflecting the collision characteristics of the benchmark vehicle models. The specific conclusions are as follows:

(1) Based on CIDAS traffic accident data, vehicle models requiring research in side collision scenarios were identified. Data on the exterior dimensions and collision stiffness of multiple SUV models were collected, and the theoretical stiffness curve of the test barrier was derived using sales volume-weighted averaging. After multiple rounds of iteration, parameters such as the appearance dimensions, internal structure, and mechanical performance curve of the test barrier were determined, and the corresponding test samples were developed.
(2) Based on the early-stage research of the project, a car-to-car side collision test method was established. Three typical SUV models were selected for car-to-car collision tests as a benchmark for the development of the side impact barrier. Three evaluation indicators for the test barrier, namely Body Structure

Deformation Index (BSI) and Structural Height Index (SHI) were developed to visually reflect the differences between the side impact barrier and actual vehicles in numerical form.

(3) By analyzing the indicator data to identify the reasons for the differences in the side impact barrier, combined with the front-end structure of the vehicle, the lower part of the side impact barrier was redesigned. The structural heights of vehicle subframe and other components were statistically analyzed, and the side impact barrier was extended downward by 50 mm. Real-vehicle collision tests were conducted for verification. The results showed that the improved side impact barrier is closer to the car-to-car test results in terms of body deformation contour and structural height. It can reflect the collision characteristics of the benchmark vehicle models, meeting the design expectations.

References

1. Huang Z, Zhang B, Wang L (2020) Experimental study on side pole collision and deformable barrier collision of car. J Phys: Conf Series Chongqing, China 1549:1–12
2. Wang LJ, Fan ZJ, Gui LJ (2008) "The design of side impact mobile deformable barrier numerical model based on GB 20071—2006", Automobile Technology. Changchun, China 7:18–21
3. Huang Z, Huang M (2014) Reflection on the development of China's car industry (part 1)—starting from the comparison between China and Europe. Bus Research China 7:58–61
4. Ellway JD (2005) The development of an advanced European mobile deformable barrier face (AE-MDB). In: 19th International technical conference on the enhanced safety of vehicles (ESV), pp1–13. Washington, DC United States
5. Li X (2024) China Automotive Test and Assessment Report. China Automotive Technology and Research Center Co., Ltd., Social Science Academic Press
6. Zhu H, Yang J, Duan B (2023) "Performance development and verification of honeycomb aluminum barrier for side impact", automobile technology. Changchun, China 2:49–55
7. Xiangshan (2022) Weighing Instrument Group "2022 China Residents' Height and Weight Health Data Report," Zhongshan, China
8. Ellway JD, Ratingen MV, Versmissen ACM (2013) The advanced European mobile deformable barrier specification for use in EuroNCAP side impact testing. In: 23rd International technical conference on the enhanced safety of vehicles (ESV), pp.1–13. Seoul, South Korea
9. Ma WJ, Cui XL (2018) The study of different dummy injury performance based on AE-MDB tests. In: Machinery design & manufacture, vol 3, pp 30–36. Shenyang, China,
10. Roussarie ML, Zeitouni R, Adalian C (2007) Enhancement of side impact protection using an improved test procedure. In: 20th International technical conference on the enhanced safety of vehicles (ESV), pp. 1–12. Lyon, France

Open Access This chapter is licensed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License (http://creativecommons.org/licenses/by-nc-nd/4.0/), which permits any noncommercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if you modified the licensed material. You do not have permission under this license to share adapted material derived from this chapter or parts of it.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.



Simulation Research on the Three-Way Catalytic Converter of a Certain Type of Automobile Exhaust System



Kun Fang, Wei Ma, Wei Chang, and Pengjie Fan

Abstract With the growth of the automobile consumption, the air pollution caused by automobiles has become a hot issue all over the world. Three-way catalytic converters are widely applied in the exhaust system to solve these problems, However, the pressure loss caused by these equipments should be discussed. In this paper, a common exhaust system is chosen as the object of study, and the internal with DOC and POC three-way catalytic converter structure is optimized by CFD numerical simulation, it is obtained that: (1)When the inlet expansion angle is 25°, the pressure loss caused by the combined three-way catalyst is relatively small; (2)When the inlet expansion angle is 25°, the exhaust back pressure of the entire exhaust system is relatively small, which has little impact on the engine; (3)The pressure loss caused by the combination type three-way catalyst directly affects the exhaust back pressure of the entire exhaust system, and the trend of the two changes is the same. The research work provides a theoretical basis and important reference for the structural optimization design of the three-way catalytic converter of the exhaust system.

Keywords Three-way catalytic converter · Pressure loss · Numerical simulation · Structure optimization

1 Introduction

With the rapid development of the automotive industry, the production and sales of cars are growing rapidly every year, nowadays, cars are indispensable in daily life, while providing great convenience for human life, the air pollution caused by cars is becoming increasingly serious[1]. The 2015 Annual Report on the Prevention and Control of Motor Vehicle Pollution in China, released in early 2015, pointed out that, one of the important causes of air pollution is the exhaust emissions from motor vehicles. The three-way catalyst is currently the most commonly used exhaust

K. Fang (✉) · W. Ma · W. Chang · P. Fan
No.723 Research Institute of CSSC, Yangzhou, Jiangsu, China
e-mail: faun1991@163.com

© The Author(s) 2026

H. Saman et al. (eds.), *The 9th International Conference on Advances in Construction Machinery and Vehicle Engineering*, Lecture Notes in Mechanical Engineering,
https://doi.org/10.1007/978-981-96-5527-4_14

purification device in automotive exhaust systems, it can effectively convert CO, HC, NOx, and PM in exhaust gas into non-toxic and harmless N_2, H_2O, and CO_2 through chemical reactions.

However, installing a three-way catalyst in the exhaust system of a car will inevitably cause performance changes in other aspects of the car, especially increasing the resistance of exhaust emissions. When the exhaust of a car encounters resistance, it is compressed to form a pressure wave that is transmitted back to the engine, invisibly increasing the exhaust back pressure of the engine [2], as a result, the amount of air introduced by the engine is reduced, and the fuel cannot be completely burned in the engine, ultimately reducing the purification efficiency of the entire catalytic device and accompanied by the occurrence of regeneration noise. Therefore, more and more enterprises, experts, and scholars have conducted in-depth research on how to reduce the flow resistance caused by catalytic devices installed in automobile exhaust systems [3, 4].

At present, a large amount of research is based on experiments, and for enterprises, the primary goal is to save costs., therefore, more and more enterprises are also using software simulation to research automobile exhaust systems. This article uses CFD numerical simulation method to optimize the structure of the automobile exhaust system equipped with DOC + POC three-way catalyst, reducing the pressure loss caused by the three-way catalyst and thus reducing the exhaust back pressure of the exhaust system.

2 Numerical Simulation Method

CFD numerical simulation technology is a technique that uses computers to solve various conservation controlled partial differential equations of fluid flow, which involves fluid mechanics, computational methods, and even computer graphics processing techniques [5].

2.1 Control Equation

One of the control equations used in numerical simulation calculations is the Navier Stokes equation in fluid mechanics.

When modeling the exhaust system, the basic equations used include the N-S equation, continuity equation, and energy conservation equation [6], as follows:

Momentum conservation equation:

$$\rho \frac{dV_i}{dt} = \rho \left[\frac{\partial V_i}{\partial t} + V_j \frac{\partial V_j}{\partial x_j} \right] \tag{1}$$

In the formula, V_i and V_j are the velocity components of each coordinate axis; ρ is the internal airflow density.

Energy conservation equation:

$$\frac{\partial}{\partial t}(\rho h) + \frac{\partial}{\partial x_j}(\rho \mu_j h) = \frac{\partial P}{\partial t} + \mu_j \frac{\partial P}{\partial x_j} + \tau_{ij} \frac{\partial V_i}{\partial x_j} \tag{2}$$

$$\tau_{ij} = 2\mu(s_{ij} - \frac{1}{3}\frac{\partial V_k}{\partial x_k}\delta_{ij}) - \overline{\rho V_i' V_j'} \tag{3}$$

$$s_{ij} = \frac{1}{2}(\frac{\partial V_i}{\partial x_j} + \frac{\partial V_j}{\partial x_i}) \tag{4}$$

In the formula, τ_{ij} is the stress tensor; s_{ij} is the fluid deformation tensor; h is the enthalpy value; μ is the viscosity coefficient; δ_{ij} is the Kroneker number; $\overline{\rho V_i' V_j'}$ is the reynolds stress tensor;P is the pressure.

Conservation of mass equation:

$$\frac{\partial \rho}{\partial t} + \rho \frac{\partial V_i}{\partial x_i} + V_i \frac{\partial \rho}{\partial x_i} = 0 \tag{5}$$

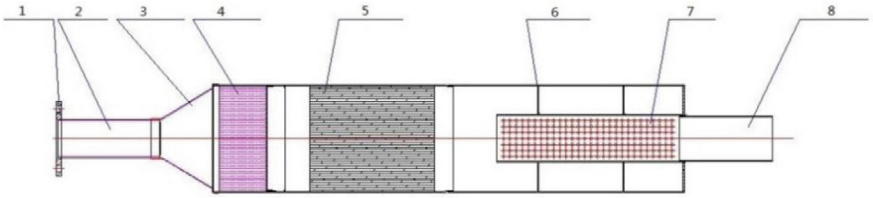
$$\begin{aligned} \varphi = &\left(\frac{\partial V_2}{\partial x_3} - \frac{\partial V_3}{\partial x_2}\right)^2 + \left(\frac{\partial V_3}{\partial x_1} - \frac{\partial V_1}{\partial x_3}\right) + \left(\frac{\partial V_1}{\partial x_2} - \frac{\partial V_2}{\partial x_1}\right) \\ &+ \frac{2}{3}\left[\left(\frac{\partial V_1}{\partial x_1} - \frac{\partial V_2}{\partial x_2}\right)^2 + \left(\frac{\partial V_2}{\partial x_2} - \frac{\partial V_3}{\partial x_3}\right)^2 + \left(\frac{\partial V_3}{\partial x_3} - \frac{\partial V_1}{\partial x_1}\right)^2\right] \end{aligned} \tag{6}$$

In the formula, φ is the dissipation function; ρ is the air flow density.

2.2 Model and Grid Division

The exhaust system of a typical car consists of an intake pipe, an expansion pipe, a three-way catalytic converter, a silencing device, and an outlet tailpipe. This article is based on a commonly used automotive exhaust system with a novel three-way catalytic converter structure, as shown in Fig. 1. The internal three-way catalyst is composed of a oxidation catalyst (DOC) and a particle oxidation catalyst (POC), which have different purification functions. The specific structure is shown in Fig. 2.

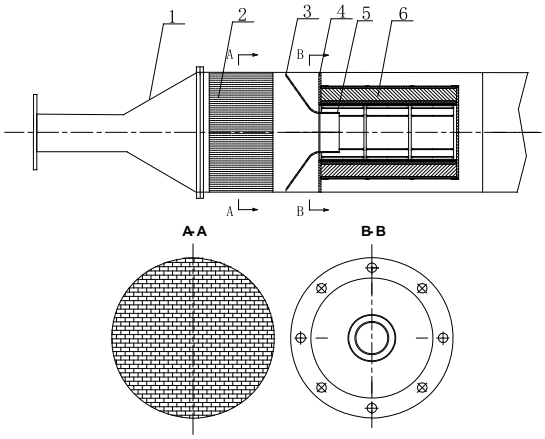
The simulation of the flow field is aimed at calculating the air domain inside the exhaust system, in order to facilitate the calculation and analysis, and improve the calculation efficiency, UG is used to simplify the modeling of the calculation model. Using ICEM software to establish a computational grid, the model structure is relatively complex with many noise reduction holes, so a non structured grid is selected, the majority of the mesh is tetrahedral, and the denser the unstructured mesh, the more accurate the calculation, but the denser the mesh, the lower the



1.Flange; 2.Intake pipe; 3.Expansion tube; 4.DOC;
5.POC; 6.Baffle; 7.Perforated silencing tube; 8.Export tail pipe

Fig. 1 Overall structure of automobile exhaust system

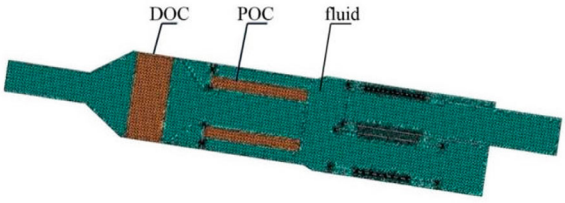
Fig. 2 Combination type three-way catalytic converter structure



1.Expansion tube; 2.DOC; 3.Guide plate;
4.Flow blocking plate; 5.Diversion straight pipe; 6.POC

computational efficiency. After comprehensive consideration, the calculation grid for the exhaust system is obtained as shown in Fig. 3.

Fig. 3 Mesh of cross-section in exhaust system model



2.3 Boundary Conditions

The fluid in the exhaust system is the high-temperature exhaust gas emitted by the engine, this article only studies the pressure loss caused by its internal structure, without considering its pollutant concentration and heat exchange, so the internal flowing gas is considered as air at room temperature and pressure. Assuming that the flow process inside the exhaust system is a steady-state process and the internal airflow is incompressible, the k—ε turbulence model is adopted. The inlet conditions are set as a mass-flow-inlet, with an exhaust mass flow rate of 590 kg/h, a room temperature of 300 K, and the outlet as the pressure-outlet, the default is atmospheric pressure, and the hydraulic equivalent diameter is selected based on the inlet and outlet diameters;the entire wall of the exhaust system is a rigid non slip adiabatic wall, and the carrier section is made of porous media.

3 Analysis of Numerical Calculation Results

3.1 Theoretical Analysis of Pressure Loss

There are many components of pressure loss in the combination type three-way catalyst. Mainly including: resistance of DOC and POC carriers, pressure loss along the pipeline, local resistance loss of expansion tube, and local resistance loss of guide plate, etc., from this, it can be seen that the overall pressure drop of the combined three-way catalyst is related to Reynolds number, pipe wall roughness, gas density, and average gas flow velocity [7]. In terms of structure, this article mainly simulates and studies the influence of intake expansion angle on the pressure loss of the entire combined three-way catalytic converter.

3.2 Calculation Results of Pressure Loss Values

The structure of the exhaust system combination type three-way catalyst is a guide plate angle of 45°,the length of the diversion straight pipe is 25 mm, and the expansion angles are 25°, 35°, 45°, and 55°.The pressure changes on the cross-section of the exhaust system are shown in Fig. 4, 5, 6, 7.

As can be seen from Fig. 4, 5, 6, 7,when the inlet expansion angle is small, the pressure field distribution at the front end of the DOC carrier is relatively uniform, and there is a significant layered pressure drop along the axis, which decreases uniformly. As the inlet expansion angle increases, a small area with high pressure appears at the front end of the DOC carrier, and the pressure field distribution becomes more uneven. Due to the larger expansion angle, the distance between the tail of the intake pipe and the DOC carrier becomes shorter, and the high-speed airflow quickly reaches

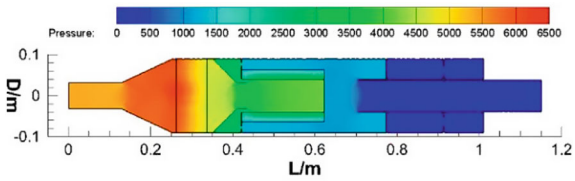


Fig. 4 Pressure distribution in the middle section when the expansion angle is 25°

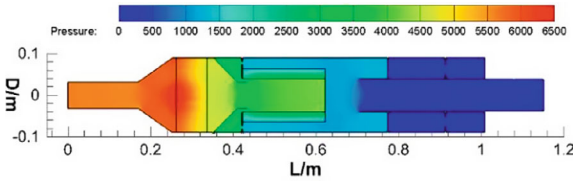


Fig. 5 Pressure distribution in the middle section when the expansion angle is 35°

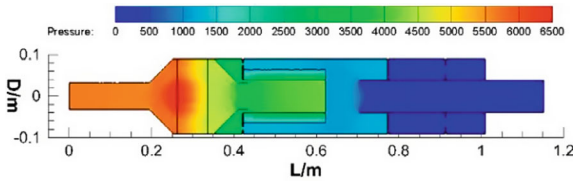


Fig. 6 Pressure distribution in the middle section when the expansion angle is 45°

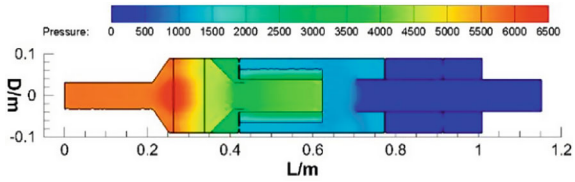


Fig. 7 Pressure distribution in the middle section when the expansion angle is 55°

the front end of the carrier, however, due to the small size of the DOC carrier channel and the high resistance of the carrier channel, high-speed airflow cannot pass through quickly and gathers at the front end of the DOC carrier, resulting in a decrease in airflow velocity, the pressure in this area has increased to a maximum value, causing significant pressure loss.

From Table 1, it can be seen that the pressure loss of the combined three-way catalyst increases with the increase of the expansion angle; when other structures in the exhaust system remain constant, the exhaust back pressure of the exhaust system increases with the increase of the inlet expansion angle.

The resistance along each small channel in the DOC carrier constitutes the pressure loss of the DOC carrier, the larger the expansion angle, the higher the average flow velocity inside the carrier, and the greater the resistance along each channel, the greater the overall pressure loss of the DOC carrier, the greater the overall pressure loss of the combined three-way catalyst.

Table 1 Pressure loss of combined three-way catalysts with different expansion angles

Expansion angle(°)	25	35	45	55
Pressure loss of combined three-way catalyst (Pa)	4094.5	4103.9	4111	4114
Exhaust back pressure (Pa)	5380.3	5633.7	5655.7	5736.4

From Table 1, it can be seen that: (1) When the inlet expansion angle is 25°, the pressure loss caused by the combined three-way catalyst is relatively small; (2) When the inlet expansion angle is 25°, the exhaust back pressure of the entire exhaust system is relatively small, which has little impact on the engine; (3) The pressure loss caused by the combination type three-way catalyst directly affects the exhaust back pressure of the entire exhaust system, and the trend of the two changes is the same.

4 Conclusion

The working conditions of automobile exhaust systems are complex, requiring both good purification effects and excellent aerodynamic performance. This article uses Icem software combined with Fluent software to simulate and analyze the internal flow field of a combined three-way catalytic converter in a certain automobile exhaust system, research on the impact of different local structures on the pressure loss of the entire exhaust system, and propose subsequent structural optimization directions based on simulation results, to achieve the goal of reducing exhaust back pressure in automobile exhaust systems. The Fluent based automotive equipment flow field simulation technology can effectively simulate the flow of air in the exhaust system, providing effective simulation data for its structural optimization design, greatly shorten the research and development cycle of automotive exhaust systems and reduce research and development costs.

References

1. Lv L, Hu J, He L et al (2021) Research progress on the impact of gasoline vehicle technology development on exhaust emissions. Environ Sci Res 34(02):286–293
2. Wu X (2024) Research and analysis on emission optimization based on the health of three-way catalytic converters. China Machinery (18):62–66
3. Chiranth S, Xiaofeng Y, Jeff S et al (2020) Conjugate heat transfer CFD analysis of an oil cooled automotive electrical motor. SAE Int J Adv Curr Pract Mobil 2(4):1741–1753
4. Zou C, Xiong B, Xue H et al (2021) The role of new energy in carbon neutral. Pet Explor Dev 48(02):480–491
5. Liu C (2023) Design and analysis of exhaust system for a certain vehicle model based on CFD method .Nanchang University
6. Long K (2011) Fluid mechanics. Higher Education Press

7. Zhang J, Shen X, Zuo Q et al (2020) Structural optimization and CFD simulation of engine catalytic converters. China Agric Mach Chem J 41(05):87–93

Open Access This chapter is licensed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License (http://creativecommons.org/licenses/by-nc-nd/4.0/), which permits any noncommercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if you modified the licensed material. You do not have permission under this license to share adapted material derived from this chapter or parts of it.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.



Comparative Study on Single-Objective and Multi-objective Evolutionary Algorithms in the Parameter Optimization of Excavation Mechanism for Hydraulic Shovel

Gongyue Xu, Yuheng Zhou, Zhiwen Zhang, and Wei Hong

Abstract There are two most common methods to deal with multi-objective optimization problems, transforming into single-objective optimization problems by aggregation function, or directly solving them using multi-objective evolutionary algorithm. For a long time, there is a lack of quantitative comparative studies for these two approaches in the optimal design of hydraulic excavator. To address this issue, the multiobjective optimization model of hydraulic excavator with ordinary shovel excavation mechanism was firstly established, which set level crowding, digging forces, crowding force and the indicator of digging atlas as objective functions. Then, twelve representative evolutionary algorithms which contained six single-objective evolutionary algorithms and six multi-objective evolutionary algorithms were applied in the optimal case of 70-ton hydraulic excavator. The performance metrics of single-objective optimization and multi-objective optimization were all considered for the comparison. The result shows that multi-objective evolutionary algorithms are significantly superior to single-objective evolutionary algorithms and proved to be the more suitable choice for multi-objective optimization problems, which can provide more quality and balanced solutions for designer.

Keywords Multi-objective optimization · Single-objective optimization · Evolutionary algorithm · Excavation mechanism · Hydraulic shovel · Comparative study

G. Xu (✉) · Y. Zhou · Z. Zhang · W. Hong
School of Mechanical Engineering, Jiangsu University, Zhenjiang, China
e-mail: xgy_ztt@163.com

© The Author(s) 2026

H. Saman et al. (eds.), *The 9th International Conference on Advances in Construction Machinery and Vehicle Engineering*, Lecture Notes in Mechanical Engineering,
https://doi.org/10.1007/978-981-96-5527-4_15

1 Introduction

Hydraulic excavator is one of the most widely used engineering machinery equipments which serves a broad field of national economy. As the core component of the hydraulic excavator, excavation mechanism often consists of boom, stick, bucket, and hydraulic cylinders. Therefore, the parameter design of excavation mechanism plays a critical role for the performance of the whole machine. In the design process of excavation mechanism, it is necessary to comprehensively consider the technical indicators and performance constraints of hydraulic excavator, such as the working range, the kinematic characteristics, the digging forces, and the digging performance. And this leads to an extremely complex multi-objective optimization problem (MOP) with multiple constraints. For the MOPs in practical engineering, multi-modal, discontinuous, non-differentiable, or even implicit expression of the objective function are very common, and traditional optimization approaches are hardly to solve. On the contrary, due to the flexible and robust framework, advanced artificial intelligence methods such as evolutionary algorithms (EAs) have become a hot research field in multi-objective optimization [1]. In single-objective optimization problems (SOPs), there is an absolute global optimal solution. But for MOPs, multiple optimization objectives are often conflict with each other, which makes it difficult to achieve an optimal solution simultaneously. Thus, the concept of Pareto optimality is introduced for comparison and trade-off, and the goal of multi-objective optimization is changed to obtain the optimal solution set in the objective space, i.e., to approach and cover the Pareto-optimal front (PF) of MOP as more as possible [2]. For the parameter optimization of hydraulic excavator, although the application studies based on MOEAs has achieved good progress in recent years, many researchers are still used to convert the MOPs into SOPs through certain prior knowledge and then solve them by single-objective EAs (SOEAs) [3, 4]. From the view of algorithm mechanism, the transformed SOP by weighted sum method can only find the intersection point of the weight vector and the PF of original MOP, while only MOEAs can obtain the complete PF which could provide the decision-makers more balanced and high-quality solutions. The superiority of MOEAs has been confirmed by researches on the convex textured friction couple [5], the traffic engineering [6], and the stormwater drainage pumping system [7], but there has been no such a comparative study in the optimization design of hydraulic excavator. Based on the above analysis, this paper takes the most popular excavation mechanism of hydraulic shovel as the research object to verify the superiority of MOEAs through comparative analysis.



Fig. 1 Hydraulic excavator with ordinary shovel excavation mechanism

2 Multi-objective Optimization Model

2.1 Ordinary Shovel Excavation Mechanism

Shovel excavation mechanism is utilized for the mining of ore material on the shut-down surface and becomes one of the most commonly used working attachment. With the development of technology, many different configurations of shovel excavation mechanism are invented, including original shovel excavation mechanism, ordinary shovel excavation mechanism, tripower shovel excavation mechanism, and so on.

As shown in Fig. 1, unlike the original shovel excavation mechanism whose bucket cylinders are hinged on the middle of stick, the bucket cylinders of ordinary shovel excavation mechanism are hinged at the end of boom. In this case, level crowding is easy to be achieved without adjusting the bucket cylinders by the coupled connection between the bucket cylinders and the boom, while the dimensional parameters are appropriate. Additionally, ordinary shovel excavation mechanism can also increase the crowding force. Based on the above advantages, ordinary shovel excavation mechanism has long been the most popular working attachment in the market.

2.2 Decision Variables

As shown in Fig. 2, the mathematical model of ordinary shovel excavation mechanism is easy to established. Considering the convenience of calculation and modeling, the position coordinates of each hinge point and the dimensions of each component are taken as decision variables. Finally, the optimization model selects the following 23 decision variables

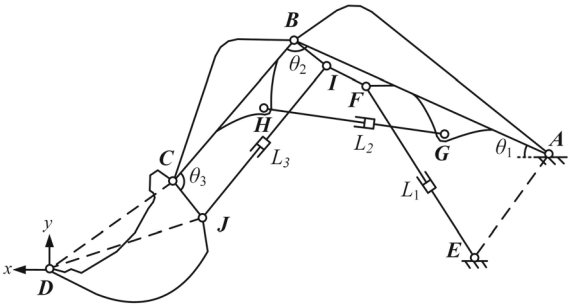


Fig. 2 The schematic diagram of ordinary shovel excavation mechanism

$$\mathbf{x} = \left\{ \begin{array}{l} x_A, y_A, x_E, y_E, L_{AB}, L_{AG}, L_{BG}, L_{AF}, L_{BF}, L_{AI}, L_{BI}, L_{BC} \\ L_{BH}, L_{CH}, L_{CJ}, L_{CD}, L_{DJ}, L_{1\min}, L_{2\min}, L_{3\min}, \lambda_1, \lambda_2, \lambda_3 \end{array} \right\} \tag{1}$$

where, $x_{(.)}$ and $y_{(.)}$ respectively represent the horizontal and vertical coordinates; $L_{(..)}$ represents the distance between the corresponding points; $L_{1\min}$, $L_{2\min}$, and $L_{3\min}$ represent the shortest lengths of cylinders; λ_1, λ_2, and λ_3 represent the extension ratio of cylinders.

2.3 Objective Function

According to the characteristics of ordinary shovel excavation mechanism and the technical parameters of hydraulic shovel, the optimization model sets the following five objective functions.

Level Crowding. Level crowding is an important operation for leveling the ground or mining in layers. During this process, the boom cylinder is floating to keep the bucket close to the ground, while the cutting angle of the bucket is required to remain unchanged. As shown in Fig. 3, if the bucket cylinder length L_3 is always adjusted to keep the crowding angle γ unchanged, the variation ΔL_3 of the bucket cylinder can be used to evaluate the performance. And the objective function for level crowding is

$$\mathbb{F}_1(\mathbf{x}) = C_1 |\Delta L_3| \tag{2}$$

where, C_1 is a constant to balance the magnitude of each objective function.

Digging Forces and Crowding Force. Digging force is the most important performance index of hydraulic excavator, which represents the maximum digging ability of hydraulic excavator. In addition, the crowding force also needs to be considered which appears as an important technical parameter in the product specifications of various hydraulic shovels. And the objective functions of digging forces and crowding force is

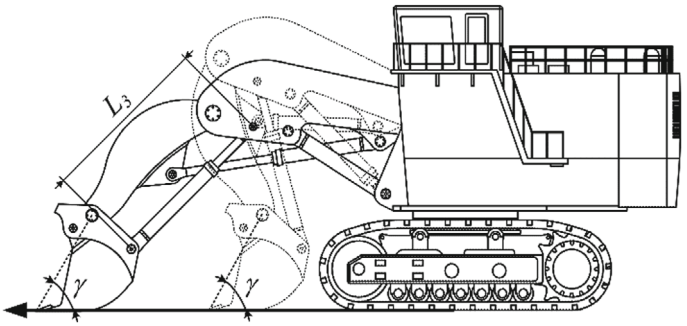


Fig. 3 Calculation for the objective function of level crowding

$$\begin{aligned} \mathbb{F}_2(\mathbf{x}) &= C_2 - \mathcal{F}_{stick} \\ \mathbb{F}_3(\mathbf{x}) &= C_3 - \mathcal{F}_{bucket} \\ \mathbb{F}_4(\mathbf{x}) &= C_4 - \mathcal{F}_{crowding} \end{aligned} \tag{3}$$

where, C_2, C_3, and C_4 are constants and used to ensure that the value of objective function is not negative; $\mathcal{F}_{(.)}$ represents the digging force or crowding force.

Digging Atlas Indicators. The digging graph is commonly used to further check the rationality of design scheme for hydraulic excavator. Based on the evaluation criteria proposed by digging atlas superposition method, the most critical indicator η_{stick} is chose as the objective function

$$\mathbb{F}_5(\mathbf{x}) = C_5(1 - \eta_{stick}) \tag{4}$$

where, C_5 is a constant.

2.4 Constraints Handling

According to the design requirements of hydraulic shovel, 61 constraint conditions are established. These constraints take into account the working range, geometric constraints, transmission angles, compact structure, mechanism characteristics, mechanical properties, and so on. Penalty function method is the most commonly used constraint handling method in engineering, and its essence is to transform constrained optimization problems into unconstrained optimization problems. Since most EAs are only suitable for unconstrained optimization problems, penalty function method is applied to make the optimization model suitable for more EAs.

3 Evolutionary Algorithm

EA is the important branch of artificial intelligence, which refers to a cluster of heuristic search algorithms based on biological evolution and natural evolution. After years of development, the connotation of EA is also gradually evolving. Broadly speaking, EA covers all the intelligent optimization algorithms that search optimization through the iteration of population.

3.1 Algorithm Selection

EAs were originally used to solve SOPs. With the deepening of research and the introduction of Pareto optimality, MOEAs came into being, and quickly showed far superior performance in multi-objective and many-objective optimization problems. The solutions of SOPs can be directly compared by the value of the objective function, while in MOPs, besides the comparison of Pareto dominance, the distribution of population should also be taken into account. Therefore, in consideration of the same recombination operator, the main difference between MOEAs and SOEAs is in the comparison and screening mechanism for solutions, namely the environmental selection operator.

In order to make a relative comprehensive comparison for SOEAs and MOEAs, the algorithms with same recombination operators are firstly chosen. For SOEAs, three classical algorithms are selected, including real-coded genetic algorithm (RCGA), differential evolution algorithm (DE), and particle swarm optimization algorithm (PSO) [8]. For MOEAs, three classical algorithms which are homologous to the above SOEAs are chosen [9], including non-dominated sorting genetic algorithm II (NSGA-II), multi-objective differential evolution algorithm based on decomposition (MOEA/D-DE), and multiobjective particle swarm optimization algorithm (MOPSO).

In addition, six novel SOEAs and MOEAs are added for comparison, including optimal foraging algorithm (OFA) [10], improved multi-operator differential evolution algorithm (IMODE) [11], PSO with an adaptive framework to tune the coordinate systems (ACoS-PSO) [12], NSGA-II with strengthened dominance relation (NSGA-II/SDR) [13], non-dominated sorting differential evolution based on object-space prediction (OSP-NSDE) [14], and novel multiobjective particle swarm optimization (NMPSO) [15].

3.2 Performance Metrics

In this paper, the performance of EAs is compared respectively from the view of single-objective optimization and multi-objective optimization. For single-objective optimization, the values of objective function are directly aggregated as the final

optimization function. For multi-objective optimization, the two popular comprehensive evaluation metrics, inverted generational distance (IGD) and normalized hypervolume (NHV) [16], are selected which can simultaneously measure the convergence and distribution of the optimization solution set. Moreover, for constrained optimization problems, success rate (SR) is implemented to describe the probability that the algorithm successfully finds a feasible solution and the running time also needs to be considered.

4 Case Study

4.1 Parameters Setting

Based on the parameters of 70t hydraulic excavator in Ref. [3], the optimization case of ordinary shovel excavation mechanism is established. Through the analogy analysis of the existing hydraulic excavators, the decision variables and constant values in the optimization model are set as shown in Tables 1 and 2, respectively.

Considering the complexity of the optimization problem, the population size is set as $NP = 1001$ and the maximum evaluation number of objective functions is set as $maxFE = 1 \times 10^6$. Except for the ACoS-PSO whose program is provided by the author, other algorithms are available in the matlab platform PlatEMO [17], which is an open-source optimization platform recommended by IEEE CEC, as shown in Fig. 4. The control parameters of EAs are set as the recommended values by original references and PlatEMO, as shown in Table 3. As the true PF of the actual

Table 1 The ranges of design variables

Variables	Value range	Unit	Variables	Value range	Unit
x_A	[50,700]	mm	L_{BH}	[1000,2500]	mm
y_A	[2000,3200]	mm	L_{CH}	[1000,2500]	mm
x_E	[700,1800]	mm	L_{CJ}	[500,800]	mm
y_E	[1500,2500]	mm	L_{CD}	[1800,2600]	mm
L_{AB}	[4300,5200]	mm	L_{DJ}	[2000,2800]	mm
L_{AG}	[600,2000]	mm	$L_{1\min}$	[2000,3000]	mm
L_{BG}	[2500,4800]	mm	$L_{2\min}$	[2000,3000]	mm
L_{AF}	[2500,4800]	mm	$L_{3\min}$	[2000,3000]	mm
L_{BF}	[800,2000]	mm	λ_1	[1.5,1.8]	–
L_{AI}	[3500,4500]	mm	λ_2	[1.5,1.8]	–
L_{BI}	[550,650]	mm	λ_3	[1.5,1.8]	–
L_{BC}	[3000,4000]	mm	–	–	–

Table 2 The values of constants in objective functions

Constant	Value
C_1	10
C_2	1000
C_3	1000
C_4	1000
C_5	10

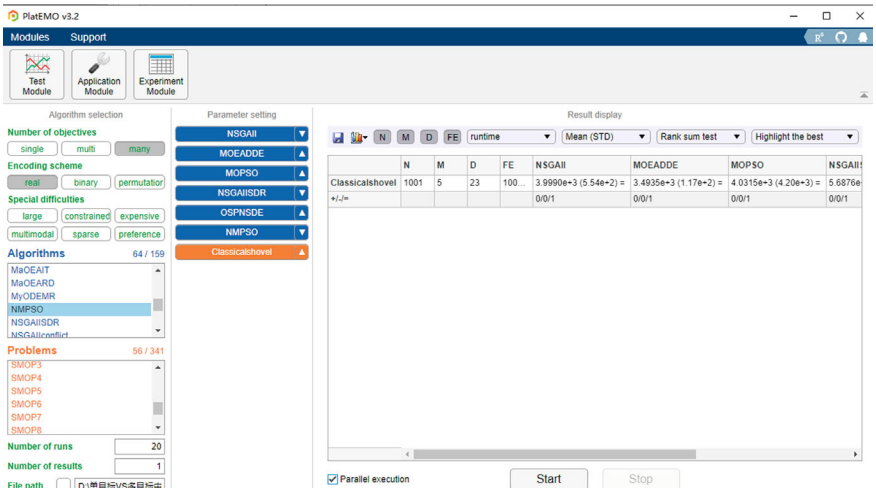


Fig. 4 The running interface of open-source optimization platform PlatEMO

engineering problem is unknown in advance, it is necessary to integrate the non-dominated solutions obtained by all algorithms, and this solution set are used as the approximate PF for the calculation of performance metrics. The whole experiments are carried out on a PC with Intel(R) Core(TM) i7-8700 K CPU.

4.2 Result Analysis

The values of performance metrics are shown in Table 4. The mean value of IGD, NHV, Runtime and other metrics is given and their standard deviation is in brackets. The running time of EAs which fails to find a feasible solution is meaningless and ignored in the table.

It can be seen from the data of performance metrics for EAs in Table 4 that the performance of MOEAs is significantly better than SOEAs on the whole. The distribution-preserving mechanism of MOEAs not only ensures the diversity of the final optimization results, but also retains the effective information in the solution set

Table 3 Parameter settings of EAs

EAs	Control parameters
RCGA	$p_c = 1, p_m = 1/D, \eta_c = 20, \eta_m = 20$
DE	$F = 0.5, CR = 0.9$
PSO	$w = 0.4$
OFA	–
IMODE	$NP^{min} = 4, A_{rate} = 2.6$
ACoS-PSO	$c_1 = 1.4962, c_2 = 1.4962, w = 0.7298$
NSGA-II	$p_c = 1, p_m = 1/D, \eta_c = 2, \eta_m = 20$
MOEA/D-DE	$F = 0.5, CR = 1.0, p_m = 1/D, \eta_m = 20, T = 0.1NP, \delta = 0.9, n_r = 2$
MOPSO	$div = 10, w = 0.4$
NSGA-II/SDR	$p_c = 1, p_m = 1/D, \eta_c = 20, \eta_m = 20$
OSP-NSDE	$F = 0.5, CR = 1.0, p_m = 1/D, \eta_m = 20, \lambda = 0.2, p = 50$
NMPSO	$p_m = 1/D, \eta_m = 20, \omega \in [0.1,0.5], c_1, c_2, c_3 \in [1.5,2.5]$

and promotes the evolution process. Even from the perspective of single-objective optimization, most MOEAs is also comparable to the best SOEAs, which demonstrates the superiority of MOEAs. It is worth noting that MOEA/D-DE performs best in the optimization problem of ordinary shovel excavation mechanism but it is not the most advanced one among these six MOEAs. On the other hand, it tells us that the performance on the standard test functions is not an absolute reference, and the key is to select the appropriate algorithm for the specific problem. And this corresponds with *No Free Lunch Theorems* [18].

In order to make a clearer understanding, the two best EAs, MOEA/D-DE and IMODE, are selected for further comparative analysis. As shown in Fig. 5, the parallel coordinate graphs of non-dominated solutions set from approximate PF, MOEA/D-DE, and IMODE are plotted, respectively. It can be founded that, as mentioned above, SOEA based on aggregate function could only obtain a narrow area of PF, and only MOEA can get the complete PF and provide more balanced solutions for decision makers.

Finally, by the multi-attribute decision making method of TOPSIS [19], the most satisfactory parameter scheme of ordinary shovel excavation mechanism is selected from the obtained 6026 non-dominant solutions, as shown in Fig. 6. The optimization scheme of ordinary shovel excavation mechanism and the main performance parameters of three 70-ton hydraulic excavators on the market are compared, as shown in Table 5. It can be found that the optimized scheme in this paper has obvious advantages in the maximum bucket digging force, maximum stick digging force, and maximum crowding force. In general, the parameter scheme of ordinary shovel excavation mechanism after optimization has strong competitiveness among existing hydraulic excavators on the market.

Table 4 Performance comparison of EAs

EAs		SR	Runtime(s)	Single-objective optimization performance metrics		Multi-objective optimization performance metrics	
				Best	Mean	IGD	NHV
Single-objective	RCGA	19/20	6.09e + 3 (3.87e + 2)	2.36e + 3	2.61e + 3 (1.64e + 2)	3.14e + 2 (4.19e + 1)	8.44e-4 (5.62e-4)
	DE	0/20	–	–	–	–	–
	PSO	0/20	–	–	–	–	–
	OFA	0/20	–	–	–	–	–
	IMODE	20/20	3.05e + 3 (1.69e + 2)	2.05e + 3	2.05e + 3 (3.06)	2.33e + 2 (2.24)	9.34e-3 (9.10e-5)
	ACoS-PSO	0/20	–	–	–	–	–
Multi-objective	NSGA-II	20/20	4.00e + 3 (5.54e + 2)	2.31e + 3	2.43e + 3 (7.58e + 1)	1.53e + 2 (3.52e + 1)	3.99e-3 (1.55e-3)
	MOEA/D-DE	20/20	3.49e + 3 (1.17e + 2)	2.05e + 3	2.07e + 3 (1.16e + 1)	2.02e + 1 (1.27)	1.84e-2 (2.40e-4)
	MOPSO	0/20	–	–	–	–	–
	NSGA-II/SDR	19/20	5.69e + 3 (8.60e + 2)	2.25e + 3	2.45e + 3 (2.04e + 2)	2.27e + 2 (1.08e + 2)	3.14e-3 (1.68e-3)
	OSP-NSDE	20/20	2.94e + 3 (4.41e + 2)	2.14e + 3	2.19e + 3 (6.19e + 1)	5.65e + 1 (2.36e + 1)	1.23e-2 (2.15e-3)
	NMPSO	13/20	4.90e + 3 (2.77e + 3)	2.22e + 3	2.49e + 3 (2.20e + 2)	1.87e + 2 (8.36e + 1)	3.61e-3 (2.95e-3)

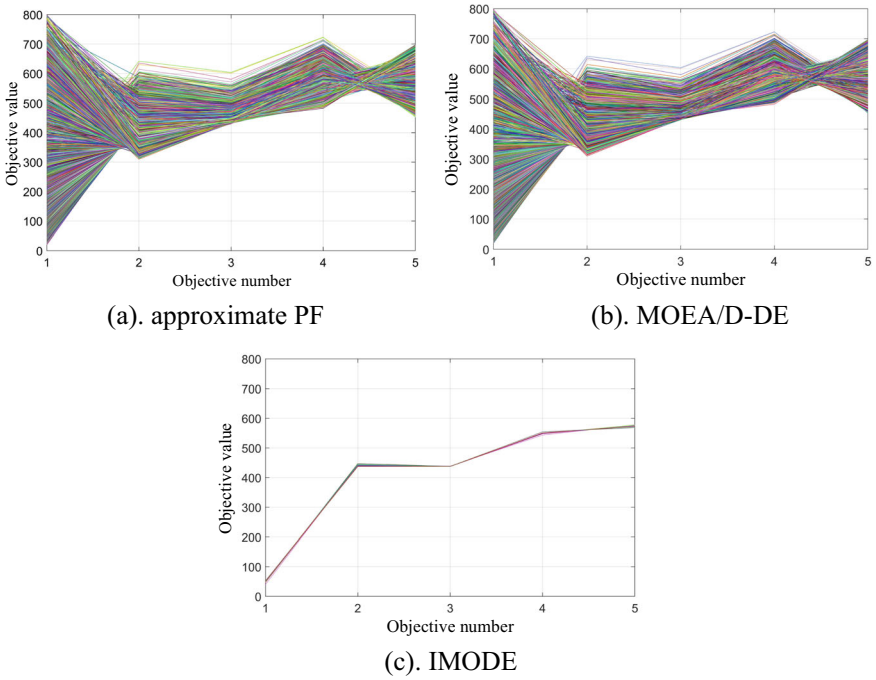


Fig. 5 The parallel coordinates of the obtained nondominated solution sets

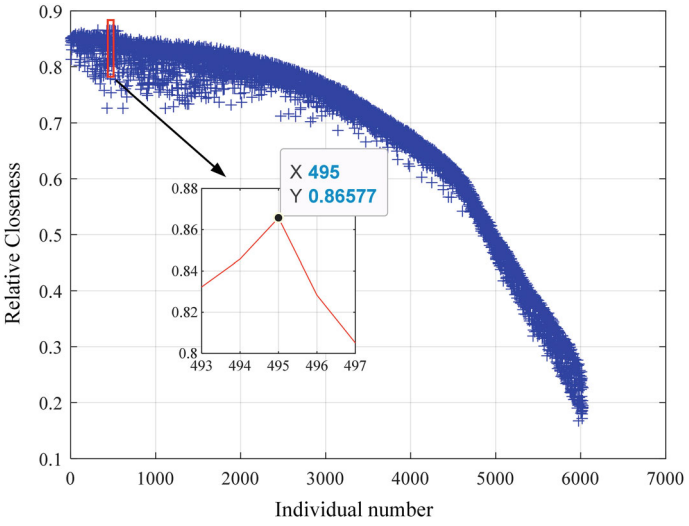


Fig. 6 The most satisfactory solution from the obtained nondominated individuals by TOPSIS

Table 5 Technical parameters comparison of 70-ton hydraulic excavators

Technical parameter	Optimized scheme	RH30E	R964	PC750
Max. digging height (mm)	10,309	8800	10,400	10,635
Max. digging reach (mm)	9493	9600	9280	10,305
Max. digging depth (mm)	2300	2000	3100	3535
Max. dumping height (mm)	8339	6500	7250	7180
Crowd distance on level (mm)	3300	3100	3610	3875
Max. stick digging force (kN)	562	450	500	404
Max. bucket digging force (kN)	562	390	370	477
Max. crowding force on the ground (kN)	451	–	410	–

'-' denotes the lack of public data from manufacturers

5 Conclusion

In this paper, the multi-objective optimization model of hydraulic excavator with ordinary shovel excavation mechanism is established. The optimization model is a really complex optimization problem which contains 5 objective functions, 23 decision variables, and 61 constraints. Two common approaches for MOPs, SOEAs with weight sum method and MOEAs are firstly compared and studied in the MOPs of hydraulic excavator. Six representative SOEAs and six MOEAs are selected to be compared in the optimization case of 70-ton hydraulic excavator. The analysis of the performance indicators from both the view of single-objective optimization and multi-objective optimization shows that MOEAs have significant advantages over SOEAs, which also verifies the effectiveness of EAs in practical engineering problem, like the parameter design of hydraulic excavator. In the future, we want to develop more optimal cases of hydraulic excavator with different tonnage and various excavation mechanisms for verifying more advanced MOEAs.

References

1. Li B, Wei Z, Wu J et al (2023) Machine learning-enabled globally guaranteed evolutionary computation. Nat Mach Intell 5(4):457–467
2. Zou J, Sun R, Liu Y et al (2024) A multipopulation evolutionary algorithm using new cooperative mechanism for solving multiobjective problems with multiconstraint. IEEE Trans Evol Comput 28(1):267–280
3. Xu G, Feng Z, Wang W et al (2024) Constrained multiobjective optimization design for ordinary shovel attachment of hydraulic excavator based on evolutionary algorithm. Eng Appl Artif Intell 135:108839
4. Xu G, Feng Z, Guo E et al (2022) Multi-objective optimization of hydraulic shovel using evolutionary algorithm. Autom Constr 142:104486

5. Yu X, Shi G, Zhao F et al (2023) Comparative study on single-objective optimization and multi-objective optimization of convex textured friction couple performance. Surf Topogr Metrol Prop 11(1):015017
6. Pereira V, Sousa P, Rocha M (2022) A comparison of multi-objective optimization algorithms for weight setting problems in traffic engineering. Nat Comput 21(3):507–522
7. Wang M, Zheng S, Sweetapple C (2022) A framework for comparing multi-objective optimization approaches for a stormwater drainage pumping system to reduce energy consumption and maintenance costs. Water 14(8):1248
8. De Lacerda MGP, Pessoa LFD, Neto FBD et al (2021) A systematic literature review on general parameter control for evolutionary and swarm-based algorithms. Swarm Evol Comput 60(10):100777
9. Patrausanu A, Florea A, Neghina M et al (2024) A systematic review of multi-objective evolutionary algorithms optimization frameworks. Processes 12(5):869
10. Ding C, Zhu GY (2024) Improved optimal foraging algorithm for global optimization. Computing 106(7):2293–2319
11. Sallam KM, Elsayed SM, Chakrabortty RK et al (2020) Improved multi-operator differential evolution algorithm for solving unconstrained problems. In: Proceedings of 2020 IEEE Congress on Evolutionary Computation. Glasgow, UK, pp 1–8
12. Liu Z, Wang Y, Yang S et al (2019) An adaptive framework to tune the coordinate systems in nature-inspired optimization algorithms. IEEE Trans Cybern 49(4):1403–1416
13. Tian Y, Cheng R, Zhang X et al (2019) A strengthened dominance relation considering convergence and diversity for evolutionary many-objective optimization. IEEE Trans Evol Comput 23(2):331–345
14. Guerrero-Peña E, Araújo AFR (2019) Multi-objective evolutionary algorithm with prediction in the objective space. Inf Sci 501:293–316
15. Lin Q, Liu S, Zhu Q et al (2018) Particle swarm optimization with a balanceable fitness estimation for many-objective optimization problems. IEEE Trans Evol Comput 22(1):32–46
16. Xu Y, Li F, Zhang H et al (2024) An adaptive reference vector guided many-objective optimization algorithm based on the pareto front density estimation. Swarm Evol Comput 88:101601
17. Tian Y, Zhu WJ, Zhang XY et al (2023) A practical tutorial on solving optimization problems via PlatEMO. Neurocomputing 518:190–205
18. Rashki M, Faes MGR (2023) No-free-lunch theorems for reliability analysis. ASCE-ASME J Risk Uncertain Eng Syst Part Civ Eng 9(3):15
19. Hooshangi N, Gharakhanlou NM, Razin SRG (2023) Evaluation of potential sites in Iran to localize solar farms using a GIS-based Fermatean Fuzzy TOPSIS. J Clean Prod 384:135481

Open Access This chapter is licensed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License (http://creativecommons.org/licenses/by-nc-nd/4.0/), which permits any noncommercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if you modified the licensed material. You do not have permission under this license to share adapted material derived from this chapter or parts of it.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.



Numerical Study of the Characteristics of Spiral Teeth of Steam Turbine Shaft End Combined Seal



Zhaojun Meng, Keyu Wang, Shuyu Meng, Da Liu, Shilong Liu, and Lei Zhang

Abstract Aiming at the spiral sealing section part of spiral-labyrinth combination seal (SLCS), the influence of different shapes of spiral teeth on the sealing effect is investigated. Select three different spiral tooth shapes of SLCS (Rectangular-SLCS, Trapezoidal-SLCS, and Serrated-SLCS) to study the pressure drop and sealing effect of 1 to 6 spiral teeth in the SLCS sealing section under identical working conditions. Simulation results indicate that the Rectangular-SLCS(R-SLCS) with three spiral teeth has the best sealing effect: the pressure drop generated by the R-SLCS steam-side spiral sealing section is 10.61% higher than that of the Trapezoidal-SLCS(T-SLCS), 17.69% higher than that of the Serrated-SLCS(S-SLCS). The leakage amount of the R-SLCS steam-side sealing section is 9.05% less than that of the T-SLCS, 18.81% less than that of the S-SLCS. This verifies the feasibility of the new combined shaft seal design and provides a reference for the optimization design of turbine shaft end seals.

Keywords Spiral-Labyrinth Combined Seal · Leakage rate · Pumping effect

1 Introduction

Steam turbines are large rotating equipment used in thermal power plants, and to prevent rubbing between rotating and stationary components, a certain clearance must be maintained between these parts during installation [1]. The presence of these clearances can lead to steam leakage during actual turbine operation; therefore, various shaft seals are implemented in power plants to reduce such leakage and ensure the stable and safe operation of the turbines. However, the utilization of shaft sealing systems only improves the working environment of steam turbines and cannot completely eliminate steam leakage. The use of shaft seal fans results in the pressure at the shaft seal exhaust port being lower than atmospheric pressure, leading to an

Z. Meng · K. Wang (✉) · S. Meng · D. Liu · S. Liu · L. Zhang
School of Energy and Power, Shenyang Institute of Engineering, Shenyang, China
e-mail: wangky_work@163.com

© The Author(s) 2026

H. Saman et al. (eds.), *The 9th International Conference on Advances in Construction Machinery and Vehicle Engineering*, Lecture Notes in Mechanical Engineering,
https://doi.org/10.1007/978-981-96-5527-4_16

increased pressure differential and consequently greater steam leakage compared to systems without shaft seals. Excessive leakage flow reduces the efficiency of the steam turbine [2]. Additionally, under the effect of the pressure differential, high-pressure steam can leak outward through the shaft seal section, adversely affecting the turbine's operating environment, while air can infiltrate the shaft seal on the low-pressure side, compromising the condenser vacuum. Therefore, researching a novel shaft end steam seal to reduce steam leakage during turbine operation is of significant necessity.

This paper proposes a Spiral-Labyrinth Combined Seal (SLCS) to enhance the sealing performance at the shaft ends of steam turbines. Since the sealing teeth of the spiral sealing section of the SLCS can create a reverse spiral airflow within the sealing section that pre-disturbs the leaking fluid from the cylinder, increasing the disorder of the leakage [3–5]. According to some previous studies, different shapes of spiral teeth are different in terms of resistance to steam leakage and pumping flow rate generated by the spiral teeth, so the shape of the spiral sealing teeth is one of the key factors affecting the sealing effect of SLCS.

To evaluate the influence of different spiral tooth shapes on the sealing effect of SLCS. By analyzing the pressure drop of the spiral sealing section on the leakage of steam and the flow rate of the leakage of steam after the sealing section to judge the sealing performance of different tooth types, and find out the best sealing performance of the spiral tooth structure. It is necessary to compare and analyze the sealing effect of different spiral tooth shapes and tooth numbers under the same operating conditions. This study aims to provide valuable insights for optimizing the design of steam turbine shaft seals.

2 Numerical Method

In this study, traditional LS is enhanced by incorporating a spiral seal, which operates as a hydrodynamic reverse-flow seal [6]. This type of seal relies on the pressure generated by the spiral grooves during shaft rotation to prevent steam from leaking outward [7], thereby achieving the sealing function. Since each spiral tooth is continuous, it generates a "pumping effect" during high-speed rotation, facilitating energy exchange between the propulsion device and the fluid. This effect of the spiral teeth creates an 'air-curtain' in the sealing channel which reduces leakage [8].

The pumping flow rate can be calculated using Eq. (1):

$$q = \frac{U_i \cos\beta}{2} \cdot s \cos\alpha \cdot hi \tag{1}$$

In Eq. (1): U_i is the peripheral speed of the spiral; β is the screw angle; s is the groove width; h is the tooth height; i is the number of spiral starts.

From the geometric relationship of the spiral, it can be expressed as Eq. (2):

$$si = \pi D \tan\beta\left(\frac{s}{s+\delta}\right) = \pi D k_1 \tan\beta \tag{2}$$

Based on Eq. (3), the pumping flow rate is shown in Eq. (4):

$$\left(\frac{s}{s+\delta}\right) = k_1 \tag{3}$$

$$q' = \frac{\Delta p}{12\eta}\left(\frac{iC^3\cos\beta}{L/\sin\beta}\right) \tag{4}$$

Assuming the pressure difference across the spiral groove as $q' = \frac{\Delta p}{12\eta}\left(\frac{iC^3\cos\beta}{L/\sin\beta}\right)$, the leakage along the spiral groove can be determined as follows:

$$q' = \frac{\Delta p}{12\eta}\left(\frac{iC^3\cos\beta}{L/\sin\beta}\right) \tag{5}$$

The leakage through the spiral groove is computed as follows:

$$q'' = \frac{\Delta p}{12\eta}\cdot\frac{i\varepsilon^3(s+\delta)\cot\beta\cos\beta}{(L/con\beta)(1-k_1)} \tag{6}$$

The total leakage of the spiral sealing system is obtained by summing the contributions from Eqs. (5) and (6), as expressed in Eq. (7):

$$Q = \frac{\pi D\varepsilon^3\Delta p}{12\eta L}\cdot\frac{\cos^2\beta}{(k_1-1)}\left[1+k_1(k_1-1)k_2^2\tan^2\beta\right] \tag{7}$$

$$k_2 = 1+\frac{h}{\varepsilon} \tag{8}$$

This study is about the viscous fluid dynamic problems, fluid flow requires adherence to the laws of conservation of mass and conservation of momentum. Any flow problem has to satisfy the mass conservation equation, which is shown in Eq. (9):

$$\frac{\partial\rho}{\partial t}+\frac{\partial(\rho u)}{\partial x}+\frac{\partial(\rho v)}{\partial y}+\frac{\partial(\rho w)}{\partial z} = 0 \tag{9}$$

In Eq. (4): u, v and w are the velocity vectors along x, y and z directions. The momentum conservation [9] is expressed in Eqs. (10) and (11):

$$\frac{\partial\rho_m}{\partial t}+\frac{\partial}{\partial x_i}(\rho_m v_{mi}) = 0 \tag{10}$$

$$\frac{\partial}{\partial x_t}(\rho_m v_{mi})+v_{mi}\frac{\partial(\rho_m v_{mi})}{\partial x_j} = -\frac{\partial p}{\partial x_i}+\mu_m\frac{\partial}{\partial x_j}\left(\frac{\partial v_{mi}}{\partial x_j}+\frac{\partial v_{mj}}{\partial x_i}\right) \tag{11}$$

ρ_m is weighted density, v_{mi} is weighted average speed, μ_m is weighted viscosity of fluid and gas, p is pressure.

3 Structural Establishment

3.1 Geometric Characterization and Model Development of SLCS Type Seal

The Spiral-Labyrinth Combined Seal is retrofitted on a sealing major shaft. Based on the traditional LS, many sealing teeth have been replaced with spiral teeth. Due to the changes in the main sealing shaft, the sealing teeth in the outer casing of the shaft seal corresponding to the spiral section also need to be removed. Figure 1 shows the overall model of the SLCS is developed based on a 600 MW generating unit.

The left-side inlet of the seal is connected to the cylinder and is designated as the steam side inlet, with the left sections of the seal collectively referred to as the steam side sealing section. The two pipes in the middle represent the shaft seal steam supply pipeline and exhaust pipeline, respectively. The right-side inlet of the seal, located near the external atmosphere, is designated as the air-side inlet. Since the right inlet of the sealing section is connected to another type of gapless seal (not show in this paper), no air leaks into the sealing section from the right side. Consequently, the spiral teeth are only installed in the steam side sealing section, with the spirals configured as left-handed.

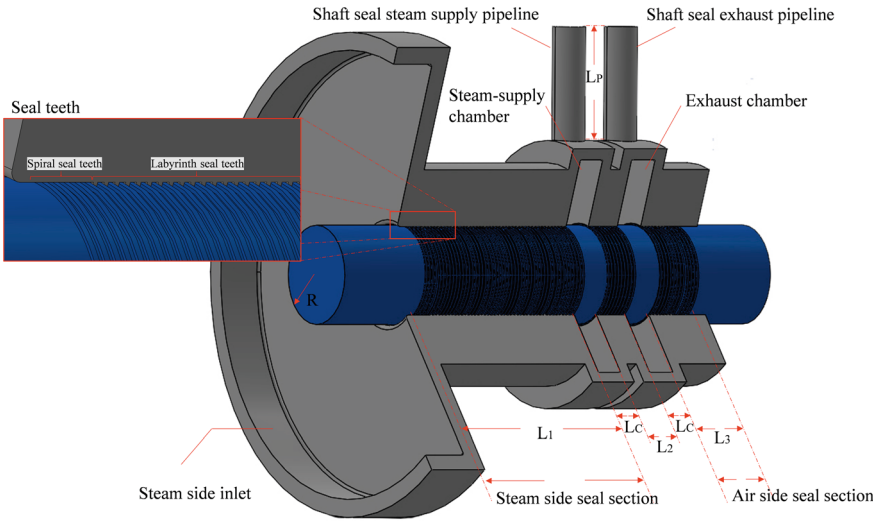
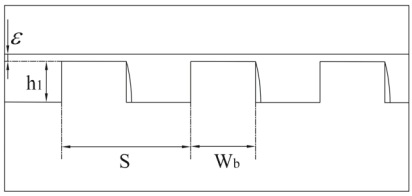
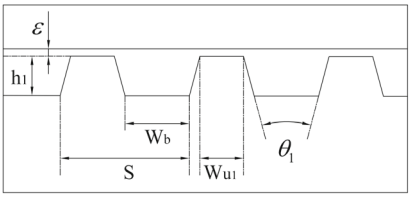


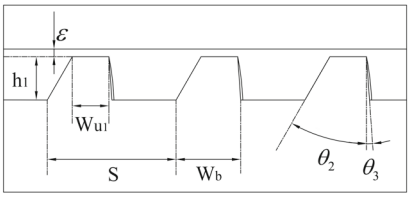
Fig. 1 Three-dimensional structure of SLCS



(a) Rectangular spiral teeth (R)-SLCS;



(b) Trapezoidal spiral teeth (T)-SLCS



(c) Serrated spiral teeth (S)-SLCS

Fig. 2 Structural parameters of 3 types of sealing teeth

In this study, Fig. 2 reveal three distinct types of spiral teeth: rectangular teeth, serrated teeth and trapezoidal teeth.

The parameters for the rectangular teeth are consistent with those of the labyrinth seal currently used in power plant at the present stage. The parameters for the serrated and trapezoidal teeth are established in accordance with relevant industry standards. This study aims to evaluate the impact of different spiral tooth shapes and number of teeth on sealing performance [10]. Consequently, SLCS with different spiral tooth shapes are established and the number of spiral teeth is set from 1 to 6. Since the parameters of gap between two teeth and gear tip clearance are kept constant for each SLCS, the analysis focuses only on the effect of spiral tooth shapes and number of teeth on the sealing performance. The overall parameters of SLCS are listed in Table 1.

4 Boundary Conditions

The simulation boundary conditions were set based on the operational parameters of a 600 MW steam turbine's low-pressure cylinder. The steam side inlet of the shaft seal, named inlet, was set with a pressure of 0.7973 MPa, while the outlet of the shaft seal's exhaust pipe, named outlet, was set with a pressure of 0.095 MPa. Since no air leakage is present, only steam was considered within the shaft seal, with the steam having a density of 82.568 kg/m^3, a specific heat capacity of 10,249 J/(kg·K), and a dynamic viscosity of 0.033 kg/(m·s). The interface between the flow field and the

Table 1 Design parameters of SLCS

Structure size parameters	Value
Sealing shaft radius R/mm	125
Length of left shaft seal section L1/mm	519
Length of middle shaft seal section L2/mm	89
Length of right shaft seal section L3/mm	146
Length of steam-supply and exhaust chamber L_C/mm	68
Radius of steam supply pipeline and exhaust pipeline r/mm	54
Length of pipeline L_P/mm	496
Number of sealing teeth	72
Gear tip clearance ε/mm	0.5
Gap between two sealing teeth S/mm	5
Height of sealing teeth h_1/mm	3
Bottom width of sealing teeth bottom W_b/mm	5
Tilt angle of trapezoidal teeth θ_1/degree	30
Apex width of trapezoidal teethW_{u2}/mm	3.4
The front bevel angle of serrated teeth θ_2/degree	30
The back bevel angle of serrated teeth θ_3/degree	3
Apex width of the serrated teethW_{u1}/mm	3.04

shaft was modeled as a rotating wall with a rotational speed of 3000 rpm, while all other surfaces were modeled as stationary walls.

The spiral flow channel of the SLCS is a spatially curved channel as shown in Fig. 3 [11]. The cartesian coordinate system is used. X–Y plane is the rotational plane of the SLCS, and the center of mass of the model is selected as the center of rotation. X-axis is the reference baseline of the sealing axis, and the positive direction of X-axis is the direction of the steam flow.

5 Results and Discussion

5.1 The Pressure and Pressure Drop of Spiral Seal Sections in Different Types of SLCS

The internal simulated flow pressure results for the sealing sections of different types of spiral tooth of SLCS which with various numbers of spiral teeth, as calculated using simulation software. The outlet pressure of the spiral section and the pressure drop across each spiral section were recorded. The compiled data is shown in Fig. 4.

From Fig. 4a, it is evident that, under a constant inlet pressure on the steam side, the outlet pressures of different spiral tooth sealing sections differ significantly.

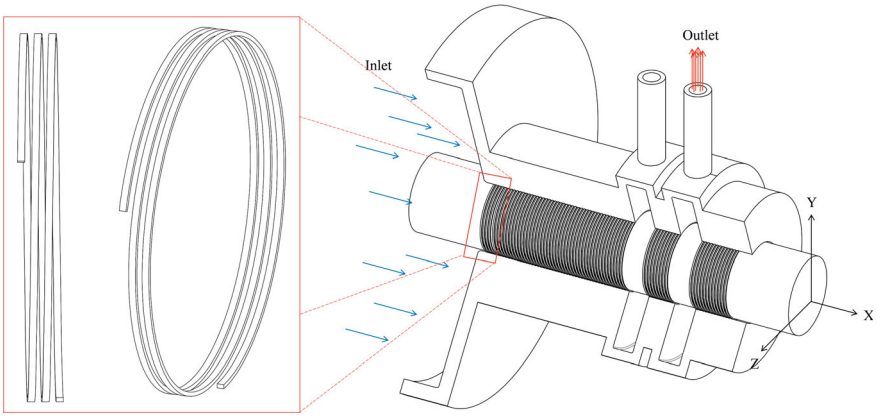
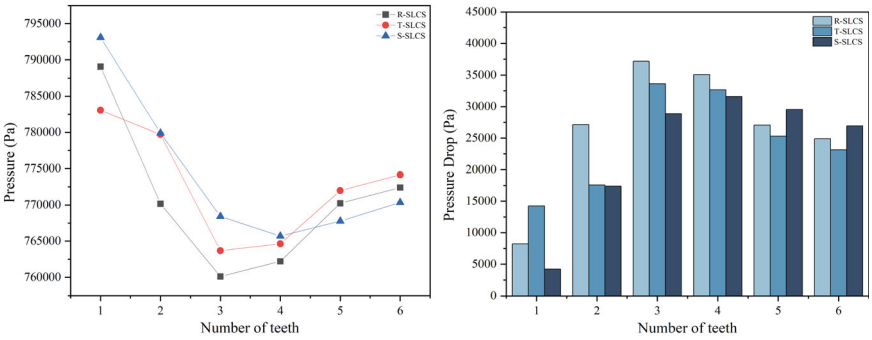


Fig. 3 Setting of SLCS inlet and outlet; Spiral flow channel magnification diagram



(a) Pressure (b) Pressure drop.

Fig. 4 The pressure and pressure drop after each screw section on the steam side are compared with LS

The over-all SLCS with three spiral tooth shapes exhibits a trend of decreasing and then increasing. Among all R-SLCS configurations, the outlet pressure of the three rectangular spiral tooth sealing sections reaches its lowest point at 760,116.53 Pa. Similarly, the T-SLCS achieves its minimum sealing outlet pressure at three spiral teeth, which is 763,680.44 Pa. The pressure situation of the S-SLCS spiral section is different from the previous two types of teeth; the outlet pressure of the sealing section of the S-SLCS only reaches its lowest point after modifying the four spiral teeth, which is 765704.84 Pa. It can be clearly seen from the figure that the pressure after the spiral sealing section of S-SLCS reaches its minimum value, the increase in outlet pressure is relatively small as the number of teeth increases. However, when the outlet pressure of the R-SLCS and T-SLCS spiral sealing sections reaches the minimum value, with the increase in the number of sealing teeth, the outlet pressure of the spiral sealing section increases slightly at first, and then shows a significant rise

when there are five spiral teeth, even exceeding the outlet pressure of the trapezoidal spiral sealing section with the same number of teeth.

From Fig. 4b, it is evident that the pressure drop generated by the three rectangular spiral tooth sealing sections is 37183.47 Pa, with a pressure drop rate of 4.66%; the pressure drop generated by the three trapezoidal spiral tooth sealing sections is 33619.56 Pa, with a pressure drop rate of 4.22%; and the pressure drop generated by the four serrated spiral sealing sections is 31595.16 Pa, with a pressure drop rate of 3.96%. Although the serrated sealing sections having more spiral teeth than the rectangular and trapezoidal teeth, the pressure drop generated is smaller than that of the other two spiral tooth shapes. A comparison of the maximum pressure drop generated by the three types of spiral sections, it can be concluded that the pressure drop generated by the steam side, R-SLCS is 10.61% higher than that of the T-SLCS spiral section and 17.69% higher than that of the S-SLCS spiral section.

5.2 *Steam Leakage After SLCS Steam Side Sealing Section of Different Spiral Teeth*

The mass flow rate after the left sealing section of each type of seal is statistically analyzed to determine the steam leakage of steam after passing through each sealing section, as illustrated in Fig. 5.

From Fig. 5, it can be seen that the trends in steam mass flow rate after the steam side sealing section of the SLCS with three different spiral tooth shapes are

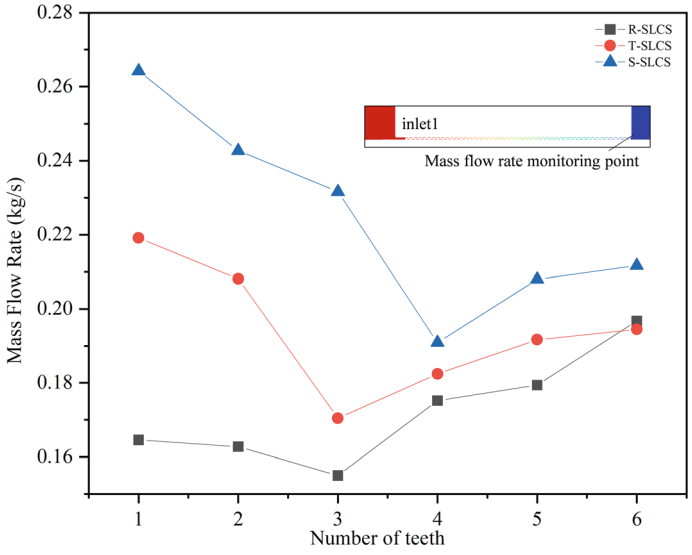


Fig. 5 Mass flow rate behind steam side of each spiral sealing section

basically the same. The mass flow rate following the steam side sealing of the R-SLCS type shaft seal with 1–5 spiral teeth is the lowest. This observation indicates that the rectangular spiral teeth are stronger than the other two types of spiral teeth in hindering the steam leakage. However, in the case with 6 spiral teeth, the sealing situation changes, and the mass flow rate of vapor after the sealing section of the T-SLCS type shaft seal becomes the lowest.

In addition, the mass flow rates of the R-SLCS and T-SLCS reach their minimum values when the number of teeth is 3, and the mass flow rate of the S-SLCS reaches its mini-mum when the number of teeth is 4. The minimum mass flow rate after the R-SLCS sealing section is 0.15501 kg/s, the minimum mass flow rate after the T-SLCS sealing section is 0.17044 kg/s, and the minimum mass flow rate after the S-SLCS sealing section is 0.19091 kg/s. Under the condition of selecting the minimum mass flow rate after each sealing section on the steam side, it is observed that the steam leakage behind the R-SLCS sealing section is 9.05% lower than that of the T-SLCS sealing section and 18.81% lower than that of the S-SLCS sealing section.

6 Conclusions

This article focuses on the spiral-labyrinth combined seal, mainly exploring the effects of different spiral tooth shapes and varying tooth counts in the spiral section on the pressure reduction and sealing performance of internal leakage steam in the SLCS. The main conclusions of this article are as follows:

(1) The SLCS with rectangular and trapezoidal spiral teeth both have lower steam leakage at three sealing teeth than all other sealing structures with the same tooth shape. The SLCS with serrated spiral teeth had lower steam leakage at 4 sealing teeth than all other sealing configurations with the same tooth shape. Among all SLCS, the sealing effect of the SLCS with 3 rectangular spiral teeth is the best.

(2) By selecting the best sealing performance shaft seals from three types of spiral tooth SLCS for comparison of steam pressure drop capability, it can be concluded that the pressure drop generated by the steam side spiral section of R-SLCS is 10.61% higher than that of T-SLCS and 17.69% higher than that of S-SLCS. The leakage amount of steam after the steam side spiral sealing section of R-SLCS is 9.05% lower than that of T-SLCS, and 18.81% lower than that of S-SLCS.

After the comparative analysis of three kinds of spiral teeth, the spiral sealing section of SLCS can use rectangular spiral teeth. This type of SLCS can significantly reduce the amount of air leakage, reduce the burden of the extractor and improve the operating efficiency of the generator unit.

References

1. Tanuma T, Tominaga J (2022) Sealing designs and analyses for steam turbines. In: Advances in steam turbines for modern power plants. Woodhead Publishing, pp 359–382
2. Yang R, Yang J, Qi M, Shi L, Pen, Z (2010) Numerical investigation of leakage and power loss for different seal types in turbine stage environment. In: Proceedings of the ASME turbo expo 2010: power for land, sea, and air, vol. 4, pp 1081–1090
3. Shujan MD, Mortazavi F, Palazzolo A (2023) Swirl brake design for improved rotordynamic vibration stability based on computational fluid dynamics system level modeling. ASME Open J Eng
4. Zhang W, Wu K, Gu C, Tian H, Zhang X, Li C (2021) Swirl brakes optimization for rotordynamic performance improvement of labyrinth seals using computational fluid dynamics method. Tribol Int 159
5. Zhang M, Xia Y (2024) Preswirl control and rotordynamic stability improvement of turbomachine seal equipped with newly designed spiral steering blades. Tribol Int 196
6. Wasilczuk F, Flaszynski P, Kaczynski P, Szwaba R, Doerffer P, Marugi K (2021) Air curtain application for leakage reduction in gas turbine shroud sealing. Aerosp Sci Technol 112:106636
7. Li X, Meng X, Zhao W, Peng X (2024) Numerical investigation on a hybrid Porous-Spiral Groove Me-chanical Face Seal. Tribol Int 198:109943
8. Auld A, Hilfer M, Hogg S, Ingram G, Messenger A (2013) Application of an air-curtain fluidic jet type seal to reduce turbine shroud leakage. Turbo Expo: Power Land, Sea, Air 55140
9. Wu Y-J, Chen H, Uzoejinwa BB, Zhao B, Xu D (2020) Hydrodynamic characteristics of cylindrical spiral grooves for pump annular seals. Int J Fluid Mach Syst 13:556–565
10. Zhou W, Zhao Z, Wang Y, Shi J, Gan B, Li B, Qiu N (2021) Research on leakage performance and dynamic characteristics of a novel labyrinth seal with staggered helical teeth structure. Alex Eng J 60:3177–3187
11. Gu Q, Wu Z, Yang J, Zhang W, Zhang M (2023) A flow control method for the leakage reduction of annular gas seals using the helical gas curtain generator. Aerosp Sci Technol 141:108528

Open Access This chapter is licensed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License (http://creativecommons.org/licenses/by-nc-nd/4.0/), which permits any noncommercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if you modified the licensed material. You do not have permission under this license to share adapted material derived from this chapter or parts of it.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.



Refined Computation of Lifting Performance Considering Two Trolleys Distance on Shipbuilding Gantry Cranes



Yahui Fan, Zhenping Jiang, and Rumin Teng

Abstract Shipbuilding gantry crane is one of the crucial equipment for enhancing shipbuilding efficiency. Most of them feature large lifting capacity, significant height and span. However, there is a lack of refined research on the lifting performance, which is an important parameter of the crane. Taking a shipbuilding gantry crane as an example, its stress and displacement are analyzed in this paper by using the geometric nonlinear method. According to the national crane design specification and based on the stress intensity of the upper and lower trolleys at different lifting distances, the relationship between the lifting capacity of shipbuilding gantry cranes and the distances of the upper and lower trolleys is obtained. A more detailed lifting performance table of the crane is provided, which provides a theoretical basis for the rational use of large-tonnage shipbuilding gantry cranes and giving full play to their performance.

Keywords Shipbuilding gantry crane · Lifting performance · Non-linear analysis · Finite element analysis

1 Introduction

Shipbuilding gantry crane is a crucial piece of equipment specifically utilized in the shipbuilding industry. It has a substantial impact on enhancing shipbuilding efficiency and can perform operations such as overturning and lifting ship segments and the entire ship. This type of crane has a large lifting capacity, a wide span, a high distance from the ground, and complex working conditions [1]. As ship tonnage gradually increases, the lifting capacity and tonnage of shipbuilding gantry cranes also become larger and larger. The largest shipbuilding gantry crane designed and

Y. Fan · R. Teng (✉)
Dalian University of Technology, Dalian, China
e-mail: tengrumin@163.com

Z. Jiang
Dalian COSCO Shipping Heavy Industry Co, Ltd, Dalian, China

© The Author(s) 2026

H. Saman et al. (eds.), *The 9th International Conference on Advances in Construction Machinery and Vehicle Engineering*, Lecture Notes in Mechanical Engineering,
https://doi.org/10.1007/978-981-96-5527-4_17

manufactured by Zhenhua Group, named 'Honghai', has a rated lifting capacity of 22,000 tons. For large shipbuilding gantry cranes, due to their heavy weight and being mostly located in harbors, wind loads and the asynchrony of large and small truck movements can lead to major accidents caused by deflection. For instance, a gantry crane in a prefabricated components factory toppled over due to wind loads, resulting in an economic loss of nearly 20 million yuan. [2] Therefore, studying the lifting performance of shipbuilding gantry cranes is of great significance.

Lifting performance is a crucial parameter in the production, manufacturing, and usage of cranes. By researching lifting performance, it is possible to maximize the lifting capacity of a crane safely, enhance productivity, and reduce the occurrence of safety accidents [3]. In the early days, the lifting performance table was mainly calculated by hand and using Excel spreadsheets [4]. Its calculation formula was obtained through relevant standards. There is a significant degree of simplification in the calculation, and there is an error between it and the actual results, which does not better reflect the changes in crane lifting performance. With the development of finite element technology, current researchers, with the help of finite element software, calculate crane lifting performance through methods such as the constant increment method, dichotomous method, proportional method, modified proportional iteration method, and Lagrange algorithm. However, compared with traditional truck cranes and crawler cranes, the shipbuilding gantry crane studied in this paper is in the form of a double lifting trolley and has a special working condition of turning and lifting, which is different from the operations of other forms of cranes. There is less research on lifting performance in this area, and there is no more detailed lifting performance table. With the development of science and technology, it is now possible to conduct more refined research on the lifting performance of shipbuilding gantry cranes through more scientific methods of lifting conditions to better ensure safe operation.

In order to comprehend the working characteristics of shipbuilding gantry cranes, scholars at home and abroad have conducted extensive research. Su Jia-Chen analyzed the cause of the overturning of a gantry crane induced by a downburst based on the fluid–solid coupling method [5]. Onur carried out theoretical analysis and numerical test on the deflection value of the main girder of shipbuilding gantry cranes under different loads [6]. Petrenko simulated the motion of a shipbuilding gantry crane under irregular vertical forces and determined the values of dynamic loads [7].

A certain shipbuilding gantry crane is taken as an example in this paper. Its overall finite element simulation model is established. It is calculated by geometrically nonlinear methods. Its stress and displacement states are analyzed. A more detailed lifting performance table of shipbuilding gantry cranes for lifting conditions is provided based on the stress intensity. A theoretical guide is thus provided for shipbuilding gantry cranes to better exert their lifting performance.

2 Overview of Shipbuilding Gantry Cranes

The shipbuilding gantry crane is mainly composed of the main beam, rigid leg, flexible leg, upper trolley, lower trolley, maintenance crane, cart travelling mechanism and so on. Its main structure is shown in Fig. 1. Among these components, the main girder, rigid leg and flexible leg have the largest mass and are the main load-bearing structures of the shipbuilding gantry crane. A set of lifting systems is respectively equipped on the upper and lower trolleys, and the lifting of the lifted objects is realized through winches. The trolleys can be moved in the direction of the main beam. The cart is made to travel on the track through the travelling mechanism, which enables the change of the position of the hoisted object. Compared to traditional single-hook or single-trolley cranes, the crane studied in this paper is a double-trolley arrangement. The upper and lower trolleys are located at different heights on the track and can work together to complete the turning or lifting conditions. The main girder of the crane and the lower support of the flexible leg are in trapezoidal box beam section. The rigid leg is in all rectangular box beam section, and the main body of the flexible leg is in round tube section.

The 200 t 120 m shipbuilding gantry cranes of a company are taken as an example in this paper to study the changes in lifting performance of shipbuilding gantry cranes. The main parameters of it are shown in Table 1.

The main girder is regarded as the main load-bearing structure of shipbuilding gantry cranes. As an example, the detailed dimensions of the main girder section are offered. The main girder of the shipbuilding gantry crane studied in this paper is arranged in a double girder structure. The two girders are of the same structure and the cross-section is in the form of a trapezoidal box beam section. Its detailed parameters are shown in Fig. 2.

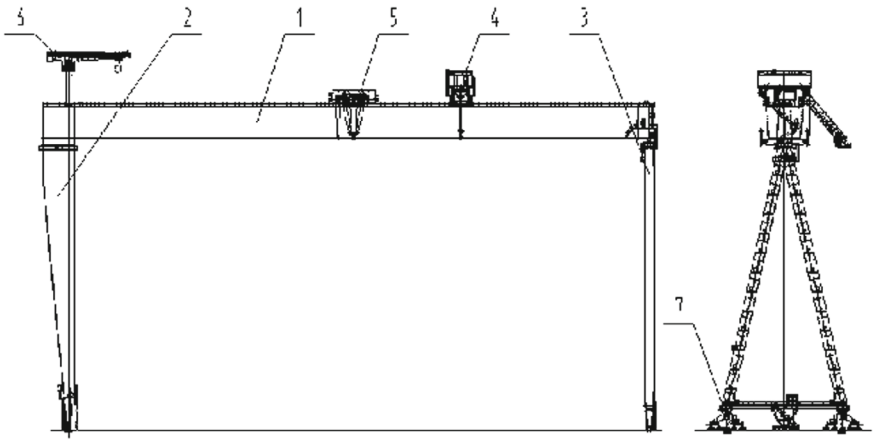
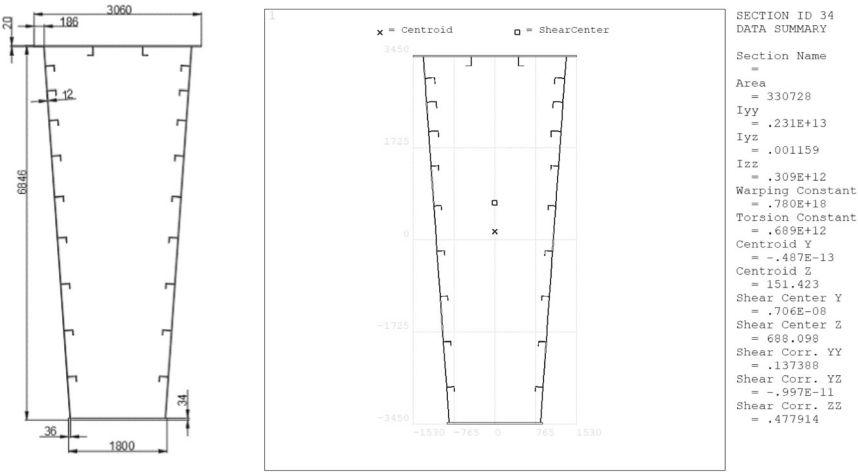


Fig. 1 Shipbuilding gantry crane structure sketch (1—main beam; 2—rigid leg; 3—flexible leg; 4—upper trolley; 5—lower trolley; 6—service crane; 7—cart)

Table 1 Shipbuilding gantry crane parameter list

Main performance parameters	Parameter values	Main performance parameters	Parameter values
Working class	A5	Self-weight of main beam	837.04 t
Rated lifting capacity	200t	Self-weight of upper trolley	86.3 t
Main beam span	120 m	Lower trolley weight	42.7 t
Lifting height	65 m	Rated overturning lifting capacity	180t
Total deadweight (without trolley)	1418.4t	Modulus of elasticity	210000 N/mm^2
Self-weight of rigid leg	269.8t	Poisson's ratio	0.3
Flexible leg deadweight	105.4t	Material Density	7850 kg/m^3



a) section view **b)** concrete size

Fig. 2 Shipbuilding gantry crane main girder section

The section is overall in the shape of an isosceles trapezoid, being wide at the top and narrow at the bottom. There are two angle steels arranged on the upper cover and ten angle steels on each side, which are used to ensure strength and stability. The main beam is bolted together with the rigid leg. Meanwhile, the main girder is connected to the flexible leg by a ball hinge, which can be rotated along the x, y, and z directions to reduce the stress of the crane.

3 Finite Element Modelling

The main girder, rigid leg, and flexible leg of the crane are all thin-walled shell structures. With length-to-section width ratios greater than 10 and shell thicknesses ranging from 14 to 34 mm, they can be simplified to shell elements or beam elements for finite element calculation. In finite element modelling, the shell element is much more complex than the beam element. Moreover, the number of elements generated by beam element modelling will be much less than that by shell element modelling. Thus, finite element calculation with beam elements is more efficient and easier to modify. In this paper, beam188 element is used for modelling. The schematic diagram of beam188 element is shown in Fig. 3. This element has two nodes, and each node has 6 to 7 degrees of freedom. It is suitable for analyzing slender to medium slender beam structures. The element is based on the theory of Timosinko beams and includes the effect of shear deformation.

When beam elements are used for modelling, the main beam, rigid leg and flexible leg are simplified as beam elements. The cart wheels are also simplified as beam elements, which are only utilized to simulate the effect of bending moments. Structures like the cross partition boarding perpendicular to the direction of the beam axis cannot be modelled with the beam element and are compensated for by having the material density adjusted. Parts such as trolley tracks, ladder platforms, etc. that penetrate through the main steel structure are also applied to the corresponding main structure by having the density adjusted. Parts such as maintenance cranes, driver's cabins, flexible hinges, etc. that are more centrally distributed and have less influence on the main body stress are applied to the corresponding nodes by having mass21 elements applied. The trolleys are connected to the main girder by having a V-shaped rigid beam structure built to simulate the effect of the bending moment of the crane weight on the main girder. The masses of the trolley and the lifting weight are applied to the corresponding nodes by means of concentrated loads according to the different operating conditions.

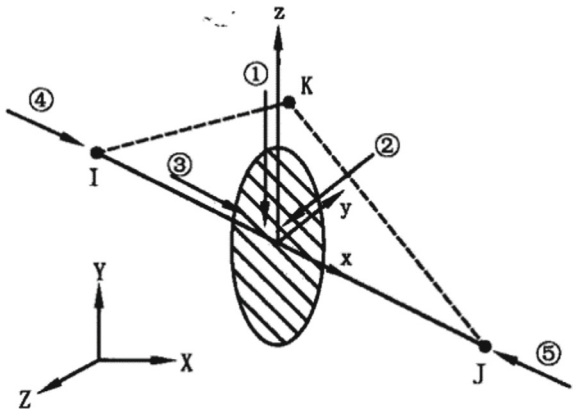


Fig. 3 Schematic diagram of beam188 element

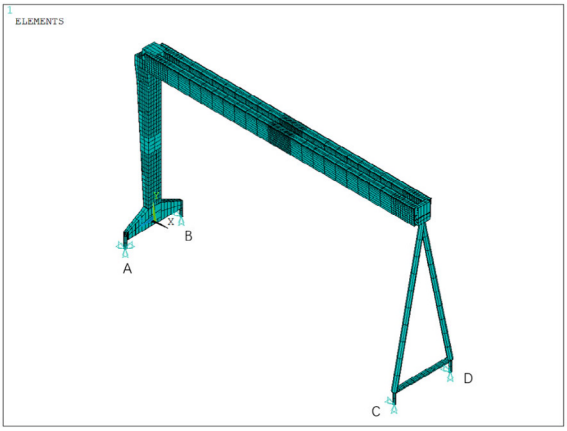


Fig. 4 Finite element modelling of shipbuilding gantry cranes

The finite element model is shown in Fig. 4:

The main node constraints in the model are shown in Table 2:

In the process of cart walking, if the speeds on both sides are not synchronous, it will cause the measurement positions on the rigid leg side and the flexible leg side to skew. According to the national standard requirements, the maximum amount of skew is 0.3%. Therefore, on the flexible leg side of the cart wheels, the degree of freedom of movement along the track direction is released. On the rigid leg side, three displacement degrees of freedom on one side are constrained, and on the other side, the displacement degree of freedom along the track direction is released. According to the characteristics of cart operation, there is a gap between the cart wheels and the track, and it will rotate. Thus, all the rotational degrees of freedom of the four nodes are released.

In terms of the connection of each main structure of the crane, the main beam is bolted to the rigid leg. In the finite element model, they are directly connected by common nodes. The flexible leg is connected to the main girder through a ball hinge. The ball hinge has three directions of rotational degrees of freedom. At the connection of the main girder and the flexible leg, the two nodes couple their three displacement degrees of freedom through cp.

According to the crane design specification [8], the main loads and their coefficients required for each working condition of finite element analysis are shown in Table 3:

Table 2 Main node constraints table

Working condition	Restricted position	X	Y	Z	RX	RY	RZ
Work strength, rigidity	A	Fixed	Fixed	Fixed	Free	Free	Free
	B	Fixed	Fixed	Free	Free	Free	Free
	C	Fixed	Fixed	Free	Free	Free	Free
	D	Fixed	Fixed	Free	Free	Free	Free

Table 3 Main load table

Working condition number	DL	TL1	TL2	TL3	az	Fx	Sk	Clarification
1	1.1	1.23			1	1	1	Rigid leg
2	1.1		1.23	1.23	1	1	1	Mid-span
3	1.1				1	1	1	Soft leg

DL: Fixed load; TL1: Trolley on rigid leg side; TL2: Trolley on mid-span; TL3: Trolley on flexible leg side; az: Acceleration dynamic load on cart; Fx: Lateral load on trolley; SK deflection load

Table 4 Allowable stress table

Plate thickness (mm)	≤ 16	> 16–40	> 40–63	> 63–80	> 80–100
σa (MPa)	228	224	220	213	206
τa (MPa)	131	130	127	123	119

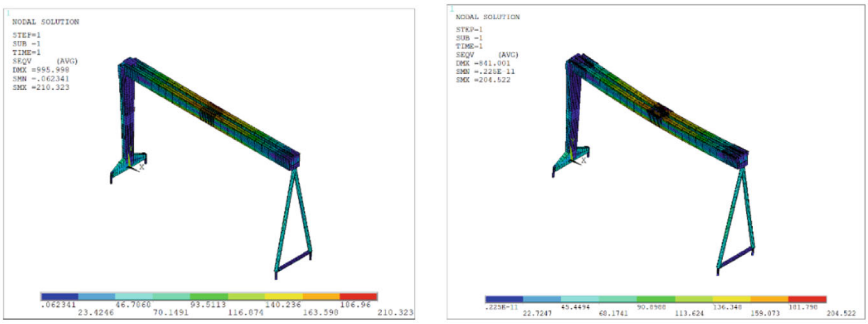
The working conditions described in this paper are class A conditions, and their permissible stresses are shown in Table 4.

4 Algorithm Simulation Case and Analysis

In the finite element calculation, since the span of the shipbuilding gantry crane is large and its rigidity is poor, resulting in large deformation under lifting load, the analysis is conducted on whether to use the geometric nonlinear method. For example, when the upper and lower trolleys are located in the middle of the span under turnover conditions, the equivalent force map and deformation cloud are shown in Fig. 5 and 6, the maximum equivalent stress obtained by nonlinear analysis is 210.3 MPa, while the maximum equivalent stress obtained by linear finite element analysis is 204.5 MPa. The nonlinear stress is 2.8% greater than the linear stress. Moreover, the geometric nonlinear displacement is 994.5 mm, and the displacement obtained by linear analysis is 841.0 mm, with a change of 18.3%. In order to better reflect the change of lifting performance and avoid the occurrence of accidents, the option of geometric nonlinearity is enabled in the finite element calculation.

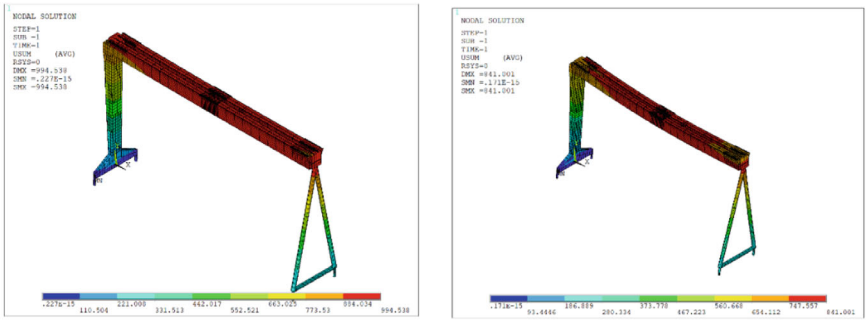
For the upper and lower trolleys located in the rigid leg side, span and flexible leg side of the three typical working conditions, static analysis is conducted to verify the reasonableness of the finite element model and provide a basis for the subsequent lifting performance analysis and research.

When the upper and lower trolleys are located at the rigid leg side, the equivalent force map is shown in Fig. 7, the maximum equivalent stress at the mid-span and the flexible leg side are 131.6 MPa, 210.3 MPa and 115.7 MPa respectively. When the upper and lower trolleys are located at the rigid leg side, the maximum equivalent stress point appears in the rigid leg side of the beam. Meanwhile, at both the mid-span and the flexible leg side, the maximum stress point appears in the mid-span of



a) nonlinear calculation **b**) linear calculation

Fig. 5 Equivalent force map

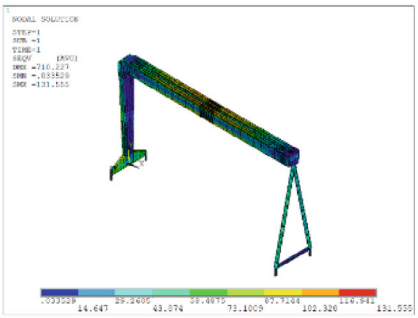


a) nonlinear calculation **b**) linear calculation

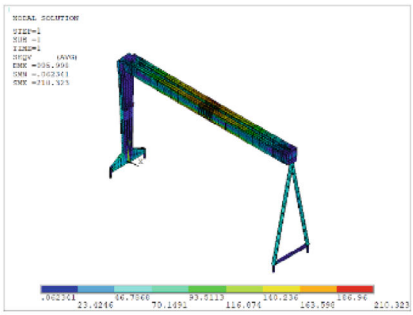
Fig. 6 Map of deformed clouds

the main girder. Weak points are defined at the middle part of the main beam span, and the thickness of the box girder at the middle part of the main girder span can be increased to enhance the overall strength of the crane.

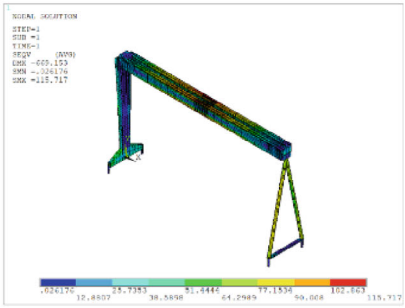
Because the upper and lower trolleys being located in the middle of the span simultaneously under turnover condition has the largest equivalent stress and is thus the most dangerous working condition. In the actual work of shipbuilding gantry crane, as the trolleys will move in the direction of the main beam, strength analyses and lifting performance studies are carried out when the trolleys are selected to be located near the span center.



a) rigid leg side



b) mid-span



c) soft leg side

Fig. 7 Equivalent force map

5 Lifting Performance Evaluation

For the evaluation of lifting performance, there are mainly methods such as strength criterion, overall stability principle and local stability principle. For shipbuilding gantry cranes, since the stability calculation is more complicated and it is difficult to quantify the evaluation index, the strength criterion is mainly adopted for analysis in this paper. In the crane design specification, the strength criterion mainly involves the stress analysis of components subjected to tensile, compressive and bending moments, which can be divided into local positive stress method and composite positive stress method.

(1) Local positive stress

When the concentrated load acts on the top surface of the member, the local compressive stresses generated at the edge of the web below the member are calculated by the following formula:

$$\sigma_m = \frac{P}{tc} \leq [\sigma] \tag{1}$$

(2) Compound positive stress

When the same calculation point of a member is subjected to a positive stress σ, a shear stress τ, and a localised positive stress σ_{m}, the composite stress at that point is calculated by the following equation.

$$\sqrt{\sigma^2 + \sigma_m^2 - \sigma\sigma_m + 3\tau^2} \leq [\sigma] \tag{2}$$

In the study of the lifting performance of shipbuilding gantry cranes, this paper starts from the strength criterion. The geometric nonlinear method is used to calculate the stresses of the cranes under different lifting capacities for lifting performance assessment.

When the crane is in the turning condition, the upper trolley wheel stress P = 20.16 t, web thickness $\tau = 12$ mm, concentrated load distribution length c = 390 mm, localized positive stress at mid-span position

$$\sigma_m = \frac{P}{\tau c} = 42.26MPa < \sigma_a = 224MPa \tag{3}$$

Based on finite element calculations, positive stress $\sigma = 197$ MPa, shear stress $\tau = 7$ MPa, then the composite stress

$$\sqrt{\sigma^2 + \sigma_m^2 - \sigma\sigma_m + 3\tau^2} = 180MPa < \sigma_a \tag{4}$$

For fixed shipbuilding gantry cranes, the other loads except the distance between the upper and lower trolleys and the lifting capacity are controlled to remain unchanged. For different distances between the upper and lower trolleys, the maximum stress of the crane is calculated by adjusting the lifting capacity. This stress is compared with the permissible stress of the working conditions. If the equivalent stress σ is less than the permissible stress (σ), then the requirement is satisfied. At this time, this lifting capacity is the rated lifting capacity at the corresponding trolley distance. If the equivalent stress σ is greater than or equal to the permissible stress (σ), then the lifting capacity is continuously adjusted to meet the requirements.

Taking the upper and lower trolleys being located in the middle of the span under turnover condition as an example, its rated turnover lifting capacity is 180 t. The weight of the lifting object is equally divided between the upper and lower trolleys. The maximum equivalent stress calculated by the finite element of this condition is taken as a standard. The upper trolley is moved to the side of the rigid leg and the lower trolley is moved to the side of the flexible leg. The two trolleys are separated by every 1 m until a total separation of 10 m. The lifting capacity is adjusted to make the stress conform to the above standard, so as to achieve the assessment and evaluation of the lifting performance of shipbuilding gantry cranes. When the distance between the upper and lower trolleys reaches 10 m, the lifting capacity reaches the rated lifting capacity of 200 t for the shipbuilding gantry crane. The variation of lifting performance with trolley distance is shown in Table 5.If the distance is continued to

Table 5 Lifting performance table

Distance between upper and lower trolley/m	Lifting capacity/t	Equivalent force/MPa	Change in lifting capacity/t
0	180.0	210.3	0
1	181.6	210.3	1.6
2	183.4	210.3	1.8
3	185.3	210.3	1.9
4	187.2	210.3	1.9
5	189.2	210.3	2.0
6	191.3	210.3	2.1
7	193.5	210.3	2.2
8	195.7	210.3	2.2
9	198.0	210.3	2.3
10	200.5	210.3	2.5

be increased, theoretically, the lifting capacity can continue to increase. However, taking into account the impact of wheel pressure, it is only calculated until the lifting performance of 200 t is reached.

The data was fitted to a quadratic term equation as follows:

$$y = 0.0404x^2 + 1.6456x + 179.97 \tag{5}$$

Fit the image to the above data as follows, the lifting performance variation chart is shown in Fig. 8.

It can be obtained that with the increase of the distance between the upper and lower trolley, the lifting capacity of the shipbuilding gantry crane also increases, and is not a linear increase, but the growth rate has become larger. When the distance between the upper and lower trolley 10 m, the lifting capacity of the shipbuilding gantry crane reaches its rated lifting capacity of 200 t. When the crane is in the rollover condition and the upper and lower trolleys are located in the middle of the span simultaneously, in the rollover process, the coordination of movements needs to be taken into account. As a result, the lifting capacity is low.

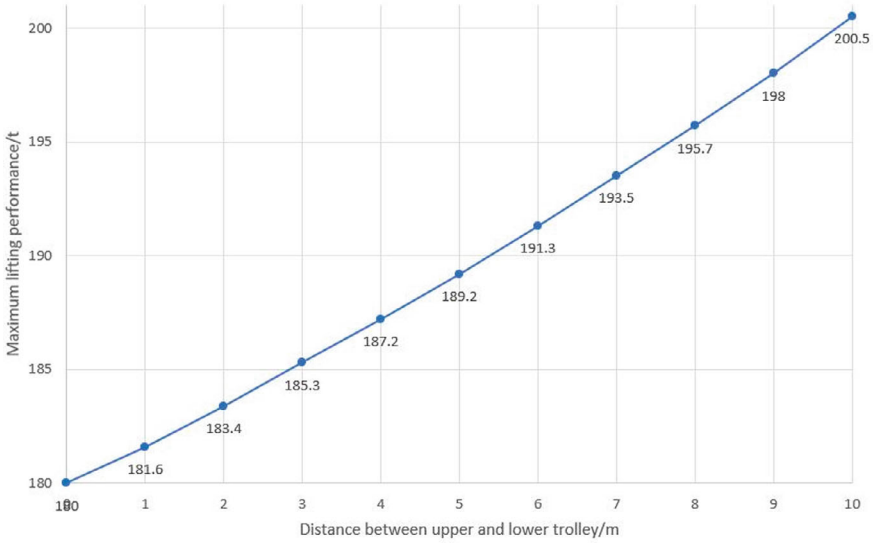


Fig. 8 Lifting performance variation chart

6 Conclusion

In this paper, a finite element simulation model based on beam elements is established for addressing the incomplete estimation of lifting performance during the working process of shipbuilding gantry cranes. The finite element model is calculated and analyzed by the geometric nonlinear method. The stress and displacement states under its typical working conditions are obtained. The lifting performance of the shipbuilding gantry crane under lifting condition is calculated by the method of fixed equivalent stress.

The lifting performance of shipbuilding gantry cranes under lifting conditions is only calculated and analyzed on the basis of stress in this paper, with fewer factors being considered. It can be analyzed from the combination of strength, stiffness and stability in the future.

Acknowledgements This work is supported by National Natural Science Foundation of China under Grant No.52275088 and the Fundamental Research Funds for the Central Universities un-der Grant (No.DUT22LAB506, No.DUT22LAB507).

References

1. Ruan Q (2013) Structural design and analysis of 600 tonne lattice type double trolley gantry crane. Harbin Institute of Technology

2. Dajiang C (2024) Crane tipping accident on the stability of gantry crane. Brand Stand 04:94–96
3. Wang T (2021) Analysis of the impact of super lifting cable pre-tensioning length on the lifting performance of telescopic boom. Dalian Univ Technol. https://doi.org/10.26991/d.cnki.gdllu.2021.002094
4. Ning W, Peng F (2016) Software development of lifting performance calculation for large tonnage truck crane. J Shaanxi Inst Technol (Nat Sci Ed) 32(06):13–18+52
5. Su J, Li L, Chan PW et al (2023) Numerical simulation research on the overturning of gantry crane by downbursts. Heliyon 9(8)
6. Onur YA, Gelen H (2020) Design and deflection evaluation of a portal crane subjected to traction load. Mater Test 62(11):1131–1137
7. Petrenko V (2024) The influence of track vertical irregularities on the crane dynamic behaviour: lecture notes in intelligent transportation and infrastructure 167–177
8. GB/T 3811-2008, Code for crane design

Open Access This chapter is licensed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License (http://creativecommons.org/licenses/by-nc-nd/4.0/), which permits any noncommercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if you modified the licensed material. You do not have permission under this license to share adapted material derived from this chapter or parts of it.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.



Smoothness Analysis of Independent Suspension Based on Interconnection Mode for a Five-Axle All-Terrain Crane



Yunbo Ma, Ruming Teng, Xifeng Zhang, Shuai Song, and Xin Wang

Abstract The interconnection mode of the hydraulic system of a 130 t five-axle all-terrain crane with independent suspension is investigated in this paper. The two-degree-of-freedom vehicle dynamics for completely independent, front and rear interconnection, and front and rear to side cross-linked hydro-pneumatic suspensions are modeled. A whole vehicle hydraulic system model of hydro-pneumatic suspension is built by using AMESim software. Two different random road spectra are inputted and the root mean square (RMS) values of body acceleration, suspension dynamic deflection, and tire dynamic load are compared for the three cases. It is demonstrated by the analysis results that a front and rear interconnected hydro-pneumatic suspension system in the independent suspension hydraulic system of all-terrain cranes is more advantageous than the other two hydraulic systems in terms of ride smoothness and comfort.

Keywords Smoothness · Interconnection mode · Independent suspension · Five-axle all-terrain crane

1 Introduction

All-terrain cranes often travel on uneven road surfaces such as mountains and hills. In the complex and changeable driving environment, the traditional form of non-independent suspension is insufficient. On the other hand, the independent suspension system, which boasts good passability, driving smoothness, and operational stability, has begun to be gradually applied to all-terrain cranes. Hydro-pneumatic suspension serves as the elastic support for the all-terrain crane body. Generally, helium is used as the elastic medium and hydraulic oil as the force transfer medium [1]. At the current

Y. Ma · R. Teng (✉) · X. Wang
Dalian University of Technology, Dalian, China
e-mail: tengrumin@163.com

X. Zhang · S. Song
Taiyuan Heavy Industry Co, LTD, Taiyuan, China

© The Author(s) 2026

H. Saman et al. (eds.), *The 9th International Conference on Advances in Construction Machinery and Vehicle Engineering*, Lecture Notes in Mechanical Engineering,
https://doi.org/10.1007/978-981-96-5527-4_18

stage, the multi-axle suspension system is usually designed as the front and rear interconnected hydro-pneumatic suspension system, so as to improve smoothness and avoid the impact of uneven loads on the front and rear wheels. Or, through the front and rear to side cross-linked mode, the lateral tilt rigidity can be increased without increasing the suspension vertical rigidity [2].

Currently, scholars both domestically and internationally have carried out research on hydro-pneumatic suspension systems in the form of coupling with different connection modes. Ma Chao from Southwest Jiaotong University analyzed the independent hydro-pneumatic suspension of dump trucks and the hydro-pneumatic suspension coupled on the same side, and concluded that the smoothness of the hydro-pneumatic suspension coupled on the same side is superior [3]. Canadian scholar Lu introduced a new interconnected suspension and obtained stability in sway and pitch through simulation [4]. Indian scholar Jayaraman analyzed independent and interconnected suspensions for four-axle heavy trucks and found that the interconnected system enhances ride comfort and suspension performance [5]. At this stage, there are relatively fewer comparative studies on various interconnection modes. Therefore, this paper takes the hydro-pneumatic suspension system of a five-axle all-terrain crane as the research object, analyzes the completely independent, front and rear interconnected, and front and rear to side cross-linked hydro-pneumatic suspension hydraulic systems. It establishes a two-degree-of-freedom suspension model for the five-axle all-terrain crane and constructs a whole vehicle simulation model using AMESim. Subsequently, it conducts simulations on the smoothness of the whole vehicle under the input of random road surfaces and compares the root mean square values of body acceleration, suspension dynamic deflection, and tire dynamic load of the three.

2 Two-Degree-of-Freedom Suspension Model

The suspension system of a vehicle is a complex multi-mass vibration system. When analyzing road smoothness, the suspension only studies the vibration characteristics in the direction of droop and does not consider the motion in the directions of lateral inclination and pitch. Under such circumstances, the two-degree-of-freedom suspension system is not different from the multi-degree-of-freedom suspension system [6]. Therefore, a vehicle two-degree-of-freedom suspension model is established, and its simplified model is shown in Fig. 1.

In Fig. 1, the sprung mass m_1 is supported by the combined elastic and damping forces of the suspension cylinder, The unsprung mass is represented by m_2. Simplifying the tire as a spring K_s and damper C_s in parallel, with q representing the road spectrum input.

The nonlinear elastic force of the suspension cylinder F_k is given by.

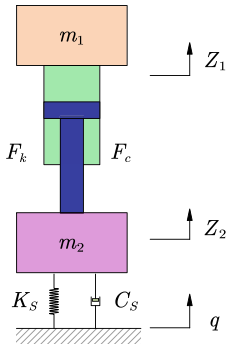


Fig. 1 Two degree of freedom vehicle suspension model

$$F_{\mathrm{k}} = \frac{P_0\left(A_{\mathrm{g}} - A_{\mathrm{h}}\right)V_0^{\mathrm{r}}}{\left[V_0 - \left(A_{\mathrm{g}} - A_{\mathrm{h}}\right)x\right]^{\mathrm{r}}} \tag{1}$$

where, P_0 is the initial inflation pressure. V_0 is the initial inflation volume. x is the relative displacement of the piston rod. A_g is the effective area of the rod cavity of the suspension. A_h is the effective area of the rodless cavity of the suspension.

The nonlinear damping force of the suspension cylinder F_d is given by.

$$F_d = \frac{\rho \cdot A_2^3 \cdot \dot{x}^2 sign(\dot{x})}{2 \cdot (c_b \cdot A_b + c_c \cdot A_c \cdot (0.5 + 0.5 sign(\dot{x})))^2} \tag{2}$$

where, ρ for the initial inflation pressure. C_b and C_c are throttle orifice, check valve flow coefficient. A_b, A_c are throttle orifice, check valve overflow area. Sign is the sign coefficient.

According to the above formula, the mathematical model is established.

$$\begin{cases} F_{\mathrm{k}} = \frac{P_0\left(A_{\mathrm{g}} - A_{\mathrm{h}}\right)V_0^{\mathrm{r}}}{\left[V_0 - \left(A_{\mathrm{g}} - A_{\mathrm{h}}\right)x\right]^{\mathrm{r}}} \\ F_d = \frac{\rho \cdot A_2^3 \cdot \dot{x}^2 sign(\dot{x})}{2 \cdot (c_b \cdot A_b + c_c \cdot A_c \cdot (0.5 + 0.5 sign(\dot{x})))^2} \\ \mathrm{m}_1 \ddot{z}_1 + F_d + F_{\mathrm{k}} = 0 \\ \mathrm{m}_2 \ddot{z}_2 - F_d - F_{\mathrm{k}} + C_S(\dot{z}_2 - \dot{q}) + K_{\mathrm{s}}(z_2 - q) \end{cases} \tag{3}$$

where: m_1 is the unsprung mass; m_2 is the unsprung mass; q is the road surface unevenness input; K_s is the tire stiffness coefficient; C_s is the tire damping coefficient, $Z_1 - Z_2$ is the suspension dynamic deflection; $K_s(Z_2 - q)$ is the tire dynamic load.

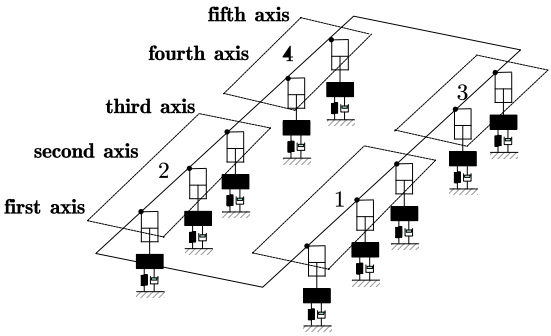


Fig. 2 Schematic diagram of four-point support

3 Interconnection Principle and Simulation Model

3.1 Principles of the Three Modes of Interconnection

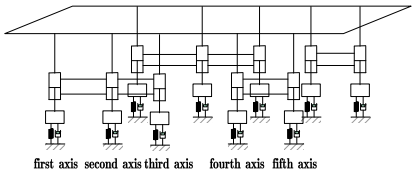
The research object of this paper is a five-axle all-terrain crane. The front three axles and the rear two axles of this five-axle vehicle are in close proximity to one another. The sprung mass and the steering center are positioned between the third and fourth axles [7]. With respect to the ten suspension cylinders, the crane will be divided into four parts. The front three wheels on the same side are considered as one part, and the rear two wheels are regarded as another part. The four hydraulic circuits act as the four support points, also known as the "four-point support", as depicted in Fig. 2.

Front and rear interconnected hydro-pneumatic suspension means that these four parts are independent of each other with no circuit connection between each part, and the rod and rodless chambers of each part are connected to each other. Front and rear to side cross-linked hydro-pneumatic suspension is characterized by the connection of the same side of the front and rear, and the staggered connection of the opposite side. That is, the rod cavities and rodless cavities of the hydro-pneumatic suspensions on both sides of the front and rear parts are connected in a staggered manner. The completely independent mode implies that the hydro-pneumatic suspension cylinders are independent of each other and do not communicate with each other [8]. The schematic diagrams of the three modes of hydro-pneumatic suspension systems are shown in Fig. 3.

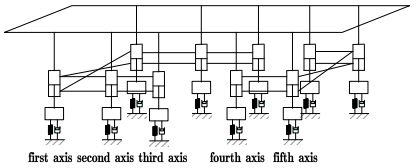
3.2 Three System Simulation Models

According to the actual model of the all-terrain crane and the hydraulic schematic diagram, the AMESim models of the hydraulic system of the three complete vehicle hydro-pneumatic suspension are established, as shown in Fig. 4.

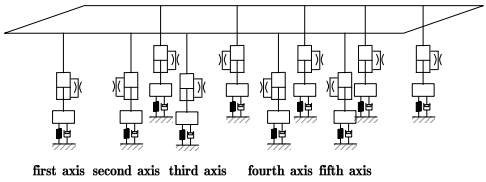
Random road spectra are fed into the model. Two random roads are selected: one with a vehicle speed of 80 km/h and pavement class B, and the other with a vehicle



a) Front and rear interconnection system

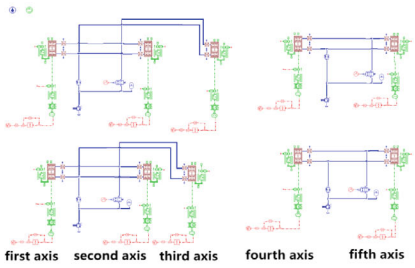


b) Front and rear to side cross-linked system

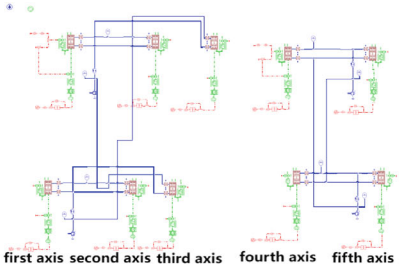


c) Completely independent system

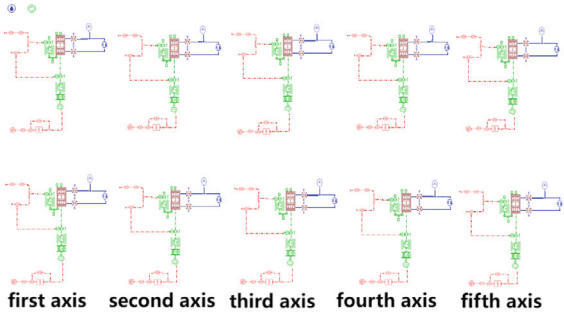
Fig. 3 Schematic diagrams of the three hydro pneumatic suspension system modes



a) Simulation model of front and rear interconnection system

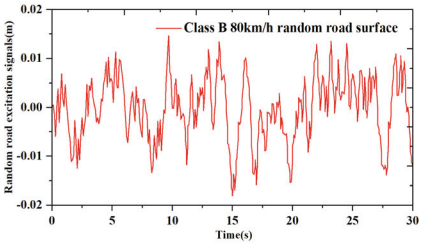


b) Simulation model of front and rear to side cross-linked system

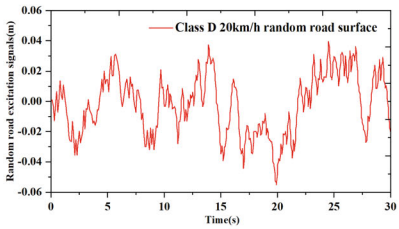


c) Simulation model of completely independent system

Fig. 4 Simulation model of three hydro pneumatic suspension systems



a) Speed 80km/h, pavement class B incentive



b) Speed 20km/h, pavement class D incentive

Fig. 5 Two random pavement excitation signals

Table 1 Main parameter settings

Main parameters	Parameter value	Unit
Single Suspension Mass	15,000	kg
Cylinder bore	125	mm
Piston rod diameter	100	mm
Unsprung mass	200	kg
Accumulator initial filling pressure	4	MPa
Accumulator filling volume	2.8/3.2	L
Tire stiffness	1×10^6	N/m
Tire Damping	10,000	N/(m/s)

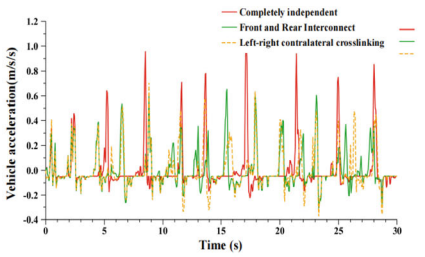
speed of 20 km/h and pavement class D. These two road spectra indicate that the all-terrain crane travels on the highway and the rugged mountain pavement respectively. The two random road excitation signals are shown in Fig. 5.

The main parameters of the AMESim model are shown in Table 1.

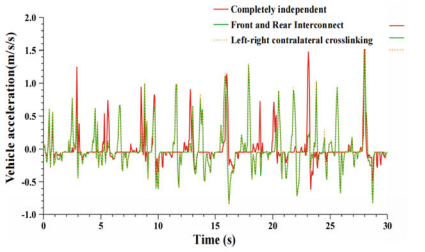
4 Comparison of Smoothness Results

Through the simulation of the three models, the comparisons of the body acceleration, the average dynamic deflection of the whole vehicle suspension, and the average dynamic load of the whole vehicle tires of the three hydro-pneumatic suspension systems under two different road excitations are obtained [9] as shown in Fig. 6 and Table 2.

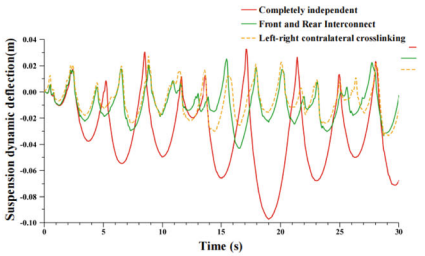
As can be seen from Fig. 6 and Table 2, among the three modes, the completely independent mode has the smallest RMS value of body acceleration, followed by the front-rear interconnection in the middle position, while the front and rear to side cross-linked mode has the largest value. Regarding the RMS value of suspension



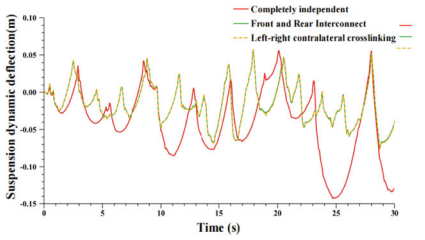
a) Comparison of body acceleration at highway speeds



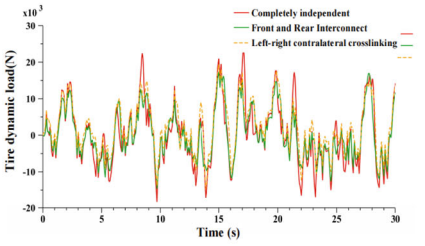
b) Comparison of body acceleration under mountain roads



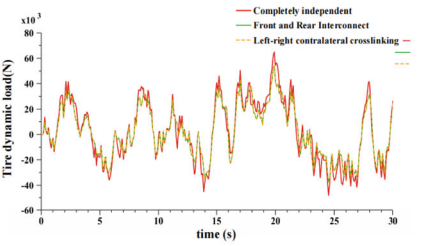
c) Comparison of suspension dynamic deflection at highway speeds



d) Comparison of the dynamic deflection of the suspension under the mountain road



e) Comparison of tire dynamic loads at highway speeds



f) Comparison of tire dynamic loads at highway speeds

Fig. 6 Comparison of smoothness indexes under two road surface spectra

dynamic deflection, the completely independent mode is the largest, with the front-rear interconnection being in the middle, and the front and rear to side cross-linked mode being the smallest. For the RMS value of tire dynamic load, the completely independent mode is the largest, the front-rear interconnection has the smallest value, and the front and rear to side cross-linked mode is in the middle. The reason for this is that the front and rear interconnections possess strong anti-pitch ability, and the front and rear to side cross-linked mode has strong anti-roll ability [10]. Thus, the front and rear interconnections have an advantage in body acceleration and tire dynamic load. The front and rear to side cross-linked mode has an advantage in suspension

Table 2 Comparison of RMS values of vehicle smoothness evaluation indicators

Road spectrum	Evaluation indicators	Completely independent	Front and rear interconnection	Front and rear to side cross-linked
Highway road spectrum	Body acceleration RMS (m/s^2)	0. 138	0. 146	0. 160
	Suspension dynamic deflection RMS (m)	0. 0343	0. 0164	0. 0145
	Tire dynamic load RMS(N)	8254	6766	7160
Mountain road spectrum	Body acceleration RMS (m/s^2)	0. 258	0. 298	0. 314
	Suspension dynamic deflection RMS (m)	0. 0577	0. 0313	0. 0309
	Tire dynamic load RMS (N)	23,900	20,214	20,227

dynamic deflection. The completely independent mode has an advantage in body acceleration, but there is a significant gap between it and the other two modes in terms of suspension dynamic deflection and tire dynamic load.

5 Conclusion

A classical two-degree-of-freedom suspension mathematical model is established to obtain suspension dynamic deflection formula, and tire dynamic load formula. AMESim software is utilized to create two interconnected modes of the whole vehicle suspension system model. Through inputting two different random road spectra, the three indexes for evaluating smoothness are compared and analyzed. The results reveal that among the three interconnected suspension structures, although the front and rear interconnected suspension system has a larger RMS value for the dynamic travel of the suspension, it has smaller RMS values for the vertical acceleration of the body and the tire dynamic load. Consequently, it holds a more distinct advantage in terms of smoothness and comfort.

The shortcoming of the paper is that it does not compare the effects in terms of sway, and in the future it will compare the effects of interconnected modes on sway.

Acknowledgements This research was supported by Taiyuan key core technology research unveiled list of marshal projects.

References

1. Yuan J, Wang C, Zhang L (2024) Design and simulation study of hydro pneumatic suspension system based on AMESim. Agric Equip Veh Eng
2. Yu Y, Chen S (2022) Simulation analysis of hydraulic interconnected suspension performance of different structures based on AMESim. Automot Pract Technol
3. Ma C (2017) Research on hydro pneumatic suspension mode of articulated dump truck. Southwest Jiaotong Univ
4. Lu Y et al (2024) An interconnected suspension with adjustable roll and pitch stiffness to enhance anti-roll and anti-dive/squat characteristics. In: Proceedings of the institution of mechanical engineers part d-journal of automobile engineering
5. Jayaraman T, Thangaraj M (2023) Standalone and interconnected analysis of an independent accumulator pressure compressibility hydro-pneumatic suspension for the four-axleheavy truck. In: Actuators, vol 12
6. Shi Y-X, Liu T-H, Cao C-Z et al (2022) Smoothness analysis and optimization design of hydrocarbon suspension for mining wide-body vehicle. Mach Tools Hydraul
7. Tian W, Du J, Xu X (2020) Effect of seven-axle hydro pneumatic suspension coupling scheme on body attitude. China Mech Eng
8. Jin H, Lei X, Zhou X et al (2024) Smoothness analysis of interconnected hydrocarbon suspension system for mining wide-body vehicles. Electromechanical Eng
9. Liu Q, Feng H-T, Liu S et al (2023) Smoothness analysis of mining dump truck suspension system based on AMESim. J Southwest Jiaotong Univ
10. Wenpeng T, Xia F, Xu X (2022) Vibration optimization and control of vehicles with seven-axle hybrid-connected oil-air suspension. Mech Strength

Open Access This chapter is licensed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License (http://creativecommons.org/licenses/by-nc-nd/4.0/), which permits any noncommercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if you modified the licensed material. You do not have permission under this license to share adapted material derived from this chapter or parts of it.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.



Research on the Smoothness of the Suspension System of a Five Axis All Ground Crane



Rui Guo, Yuang He, Shuai Song, Rumin Teng, and Xifeng Zhang

Abstract A new independent suspension system was designed to meet the smoothness requirements of all terrain cranes, and smoothness research was conducted. Based on the GB/T 7031–2005 standard, a road surface power spectral density was used to describe random road surface characteristics, and a road surface simulation model was established. Simulation studies were conducted on multiple working conditions at different vehicle speeds and road surfaces. The main focus was on optimizing the smoothness of the acceleration response at the three points of the vehicle center of mass, the bottom of the driver's cab, and the same tonnage rigid bridge. In addition, a comparative analysis of suspension dynamic deflection and tire dynamic displacement was also provided. The results show that under the conditions of 10 km/h E-Class road surface, 20 km/h D-Class road surface, 40 km/h C-Class road surface, and 80 km/h B-Class road surface, the independent suspension system is superior to non independent suspension systems of the same tonnage in terms of RMS values at the center of gravity, cab bottom, and cab, with an optimization rate of over 12%. Among them, the optimization rate of cab bottom and cab under harsh ground conditions reaches 30%, which can more effectively reduce driver fatigue driving. The optimization rate of suspension dynamic deflection and tire dynamic displacement also reaches over 10%, which can more efficiently reduce the probability of oil and gas suspension hitting limit blocks and improve driving safety. This study provides a certain theoretical basis for the design and smoothness research of the suspension system of all terrain cranes.

R. Guo (✉)
National Key Laboratory of Key Technologies of Lifting Appliances, Qinhuangdao, Hebei, China
e-mail: guorui@ysu.edu.cn

R. Guo · Y. He
School of Mechanical Engineering, Yanshan University, Qinhuangdao, Hebei, China

R. Teng
School of Mechanical Engineering, Dalian University of Technology, Dalian, Liaoning, China

S. Song · X. Zhang
Taiyuan Heavy Industry Co. Ltd., Taiyuan, Shanxi, China

© The Author(s) 2026

H. Saman et al. (eds.), *The 9th International Conference on Advances in Construction Machinery and Vehicle Engineering*, Lecture Notes in Mechanical Engineering,
https://doi.org/10.1007/978-981-96-5527-4_19

Keywords All terrain crane · Hydraulic suspension · Smoothness · Simulation analysis

1 Introduction

Because the all-ground crane often works in the open-air non-highway environment, its driving condition is complex and changeable, the suspension bears a large load and the load sudden change frequently, which puts forward high requirements for the ride comfort of the crane. As an indispensable part of the vehicle structure, the function of the suspension system is not only limited to the connection between the frame and the axle, but also lies in the effective absorption and attenuation of the shock and vibration in the process of vehicle driving. It can effectively improve the driver's ride experience, and then significantly affect the ride comfort of the vehicle. Therefore, the design and characteristics of the suspension system determine the overall dynamic performance and comfort of the vehicle to a great extent.

Scholars at home and abroad have done a lot of research on hydro-pneumatic suspension, Zheng [1] constructed a complete three-dimensional multi-body dynamic model to simulate the dynamic behavior of wheeled tractors and their supporting equipment systems. The model verification is based on the evaluation of the vertical acceleration root mean square (RMS) of the driver's seat, cab, chassis and equipment under different speed and road conditions. The effectiveness of the model is verified by comparing the simulation analysis with the experimental data. Wang [2] and others studied the hydro-pneumatic suspension of mining dump trucks. Through the AMESim model, compared with the single and double-chamber suspension, it was found that when the double-chamber suspension passed the 120 mm barrier, the peak vibration acceleration decreased by 26.10%. Liu [3] in order to solve the problem of cab jitter of domestic truck crane at specific speed, a whole vehicle power model is established for simulation test, which successfully improves the ride comfort of the vehicle, especially at 40 km/h speed. Ma [4] and others established the simulation model of QAY220 all-ground crane including driving room, tire and road surface to improve its ride comfort by 47.71%. In order to solve the ride comfort problem of a heavy dump truck.

In this paper, according to the ride comfort requirements of all-ground crane, a new independent suspension system is designed, and the ride comfort is studied. Under a variety of different speeds and road conditions, the ride comfort of independent suspension compared with the same tonnage rigid bridge in five different indexes: body centroid, cab bottom, cab, suspension deflection and tire displacement, in order to improve the ride comfort of the suspension system under various working conditions.

2 Pavement

In the process of vehicle driving, road excitation is the main external source of vehicle vibration. The construction of road excitation model is mainly divided into random road model and pulse road model. Considering the road conditions of the all-ground crane, this paper chooses the random road model.

According to the standard GB/T7031-2005, the road power spectral density is used to describe the random road characteristics.

$$G_q(n) = G_q(n_0)\left(\frac{n}{n_0}\right)^{-p} \tag{1}$$

In the formula, p-is the frequency index; n-spatial frequency n0-reference spatial frequency,and selects n0 = 0.1 m-1; Gq(n0)-the power spectral density of displacement under the reference spatial frequency n0, which is called road roughness coefficient.

According to different pavement grades, the root mean square value of Gq(n0) has different values, and the pavement is divided into A ~ H eight grades [5], Table 1 shows A ~ F.

When the frequency index p = 2 is brought into the above formula and the spatial power density is converted into time frequency power density, it can be obtained.

$$G_q(n) = (2\pi n)^2 G_q(n) = (2\pi n_0)^2 G_q(n_0) \tag{2}$$

$$G_q(f) = (2\pi n)^2 G_q(f) = (2\pi n_0)^2 G_q(n_0)u \tag{3}$$

Table 1 Schematic diagram of classification standards for road surface roughness

Pavement grade	$G_q(n_0)(n_0 = 0.1m^{-1})(10^{-6}m^3)$			Pavement description
	Lower limit	Geometric mean	Upper limit	
A	8	16	32	Highway, national highway, provincial highway and other paved roads
B	32	64	128	Flat dirt roads, gravel and other paved roads
C	128	256	512	Lots of dirt roads, gravel, continuous winding roads
D	512	1024	2048	There are many uneven and convex stones of about 10 cm
E	2048	4096	8192	20 cm of snow and long dirt roads
F	8192	16,384	32,768	There are 30 cm high rocks and 50 cm hard bottom wading

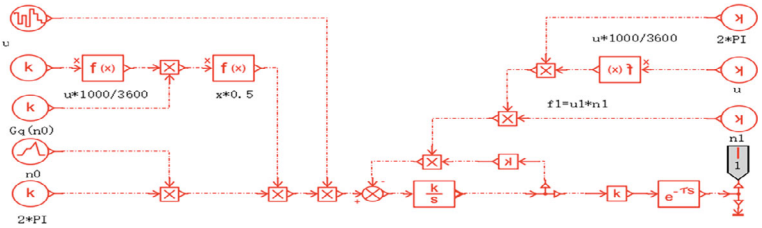


Fig. 1 Mathematical modeling of road surface excitation

It is known that the road roughness coefficient $G_q(n_0)$ is related to the speed u, when the road roughness coefficient and vehicle speed are constant, it is a constant in the whole frequency norm, so the "white noise model" can be used to model the random road surface. As shown in Fig. 1.

The excitation input can be obtained according to different road grades, as shown in Figs. 2, 3, 4 and 5.

Due to the low speed and long five-axle crane, delay time excitation is adopted from one axis to five axes. The delay parameter table is shown in Table 2. With the increase of vehicle speed, the delay time decreases, but the 80 km/h delay time is ignored when the vehicle speed is increased.

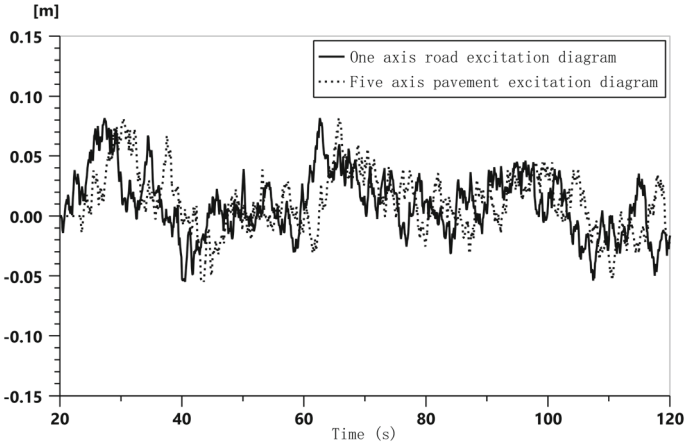


Fig. 2 Comparison diagram of E-level road excitation delay at aspeed of 10 km/h

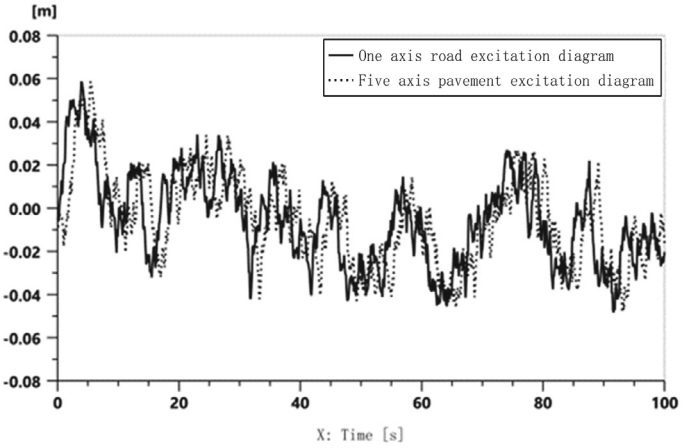


Fig. 3 Comparison diagram of D-level road excitation delay at a speed of 20 km/h

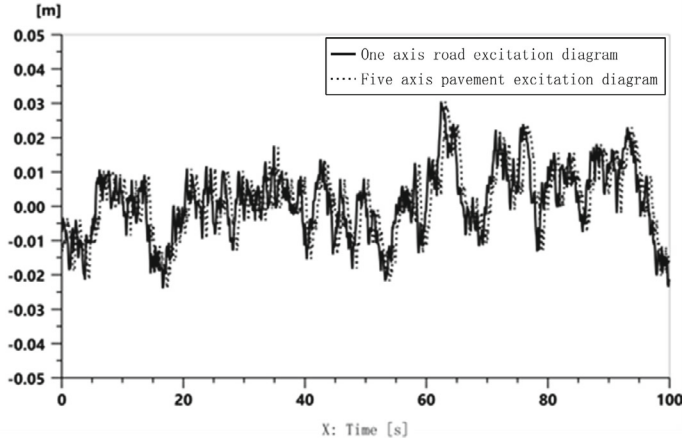


Fig. 4 Comparison diagram of C-level road excitation delay at a speed of 40 km/h

3 Vehicle Model

Aiming at the modeling and simulation analysis of the ride comfort of the all-ground crane under the condition of different speeds and different road grades, the whole vehicle model of independent suspension with five axes is established as shown in Fig. 6.

Aiming at the ride comfort acceleration simulation, the whole vehicle model is optimized to get a single-point two-axis model as shown in Fig. 7. The model is mainly composed of frame, suspension cylinder, accumulator, one-way throttle valve, relief valve, pipeline and tire. As the main bearing component, the suspension cylinder

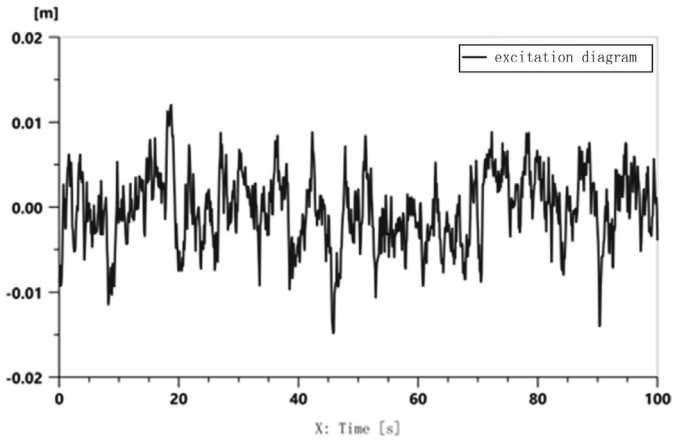


Fig. 5 Comparison diagram of B-level road excitation delay at a speed of 80 km/h

Table 2 Delay parameters for various operating conditions

	Five-axis (s)	Four-axis (s)	Three-axis (s)	Two-axis (s)	One axis
10 km/h E-class pavement	3.386	2.592	1.548	0.954	0
20 km/h Class D pavement	1.693	1.296	0.7d74	0.477	0
40 km/h C-class pavement	0.846	0.648	0.387	0.238	0

bears the weight of the car body; Effectively cushioning the excitation caused by the road surface. The pipeline is connected with a damping element and an one-way valve, and the oil can reduce the vibration through the reciprocating motion in the damping element. Through the synergistic action of the key components, the purpose of reducing vehicle vibration is finally achieved.

4 Comparative Analysis of Ride Comfort Index

The ride comfort of a vehicle can be analyzed by the “road-vehicle-human” system block diagram shown in Fig. 8 [6, 7]. The ride comfort study mainly evaluates the comfort feedback of the suspension system to evaluate the performance of the damping system and to explore the damping performance of the suspension performance. The lower the weighted acceleration of the car body excited by the uneven road, the smaller the impact on people. The higher the comfort of the human body. In the crane, not only the acceleration of the whole vehicle, but also the dynamic deflection of the suspension and the dynamic displacement of the tire relative to the

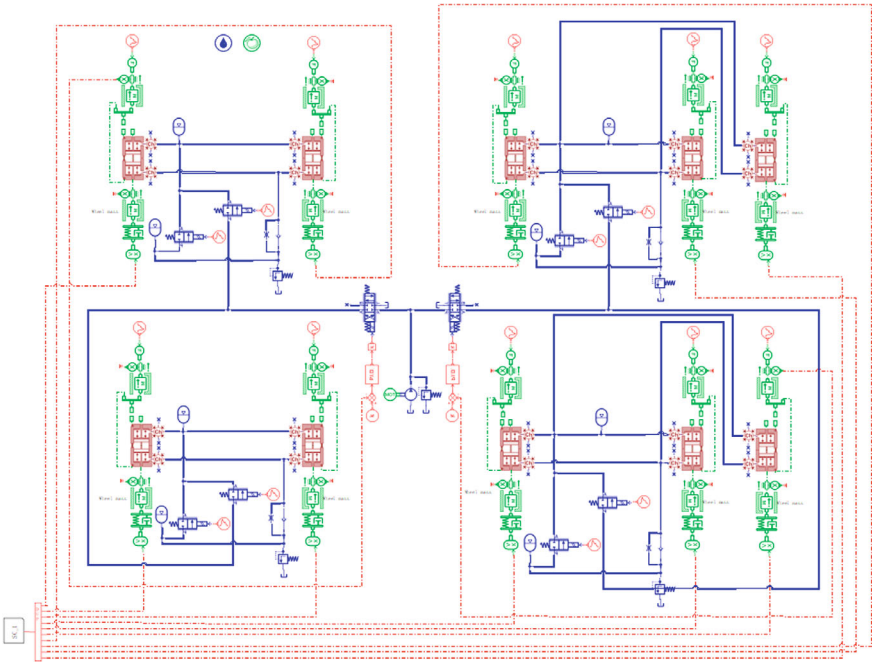
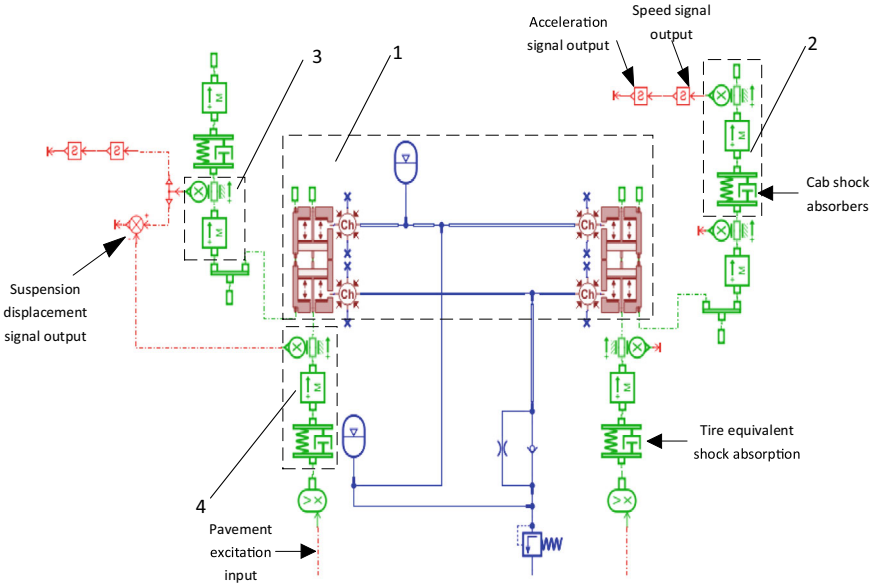


Fig. 6 Five axis independent suspension vehicle modeling diagram



1-hydro-pneumatic suspension 2-cab mass 3-body mass 4-tyres

Fig. 7 Single point independent suspension modeling diagram

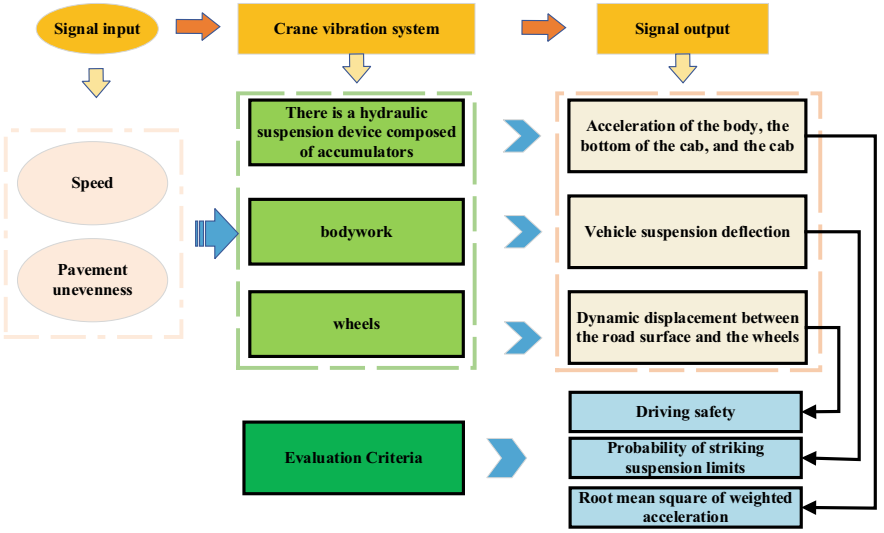


Fig. 8 Road vehicle human system block diagram

uneven road are considered, which affect the "driving safety" and the probability of hitting the suspension limit respectively.

According to the standard GB/T4970-2009 vehicle ride comfort test method, the root mean square of total weighted acceleration is usually regarded as a quantitative method to measure ride comfort [7]. Reducing the vibration amplitude and frequency of acceleration is beneficial to improve the ride comfort and human comfort of the whole vehicle.

In this paper, four different speeds and road grades are used to simulate and analyze the whole ground surface AMESim, and the body centroid, cab bottom, cab, three-point acceleration response, suspension deflection and tire displacement are compared with the existing rigid bridges of the same tonnage.

Through simulation and mathematical calculations, the comparison and optimization of non independent suspension and independent suspension with the same tonnage in terms of vehicle center of mass acceleration, cab acceleration, cab bottom acceleration, suspension dynamic deflection, and tire dynamic displacement are shown in Figs. 9, 10, 11, 12, and 13.

The comparison diagram of body centroid acceleration root mean square shows that the optimization effect of C class pavement of 40 km/h is the best, the root mean square value of D class pavement of 20 km/h and C class pavement of 40 km/h is lower, and the root mean square of acceleration is reduced to a certain extent under other working conditions. This shows that the independent suspension system can absorb the road roughness more effectively.

The root mean square diagram of the acceleration at the bottom of the cab shows that the acceleration of the independent suspension system is lower than that of the non-independent suspension system of the same tonnage, and the data show that the

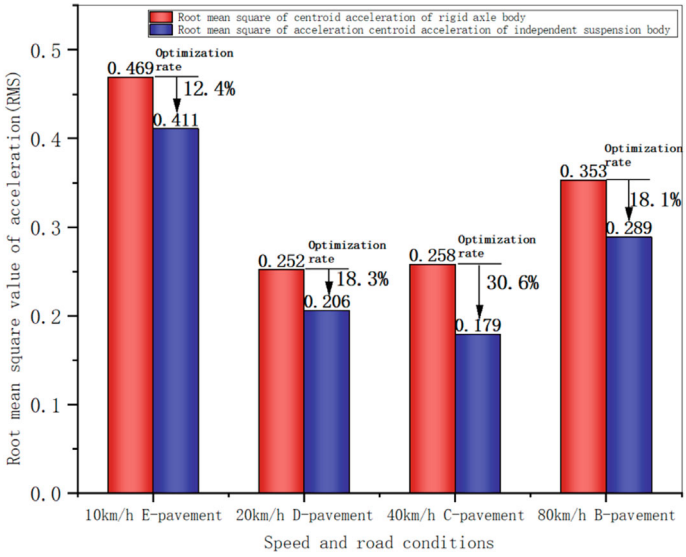


Fig. 9 Root mean square comparison of vehicle acceleration

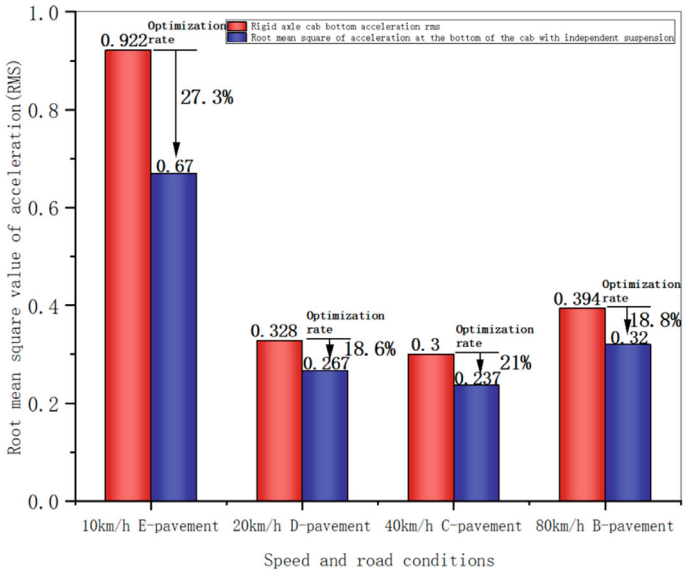


Fig. 10 Comparison of root mean square acceleration at the bottom of the cab

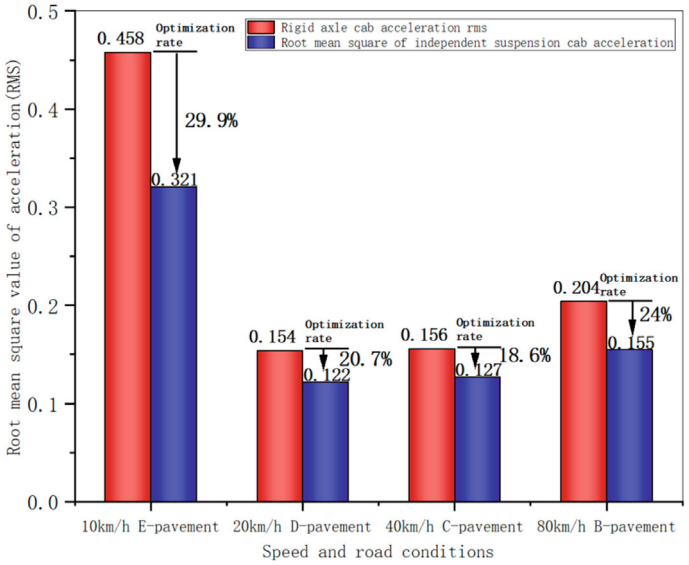


Fig. 11 Comparison of root mean square acceleration in the driver's cab

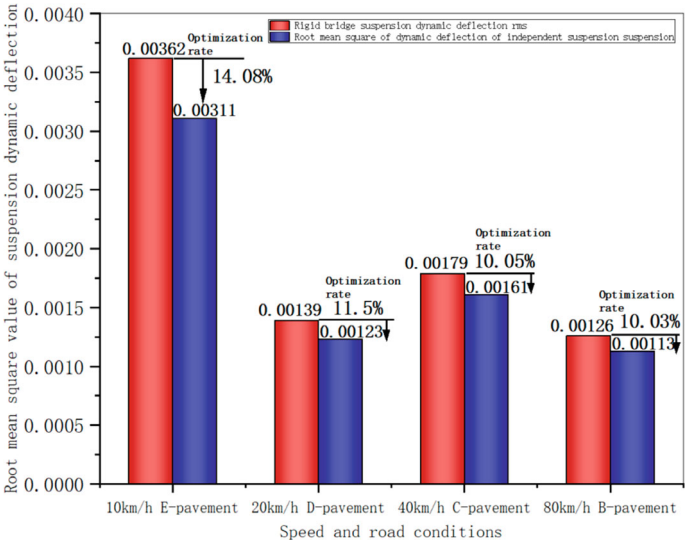


Fig. 12 Root mean square comparison of suspension dynamic deflection

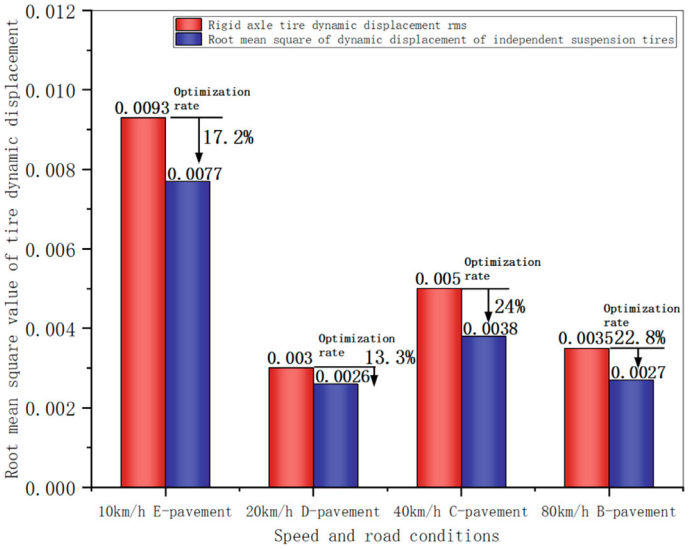


Fig. 13 Root mean square comparison of tire dynamic displacement

E-class pavement optimization effect of 10 km/h is the best. The root mean square (RMS) of D grade pavement of 20 km/h, C grade pavement of 40 km/h and B grade pavement of 80 km/h are lower. This shows that the independent suspension system can absorb the road roughness more effectively under the bad road noodles, so as to provide better ride comfort.

The acceleration comparison diagram of the cab reveals the damping effect of the independent suspension system in the cab position. The comparative images show that the E-grade road surface optimization rate of 10 km/h is the best, and it also shows that the damping effect of the cab is the best under harsh conditions, which helps to reduce the driver's fatigue during driving and improve the safety of operation.

From the root mean square comparison diagram of the dynamic deflection of the suspension, it can be seen that the independent suspension is lower than the rigid bridge in varying degrees, among which the E grade road surface of 10 km/h is more obvious, which can effectively reduce the probability of the cylinder piston rod hitting the working cavity and increase the service life of the parts. in addition, the dynamic disturbance of the suspension system should be less than or equal to 1/3 of the allowable suspension dynamic deflection, in order to ensure the smoothness performance. The moving stroke of the suspension needs to be limited.

From the root-mean-square comparison diagram of tire dynamic displacement, we can see that the optimization rate of C-class pavement in 40 km/h is the most obvious, and it decreases in different degrees under other working conditions. Among them, the root mean square of dynamic displacement of E-grade road tire in 10 km/h is larger, which indicates that its road grade has a great influence on tire dynamic displacement. The greater the dynamic displacement of the tire, the greater the

dynamic load of the tire. The dynamic load of the tire reflects the reaction force of the uneven road surface on the tire, the excessive vibration shock leads to the vertical jump of the wheel, the greater the dynamic load of the tire, the longer the tire is off the ground, and it is difficult for the driver to operate.

Perform Fourier transform on the acceleration vibration of the entire vehicle's cab to obtain the frequency domain diagram of the cab vibration, as shown in Figs. 14, 15, 16, and 17.

For ergonomics, in the vertical direction, the human body is sensitive to vibration frequencies of 4-8 Hz. As shown in Figs. 14, 15, 16, and 17, the vibration frequencies under different working conditions are mostly concentrated between 1-10 Hz, with low peak values. The frequency at the highest peak is concentrated between 1.5–2.5 Hz, which is good for human health. Table 3 shows the peak vibration frequencies

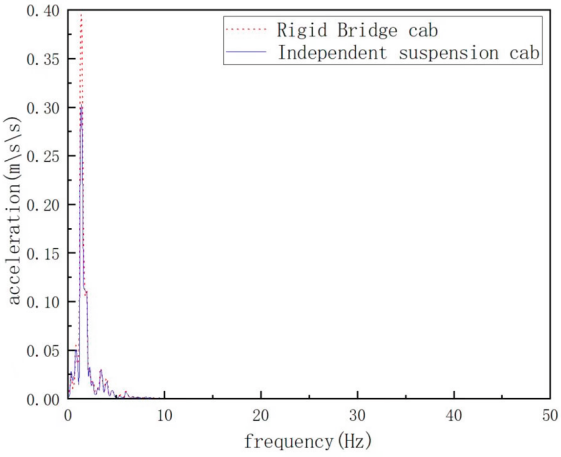


Fig. 14 Frequency domain diagram of 10 km/h E-class pavement vibration

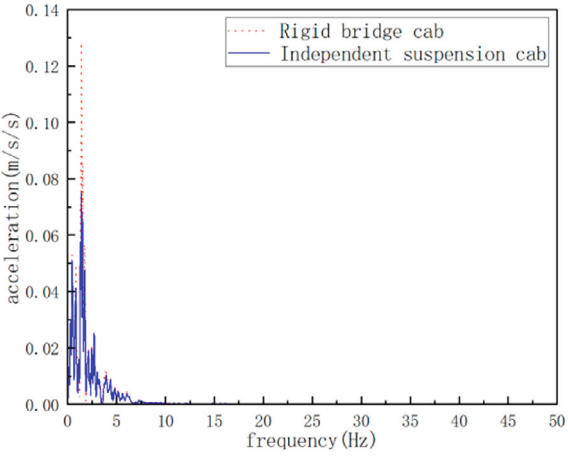


Fig. 15 Frequency domain diagram of 20 km/h D-class pavement vibration

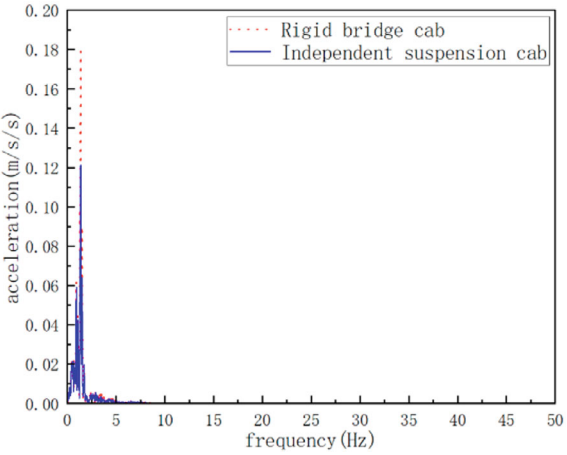


Fig. 16 Frequency domain diagram of 40 km/h C-class pavement vibration

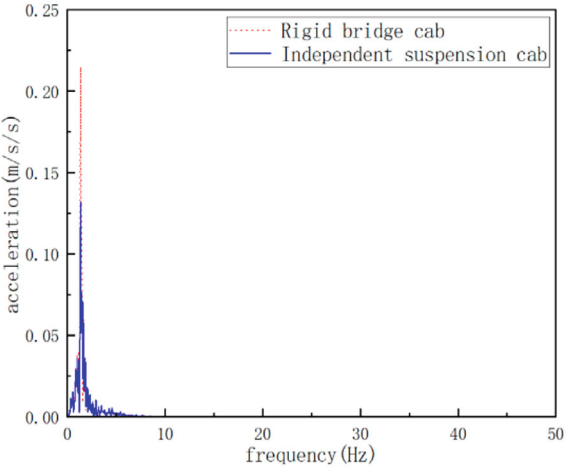


Fig. 17 Frequency domain diagram of 80 km/h B-class pavement vibration

under various operating conditions. It can be seen from Table 3 that the optimization of independent suspension compared to non independent suspension at the vibration peak reaches more than 20%, which is a significant improvement.

Table 3 Peak vibration frequency table of the cab under different operating conditions

	10 km/h E-class road surface,	20 km/h D-class road surface,	40 km/h C-class road surface,	80 km/h B-class road surface
Independent suspension frequency peak (m/s/s)	0.283	0.075	0.123	0.133
Peak frequency of non independent suspension of the same tonnage (m/s/s)	0.362	0.128	0.182	0.215
Peak optimization rate	21.8%	41.4%	32.5%	38.13%

5 Conclusion

We conducted in-depth simulation analysis and optimization of the oil air suspension system for all ground cranes. Using road power spectral density to describe random road surfaces, a road simulation model is established. By constructing a simulation model of an independent suspension system, the acceleration, suspension deflection, tire displacement, and cab vibration frequency are compared and analyzed.

The results showed that independent suspension had significant optimization in the vehicle center of mass, bottom of the cab, and root mean square (RMS) value of the cab under four operating conditions, with optimization rates of over 12%. The optimization rates of the bottom of the cab and the cab under harsh ground conditions reached over 30%, indicating that independent suspension can more effectively absorb road roughness under harsh road conditions. The suspension displacement and tire displacement have also been reduced by more than 10%. In terms of cabin vibration frequency, the vibration peak has been optimized by more than 20%, and the frequency at the highest peak is concentrated in 1.5–2.5 Hz, which is good for human health. This provides a theoretical basis for the study of the smoothness of the ultimate multi axle all ground suspension system.

Acknowledgements Supported by the National Natural Science Foundation of China (Grant No. 52075469, 12173054), This thesis is also funded by the Key Core Technology Research Project of Taiyuan City.

References

1. Zheng E, Zhong X, Zhu R et al (2019) Investigation into the vibration characteristics of agricultural wheeled tractor-implement system with hydro-pneumatic suspension on the front axle. Biosyst Eng 186:14–33

2. Wang G, Liu X, Du T et al (2022) Characteristic analysis of two-stage pressure hydro-pneumatic suspension of mine dump truck. Hydraul Pneum 46(09):92–98
3. Liu Y (2022) Analysis and optimization of cab jitter of a truck crane at specific speed. Hunan University
4. Ma X, Fu L, Sun S (2015) Simulation analysis and optimization of ride comfort of all-road crane. Lift Transp Mach 01:32–35
5. Jie D, Chen X, Yang H et al (2024) Study on ride comfort of hydro-pneumatic suspension based on RBF neural network. Equip Environ Eng 721(03):113120
6. Yu Z (2009) Automobile theory. Machinery Industry Press, Beijing, vol 203, pp 246
7. Xia X, He D, Chen D et al (2018) Test and analysis of ride comfort of car copilot seat. Pract Automob Technol 20:55–57

Open Access This chapter is licensed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License (http://creativecommons.org/licenses/by-nc-nd/4.0/), which permits any noncommercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if you modified the licensed material. You do not have permission under this license to share adapted material derived from this chapter or parts of it.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.



Optimization of 3D Printing Process for Hydrogel-Based Thermal Management Patches in Engineering Machinery



Bowen Li, Zhen Wang, Chuanzhen Huang, Longhua Xu, Shuiquan Huang, Meina Qu, Zhengkai Xu, Dijia Zhang, Baosu Guo, and Chunhui Ji

Abstract In mechanical engineering applications, components such as electric motors, gear reducers, and pumps generate substantial heat during operation. Insufficient thermal management can result in equipment overheating, which adversely affects both performance and lifespan. Recently, hydrogels have emerged as an ideal material for thermal management due to its exceptional thermal conductivity and lightweight properties. This study employs computational fluid dynamics to select a nozzle with a diameter of 0.6 mm and a length of 13 mm for 3D printing. This nozzle demonstrates optimal performance in balancing wall shear stress (165.4 Pa) and printing efficiency. Subsequent optimization of printing parameters revealed that the best printing results are achieved under conditions of a pressure of 15 kPa, a printing height of 0.48 mm, and a printing speed of 180 mm/min. This research provides a novel solution for the efficient thermal management of mechanical engineering components and presents an optimized example of 3D printing technology for hydrogel-based thermal management patches.

B. Li · Z. Wang (✉) · C. Huang (✉) · L. Xu · S. Huang · M. Qu · Z. Xu · D. Zhang · B. Guo
School of Mechanical Engineering, Yanshan University, Qinhuangdao, China
e-mail: wangzhen@ysu.edu.cn

C. Huang
e-mail: wangzhen@ysu.edu.cn

Z. Wang · C. Huang
State Key Laboratory of Crane Technology, Yanshan University, Qinhuangdao, China

L. Xu
Guangdong Key Laboratory of Minimally Invasive Surgical Instruments and Manufacturing Technology, School of Electromechanically Engineering, Guangdong University of Technology, Guangzhou, China

Z. Wang
College of Mechanical and Electronic Engineering, Shandong University of Science and Technology, Qingdao, China

C. Ji
School of Mechanical Engineering, Tianjin University, Tianjin, China

© The Author(s) 2026

H. Saman et al. (eds.), *The 9th International Conference on Advances in Construction Machinery and Vehicle Engineering*, Lecture Notes in Mechanical Engineering,
https://doi.org/10.1007/978-981-96-5527-4_20

Keywords Components of engineering machinery · Hydrogel-based thermal management patches · Computational fluid dynamics · Optimization of 3D printing process

1 Introduction

Key components in engineering machinery, such as motors, reducers, and pumps, generate significant amounts of heat during operation. This heat can adversely affect the lifespan of the equipment [1, 2]. Effective heat dissipation from these components is essential for ensuring stable operation. It also extends the lifespan of the equipment and enhances overall performance [3, 4].

The most commonly used heat dissipation methods in engineering machinery include air cooling and liquid cooling. However, these traditional techniques have certain limitations. Air cooling technology has a simple structure, low cost, high reliability, ease of integration, and convenient maintenance and management. Therefore, it has been widely adopted for heat dissipation in engineering machinery [5]. However, the heat transfer medium in air-cooled technology is air. Consequently, its cooling effectiveness is influenced by various environmental factors, including climate and geographical location. Additionally, the low specific heat capacity of air leads to a slower temperature control response time. Compared with air-cooled technology, liquid cooling technology offers significant advantages in heat dissipation capacity and response speed. The operation of liquid cooling involves circulating a cooling liquid through the system, where it absorbs heat from the equipment. Then, the liquid is cooled by a radiator to meet the thermal management requirements of engineering machinery components [6]. Although liquid cooling is effective at dissipating heat, it requires an additional cooling system. This increases the complexity and cost of the equipment.

With the advancement of new materials and technologies, hydrogels have emerged as a prominent research focus in the field of heat dissipation materials. This is due to their exceptional thermal conductivity and lightweight properties. Liu et al. investigated the vapor pressure regulation capabilities of various hydrogels at a consistent water content ratio [7]. The results indicated that a higher Young's modulus correlates with a stronger ability to regulate vapor pressure. Pu et al. developed and investigated the heat dissipation technology using passive hydrogel films [8]. These films are designed to cover a heat source, allowing moisture within the film to evaporate rapidly, which reduces the temperature of the heat source. When the heat source ceases operation, the hydrogel film absorbs water vapor from the surrounding environment. This replenishes its moisture content and enables a recycling process. It is evident that the application of hydrogel in the field of heat dissipation not only has a robust theoretical foundation but also demonstrates significant practical advantages.

In the field of hydrogel product manufacturing, 3D printing technology is undoubtedly the superior choice. This technology not only facilitates the precise fabrication of complex geometries but also allows for highly customized production tailored to

specific requirements. The application of 3D printing in hydrogels is predominantly observed in biomedical fields, such as in vitro organoid manufacturing [9] and the production of hydrogel wearable sensors [10]. Utilizing 3D printing technology to create hydrogel-based thermal management patches enables the design and optimization of intricate geometries. It also enhances material utilization and production efficiency. This makes it an ideal solution for improving the heat dissipation performance of construction machinery components.

The purpose of this paper is to develop hydrogel patches for the heat dissipation of engineering machinery components by using 3D printing technology. Additionally, it aims to optimize the printing parameters through fluid simulation. This approach enables the precise design and fabrication of hydrogel-based thermal management patches, thereby enhancing heat dissipation efficiency. The significance of this research lies in its exploration of an innovative heat dissipation method. This method not only improves the thermal performance of engineering machinery components but also minimizes material waste, reduces costs, and enhances production efficiency. Ultimately, it increases the reliability and service life of construction machinery through optimized printing parameters.

2 Materials and Methods

2.1 Synthesis and Preparation of Hydrogels

The gelatin powder was dissolved in an acetic acid solution with a mass fraction of 1 wt% at 60 °C. It was mixed and stirred at 1000 rpm for 30 min. After homogenization, chitosan powder was added to the well-mixed gelatin precursor. The mixture was then homogenized for an additional 30 min at the same temperature and rotation speed to obtain the ink material for the hydrogel-based thermal management patches. The mass fraction of gelatin was 10 wt%, and the mass fraction of chitosan was 1.5 wt%.

After printing, 1-ethyl-3-(dimethylaminopropyl)carbodiimide (EDC) and N-hydroxy succinimide (NHS) were used as crosslinkers to enhance the thermal stability of the three-component hydrogel patch material. The mass fractions of EDC and NHS were 2 wt% and 0.2 wt%, respectively.

2.2 Optimization of 3D Printing Process for Hydrogel-Based Thermal Management Patches

The prepared hydrogel precursor solution was transferred to a barrel and cooled to 4 °C. It was then moved to an extruded bio-3D printer (EFL-BP6602P) set to a printing temperature of 27 °C. The effectiveness of the printing process was assessed by measuring the diameter of the extruded filament.

Table 1 The values of the printing parameters utilized in the optimization experiment for printing parameters

Sample number	Printing pressure (kPa)	Printing speed (mm/min)	Printing height (mm)
1	15	120	0.48
2	15	180	0.48
3	15	240	0.48
4	15	300	0.48
5	15	180	0.12
6	15	180	0.3
7	15	180	0.6

The values of the printing parameters utilized in the optimization experiment for printing parameters is illustrated in Table 1.

2.3 Fundamentals of Fluid Mechanics and Simulation Analysis of Extrusion Runners

The following assumptions must be considered regarding the flow state before conducting fluid dynamics simulations: (1) the material is incompressible; (2) there is no wall slippage during the printing process; (3) the flow of the material in the extrusion flow channel is fully developed laminar flow; (4) the material is highly viscous, and its inertial effects are significantly smaller than those of the viscous resistance; (5) the entrance effects are negligible; and (6) during the extrusion process, the heat generated by viscous resistance does not alter the temperature of the material.

Hydrogel inks are typically characterized as power-law fluids, with the specific form represented in Eq. (1),

$$\tau = K\gamma^n \tag{1}$$

the consistency index, and n signifies the power-law index. To facilitate numerical fitting, the logarithm of Eq. (1) is taken,

$$\log(\tau) = \log(K) + \log(\gamma) \tag{2}$$

The radial velocity distribution of a power-law fluid in a circular pipe is as follows,

$$\mu(r) = (\Delta P/2LK)^{(1/n)} \cdot ((n+1)/n) \cdot R^{((1+n)/n)} \cdot \left[1 - (r/R)^{((1+n)/n)}\right] \tag{3}$$

where, $\triangle$P is the pressure drop, L is the nozzle length, and R is the nozzle radius.

The expression for wall shear stress in a power-law fluid is,

Table 2 Nozzle specifications used in fluid simulation analysis

Sample number	Diameter of nozzle (mm)	Length of nozzle (mm)
1	1.2	25
2	1.2	13
3	1.2	6.5
4	0.6	25
5	0.6	13
6	0.6	6.5
7	0.16	25
8	0.16	13
9	0.16	6.5

$$\tau_w = (R/2)/\Delta P \tag{4}$$

where, τ_w is wall shear stress.

The specifications of the nozzle for fluid simulation is shown in Table 2.

2.4 Rheological Properties for Hydrogel Materials

The materials were subjected to rheological tests using a rotational rheometer (HAAKE MARS40). These tests were conducted to determine the conditions for their 3D printing process and to establish the fluid dynamics underlying the simulation.

During the temperature scanning test, the temperature was varied from 20 °C to 50 °C at a heating rate of 5 °C/min. The oscillation test was conducted under conditions of fixed stress at 1 Pa and a frequency of 1 Hz. Rotational shear viscosity testing was performed at 27 °C, using shear rates ranging from 0.1 s^{-1} to 500 s^{-1}.

In the three-stage thixotropy test, the first stage employed a shear rate of 0.1 s^{-1} to simulate the initial resting state of the material without shearing at 25 °C. The second stage applied a shear rate of 50 s^{-1} to simulate the shear experienced during nozzle extrusion at 25 °C. The third stage reverted to a shear rate of 0.1 s^{-1} to simulate the static state after printing on the platform at 4 °C. Prior to each rheological test, the materials were maintained at a warm temperature for 10 min.

3 Results and Discussions

3.1 Result of Rheological Test

The effect of temperature on the storage modulus and loss modulus of the hydrogel-based thermal management patches is shown in Fig. 1a. In the figure, temperature is represented on the horizontal axis, while the storage modulus (G′) and loss modulus (G″) are plotted on the vertical axes.

The data indicate that as the temperature increases, both the storage modulus and loss modulus of the hydrogel-based thermal management patches gradually decrease. At a certain critical temperature, the two curves intersect. This temperature is referred to as the gel-sol transition temperature (Tgel).

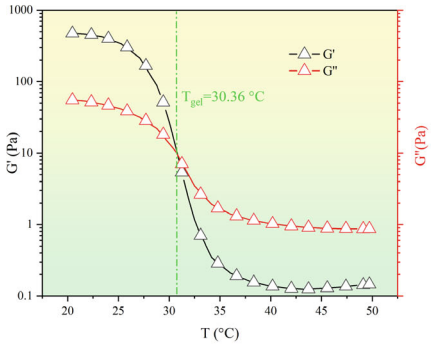
For the hydrogel patch material, the gel-sol transition temperature is 30.36 °C, indicating that the material begins to flow as it approaches 30 °C. To achieve optimal results during the printing process, it is often necessary for the material to remain in a gel state. Therefore, 27 °C is selected as printing temperature.

Thixotropy determines the ability of the structure to stabilize after printing. The thixotropy test results of the hydrogel-based thermal management patches is shown in Fig. 1b. In the figure, time is used as the horizontal axis, and viscosity values are used as the vertical axis.

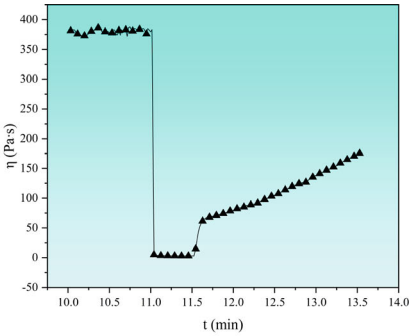
As shown in Fig. 1b, the initial viscosity of the material is high and stable. After shear stress is applied, the viscosity drops sharply. Once the shear stress is removed, the material's structure quickly recovers, and the viscosity increases. The material exhibits good thixotropy, with the viscosity returning to 31.83 Pa•s within 3 s after experiencing shear stress.

The stress-shear rate in logarithmic coordinates in a rotational shear test is illustrated in Fig. 1c. In this figure, the vertical axis represents log(τ), while the horizontal axis represents log(γ). Figure 1c reveals that the intercept is −0.05609, the slope is 0.16335, and the goodness-of-fit value is 0.97269. With Eq. (2), the consistency index K of the material is determined to be 0.8788, and the non-Newton index is found to be 0.16335. Empirical research indicates that hydrogel stickers are typical power-law fluid materials, making them suitable for extrusion 3D printing technology.

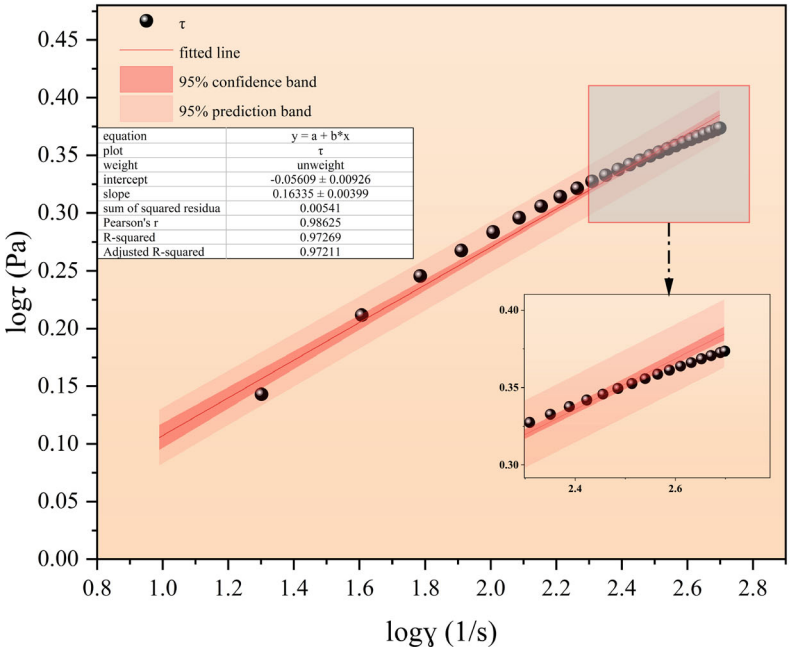
In summary, the rheological properties of hydrogel-based thermal management patches were investigated through rheological testing. The results indicate that the composite power-law fluid of the material can flow under pressure, satisfying the requirements of the printing printhead. Additionally, the material returns to its original state once the shear stress is removed. This behavior ensures the stability of the printed structure.



(a)effect of temperature on storage modulus and loss modulus of hydrogels



(b)thixotropy test results of hydrogels



(c)fitting plot of stress-shear rate at logarithmic coordinates

Fig. 1 Results of rheological test

3.2 Fluid Dynamics Simulation Analysis of Extrusion Channel

To elucidate the impact of nozzle variations on hydrogel material flow properties and print quality, this study utilizes computational fluid dynamics (CFD) simulations. The



Fig. 2 3D model of the extrusion channel

3D model of the extrusion channel is illustrated in Fig. 2. As depicted in the figure, the entire extrusion channel consists of a series of fundamental geometric shapes.

The material information utilized in computational fluid dynamics (CFD) is shown in Table 3.

The magnitude and distribution of wall shear stress for nozzles with varying specifications is illustrated in Fig. 3. As shown in this figure, the wall shear stress is low before the material enters the nozzle. At the junction between the barrel and the nozzle, the flow channel narrows abruptly due to the smaller nozzle diameter, causing a significant increase in wall shear stress.

For effective comparison, uniform wall shear values are used for research and analysis, ignoring the sharp changes at the connections. The uniform wall shear values for nozzles of various specifications is shown in Table 4.

As shown in Table 4, it is clear that for nozzles of the same diameter, a shorter length results in higher wall shear stress. For example, when the nozzle diameter is 1.2 mm, the wall shear stresses for lengths of 25 mm, 13 mm, and 6.5 mm are 177.3 Pa, 317.6 Pa, and 556.8 Pa, respectively. As shown in Table 4, it is clear that for nozzles of the same length, a smaller diameter results in lower wall shear stress. For example, when the nozzle length is 25 mm, the wall shear stresses for diameters of 1.2 mm, 0.6 mm, and 0.16 mm are 177.3 Pa, 92.06 Pa, and 33.23 Pa, respectively.

Table 3 The material information utilized in computational fluid dynamics

Density (kg/m^3)	n	K	Pressure (kPa)
986.2	0.16335	0.8788	15

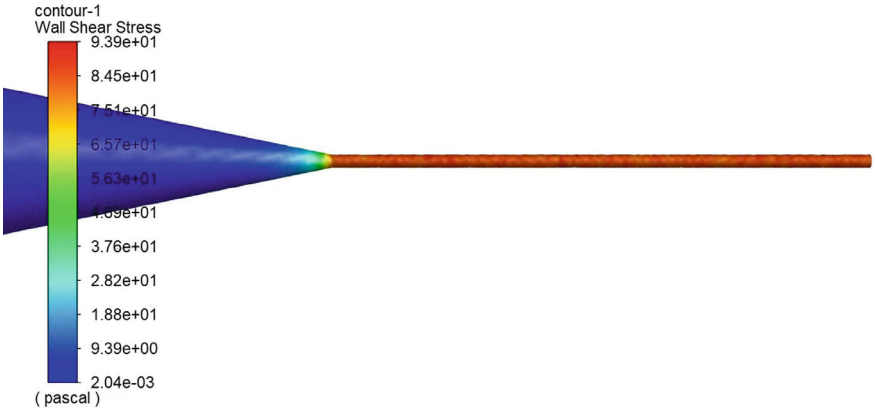


Fig. 3 Simulation result of wall shear stress (L = 6.5mm, d = 0.16mm)

Table 4 The uniform wall shear values for nozzles of various specifications

Nozzle size	Wall shear stress (Pa)
L = 25 mm, d = 1.2 mm	177.3
L = 13 mm, d = 1.2 mm	317.6
L = 6.5 mm, d = 1.2 mm	556.8
L = 25 mm, d = 0.6 mm	92.06
L = 13 mm, d = 0.6 mm	165.4
L = 6.5 mm, d = 0.6 mm	310.8
L = 25 mm, d = 0.16 mm	33.23
L = 13 mm, d = 0.16 mm	42.18
L = 6.5 mm, d = 0.16 mm	83.15

In addition, the combination of nozzle diameter and length significantly impacts wall shear stress. Therefore, it is essential to consider both factors comprehensively when designing nozzles. For example, a nozzle with a diameter of 0.16 mm shows a low wall shear stress of 33.23 Pa, even when it is longer (such as 25 mm). Conversely, a nozzle with a diameter of 1.2 mm demonstrates a high wall shear stress of 556.8 Pa, despite being shorter (such as 6.5 mm).

Research has shown that a nozzle with a diameter of 0.6 mm and a length of 13 mm provides optimal performance. This specification ensures stable material extrusion with sufficient wall shear stress (165.4 Pa) without compromising print quality.

Compared to other diameters, a 0.6 mm diameter enhances printing speed while maintaining good interlayer adhesion and surface smoothness. The 13 mm length improves material fluidity and extrusion speed while reducing the risk of excessive shear stress. Therefore, this nozzle specification will be used in future experiments to optimize printing parameters.

3.3 Optimization of 3D Printing Process Parameters of Hydrogel-Based Thermal Management Patches

Table 5 is the extrusion filament diameters at a printing pressure of 15 kPa, a printing height of 0.48 mm, and printing speeds of 120 mm/min, 180 mm/min, 240 mm/min, and 300 mm/min, respectively. As shown in the table, the extrusion filament diameters for printing speeds of 120 mm/min, 180 mm/min, 240 mm/min, and 300 mm/min are 0.726 mm, 0.463 mm, 0.379 mm, and 0.283 mm, respectively.

The microscopic images of extruded filaments at various printing speeds is shown in Fig. 4. As is shown in Fig. 4a, at a printing speed of 120 mm/min, there is significant fiber accumulation, causing the central mesh to contract due to diffusion. As is shown in Fig. 4b, at 180 mm/min, the printed grid is clear and regular, with no stretching or excessive fiber deposition. As is shown in Fig. 4c and d, at printing speeds of 240 mm/

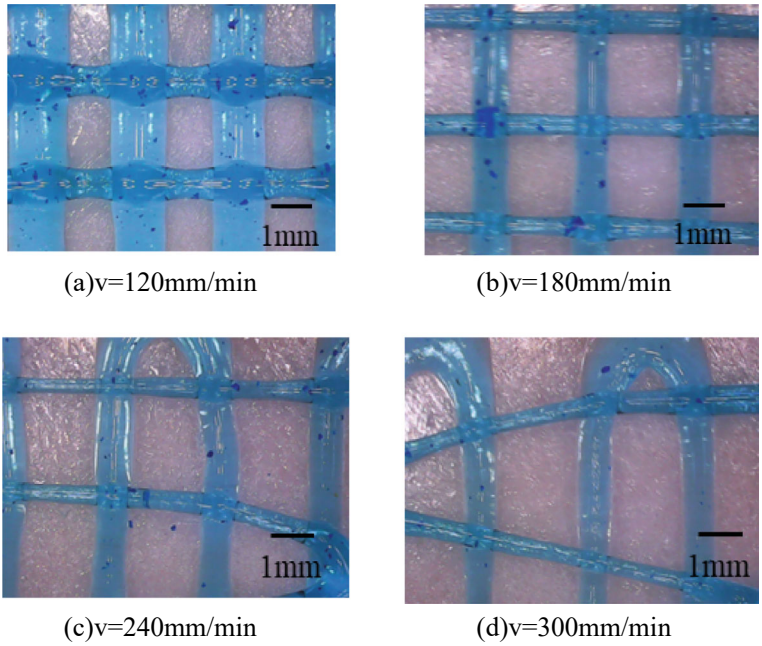
Table 5 Extrusion filament diameter at varying printing speeds

Printing speed (mm/min)	Sample 1 (mm)	Sample 2 (mm)	Sample 3 (mm)	Average (mm)
120	0.737	0.747	0.695	0.726
180	0.484	0.463	0.442	0.463
240	0.365	0.376	0.398	0.379
300	0.248	0.273	0.33	0.283

min and 300 mm/min, the extrusion becomes elongated, deforming the printed grid shape.

Table 6 presents the extrusion filament diameters at a printing pressure of 15 kPa, a printing speed of 180 mm/min, and printing heights of 0.12 mm, 0.3 mm, and 0.6 mm, respectively. As indicated in the table, when the printing heights are 0.12 mm, 0.3 mm, and 0.6 mm, the corresponding extrusion wire diameters are 2.06 mm, 0.77 mm, and 0.381 mm.

The extrusion process during the printing layer height experiment is illustrated in Fig. 5. As depicted in Fig. 5a, b, when the printing layer height is set to 0.12 mm and 0.3 mm, the printing filament becomes compressed, leading to reduced shape and size accuracy of the printed square hole. In Fig. 5c, when the printing layer height



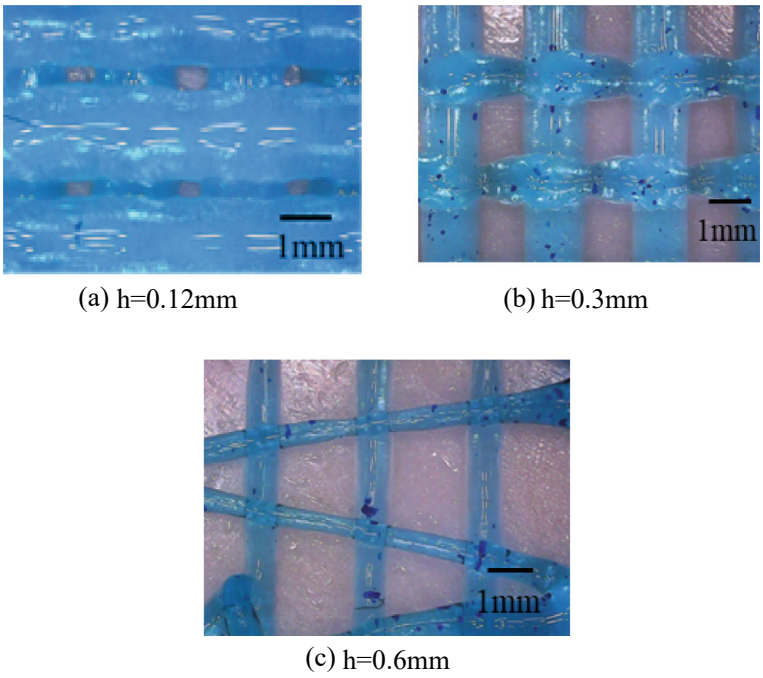
(a)v=120mm/min (b)v=180mm/min

(c)v=240mm/min (d)v=300mm/min

Fig. 4 Microscopic images of filament at various printing speeds

Table 6 Extrusion filament diameter at varying printing heights

Printing height (mm)	Sample 1 (mm)	Sample 2 (mm)	Sample 3 (mm)	Average (mm)
0.12	2.14	2.02	2.03	2.06
0.3	0.76	0.75	0.78	0.77
0.6	0.34	0.42	0.39	0.38



(a) h=0.12mm (b) h=0.3mm

(c) h=0.6mm

Fig. 5 Microscopic images of filaments at varying printing heights

is increased to 0.6 mm, the filament is significantly stretched, resulting in a notable decrease in the dimensional accuracy of the printed square hole.

4 Discussion and Future Prospects

With the advancement of construction machinery towards higher efficiency and energy conservation, effective heat dissipation has become increasingly important. There is an urgent need for customized and lightweight heat dissipation solutions. Hydrogels, known for their excellent thermal conductivity and lightweight properties, offer a promising approach.

This article evaluated nozzle specifications using CFD simulations. The results indicate that for nozzles of the same length, a smaller diameter corresponds to lower wall shear stress. Conversely, for nozzles of the same diameter, a shorter length results in higher wall shear stress. Based on these findings, a nozzle with a diameter of 0.6 mm and a length of 13 mm was selected for printing experiments. Comparing various printing effects, the optimal printing quality was achieved with a printing pressure of 15 kPa, a printing height of 0.48 mm, and a printing speed of 180 mm/min.

The significance of this study lies in exploring the potential application of hydrogels as high-performance heat dissipation materials, addressing the dual needs for lightweight construction and improved heat dissipation efficiency.

Acknowledgements This work is supported by National Natural Science Foundation of China (52305513, 32271432), Hebei Natural Science Foundation (E2022203136, E2022203125, E2022203123, No. E2023203175), the Open Fund of Guangdong Key Laboratory of Minimally Invasive Surgical Instruments and Manufacturing Technology (MISIMT-2022-6, MISIMT-2022-3), and Shandong Provincial Natural Science Foundation (ZR2021QE253).

References

1. Li L, Chen X, Liu D, Zhou HG, Wu H (2020) An intelligent loader-digger heat balance system based on simulation and wireless sensor networks. EURASIP J Wirel Commun Netw 2020:1–15
2. Lipnický M, Brodnianská Z (2023) Enhancement of heat dissipation from the hydraulic system using a finned adaptive heat exchanger. Appl Sci 13(9):5479
3. Fang G, Yuan W, Yan Z, Sun Y, Tang Y (2019) Thermal management integrated with three-dimensional heat pipes for air-cooled permanent magnet synchronous motor. Appl Therm Eng 152:594–604
4. Xiao N, Khonsari MM (2013) A review of mechanical seals heat transfer augmentation techniques. Recent Pat Mech Eng 6(2):87–96
5. Kawano T, Fuchiwaki M (2021) Pressure drop mechanisms generated in a cooling system enclosure of construction machinery. In: Fluids engineering division summer meeting American society of mechanical engineers, vol 85291 V002T03A018
6. Twelvetrees WN (1939) The development of liquid cooling: a historical summary of progress to the present-day methods. Aircr Eng Aerosp Technol 11(3):107–110
7. Liu Y, Liu X, Duan B, Yu Z, Cheng T, Yu L, Liu K (2021) Polymer–water interaction enabled intelligent moisture regulation in hydrogels. J Phys Chem Lett 12(10):2587–2592
8. Pu S, Fu J, Liao Y, Ge L, Zhou Y, Zhang S, Chen J (2020) Promoting energy efficiency via a self-adaptive evaporative cooling hydrogel. Adv Mat 32(17):1907307
9. Murphy SV, Atala A (2014) 3D bioprinting of tissues and organs. Nat Biotechnol 32(8):773–785
10. Zheng J, Chen G, Yang H, Zhu C, Li S, Wang W, Fu J (2023) 3D printed microstructured ultra-sensitive pressure sensors based on microgel-reinforced double network hydrogels for biomechanical applications. Mater Horiz 10(10):4232–4242

Open Access This chapter is licensed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License (http://creativecommons.org/licenses/by-nc-nd/4.0/), which permits any noncommercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if you modified the licensed material. You do not have permission under this license to share adapted material derived from this chapter or parts of it.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.



Digital Twin-Based Structural Performance Prediction Method for Telescopic Boom Cranes



Xuyang Cao, Kai Lin, Baokang Yan, and Chengjie Zhang

Abstract Telescopic boom cranes make it difficult to monitor the full-domain stress in real operations due to the complex structure and randomly variable loads and environmental conditions. A structural performance prediction method based on digital twin is proposed, which combines real-time monitoring data with finite element simulation data and incorporates telescopic boom expertise with feature derivation, utilizes the XGBoost algorithm to establish a real-time stress prediction agent model, realizes the structural stress prediction of telescopic boom cranes with the boom and rotary table as the main key components of the whole life cycle, and builds a digital twin system of telescopic boom crane structural performance using Unity 3D to realize real-time stress control and real-time stress prediction, and builds a digital twin system of telescopic boom crane structural performance. Unity 3D is used to build a digital twin system of telescopic boom crane structural performance, realizing mutual control of reality and real-time stress cloud display to ensure its safety performance.

Keywords Telescopic boom crane · Digital twin · Real-time stress prediction · XGBoost

1 Introduction

Telescopic boom cranes are widely used in heavy lifting tasks due to their flexibility, and their reliability and safety are crucial [1]. Although the structural health monitoring technology has made great progress in recent years [2], the existing methods are still difficult to achieve real-time monitoring of the whole domain of the stress of key components of telescopic boom cranes, for example, the traditional methods such as strain gauges can only detect the local stress points [3], and the real-time performance of the finite element method is still very insufficient under complex

X. Cao · K. Lin (✉) · B. Yan · C. Zhang
Dalian University of Technology, Dalian, China
e-mail: link123@mail.dlut.edu.cn

© The Author(s) 2026

H. Saman et al. (eds.), *The 9th International Conference on Advances in Construction Machinery and Vehicle Engineering*, Lecture Notes in Mechanical Engineering,
https://doi.org/10.1007/978-981-96-5527-4_21

working conditions [4], and the digital twin technology can effectively solve similar problems. Many scholars have done a lot of research on the digital twin system of major mechanical equipment [5–8], but there are relatively few studies on the digital twin system of telescopic boom crane, which is a major mechanical equipment with variable structure.

In this paper, digital twin technology is combined with XGBoost algorithm to establish an agent model through real-time monitoring data and finite element simulation data to achieve real-time stress prediction of key components of telescopic boom cranes (mainly booms and slewing platforms), and a digital twin platform is established to achieve the mutual control of virtual reality and real-time stress maps to ensure the operational safety.

2 System Framework

The system framework is shown in Fig. 1 and consists of the following four parts: the physical space includes the telescopic boom crane, the perception system composed of various types of sensors, industrial control machine and other hardware equipment. The twin space includes the finite element simulation model, stress prediction agent model, telescopic boom crane twin model and other virtual models. Twin data includes all kinds of data generated and derived from the fusion of physical space and twin space. The data communication includes two-way transmission within the twin data to realize virtual and real mutual control.

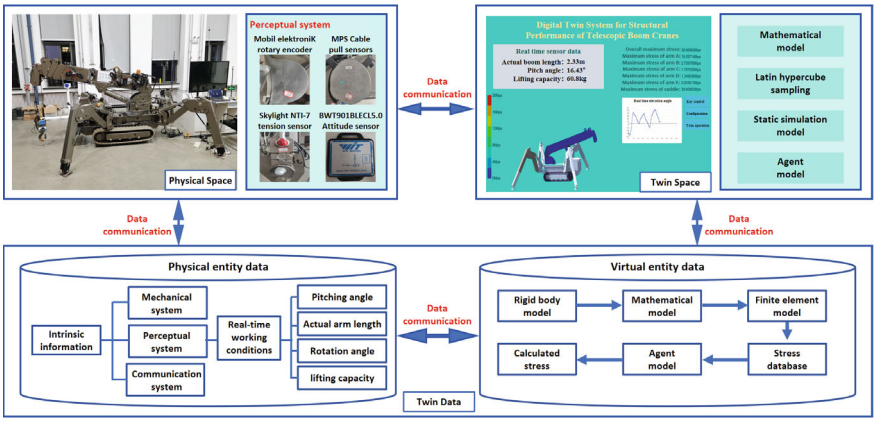


Fig. 1 System framework diagram

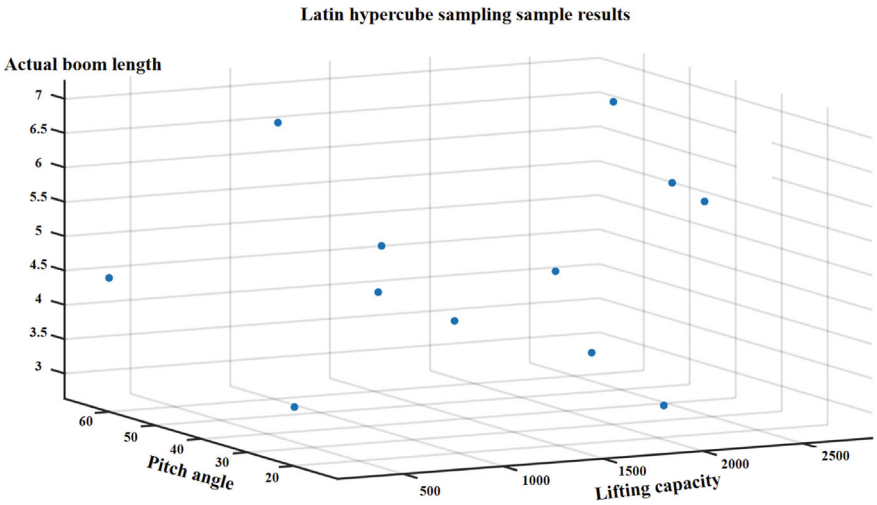


Fig. 2 Latin hypercube sampling results

3 Stress Database

3.1 Latin Hypercube Sampling

Latin hypercube sampling is a stratified Monte Carlo sampling method. The basic working condition parameters of the telescopic boom crane are: actual boom length, pitch angle, lifting capacity of a total of three characteristic parameters, using Matlab to normalize the data, randomly selected 12 sample points, the distribution is shown in Fig. 2.

3.2 Static Simulation

According to the sampling results, ANSYS is used to establish the static simulation model. Focusing only on the force of the boom and rotary table, the model is simplified, and the boom and rotary table are adopted as shell units, while the slider and bearing are adopted as solid units, as shown in Fig. 3. The internal meshes of the components are set to be shared, binding contact is applied between the components, fixed support is applied to the bottom surface of the rotary table, and a mass point is coupled at the head of the boom to simulate the lifting weight. The pitching hydraulic cylinder is only considered to support, and the beam connection is used to replace the pitching hydraulic cylinder. The stress cloud diagram is obtained by solving, as shown in Fig. 4, and the stress database is obtained.

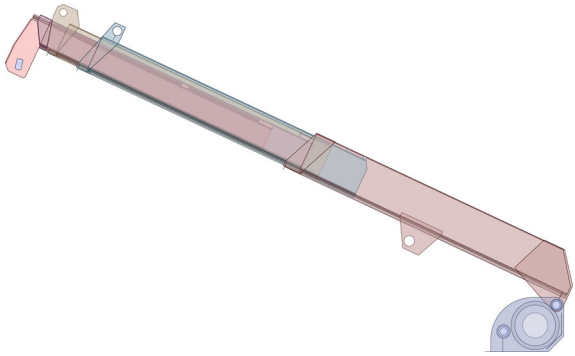


Fig. 3 Geometric model pre-processing

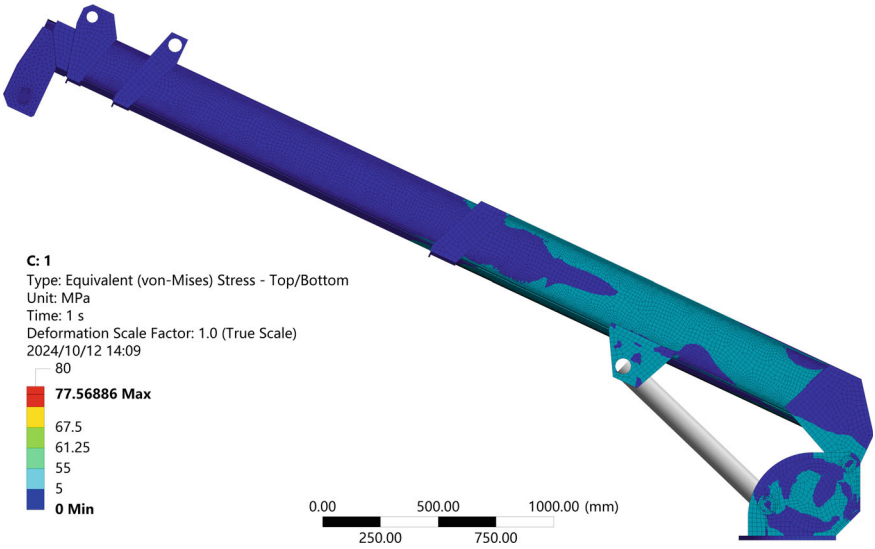


Fig. 4 Stress cloud for Case 1

4 Agent Modeling

Stress prediction is a multi-input, multi-output regression problem, as shown in Fig. 5 the model inputs are three basic features: actual arm length, pitch angle, and lifting capacity, and the model outputs are the stress values at each node.

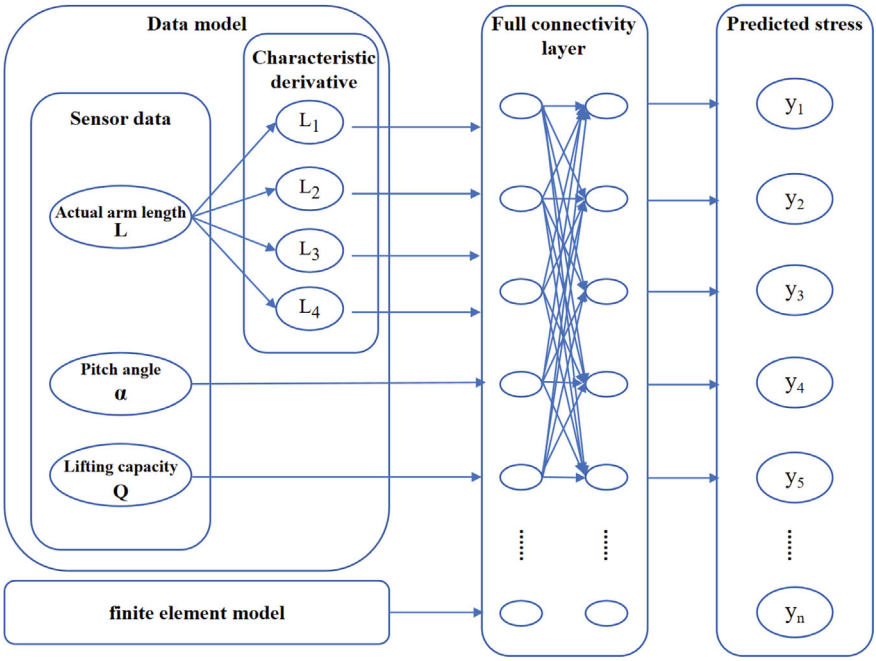


Fig. 5 Schematic diagram of stress prediction agent model

4.1 Mathematical Modeling

The telescopic boom crane used in this paper has one and two sections for sequential telescoping, three and four sections for synchronized telescoping, and the fifth section is a manual telescopic boom with only four positions. Before using i section of the arm maximum working radius of $R_i(1 \leq i \leq 5, i \in Z)$, set the telescopic length of each section of the arm for $L_i(2 \leq i \leq 5, i \in Z)$, when $R_1 \leq L \leq R_2$, only the second section of the arm telescopic, telescopic length, the same reason can be obtained from the telescopic length of each arm, as shown in Table 1.

Table 1 Calculation of boom telescopic length

Actual arm L	Section 2 Arm L_2	Section 3 Arm L_3	Section 4 Arm L_4	Section 5 Arm L_5
$R_1 \leq L \leq R_2$	$L - R_1$	0	0	0
$R_2 \leq L \leq R_4$	$R_2 - R_1$	$(L - R_2)/2$	$(L - R_2)/2$	0
$R_4 \leq L \leq R_5$	$R_2 - R_1$	$(R_4 - R_2)/2$	$(R_4 - R_2)/2$	4 Types

4.2 Feature Derivation

Feature derivation refers to constructing new features from the original data. The data directly measured by the sensor is the actual arm length L, due to the telescopic crane each section of the arm retracts in a different way, based on Table 1 will be an actual arm length L, derived from four new features with greater relevance, that is, the length of each section of the arm retracts $L_2 \sim L_5$, feature derivation makes the agent model can be integrated into the telescopic arm professional domain knowledge, so that the model is more in line with the laws of physics to improve the performance of the model.

4.3 Comparison of Stress Prediction Agent Models

In this paper, RandomForestRegressor and XGBoost are used to build agent models to predict real-time stress and compare their prediction effects, respectively.

RandomForestRegressor. RandomForestRegressor improves the prediction performance by constructing multiple decision trees. For the regression prediction of each node stress, firstly, we need to select the appropriate division point and find the optimal cut-off variable to build the decision tree, if M the nodes $R_1, R_2, R_3 \cdots R_m$ have a fixed output value c_m on each node R_m, the decision tree model can be expressed as:

$$f(x) = \sum_{m=1}^{M} \hat{c}_m I(x \in R_m) \tag{1}$$

The model is trained using the stress dataset and the stress value of the first node is used to evaluate the model, the final coefficient is determined $R^2 = 0.878$, and the model accuracy is good, Fig. 6 shows the comparison between the predicted value and the training value, and the predicted data basically matches.

XGBoost. XGBoost is a machine learning algorithm based on the gradient boosting framework, which improves the prediction accuracy by continuously building new regression trees to optimize the objective function, and ultimately using multiple decision trees to make joint decisions. The objective function of XGBoost consists of two parts: a loss function that measures the prediction error and a regularization term that prevents the algorithm from overfitting. The objective function is denoted as:

$$Obj(\theta) = \sum_{i=1}^{n} L(y_i, \hat{y}_i) + \sum_{k=1}^{K} \Omega(f_k) \tag{2}$$

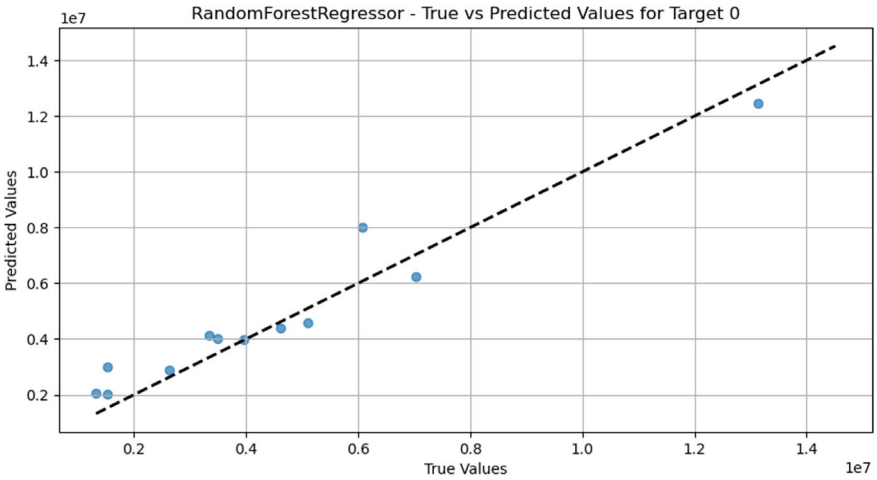


Fig. 6 Comparison of RandomForestRegressor errors

where n is the number of samples, $L(y_i, \hat{y}_i)$ is the loss function, $\hat{y}_i$ is the predicted value of the i sample, K is the number of trees, and $\Omega(f_k)$ is the regularization term. The expression of $\Omega(f_k)$ is:

$$\Omega(f_k) = \gamma T + \frac{1}{2}\lambda\omega_k^2 \tag{3}$$

where T is the number of leaf nodes inside the tree, ω_k^2 is the score of the leaf nodes in the tree, and γ and λ are the weight coefficients of the penalty term.

The stress dataset is used to train the model and the stress value of the first node is used to evaluate the model and finally $R^2 = 0.935$, which indicates that the model has excellent accuracy. XGBoost is able to provide better performance. Figure 7 shows the error comparison between the predicted values and the training values, the predicted data basically matches, and the accuracy is greatly improved compared to RandomForestRegressor.

5 Digital Twin System for Structural Performance of Telescopic Boom Cranes

The digital twin system established in this paper is developed based on Unity 3D, focusing on the realization of real-time stress display function.

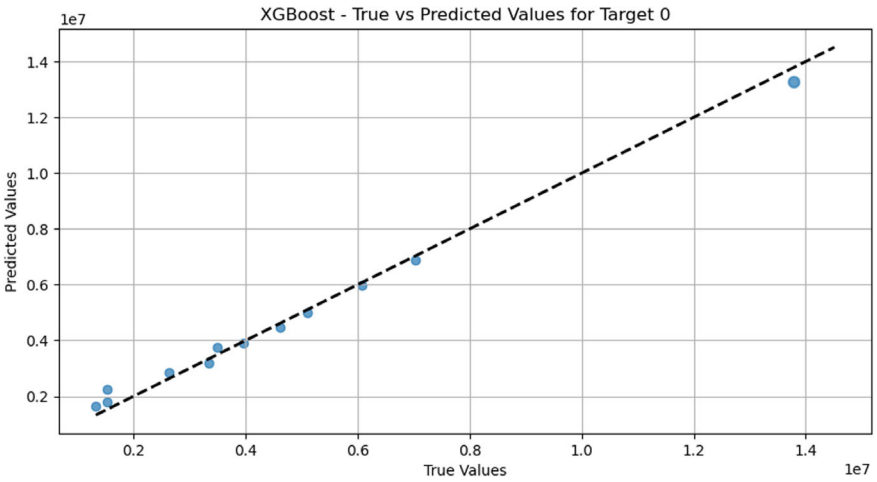


Fig. 7 Comparison of XGBoost errors

5.1 Cloud Model Reconstruction

Unity 3D imports the model, the number of triangular surfaces is not consistent with the number of ANSYS finite element meshes, Matlab writes a code to read the STL file of each part exported by ANSYS, and obtains the STL node coordinates and the sequence of STL indexes to complete the reconstruction of the cloud model.

5.2 Real-Time Stress Data Acquisition

ANSYS and Matlab have computational errors, the node coordinates obtained by the two software are not the same, so it is not possible to match the stress data directly. It is necessary to use the stress data set exported by ANSYS to establish an RBF agent model, predict the stress value of the cloud map through the STL node coordinates, then get the corrected stress database, and based on the corrected stress database and the real-time data of the sensors, the XGBoost model is used to predict the real-time stress, and the real-time stress data is sent back to Unity 3D through TCP communication. Based on the modified stress database and the sensor real-time data, the XGBoost model is used to predict the real-time stress, and the real-time stress data is sent back to Unity 3D through TCP communication, and through the Mesh component, the vertices are directly assigned with values, the color value bands are set, and the color mapping is completed to finally realize the display of the real-time stress cloud map, as shown in Fig. 1.

6 Conclusion

This article compares XGBoost and random forest classifier through experiments, and XGBoost shows better prediction accuracy, which can more accurately predict stress distribution under different working conditions. Based on the Unity platform, a digital twin system for the structural performance of a telescopic boom crane was constructed. The twin data achieved virtual reality interactive control and real-time stress cloud display of the crane, greatly improving the structural performance monitoring capability of the telescopic boom crane in actual operation. However, the structural performance evaluation indicators of the constructed digital twin system are still relatively single, laying the foundation for the subsequent completion of life prediction and fault diagnosis based on predicted stress.

Acknowledgements This work is supported by National Natural Science Foundation of China under Grant No.52275088 and the Fundamental Research Funds for the Central Universities under Grant (No.DUT22LAB506, No.DUT22LAB507).

References

1. Zhang H, Jing WZ, Qin XR et al (2020) Research and application of digital twin system for harbor crane based on VTK visualization engine technology. Crane Transp Mach Mach 17:69–74
2. Moi T, Cibicik A, Rølvåg T (2020) Digital twin based condition monitoring of a knuckle boom crane: an experimental study. Eng Fail Anal 112:104517
3. Li DY (2023) Research on real-time monitoring of crane main girder mechanical state based on digital twin technology. Henan University of Technology
4. Fan Y (2024) A high-fidelity model-driven life prediction method for crane metal structures. Taiyuan University of Science and Technology
5. He XW, Lai XN, Yang LL et al (2021) Mechanism and data joint-driven digital twin design of gantry crane. In: Proceedings of the 17th annual China CAE engineering analysis technology conference, pp 21–25+33
6. Ren XY (2023) Research on digital twin model driving technology for roadheader. Yanshan University
7. Dong Q, Hu SX, Xu GN et al (2024) Digital twin system construction method for evaluating load carrying capacity of crane structures. Comput Integr Manuf Syst, CIMS 30(3):837–851
8. Ding H, Yang LL, Yang ZJ et al (2020) Health state prediction of coal mining machine driven by fusion of digital twin and deep learning. China Mech Eng 31(7):815–823

Open Access This chapter is licensed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License (http://creativecommons.org/licenses/by-nc-nd/4.0/), which permits any noncommercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if you modified the licensed material. You do not have permission under this license to share adapted material derived from this chapter or parts of it.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.



Research on the Trajectory of Hydraulic Boom in Medium Sized Excavators Based on D-H Parameter Method



Tingting Luo, Chengqiang Liu, Xiliang Ma, and Nianning Luo

Abstract Electrification is a current research hotspot in construction machinery and has been exploratory applied in small-sized excavators. There are few research reports on the electrification of medium and large-sized excavators with higher cost-effectiveness. In this thesis, based on the electro-hydraulic drive tooling structure which is more suitable for the operating condition of medium-sized excavator, cubic polynomial is used to fit the trajectory of excavator in sections. The D-H parameter method is used to study the relationship between operating condition, tooling attitude of medium-sized excavator and the pressure and flow rate of driving cylinder, and the whole machine model of medium-sized electric hydraulic drive excavator is established. Through simulation studies and experimental validation on a 37 t prototype, the discussion explored the motion laws of the tool under the constraints of the expected working trajectory, offering a reference for the design selection and trajectory control of the electric hydraulic drive system for medium-sized excavators.

Keywords Electric excavator · DH parameter method · Trajectory tracking

1 Introduction

Currently, China's engineering machinery is accelerating its transition to electrification. According to a sales report released by the China Construction Machinery Industry Association, the penetration rate of electric loaders has reached 13%. There have been exploratory applications of small electric excavators, but research literature on the electrification of medium and large excavators, which are widely used and cost-effective, is still scarce. Although traditional medium-sized excavators have achieved a degree of energy conservation and emission reduction through trajectory

T. Luo · N. Luo (✉)
Harbin Institute of Technology, Weihai, Shandong, China
e-mail: luo@hit.edu.cn

C. Liu · X. Ma
Xuzhou University of Technology, Xuzhou, Jiangsu, China

© The Author(s) 2026

H. Saman et al. (eds.), *The 9th International Conference on Advances in Construction Machinery and Vehicle Engineering*, Lecture Notes in Mechanical Engineering,
https://doi.org/10.1007/978-981-96-5527-4_22

control optimization, engine-main pump matching, pump-valve operation matching, and energy recovery methods, experiences from the automotive industry suggest that electrification is the ultimate solution for energy conservation and emission reduction.

Medium-sized excavators are efficient, versatile, well-equipped, and offer a high cost-performance ratio, making them the mainstay of the entire excavator market. From an application perspective, the electric hydraulic drive topology for medium-sized excavators may be more suitable for their typical operating conditions. Compared to traditional excavators and small electric excavators, they coordinate the movements of the arm and rotation mechanism to achieve the operator's desired trajectory; the key difference lies in the need for higher power matching and adaptability to high-power operating conditions for medium electric excavators.

1.1 Analysis of Research Status

The Denavit-Hartenberg (D-H) parameter method regards the boom, bucket, and arm operating mechanism of the excavator as a planar linkage mechanism, which is suitable for modeling a serial linkage boom. It has been widely used in excavator trajectory planning, mining vehicle-mounted booms, four-link heavy-duty transfer robots, and 7-DOF boom mechanism control accuracy research [1–3]. Yin and others studied the active decoupling method of 6R robots, using the D-H parameters containing coupling quantities to solve the forward kinematics, and used algebraic solutions to find the inverse solution [4]. Yang and others studied the trajectory control of excavators using hybrid interpolation methods [5]. Haga and others studied the automatic ditch digging of excavators, based on trajectory planning control to achieve automatic digging of horizontal ditches by controlling the depth of the bucket [6]. Huang and others studied the multi-objective joint space optimal trajectory with operation time and operation energy consumption as optimization objectives, using piecewise polynomial interpolation to ensure smooth motion of each joint [7]. Qiu and others established the inverse kinematics of a working device with closed chain for a multi-joint control system considering time lag, indicating that the multi-joint control system considering time lag can track the planned trajectory more accurately, quickly and smoothly [8]. And others studied a trajectory planning method that achieves a certain comprehensive optimization of the operating time and joint pulsation of excavator joints [9].

Based on the D-H parameter method, this thesis studies cubic polynomials for piecewise fitting of excavator trajectories. It investigates the motion laws of the tooling and energy consumption distribution for medium-sized excavators, exploring the dynamic adjustment patterns of hydraulic system parameters required for driving the tooling, offering a reference for the design selection and trajectory control of the electric hydraulic drive system for medium-sized excavators.

2 Modeling and Simulation Based on D-H Parameter Method

The hydraulic boom of excavator can be equivalent to a three-link series mechanism. The length of three cylinders determines the angle of connecting rods, and the length of three cylinders determines the posture of bucket. D-H parameter method can be used to calculate the posture of excavator bucket.

The D-H parameter method is used to describe the relationship between translation and rotation between adjacent links. The posture matrix between adjacent links can be established, and the posture of the end mechanism in the base coordinate system can be obtained by matrix transformation. The D-H parameter method in the front coordinate system is to establish the connecting rod coordinate system at the front end of the connecting rod, which has the characteristics of clearer physical meaning. The D-H parameter method uses four parameters to describe the tandem linkage mechanism.

The hinge points and coordinate system numbers of the mechanical structure of the excavator boom are shown in Fig. 1.

The length L_i of connecting rod i. The connecting rod length is defined as the length of the common perpendicular line of adjacent joint axes. The connecting rod length is equal to the distances AG, GK, and KP between the hinge points respectively.

(1) Torsion angle R_i of connecting rod i. The link torsion angle is defined as the angle between adjacent joint axes. The torsion angle is 0°.
(2) Offset d_i of link i. It is defined as the distance between the intersection of the joint axis and the plane where the link is located. Ignoring the error caused by manufacturing accuracy, the three links of the excavator operating mechanism are considered to be on the same plane.
(3) The rotation angle θ_i of joint i. It is defined as the rotation angle of the link around the axis, which is the rotation angle of link i relative to link i−1, that is, the angle between the connecting rods. The angle between the connecting rods of the excavator hydraulic boom is determined by the length of the cylinder, as shown in Table 1.

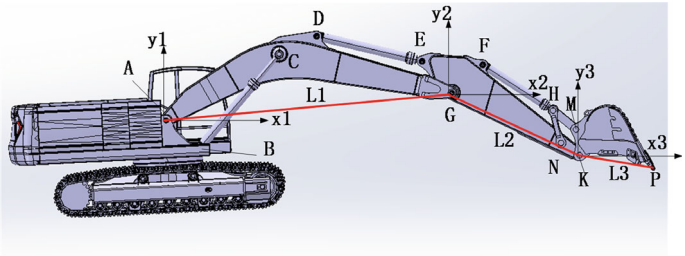


Fig. 1 Hinge points and coordinate system of excavator working mechanism

Table 1 D-H parameter table of excavator operating mechanism

i	R_i	L_i (mm)	d_i	θ_i
Link 1	0°	5613	0	$\theta_{1\,\min}$ $\theta_{1\,\max}$
Link 2	0°	2971	0	$\theta_{2\,\min}$ $\theta_{2\,\max}$
Link 3	0°	1295	0	$\theta_{3\,\min}$ $\theta_{3\max}$

According to the coordinate system defined in Fig. 1, the angle between the boom link L1 and the coordinate system × 1 is θ_1, determined by the boom cylinder length,

$$\theta_1 = \arccos\left(\frac{l_{AB}^2 + l_{AC}^2 - l_{BC}^2}{2l_{AB}l_{AC}}\right) - \angle \mathrm{BAx1} \tag{1}$$

In Eq. (1), ∠BAx1 is the angle between the line connecting the hinge point A and hinge point B of the excavator and the horizontal coordinate axis, l_{AB} is the distance between points A and B, l_{AC} is the distance between points A and C, l_{BC} is the length of the boom cylinder, which are determined by the structural parameters of the excavator.

The angle between the boom link L2 and the boom link in the coordinate system × 2 is θ_2, determined by the length of the boom cylinder:

$$\theta_2 = 2\pi - \arccos\left(\frac{l_{\mathrm{DG}}^2 + l_{\mathrm{EG}}^2 - l_{\mathrm{DE}}^2}{2l_{\mathrm{DG}}l_{\mathrm{EG}}}\right) - \angle \mathrm{XGX2} \tag{2}$$

θ_3 is a function of the length of the bucket cylinder:

$$\theta_3 = 2\pi - \angle \mathrm{MKG} - \angle \mathrm{XKX3} \tag{3}$$

In Eq. (3), ∠MKG is the value determined by the length of the bucket cylinder. According to the coordinate transformation, the bucket end posture matrix is:

$$\begin{bmatrix} c_{123} & -s_{123} & s_0 & \mathrm{L}_3c_{123} + \mathrm{L}_2c_{12} + \mathrm{L}_1c_1 \\ 0 & 0 & -1 & 0 \\ s_{123} & c_{123} & 0 & \mathrm{L}_3s_{123} + \mathrm{L}_2s_{12} + \mathrm{L}_1c_1 \\ 0 & 0 & 0 & 1 \end{bmatrix} \tag{4}$$

In Eq. (4), c_{123} is abbreviation of $\cos(\theta_1+\theta_2+\theta_3)$, s_{123} is abbreviation of $\sin(\theta_1+\theta_2+\theta_3)$; The structural parameters of the three cylinders of the excavator are shown in Table 2.

According to the stroke of the cylinder, the angle intervals of the three joint angles are shown in Table 3.

Table 2 Parameter table of excavator operating mechanism

Name	Parameter	Numerical value	Unit
Boom cylinder	Cylinder bore	120	mm
	Rod diameter	85	mm
	Journey	1220	mm
Arm cylinder	Cylinder bore	140	mm
	Rod diameter	100	mm
	Journey	1470	mm
Bucket cylinder	Cylinder bore	120	mm
	Rod diameter	85	mm
	Journey	1060	mm

Table 3 Motion range of three joint angles of excavator

Cylinder	Angle	Angle Value (°)
1 boom cylinder	$\theta_{1\,\min}$	−30
	$\theta_{1\,\max}$	60
2 arm cylinder	$\theta_{2\,\min}$	30
	$\theta_{2\,\max}$	140
3 bucket cylinder	$\theta_{3\,\min}$	−44
	$\theta_{3\max}$	110

The operating space of the bucket teeth of the excavator is obtained, as shown in Fig. 2, and the position that the bucket teeth can reach in the operating plane when the excavator does not rotate.

The operating space of excavator bucket teeth is the operating position that the bucket teeth can reach, and the actual operating condition is related to the initial posture of the bucket, as shown in Fig. 3.

It can be seen from Fig. 3 that the circle in the figure represents the operating space of the bucket hinge point K, and the line segment represents the initial posture of the bucket. The initial posture of the bucket is different at different operating points. The initial posture determines the range of motion of the bucket and whether the excavator can realize excavation, lifting and unloading when the hinge point K of the bucket is here.

Different postures of the operating mechanism of the excavator affect the posture of the bucket, and the displacement of the bucket cylinder affects the posture of the bucket, thus affecting the force of the bucket cylinder. The effect of excavator bucket gravity on bucket cylinder is shown in Fig. 4. Different positions of the center of gravity of the bucket relative to the hinge point K produce clockwise or counterclockwise torque, which produces pulling or pressure on the bucket cylinder. It can be seen that the pressure of the rodless cavity of the bucket cylinder will be affected by the position and attitude of the bucket, and there will be positive pressure

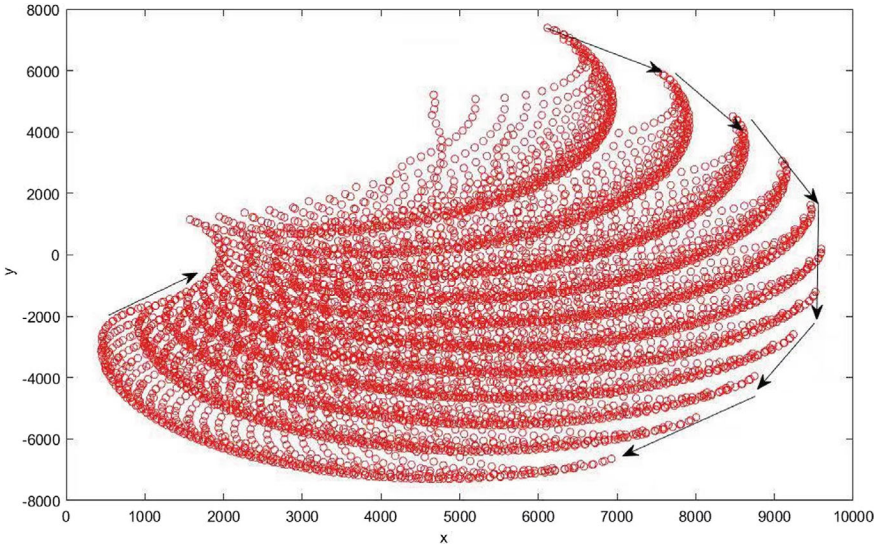


Fig. 2 Operating space of bucket teeth

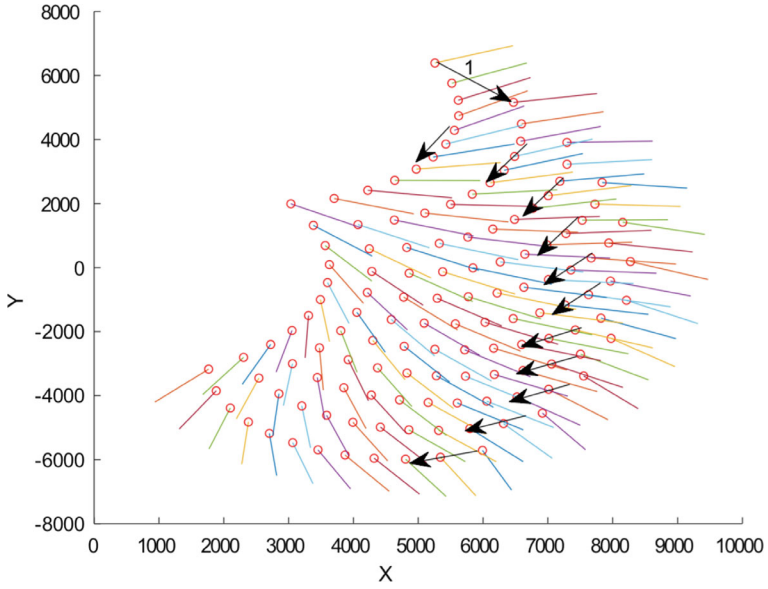


Fig. 3 Bucket position posture in workspace

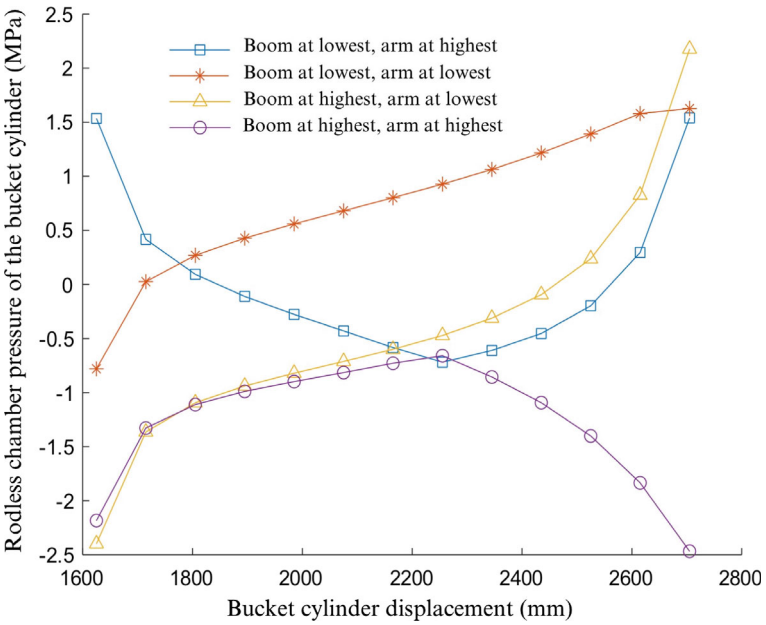


Fig. 4 Pressure of the bucket cylinder rodless cavity

or negative pressure, which will drive the load when the positive pressure is positive, and it will be negative load when the negative pressure is negative.

Boom position, bucket attitude and arm attitude affect the force of arm cylinder, and the force of arm cylinder within the operating range is shown in Fig. 5.

A positive value means that the arm cylinder is subjected to tension, and the large cavity of the arm cylinder will generate negative pressure. The negative value means that the bucket cylinder is under pressure, and the bucket cylinder will have a zero crossing point during the operating process. Under the action of gravity, the bucket cylinder will be switched between tension and pressure.

The attitude of the operating mechanism of the excavator affects the force on the cylinder of the boom, and the force on the cylinder of the boom under different attitudes of the excavator is shown in Fig. 6. When the arm is deployed to the maximum, the influence of bucket attitude on the force of boom cylinder is relatively small. When the arm is retracted, the posture of the bucket has a relatively large influence on the force of the boom cylinder. Under all postures, the boom cylinder is the most stressed when it is at the lowest position, and with the lifting of the boom, the boom cylinder is stressed smaller and smaller. The force on the cylinder of the boom is mainly generated by the gravity of the operating mechanism, and the cylinder of the boom is always under pressure.

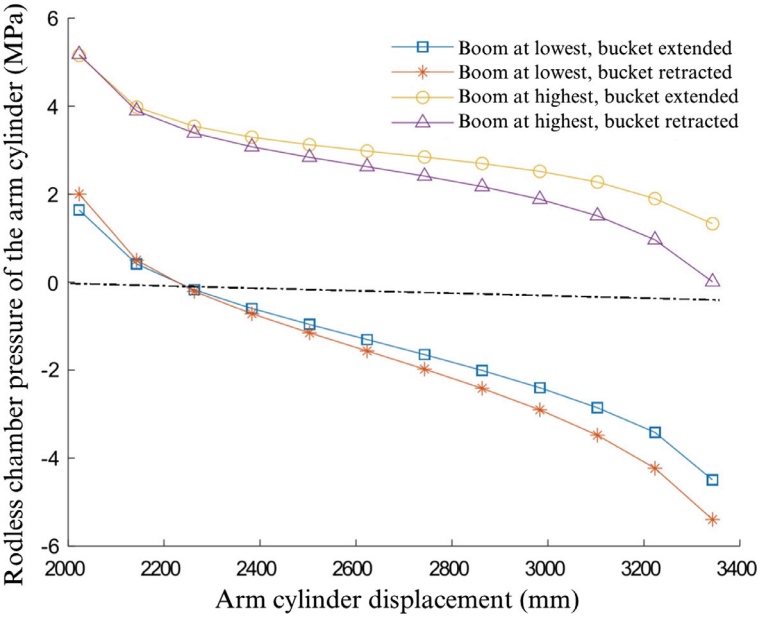


Fig. 5 Pressure of the bucket cylinder rodless chamber

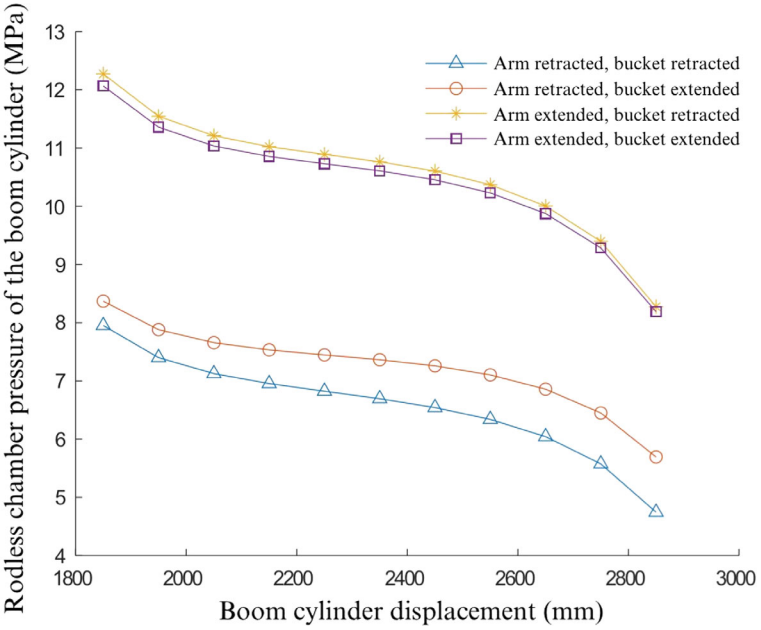


Fig. 6 Pressure in the rodless chamber of the boom cylinder

3 Trajectory Planning of Excavator Hydraulic Boom

The excavator trajectory is fitted by piecewise cubic polynomial. As shown in Eq. 5, there are 4 unknowns of cubic polynomial, and the fitted polynomial can be obtained by obtaining the coefficient value.

$$\begin{cases} \theta_{1\,i} = a_{10\,i} + a_{11\,i}t + a_{12\,i}t^2 + a_{13\,i}t^3 \\ \theta_{2\,i} = a_{20\,i} + a_{21\,i}t + a_{22\,i}t^2 + a_{23\,i}t^3 \\ \theta_{3\,i} = a_{30\,i} + a_{31\,i}t + a_{32\,i}t^2 + a_{33\,i}t^3 \end{cases} \tag{5}$$

In Eq. (5), $\theta_{1\,i}$ is the boom joint angle, $\theta_{2\,i}$ is the arm joint angle, $\theta_{3\,i}$ is the bucket joint angle. The starting position and the end position are given values, that is, the starting point and the end point of the given bucket end trajectory. The starting point of the first segment is:

$$\begin{cases} \theta_{110} = a_{101} \\ \theta_{210} = a_{201} \\ \theta_{310} = a_{301} \end{cases} \tag{6}$$

The end angle of the first trajectory is given, that is, through the given interpolation point, the equation is obtained as follows:

$$\begin{cases} \theta_{1\,t1} = a_{101} + a_{111}t_1 + a_{121}t_1^2 + a_{131}t_1^3 \\ \theta_{2\,t1} = a_{201} + a_{211}t_1 + a_{221}t_1^2 + a_{231}t_1^3 \\ \theta_{3\,t1} = a_{301} + a_{311}t_1 + a_{321}t_1^2 + a_{331}t_1^3 \end{cases} \tag{7}$$

Given that the velocity at the starting point of the trajectory is zero, the coefficient equation is obtained by deriving Eq. (8):

$$\begin{cases} a_{111} = 0 \\ a_{211} = 0 \\ a_{311} = 0 \end{cases} \tag{8}$$

In order to keep the velocity at the interpolation point continuous, given that the angular velocity at the end of the first segment is equal to the angular velocity at the starting point of the second segment, the coefficient equation is obtained:

$$\begin{cases} a_{120} = a_{111}t_1 + 2a_{112}t_1^1 + 3a_{113}t_1^2 \\ a_{220} = a_{211}t_1 + 2a_{212}t_1^1 + 3a_{213}t_1^2 \\ a_{220} = a_{311}t_1 + 2a_{312}t_1^1 + 3a_{313}t_1^2 \end{cases} \tag{9}$$

In order to keep the acceleration at the node continuous, the angular acceleration at the end of the first segment is equal to the acceleration at the starting point of the

second segment

$$\begin{cases} a_{120} = a_{112} + 3a_{113}t_1 \\ a_{220} = a_{212} + 3a_{213}t_1 \\ a_{320} = a_{312} + 3a_{313}t_1 \end{cases} \tag{10}$$

The remaining segments should also meet the equal velocity and acceleration at the interpolation point, which is similar to the equation between the first and second segments. According to the above equation, the coefficient matrix is as follows:

$$\begin{bmatrix}
t_1^3 & t_1^2 & t_1 & 1 & 0 & 0 & 0 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
3t_1^2 & 2t_1 & 1 & 0 & 0 & 0 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
6t_1 & 2 & 0 & 0 & 0 & -2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & t_2^3 & t_2^2 & t_2 & 1 & 0 & 0 & 0 & 0 & 0 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 3t_2^2 & 2t_2 & 1 & 0 & 0 & 0 & 0 & 0 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 6t_2 & 2 & 0 & 0 & 0 & 0 & 0 & -2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & t_3^5 & t_3^4 & t_3^3 & t_3^2 & t_3 & 1 & 0 & 0 & 0 & -1 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 5t_3^4 & 4t_3^3 & 3t_3^2 & 2t_3 & 1 & 0 & 0 & 0 & -1 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 20t_3^3 & 12t_3^2 & 6t_3 & 2 & 0 & 0 & 0 & -2 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & t_4^3 & t_4^2 & t_4 & 1 & 0 & 0 & 0 & -1 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 3t_4^2 & 2t_4 & 1 & 0 & 0 & 0 & -1 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 6t_4 & 2 & 0 & 0 & 0 & -2 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & t_5^3 & t_5^2 & t_5 & 1 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 3t_5^2 & 2t_5 & 1 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 6t_5 & 2 & 0 & 0 \\
0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1
\end{bmatrix} \tag{11}$$

The given interpolation points are shown in Table 4.

According to the multinomial coefficient equation of cubic fitting, given the initial point, end point and path point, the joint angle of cubic polynomial fitting is obtained, as shown in Fig. 7. Fitting curve 7 shows that the change amplitude of arm joint angle

Table 4 Given path points for the excavator

	Path point	Boom angle (°)	Arm angle (°)	Bucket angle (°)
Excavation starting point	P1	25	60	10
Mining path points	P2	24	65	30
	P3	22	70	50
	P4	20	75	70
	P5	22	85	90
Excavation endpoint	P6	22	96	95

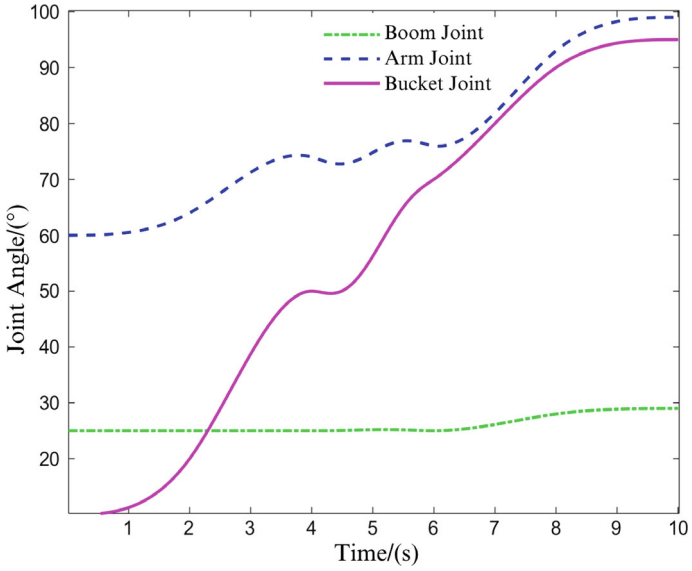


Fig. 7 Curve of joint angle

is small, the change of arm joint angle is middle, and the change of bucket joint angle is the largest.

The joint angular velocity fitted by the third degree polynomial is shown in Fig. 8. It can be seen that the angular velocity of the arm joint changes smoothly, the angular velocity of the bucket joint changes greatly, and the angular velocity of the arm joint changes in the middle.

The angular acceleration of the joint fitted by the third degree polynomial is shown in Fig. 9. It can be seen that the angular acceleration of the arm joint changes little, while the angular acceleration of the bucket joint and the angular acceleration of the arm joint are large, which have the same change trend.

4 Excavator Hydraulic Boom Experiment

Typical medium-sized hydraulic excavators use a double-union piston-type variable pump for their hydraulic power systems, which allows for relatively independent oil supply to two actuating elements, and can also achieve combined flow driving according to the operational requirements. This paper conducts experimental tests on a medium-sized excavator, with the main parameters shown in Table 5.

Operate the excavator handle to control the compound action of boom cylinder and arm cylinder. The data curve of test is shown in Fig. 10.

It can be seen from Fig. 10 that the boom cylinder and the arm cylinder achieve a good linear compound action. The flow rate of pump 1 is supplied to the boom

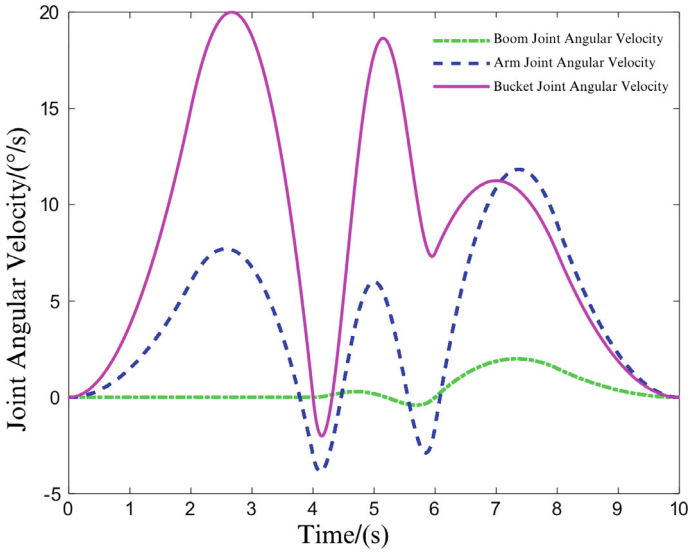


Fig. 8 Velocity curve of joint angle

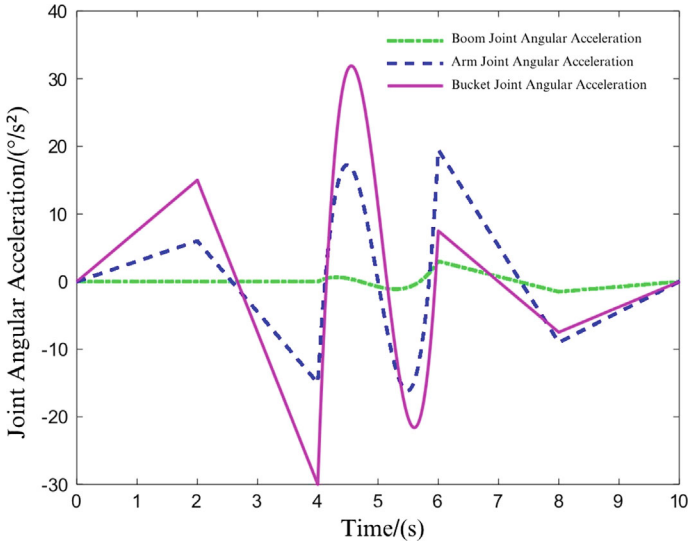


Fig. 9 Angular acceleration of joint angle

Table 5 Main technical parameters of the power system of medium-sized excavator

Name	Parameter	Value	Unit
Hydraulic pump	Flow	217	L/min
	Pressure	37.3	MPa
Multiple valve	Flow	230	L/min
	Pressure	34.3	MPa
Bucket	Volume	1	m^3
Engine	Maximum torque	740	Nm
	RPM	1250	r/min
Excavator	Mass	23,000	Kg

cylinder, and the flow rate of pump 2 is supplied to the arm cylinder. The flow rate matches the action, and the load interference is avoided due to the adoption of double-pump power system. The instantaneous pressure of boom cylinder and bucket cylinder is related to the attitude and action state of excavator hydraulic boom. The flow rate of the pump is matched to the load demand. The pressure in the cylinder is related to the operation speed and the movement posture. According to the displacement measured by the boom cylinder and the arm cylinder, the trajectory curve of the bucket is calculated as shown in Fig. 11. The inconsistency of the signal of the artificial handle causes the trajectory curve of the bucket to not overlap, the change trend is consistent, and the trajectory curve is smooth. The bucket trajectory falls within the bucket workspace simulation Fig. 2.

Operate the excavator handle to control the compound action of boom cylinder and bucket cylinder. The data curve of test is shown in Fig. 12. It can be seen from Fig. 12 that the boom cylinder and bucket cylinder displacement can achieve a good linear compound action. The flow rate of pump 1 is supplied to the boom cylinder, and the flow rate of pump 2 is supplied to the bucket cylinder, and the flow rate matches the action. The pressure curves of boom cylinder and bucket cylinder are related to the posture and action of excavator hydraulic boom. When the pressure of the bucket cylinder is extended, the pressure of the rodless cavity rises first and then suddenly decreases. The experimental results show that the simulation of the bucket posture is correct.

According to the displacement measured by boom cylinder and bucket cylinder, the bucket trajectory curve is calculated as shown in Fig. 13. The trajectory trend of the bucket is the same, but the trajectory does not coincide completely due to the inconsistent signal of the operation handle.

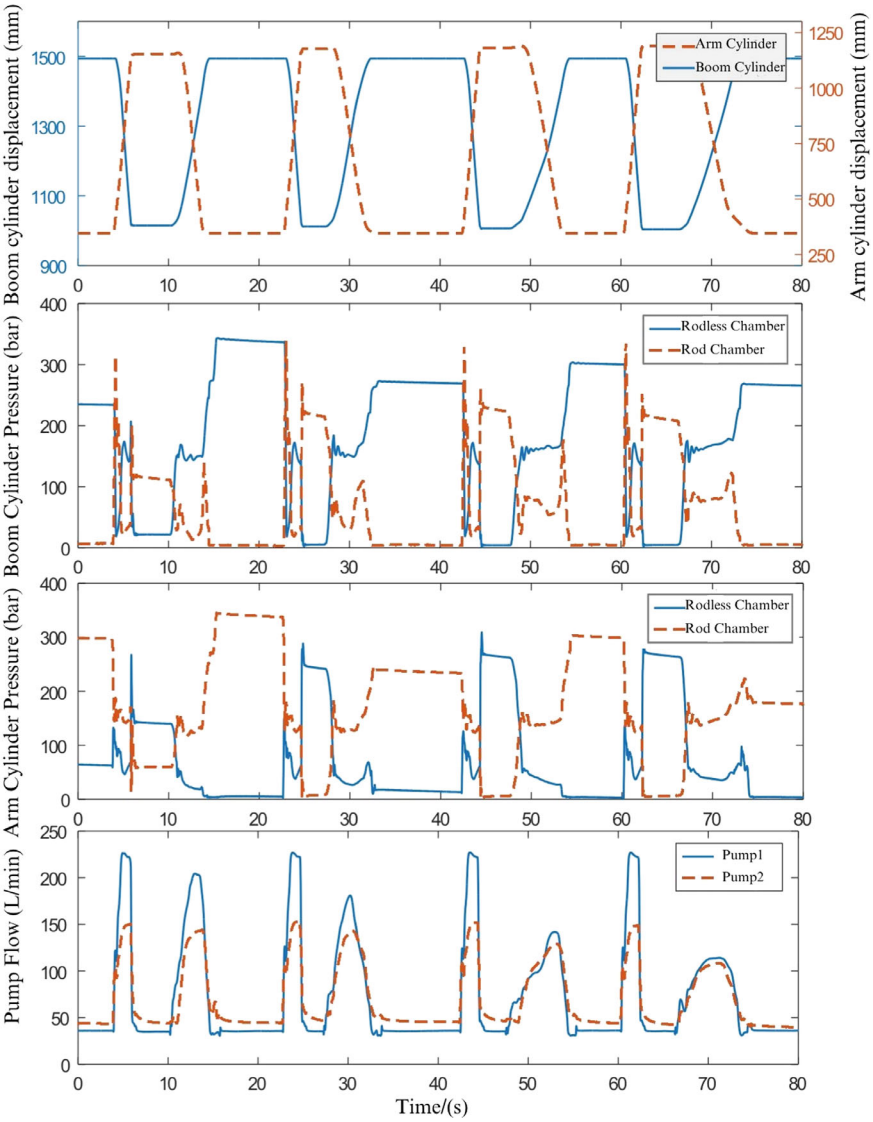


Fig. 10 Test displacement of boom and arm cylinder

5 Conclusion

(1) The medium-sized excavator is self-heavy, and the position and attitude of the operating mechanism affect the force of the three cylinders. The rodless cavity of the boom cylinder is always under pressure, and the static pressure varies between 5 and 12 MPa. The arm cylinder armless cavity will switch between

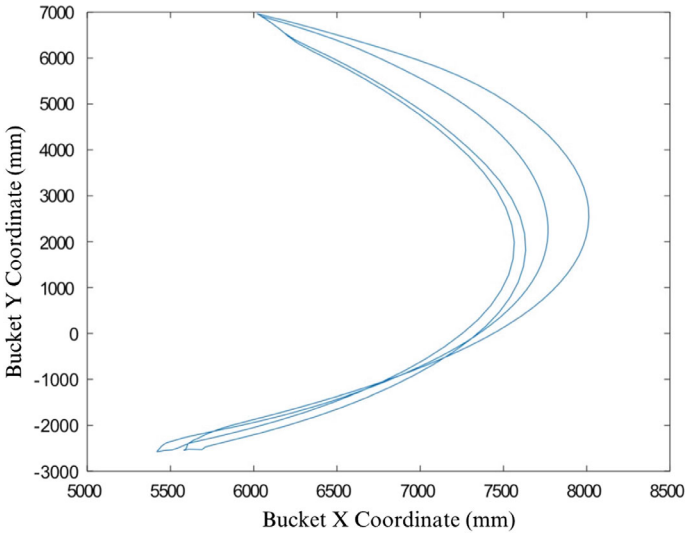


Fig. 11 Bucket trajectory of boom and arm simultaneous action

positive load and negative load, and the static pressure varies between −5 MPa and 5 MPa. The rodless cavity of the bucket cylinder will switch between positive load and negative load, and the static pressure varies between −2 MPa and 2 MPa. The simulation results conform to the variation law of experimental test data.

(2) The D-H parameter method is used to model and simulate the hydraulic boom of medium excavator. The cubic polynomial piecewise fitting the given trajectory, and the joint angle, joint angular velocity and joint angular acceleration curves of the three joints of the hydraulic boom are obtained, which provides a reference for the closed-loop control of the hydraulic boom of excavator.

(3) Experimental test is carried out on the principle prototype of medium-sized excavator, and the bucket trajectory is calculated according to test data, which verifies the simulation results of bucket workspace. When the bucket is retracted, it is affected by the negative load generated by the dead weight of the bucket, and the pressure of the rodless cavity is low. When braking, there is a pressure peak in the rod cavity, which can realize the rapid braking of the bucket. Improving the braking acceleration of the bucket can help the clay to break away from the bucket. Test data reflects the real-time dynamic hydraulic parameters of excavator hydraulic boom, which can provide reference for the real-time power matching of medium-sized excavator hydraulic boom by attitude.

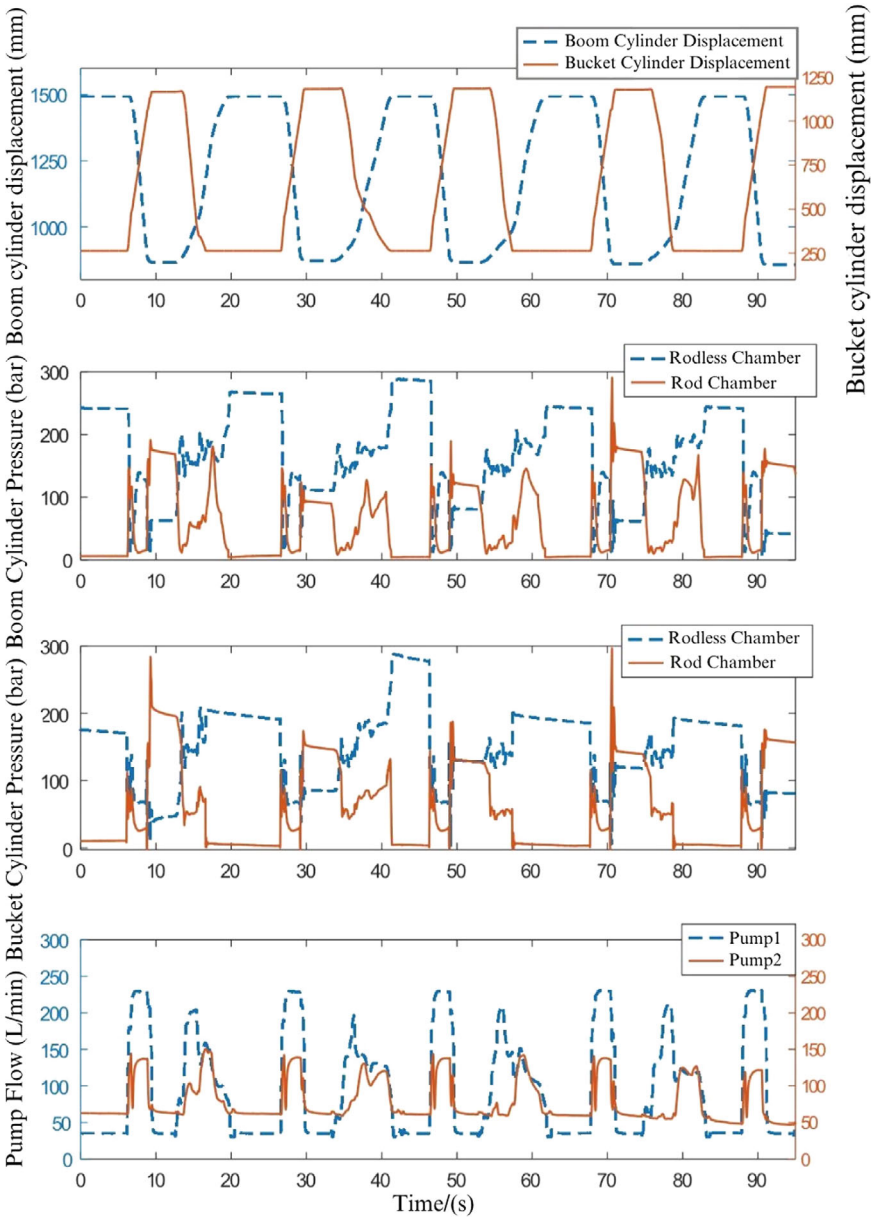


Fig. 12 Test displacement of boom and bucket cylinder

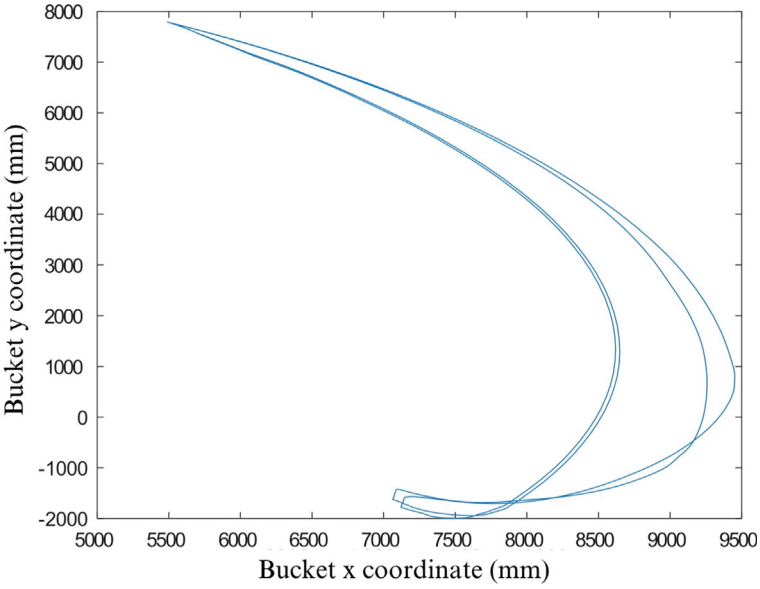


Fig. 13 Bucket trajectory of boom and bucket simultaneous action

Acknowledgements This research was supported by the National Key R&D Program of China (2021YFB2011905) and the Teaching and Research Project of Xuzhou University of Technology (YGJ2237).

References

1. Jing Z, Dongbao W, Guangping WU et al (2022) Mechanism design and motion analysis of heavy-load transfer robot with parallel four-bar mechanism. Trans Nanjing Univ Aeronaut Astronaut 39(5)
2. Zhangjing WD, Wu GP et al (2022) Mechanism design and motion analysis of heavy-load transfer robot with parallel four-bar mechanism. Trans Nanjing Univ Aeronaut Astronaut 39(5):606–618
3. Gao Y, Fang LJ, Jiang XJ et al (2022) Research on the precision synthesis method of a 7 DOF manipulator based on D-H parameters. Chin J Sci Instrum 43(4):137–145
4. Yin J, Xi WQ (2023) An active decoupling method for 6R robots based on Denavit-Hartenberg Parameters. Journal of Mechanical Transmission 47(12):89–96
5. Yang J, Gao Y, Guo R et al (2023) Research on excavator trajectory control based on hybrid interpolation. Sustainability 15(8):6761
6. Haga M, Hiroshi W, Fujishima K (2001) Digging control system for hydraulic excavator. Mechatron: Sci Intell Mach 6(11):665–676
7. Huang L, Xu WB, Yang YH et al (2023) Optimal trajectory planning for multi-objective joint space of excavator. Mod Manuf Eng (1):83–89+103
8. Qiu QY, Dou FJ, Guan C et al (2024) Research on trajectory tracking of hydraulic excavators considering multi joint time delay. J Harbin Eng Univ 5:1–8

9. Sun ZY, Liu S, Wang Y et al (2024) Time optimal and pulsation optimal trajectory planning of excavator. Mach Des Manuf 313(1):304–307

Open Access This chapter is licensed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License (http://creativecommons.org/licenses/by-nc-nd/4.0/), which permits any noncommercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if you modified the licensed material. You do not have permission under this license to share adapted material derived from this chapter or parts of it.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.



Weak Signal Detection of Construction Machines Based on Van Der Pol Oscillator



Ke-Feng Li and Qing Wang

Abstract This paper introduces a new approach based on the van der Pol oscillator for detecting weak signals in powertrain fault diagnosis. In contrast to the detection technique relying on the Duffing oscillator, this proposed approach focuses on identifying weak signals by examining state transitions from non-periodic to periodic states. Numerical simulations demonstrate that the van der Pol oscillator-focused approach offers significant advantages over the Duffing oscillator-based one, particularly regarding its enhanced sensitivity to weak signals under strong noise disturbance. Extensive investigations into the proposed approach further confirm the effectiveness of the van der Pol oscillator in weak signal detection.

Keywords Weak signal detection · Van der Pol oscillator · Bifurcation · Phase diagram · Powertrain fault diagnosis

1 Introduction

In the sectors of modern automobiles and heavy construction machinery, with the rapid development of intelligent transportation systems and autonomous driving technology, the importance of signal processing and fault detection has become increasingly prominent. Vehicles are affected by many factors during operation, including environmental noise, mechanical vibration, and electrical interference, which may cause signal weakening or distortion. Therefore, it is particularly important to develop an effective weak signal detection method. Weak signal detection entails extracting weak signals against a backdrop of strong noise. Conventional signal processing methods often have difficulty in effectively separating useful signals from noise in complex vehicle environments. Therefore, researchers have turned their attention to

K.-F. Li
Beijing Institute of Construction Mechanization Co., Ltd., Beijing, China

Q. Wang (✉)
Hygea Medical Technology Co., Ltd., Beijing, China
e-mail: wqhit084@126.com

© The Author(s) 2026

H. Saman et al. (eds.), *The 9th International Conference on Advances in Construction Machinery and Vehicle Engineering*, Lecture Notes in Mechanical Engineering,
https://doi.org/10.1007/978-981-96-5527-4_23

signal processing methods based on nonlinear dynamics, with the van der Pol oscillator emerging as a notable candidate thanks to its advantageous nonlinear characteristics. This oscillator operates as a self-excited oscillation system, adept at effectively capturing and amplifying weak signals. These signals can be effectively extracted in signal processing by taking advantage of such characteristics. This proposed method enhances the signal detectability and maintains high robustness in challenging noise environments. Recently, a growing focus on weak signal detection methods utilizing chaotic oscillators has emerged, as these oscillators often demonstrate effectiveness than conventional techniques (such as wavelet-based methods) due to their sensitivity to specific signals while maintaining immunity to noise interference.

Duffing oscillator has demonstrated its effectiveness in various applications for weak signal detection [1–5]. When adjusted to a critical state, this oscillator can be activated by weak periodical signals to shift from a state of chaos to one of periodicity. Wang [2] proposed a detection method that leverages the bifurcation behavior of the Duffing oscillator. Li [3] discussed the impact of the oscillator's restoring force on its capability to differentiate between signals and noise. Hu [4] employed the Duffing oscillator to identify characteristics indicative of early mechanical malfunction. Furthermore, Deng [5] introduced a sophisticated Duffing oscillator-based system specifically designed for detecting complex signal patterns.

Despite the substantial research conducted on the Duffing oscillator for weak signal detection, little attention has been paid to fully exploiting the potential of other nonlinear oscillators in this field. Further research is needed to determine whether other chaotic oscillators can also be effective for this purpose [4].

The van der Pol equation, since its inception in 1920, has emerged as a pivotal model for illustrating self-excited limit cycle oscillatory systems. This equation serves to characterize a diverse range of physical and biological phenomena. The periodically forced van der Pol oscillator [6] exemplifies the understanding of limit cycle behavior in response to periodic disturbances, exhibiting a spectrum of nonlinear behaviors. It is delineated by a three-dimensional parameter space comprising driving amplitude, driving frequency and damping factor, all of which contribute to the system's nonlinearity. In the absence of external driving forces, a van der Pol oscillator with positive damping exhibits stable limit cycles in phase space [7]. Conversely, the introduction of external driving forces can lead the system to exhibit periodic, quasiperiodic and aperiodic motions. To investigate global bifurcation characteristics of the forced van der Pol equation under conditions of origin offset, numerical studies have been conducted [8].

Despite extensive research conducted by many scientists and mathematicians [9, 10] on the van der Pol oscillator, its potential applications in weak signal detection have not been thoroughly investigated. This paper serves as the foundation for a series of studies aimed at exploring the effectiveness of the van der Pol oscillator in weak signal detection. The primary distinctions between this proposed method and the Duffing oscillator-focused one are as follows: (1) The former one is predicated on the transition from a non-periodical to a periodical one, while the latter one is from a chaotic state to a non-chaotic one; (2) The limit cycle state in the van der Pol oscillator enhances its robustness against noise; (3) The proposed oscillator

exhibits greater sensitivity to variations in forcing frequency compared to the Duffing oscillator, a characteristic attributable to the fundamental differences between the two systems. Consequently, the van der Pol oscillator presents advantages over the Duffing oscillator in weak signal detection.

This paper introduces a method for detecting weak signals using the van der Pol oscillator. Firstly, the dynamical behavior of the oscillator is analyzed. Secondly, the signal detection process is carried out by examining changes in phase trajectories through the van der Pol framework. Finally, challenges inherent to this signal detection method are discussed. Numerical simulations indicate the effectiveness of the proposed method in identifying weak signals.

2 Dynamical Analysis of Van Der Pol Oscillator

The van der Pol oscillator is an oscillator with nonlinear damping governed by the second-order differential equation:

$$\ddot{x} + \mu(x^2 - 1)\dot{x} + x = 0 \tag{1}$$

where x represents the dynamical variable and μ ($\mu > 0$) serves as a parameter. This model was proposed by Balthasar van der Pol in 1920. Van had examined the response of the van der Pol oscillator to a periodic forcing in his paper, which can be formulated as:

$$\ddot{x} + \mu(x^2 - 1)\dot{x} + x = F\cos(\omega t) \tag{2}$$

The system contains two frequencies: the self-oscillation frequency determined by μ, and the frequency of the periodic forcing. Chaos can also emerge in the system when its nonlinearity becomes sufficiently strong.

The classic state-space form of the van der Pol oscillator is:

$$\begin{aligned} \dot{x} &= y \\ \dot{y} &= \mu \cdot y(1 - x^2) - x \end{aligned} \tag{3}$$

The fixed points of Eq. (3) are:

$$\begin{aligned} y^* &= 0 \\ \mu \cdot y^*(1 - (x^*)^2) - x^* &= 0 \end{aligned} \tag{4}$$

Hence, the fixed points of the system are (xP^*P, yP^*) = (0,0). Evaluating the Jacobian of Eq. (3) at the equilibrium points, we get Eq. (5).

$$J = \begin{pmatrix} 0 & 1 \\ -1 & \mu \end{pmatrix} \tag{5}$$

When $\mu > 0$, the system exhibits complex conjugate eigenvalues with positive real parts; when $\mu < 0$, it exhibits those with negative real parts. According to [11], a Hopf bifurcation is observed when $\mu = 0$.

Guo's study [12] show the forced van der Pol oscillator exhibits diverse nonlinear phenomena, including chaos, under different parameter setting in Eq. (2). Specifically, for $\mu = 6$, $F = 2.5$, and ω=3, the oscillator was shown to operate chaotically.

When $\mu = 3$ and $\omega = 1$, the state of the van der Pol system state evolves as the parameter F increases. Figure 1 depicts the phase portrait for $F = 0.9$, while Fig. 2 represents the phase portrait for $F = 2.0$. Simulations were conducted over 600 s, with a step size of 0.01, disregarding the initial 100 s of data. The results indicate that as F increases, the system transitions from a non-periodic state to a periodic state.

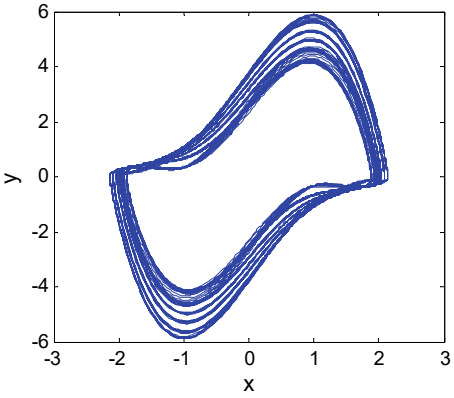


Fig. 1 Phase plane diagram of non-periodic motion when $F = 0.9$

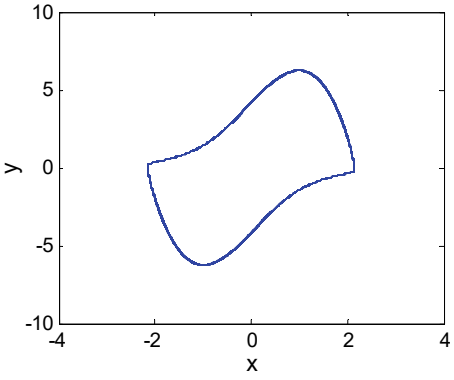


Fig. 2 Phase plane diagram of periodic motion when $F = 2.0$

3 Proposed Weak Signal Detection Method

In a forced van der Pol oscillator, the system can undergo a transition in its state when subjected to a small external periodical force, particularly when it is in a critical state. This state is typically linked to the bifurcation value of the parameter F. A bifurcation diagram is utilized to calculate the bifurcation values of the forced van der Pol oscillator. For the subsequent analysis, the parameters are set to $\mu = 3$ and $\omega = 1$.

To depict global dynamics of the Eq. (2) and its dependence on the parameter F, bifurcation diagrams have been constructed. These diagrams plot points from the Poincaré section against a specific system parameter. Figure 3 displays the bifurcation diagram for Eq. (2), with 50 points from the Poincaré section used for each F value. The parameter F ranges from 0.01 to 8, increasing steps of 0.01 s. As illustrated in Fig. 3, the forced van der Pol oscillator exhibits bifurcation, with the bifurcation value F_b approximately to 1.2. When the parameter F surpasses F_b, the system undergoes a transition from a non-periodic state to a periodic state.

The fundamental concept underlying the proposed method, akin to that of the Duffing oscillator, posits that the forced van der Pol oscillator can detect a small periodic signal in noise by transitioning from a non-periodic motion to a periodic motion. The bifurcation diagram provides an approximate value of the critical state, while the numerical simulation method is used to determine the precise bifurcation value [13–15].

When applying the forced van der Pol oscillator in weak signal detection, the equation is:

$$\begin{aligned} \dot{x} &= y \\ \dot{y} &= \mu \cdot y(1 - x^2) - x + F\cos(\omega t) + s(t) \end{aligned} \tag{6}$$

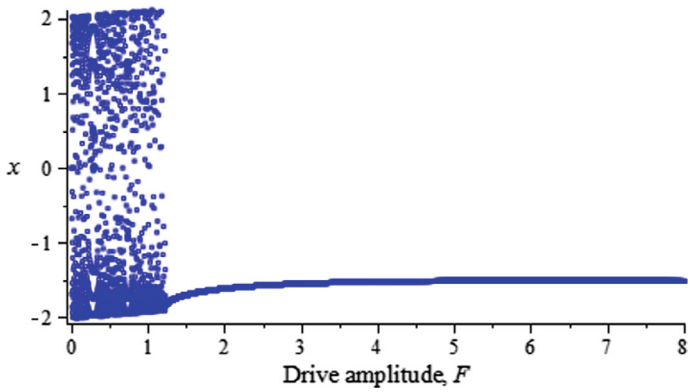


Fig. 3 Bifurcation diagram

where $s(t)$ represents an external perturbation, also expressed as $F_1 \cos(\omega_1 t + \varphi) + n(t)$, and $F_1 \cos(\omega_1 t + \varphi)$ is the target signal. $n(t) = \sigma \bullet \varepsilon(t)$ refers to Gaussian noise with a variance of 1 and a mean of 0.

The threshold of the van der Pol oscillator was determined using a bifurcation diagram in conjunction with numerical experiments. When $h = 0.01$, the threshold value F_b is calculated as 1.22 in the absence of external forces. Figure 4 shows the phase plane trajectories at this critical state. Upon introducing a weak periodic signal with an amplitude of $F_b = 0.01$, the resulting phase plane trajectories is shown in Fig. 5a. As illustrated, the forced van der Pol oscillator transitions from a non-periodic state to a periodic state, demonstrating its sensitivity to weak periodical signals. In contrast, when pure Guassian noise with $\sigma = 1.2$ is applied, the phase plane trajectories, as shown in Fig. 5b, indicate that the oscillator maintains a non-periodic state, suggesting its resilience to noise perturbations.

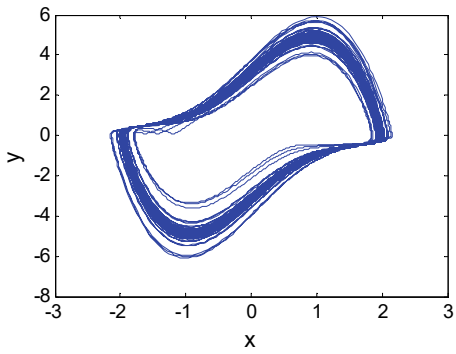
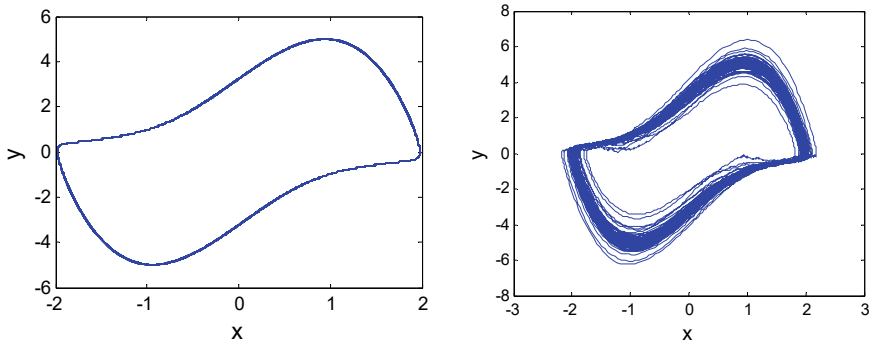


Fig. 4 Phase plane diagram when $F = 1.22$



(a) F_b =0.01, σ=0 (b) F_b =0, σ=1.2

Fig. 5 Sensitivity of van der Pol oscillator to weak signals and immunity to noise ($F = 1.22$) **a** $F_b = 0.01$, $\sigma = 0$; **b** $F_b = 0$, $\sigma = 1.2$

4 Further Discussions

This section discusses weak signal detection using the van der Pol oscillator, aiming to identify a weak sine signal that matches the oscillator's forcing frequency.

$$\ddot{x} + 3(x^2 - 1)\dot{x} + x = 1.22\cos t \tag{7}$$

The van der Pol system is in a critical state. Adding an external signal $s(t)$ yields:

$$\ddot{x} + 3(x^2 - 1)\dot{x} + x = 1.22\cos t + s(t) \tag{8}$$

4.1 Impact of Noise

When the external signal is Gaussian noise $s(t) = \sigma \bullet \varepsilon(t)$, its ability to induce a phase state shift is investigated. By considering ε(t) ~ N(0,1) with the σ values of 0.9, 2.0, 5.0 and 9.0, the solutions of Eq. (8) are analyzed. As illustrated in Fig. 6, the system maintains in a non-periodic state, indicating that Gaussian noise does not cause a phase state shift [16, 17].

4.2 The Ability to Detect Periodical Signals in Strong Noise

The external signal is given by $s(t) = 0.01\cos(t) + \sigma \bullet \varepsilon(t)$, with σ increasing from 0.01 to examine weak signal detection in strong noise. Numerical simulations show that the weak sine signal is detectable when $\sigma \leq 0.12$. Figure 7 shows the phase plane diagrams of the forced van der Pol oscillator when $\sigma = 0.12$ and $\sigma = 0.18$. Consequently, the threshold of the signal-to-noise ratio (SNR) for weak signal is greater than:

$$SNR = 10\log\left(0.5\frac{a^2}{\sigma^2}\right) = 10\log\left(0.5\frac{0.01^2}{0.12^2}\right) = -56.6296$$

4.3 Impact of Different Frequency Signals

When the frequency of an external signal deviates from that of a reference signals, as exemplified by the equation $s(t) = 0.3\cos[(1 \pm \Delta\ \omega)t]$, with ω values of ± 0.001, ± 0.002, ± 0.003, ± 0.005, ± 0.01, ± 0.02, and ± 0.05, the Eq. (8) is employed

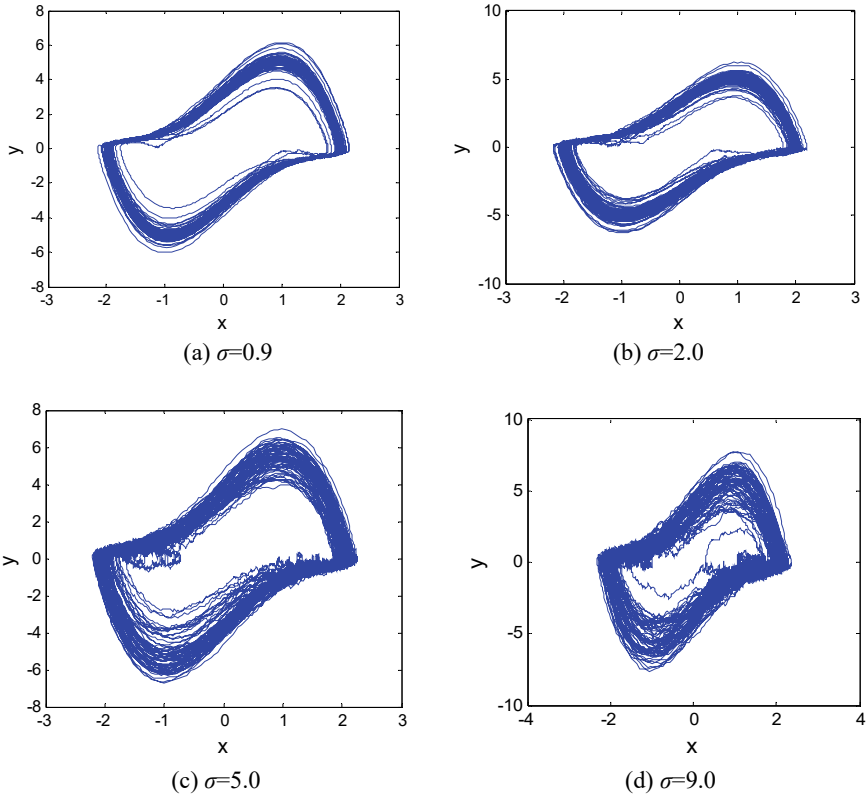


Fig. 6 State of the system with perturbation of Gaussian noise: **a** $\sigma = 0.9$; **b** $\sigma = 2.0$; **c** $\sigma = 5.0$; **d** $\sigma = 9.0$

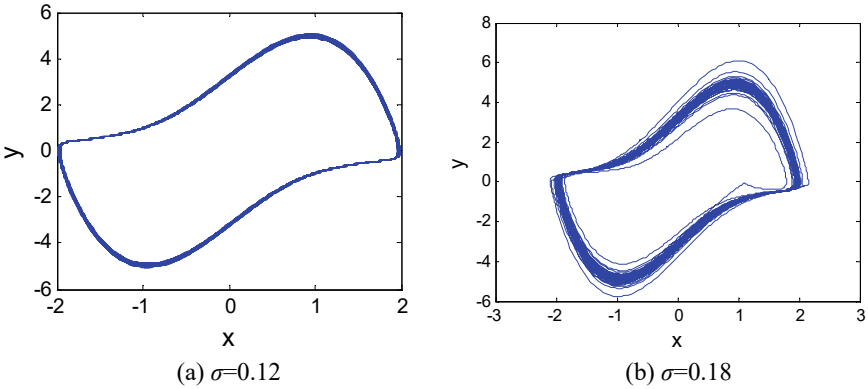
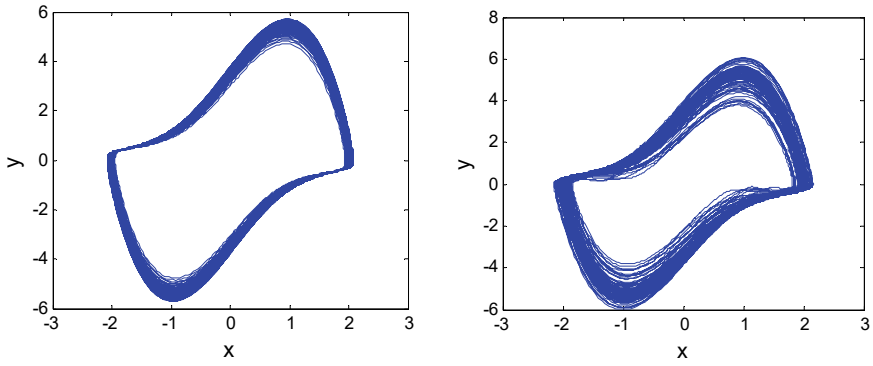


Fig. 7 Phase plane diagram for $F1 = 0.01$, **a** $\sigma = 0.12$, **b** $\sigma = 0.18$



(a) $\Delta\ \omega$=0.003 (b) $\Delta\ \omega$= - 0.008

Fig. 8 Phase plane diagram **a** $\Delta\ \omega = 0.003$; **b** $\Delta\ \omega$ = -0.008

to assess the impact on the orbits. The phase plane diagrams corresponding to $\Delta\ \omega = 0.003$ and $\Delta\ \omega = -0.008$ are demonstrated in Fig. 8. The results reveal that the system maintains in a non-periodic state. Numerical simulations reveal that the state transitions occur within the range $|\Delta\ \omega| \leq 0.002$. This characteristic signifies an enhancement over the Duffing oscillator, which experiences state transitions within the broader range $|\Delta\ \omega| \leq 0.05$, as referenced in [4]. Furthermore, as discussed in [7], alterations in the driving frequency ω and damping parameter μ result in modifications to the bifurcation structure of the van der Pol oscillator. This observation implies that the system's behavior is sensitive to variations in the driving frequency ω.

4.4 Impact of the Initial Phase

To evaluate the initial phase impact of the external signal, consider the reference signal as $F_b \cos(\omega t)$ and the weak signal as $a\cos(\omega t + \varphi)$. The combined periodic excitation force can then be expressed as:

$$\begin{aligned} A(t) &= F_b \cos t + a \cos(t + \varphi) \\ &= (F_b + a \cos \varphi) \cos t - a \sin \varphi \sin t \\ &= f(t) \cos(t + \phi(t)) \end{aligned} \tag{9}$$

where $f(t) = \sqrt{F_b^2 + 2F_b a \cos \varphi + a^2}$, $\phi(t) = arctg(a \sin \varphi / (F_b + a \cos \varphi))$

Equation (9) shows that the phase shift is influenced by the initial phase of the periodic exciting force. When $\pi - \arccos(a/2F_b) \leq \varphi \leq \pi + \arccos(a/2F_b)$ and $f(t) \leq F_b$, the orbit transition does not occur.

5 Conclusion

Weak signal detection represents a vital technique with broad applications across various fields. This paper presents an innovative approach for weak signal detection based on the van der Pol oscillator by identifying transition from a non-periodic state to a periodic state. Furthermore, significant challenges associated with the application of the van der Pol system for weak signal detection have been examined. Unlike the weak signal detection method based on the Duffing oscillator, the proposed method demonstrates notable advantages, including enhanced sensitivity to frequency variations and robust resilience to high-noise environments. A comprehensive analysis of various factors influencing this proposed method, such as noise impact, frequency discrepancies, and phase shifts, has been conducted. The findings substantiate the feasibility of employing the van der Pol oscillator for weak signal detection and demonstrate its effectiveness in practical applications, particularly in construction machine.

References

1. Liu HB, Wu DW, Jin W et al (2013) Acta Phys Sin 62(5)
2. Wang G, Chen D, Lin J et al (1999) IEEE Transactions on 46(2):440–444
3. Li Y, Yang BJ, Yuan Y et al (2007) Chin Phys 16(4):1072
4. Hu NQ, Wen XS (2003) J Sound Vib 268(5):917–931
5. Deng XY, Liu HB, Long T (2012) Chin Sci Bull 57(17):2185–2191
6. Van Der Pol (1927) Lond, Edinb, Dublin Philos Mag J Sci 3(13):65–80
7. Mettin R, Parlitz U, Lauterborn W (1993) Int J Bifurc Chaos 3(06):1529–1555
8. Xu JX, Jiang J (1996) Chaos, Solitons Fractals 7(1):3–19
9. Guckenheimer J, Hoffman K, Weckesser W (2003) SIAM J Appl Dyn Syst 2(1):1–35
10. Xu Y, Agrawal OP (2013) Commun Nonlinear Sci Numer Simul 18(12):3575–3589
11. Hale JK, Koçak H (1991) Springer-Verlag, New York
12. Guo CX, Jiang QY, Cao YJ (2007) Chaos, Solitons Fractals 34(3):1014–1019
13. Lu S, He Q, Wang J (2019) A review of stochastic resonance in rotating machine fault detection. Mech Syst Signal Process 116:230–260
14. Gong S, Li S, Wang H et al (2021) Multi-frequency weak signal detection based on wavelet transform and parameter selection of bistable stochastic resonance model. J Vib Eng Technol 9(5):887–906
15. Li Z, Zhang P, Mu X et al (2019) Guiding and routing of a weak signal via a reconfigurable gravity-like potential. Photonics Res 9:1087–1093
16. Guo L, Kan E, Wu Y et al (2020) Noise suppression ability and its mechanism analysis of scale-free spiking neural network under white Gaussian noise. PLoS ONE 15(12)
17. Wang Z, Shi X (2020) Electric activities of time-delay memristive neuron disturbed by Gaussian white noise. Cogn Neurodynamics 14(1):115–124

Open Access This chapter is licensed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License (http://creativecommons.org/licenses/by-nc-nd/4.0/), which permits any noncommercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if you modified the licensed material. You do not have permission under this license to share adapted material derived from this chapter or parts of it.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.



Generating Adversarial Examples for Pedestrian Detectors in Dynamic Driving Scenarios



Zhuang Zhang, Lijun Zhang, Dejian Meng, Wei Tian, Ye Han, Kaikun Pei, and Zhou You

Abstract Pedestrian detection is a key component of vision-based systems in autonomous driving, critical for ensuring traffic safety. Recent research has shown that adversarial examples can mislead these systems, highlighting the need to evaluate and improve their robustness. However, most prior work assumes that adversarial examples can be clearly captured, overlooking a common real-world scenario: adversarial examples can be blurred by vehicle motion. This motion blur, frequently encountered in dynamic driving scenarios, can obscure adversarial examples and reduce the effectiveness of attacks. To address this issue, this paper proposes an adversarial example generation method based on an image motion blur model for driving scenarios. First, we develop the model using vehicle kinematics, revealing the non-uniform distribution of motion blur parameters for pedestrians at different positions. Based on this, we propose a motion blur model-guided adversarial example generation method that adapts to both the driving scenario and the pedestrian's position, ensuring that adversarial examples remain effective despite motion blur. Finally, we validate the proposed method through real-world vehicle experiments, demonstrating a 19.67% improvement in the average attack success rate across various vehicle speeds compared to existing methods. This method provides a practical tool for testing pedestrian detection robustness in autonomous driving.

Keywords Pedestrian detection · Adversarial examples · Autonomous driving · Motion blur · Vehicle kinematics

Z. Zhang · L. Zhang · D. Meng · W. Tian · Y. Han · K. Pei
School of Automotive Studies, Tongji University, Shanghai, China

Z. You (✉)
School of Mechanical Engineering, Tongji University, Shanghai, China
e-mail: youzhou3308@163.com

© The Author(s) 2026

H. Saman et al. (eds.), *The 9th International Conference on Advances in Construction Machinery and Vehicle Engineering*, Lecture Notes in Mechanical Engineering,
https://doi.org/10.1007/978-981-96-5527-4_24

1 Introduction

With the rapid development of deep learning technologies, particularly propelled by advanced models such as convolutional neural networks (CNN), image-based pedestrian detection techniques have also made significant progress, becoming a crucial component of modern autonomous driving systems to ensure safe driving [1, 2]. These systems rely on pedestrian detection to identify pedestrians in the surrounding environment, thereby avoiding potential collision risks. However, as research advances, an increasing number of studies have revealed the high vulnerability of deep learning models to adversarial examples [3, 4]. Adversarial examples are generated by applying small perturbations to input images, which can lead to incorrect predictions by pedestrian detection models, thereby posing potential safety risks. This vulnerability is particularly concerning in autonomous driving scenarios, as erroneous decisions can directly result in accidents.

Researchers [5] introduced the concept of adversarial examples by making subtle pixel-level modifications to digital images, successfully misleading image classification networks and exposing the vulnerabilities of deep learning models. Following this, researchers [6, 7] leveraged gradient information from these models to develop more efficient adversarial perturbations. Iterative approaches [8–11] further enhanced attack effectiveness by refining the gradient-based perturbations across multiple steps, improving both the robustness and precision of adversarial examples.

Despite the success of these methods in digital spaces, applying them to real-world pedestrian detection remains challenging. To enhance the attack performance in actual driving scenarios, researchers introduced the Expectation Over Transformation (EOT) framework. EOT involves incorporating random image transformations—such as rotation, scaling, and brightness adjustments—during the generation of adversarial examples, allowing these examples to adapt to various physical conditions while maintaining their attack effectiveness. For instance, Thys et al. [4] adjusted the size and brightness of adversarial patches, successfully evading pedestrian detection systems, while Hu et al. [12] considered the deformation of patches adhered to pedestrians, utilizing non-rigid transformations to enhance attack adaptability. Wang et al. [3] examined the target size distribution in driving scenarios, which is primarily characterized by smaller targets, and imposed constraints on the size distribution range.

However, existing studies [3, 4, 12] typically assume that adversarial examples can be clearly captured, overlooking the impact of high-speed motion and the resulting motion blur. In high-speed driving situations, rapid vehicle movement can lead to substantial motion blur in captured images, obscuring the carefully crafted perturbation patterns of adversarial examples. This limitation reduces their effectiveness, making them less applicable in real-world driving conditions.

To address this issue, this paper proposes a motion model-guided adversarial example generation method, aiming to improve the attack performance in driving scenarios.

The key contributions of this paper are as follows:

- Image motion blur model based on driving scenarios: An image motion blur model is developed based on vehicle kinematics to simulate vehicle motion under multi-dimensional driving conditions. This model captures the non-uniform distribution of motion blur parameters for pedestrians at different positions, providing a theoretical foundation for analyzing and introducing variations in motion blur effects on adversarial patches in dynamic driving scenarios.
- Adversarial example generation considering motion blur: An adversarial example generation method that uses a dynamic driving scenario model is proposed to generate and apply motion blur based on the driving scenario, ensuring the attack performance in real-world driving conditions. To the best of our knowledge, this is the first method to introduce a driving scenario model into the adversarial example generation pipeline.
- Improved attack performance in dynamic driving scenarios: The proposed method accounts for the effects of motion blur in real-world driving conditions, particularly at high speeds, ensuring that adversarial examples remain effective under varying vehicle speeds and driving scenarios.

The rest of this paper is organized as follows: Sect. 2 shows the image motion blur model and motion blur analysis. Section 3 shows the specific steps of the proposed adversarial example generation method. In Sect. 4, drive-by tests and ablative analysis are conducted to verify the effectiveness of the proposed method. Finally, Sect. 5 concludes this paper.

2 Image Motion Blur Model and Motion Blur Analysis

2.1 Image Motion Blur Model Based on Driving Scenarios

For the typical smart car with a forward-looking camera, we ignore the vertical movement of the vehicle and assume that the turning radius of the vehicle is much greater than the wheel base, then the difference between steering angles on the outside and inside front wheels is negligible. Thus, the vehicle with a forward-looking camera can be simplified to the bicycle-camera model shown in Figs. 1 and 2.

In the image motion blur model, we define three right-handed orthogonal coordinate systems. The origin of the world coordinate system is O_w, which is fixed on the ground. The origin of the vehicle coordinate system is O_v, which is fixed at the center of the rear axle of the vehicle. The X_v axis of the vehicle coordinate system is forward and on the longitudinal plane of symmetry, and the Y_v axis is lateral out the left side of the vehicle, and the Z_v axis is upward to the vehicle. The origin of the camera coordinate system is O_c, which is fixed to the optical center of the camera. The X_c and Y_c axes of the camera coordinate system are aligned parallel to the plane of the image.

The position of vehicle's CoG (Center of Gravity) at time t can be expressed as:

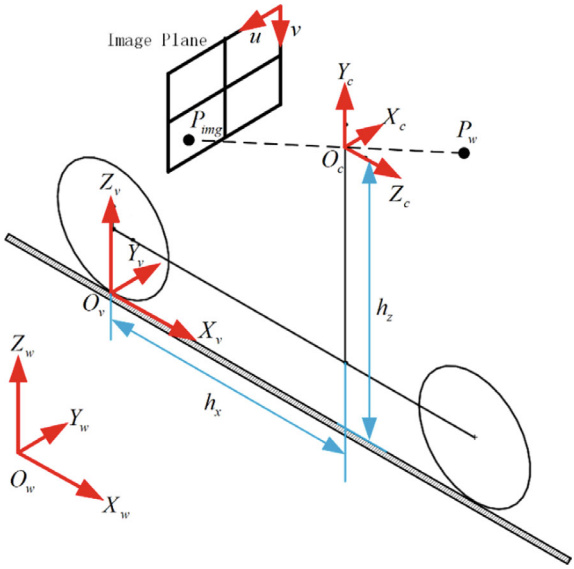


Fig. 1 Image motion blur model based on driving scenarios (axonometric view)

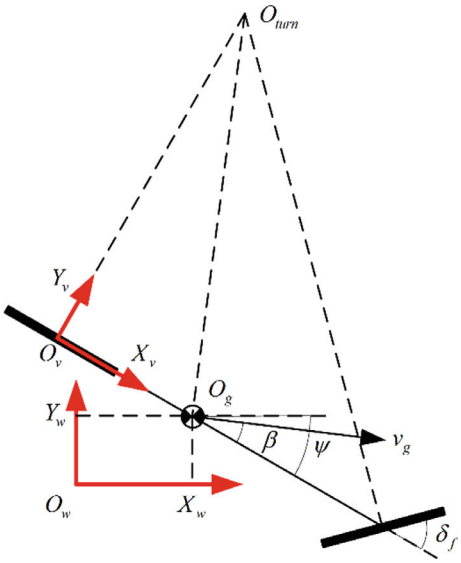


Fig. 2 Image motion blur model based on driving scenarios (bird-eye view)

$$x_{g,t} = x_{g,t-1} + v_{g,t-1} \cos\left(\psi_{g,t-1} + \beta\right) \times dt \tag{1}$$

$$y_{g,t} = y_{g,t-1} + v_{g,t-1} \sin\left(\psi_{g,t-1} + \beta\right) \times dt \tag{2}$$

$$\psi_{g,t} = \psi_{g,t-1} + \left(v_{g,t-1}/L_r\right) \sin(\beta) \times dt \tag{3}$$

$$v_{g,t} = v_{g,t-1} + a \times dt \tag{4}$$

where $O_{g,t}\left(x_{g,t}, y_{g,t}\right)$ is the coordinate of the vehicle's CoG at time t, $\psi_{g,t}$ is the yaw angle of the vehicle's CoG at time t and counterclockwise means positive, $v_{g,t}$ is the vehicle speed of the vehicle's CoG at time t, L is the wheelbase, L_r is the distance between the rear axle and vehicle's CoG, $\delta_{f,t}$ is the front wheel steering angle at time t, $\beta = \tan^{-1}\left((L_r/L)\tan\left(\delta_f\right)\right)$, a is the acceleration of the vehicle.

Coordinate of $O_v(x_v, y_v)$ can be expressed as:

$$x_v = x_{g,t} - L_r \cos\left(\psi_{g,t}\right) \tag{5}$$

$$y_v = y_{g,t} - L_r \sin\left(\psi_{g,t}\right) \tag{6}$$

For the camera, the camera internal parameters can be expressed as (7).

$$M_{in} = \begin{bmatrix} fW/w & 0 & W/2 \\ 0 & fH/h & H/2 \\ 0 & 0 & 1 \end{bmatrix} \tag{7}$$

where f is the focal length, $w \times h$ is the sensor size, and $W \times H$ is the image resolution.

Assuming the axes of the camera coordinate system align with the corresponding axes in the vehicle coordinate system, the external parameter matrix is presented in (8).

$$M_{\text{ex}} = \begin{bmatrix} 1 & 0 & 0 & -y_v \\ 0 & 1 & 0 & -h_z \\ 0 & 0 & 1 & -(x_v + h_x) \\ 0 & 0 & 0 & 1 \end{bmatrix} \tag{8}$$

where h_x and h_z are the installation offset distances of the camera in the X_v and Z_v directions.

According to the above relationship, for an object $P(x_w, y_w, z_w)$ in the world system, the corresponding coordinate in image plane $P_{img}(u, v)$ is:

$$\begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \begin{bmatrix} fW/w & 0 & W/2 & 0 \\ 0 & fH/h & H/2 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 & -y_v \\ 0 & 1 & 0 & -h_z \\ 0 & 0 & 1 & -(x_v + h_x) \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} y_w \\ z_w \\ x_w \\ 1 \end{bmatrix} \tag{9}$$

where x_v and y_v can be solved by (5) and (6).

The motion length *Len* and direction *Ang* in frame i can be calculated in (10) and (11).

$$Len = \sqrt{(du)^2 + (dv)^2} \tag{10}$$

$$Ang = \tan^{-1}(dv/du) \tag{11}$$

where f_r is the camera frame rate, f_s is the shutter speed, du and dv can be defined as $du = u(i/f + 1/fs) - u(i/f)$, and $dv = v(i/f + 1/fs) - v(i/f)$.

2.2 Motion Blur Analysis Based on Image Motion Blur Model

To conduct a thorough analysis of the motion blur characteristics in driving scenarios, this study employs the Monte Carlo simulation method based on the image motion blur model established in the previous section. The simulation aims to reveal the motion blur characteristics in driving scenarios by mimicking dynamic interactions in real-world road environments, laying a foundation for the subsequent development of adversarial example generation algorithms that take motion blur into account.

Specifically, we designed a representative simulation scenario: a pedestrian standing on the side of the road while a driver passes by in a straight line. This scenario is not only typical and common but also effectively reflects various complex factors in road traffic, such as variations in vehicle speed, adjustments to camera shutter speed, and diversity in pedestrian positions. By simulating this scenario, we can gain a deeper understanding of the manifestations and influencing factors of motion blur in real-world driving situations.

During the simulation process, we meticulously selected multiple vehicle speeds, camera shutter speeds, and pedestrian positions as key scenario factors (detailed factors are provided in Table 1). The selection of these factors is based on statistical analysis of actual driving data and theoretical derivations, ensuring that the simulation results comprehensively and accurately reflect the combined effects of vehicle motion, in-vehicle equipment performance, and environmental target characteristics on motion blur.

The simulation results are presented through Figs. 3 and 4. Specifically, Fig. 3 illustrates the distribution of target motion blur lengths. In this figure, the horizontal

Table 1 Factors in driving scenario model

Symbol	Meaning	Distribution range
$v_{g,0}$	Initial velocity of the car	0 to 60 (km/h)
f_s	Shutter speed	30 to 60 (Hz)
x_w	X coordinate of object	1 to 21 (m)
y_w	Y coordinate of object	−1.5 to −3 (m)
z_w	Z coordinate of object	0 to 2 (m)

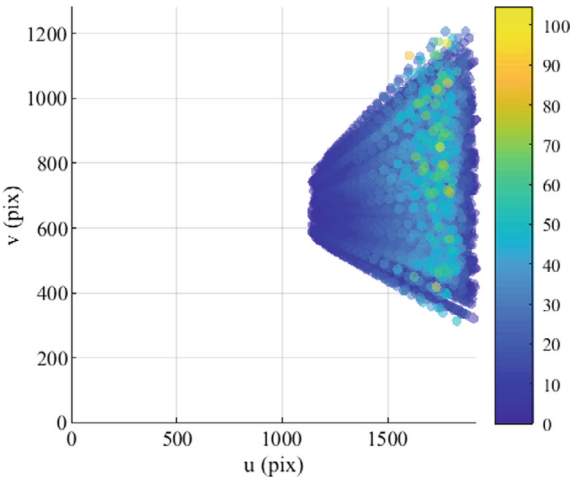


Fig. 3 Motion length

and vertical axes represent the pixel positions of the target in the image, while the color variation of the scatter points reflects the motion blur length at the corresponding positions. As shown in this figure, the motion blur lengths vary unevenly across pixel positions: targets located near the center of the image exhibit shorter motion blur lengths, whereas the blur intensity increases substantially closer to the image boundaries. Figure 4, on the other hand, focuses on the distribution of target motion blur angles. Similarly, the horizontal and vertical axes indicate the pixel coordinates of the target in the image, and the scatter point colors are used to distinguish different motion blur angles. This figure reveals that there are also differences in the motion blur angles of targets at different pixel positions. Through a comprehensive analysis of Figs. 3 and 4, we uncover an important new finding regarding the characteristics of target motion blur: both the length and angle of motion blur vary significantly with the target's position, exhibiting a non-uniform distribution.

The discovery of the non-uniformity of motion blur parameters in this section provides important guidance for the development of adversarial example generation strategies. In traditional adversarial example generation pipelines, the impact of motion blur on attack effectiveness is often overlooked. A simple method to improve the robustness of adversarial examples under motion blur conditions is to apply multiple motion blur transformations to them. This approach ensures that the resulting adversarial examples retain strong attack effectiveness even when motion blur is present. However, our experimental results indicate that the distribution of motion blur parameters exhibits specific non-uniform characteristics. This finding implies that in the process of generating adversarial examples, we cannot simply assume that motion blur is uniformly distributed; instead, we must precisely apply transformations based on the actual non-uniform distribution. Therefore, our research provides effective guidance for the development of adversarial example generation strategies in the next section.

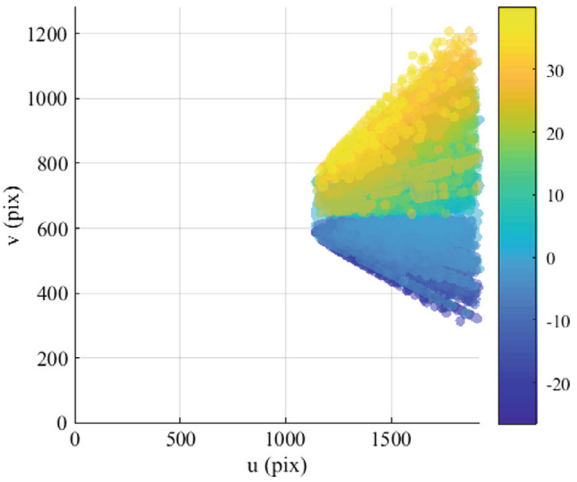


Fig. 4 Motion angle

3 Adversarial Example Generation Driven by a Motion Blur Model in Dynamic Driving Scenarios

Building upon a comprehensive analysis of motion blur characteristics within dynamic driving scenarios, this section innovatively proposes a framework for adversarial example generation guided by a motion blur model in dynamic driving scenaris. By integrating the spatiotemporal dynamics of driving environments into the generation process of adversarial patches, this framework significantly enhances the attack performance of adversarial examples in complex real-world environments, effectively overcoming the limitations of existing methods that rely on static scene assumptions.

As illustrated in Fig. 5, the proposed method first applies a series of transformations to the adversarial patch, including color transformation, perspective transformation, and motion blur. These transformations not only simulate the image distortions caused by equipment and environmental changes in real driving scenarios but also set adaptive blur parameters based on the motion characteristics of the driving scene. Unlike traditional methods that rely on random parameter settings, our approach precisely captures the non-uniform distribution of motion blur through scenario modeling, making the generated adversarial patches more realistic and effective. Subsequently, the transformed patches are superimposed onto the scene images and fed into the target detector. Based on the detector's output, the method designs an attack target function that considers both detection confidence and false positive rates. This attack target function is optimized via backpropagation to iteratively adjust the patch, thereby enhancing the attack effectiveness and robustness.

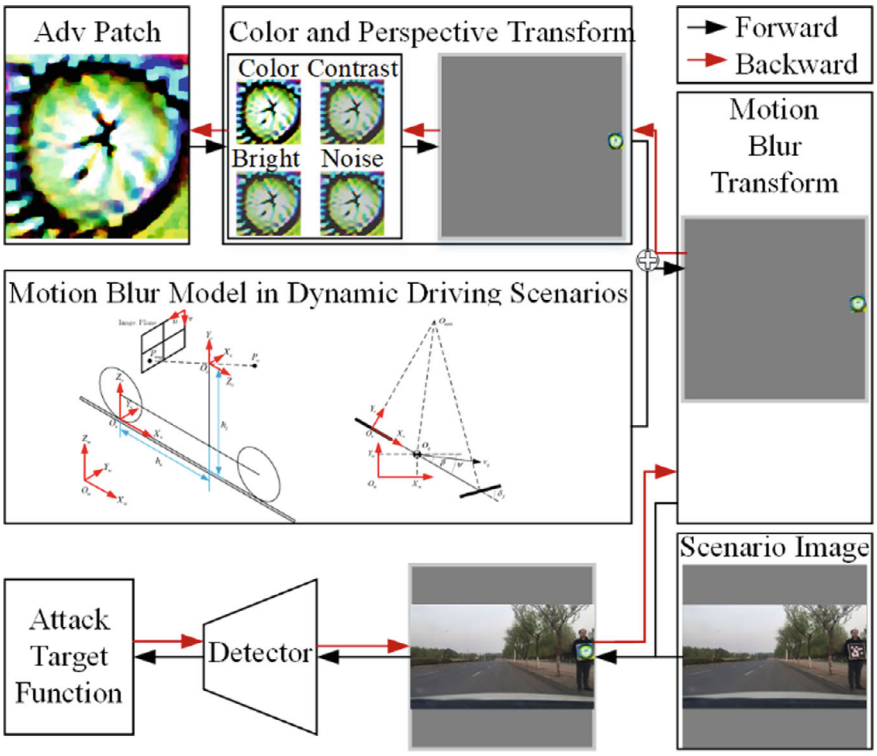


Fig. 5 Pipeline of adversarial example generation method based on a motion blur model

3.1 Color and Affine Transformations

To ensure that the adversarial patch retains its attack performance under varying lighting conditions and imaging device distortions in real-world driving environments, this method adopts a multi-level strategy for color and affine transformations. First, to address the color deviations caused by printing devices, this section captures a color calibration board [13, 14] using a camera and utilizes a neural network to establish a fine-grained mapping between standard colors and printed colors, simulating real-world color distortions. In addition, by applying contrast, brightness, and noise transformations, the method further simulates the diverse lighting conditions encountered in driving scenarios. To model potential changes in the adversarial patch's position during vehicle movement, the method introduces an affine transformation strategy based on the geometric properties of the scene, ensuring that the patch remains effective across different perspectives.

3.2 Motion Blur Transformations

Unlike traditional adversarial example generation methods that ignore the high dynamics of driving scenarios, the proposed method integrates the physical characteristics of dynamic driving environments by automatically adjusting motion blur parameters through the scenario model. Specifically, this section leverages the distribution of scene parameters shown in Table 1 to compute the mapping between target positions and motion blur parameters (including blur length and angle), as illustrated in Figs. 3 and 4. Based on the position of the patch after the affine transformation, the method accurately applies motion blur effects that correspond to the scene dynamics, ensuring that the adversarial patch can still evade the detector's defenses in real-world dynamic scenarios.

3.3 Attack Target Function

After various transformations in the above part, input the adversarial example to object detector, and then update the adversarial example based on attack target function.

L_{nps} is the non-printability score, which is an index used to measure the degree to which the color in the patch can be expressed by the printer. Denote P_{patch} as the colors in the patch and C_{print} as colors in the printable color set, L_{nps} can be calculated as:

$$L_{nps} = \sum_{P_{patch} \in p} \min_{C_{print} \in C} \left| P_{patch} - C_{print} \right| \tag{12}$$

L_{tv} is total variation in the image. This loss guarantees that our optimizer tends to smooth pictures with color changes while avoiding noisy pictures. It can be calculated as:

$$L_{tv} = \sum_{i,j} \sqrt{\left(\left(p_{i,j} - p_{i+1,j} \right)^2 + \left(p_{i,j} - p_{i,j+1} \right)^2 \right)} \tag{13}$$

L_{obj_cls} is the maximum objectness in the image. Our goal is to hide the pedestrians in the image, so we need to minimize the objectness and class score.

The attack target function is composed of the above three parts:

$$L = \alpha L_{nps} + \beta L_{tv} + L_{obj_cls} \tag{14}$$

In this paper, the weight coefficient α=0.01 and β=2.5, following the parameter configuration in [4], which has demonstrated effectiveness in similar scenarios.

4 Evaluation

4.1 Experimental Setup

To evaluate the attack performance of the proposed method in real driving scenarios, a real-world vehicle experiment was conducted in this study. The real-world vehicle experimental platform used is shown in Fig. 6, where an iPhone 6 s smartphone is placed in the center of the dashboard to record real car images.

The speeds of 5 km/h, 20 km/h, 40 km/h, and 60 km/h are chosen to represent a range of typical urban and suburban driving conditions, capturing scenarios from low-speed maneuvering to moderate and higher-speed cruising. In each experiment, recording starts at a distance of 18 m from the pedestrian and continues until the pedestrian is entirely out of the image. The real-world scene images at different speeds are shown in Fig. 7. It is shown that the captured scene images contain different intensity of motion blur, which is conducive to fully evaluating the adversarial attack effect under the influence of motion blur.

The metric is attack success rate. Following the approach in [4], the attack success rate is defined as the ratio of frames where a Stop sign went undetected to the total number of frames.

$$ASR = \frac{N_{attacked}}{N_{all}} \tag{15}$$

To comprehensively assess our method's effectiveness, this section presents a comparison with two types of methods: (1) a baseline without a patch and a white patch to demonstrate our method's advantage over non-attacked cases, and (2) adversarial patches from [4] to highlight the improvements our approach offers over existing methods.



Fig. 6 Vehicle and onboard camera

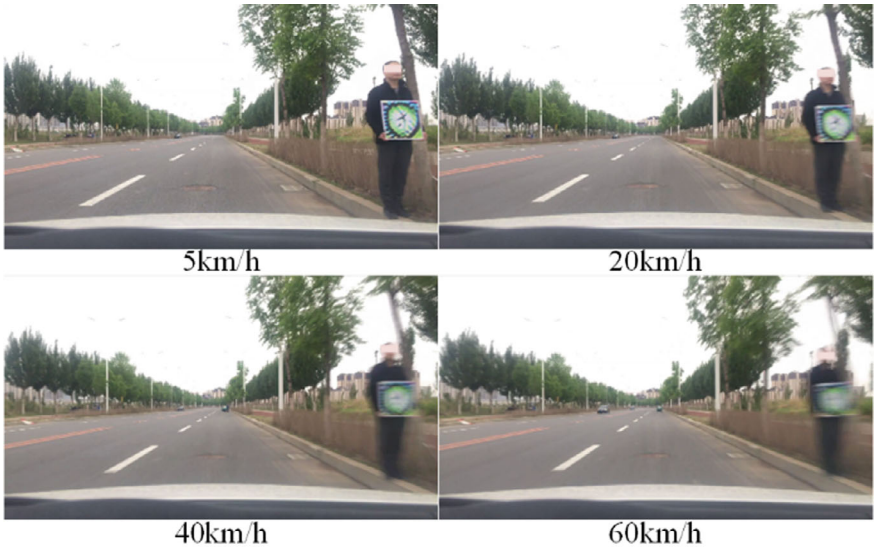


Fig. 7 Typical images in drive-by tests

4.2 Experimental Results

The average attack success rate of each adversarial example under four working conditions is calculated, and the results are shown in Table 2.

In Table 2, the ASR without attack (the no-patch and white-patch conditions) is about 40%, while the setting of [4] can increase the ASR to about 68% and 53%, which is about 28% and 13% higher than the case without attack, verifying that the existing method has a strong attack effect. The attack success rate of the proposed method reaches about 87.96%, which is about 40% higher than the case without attack, and 19.67% and 34.95% higher than the existing research, proving that the proposed method can effectively improve the attack performance.

This paper also fully compares the attack performance at different vehicle speeds. The results are shown in Table 3. It is shown that the ASR of the method proposed in this paper exceeds that of other methods at all vehicle speeds, and the ASR at all vehicle speeds exceeds 80%. Notably, under high-speed conditions of 60 km/h, the proposed method attains a 95% attack success rate, marking a 16.05% improvement over existing approaches. This result confirms the method's effectiveness in high-speed driving scenarios. Therefore, our research not only reveals new patterns in

Table 2 Average attack success rate (%)

Method	None	White	Obj	ObjCls	Ours
ASR (%)	41.17	39.40	68.29	53.01	**87.96**

Table 3 Attack success rate at different vehicle speeds

Method	ASR (%)			
	5 km/h	20 km/h	40 km/h	60 km/h
None	31.06	32.69	53.57	47.37
White	29.17	27.27	46.15	55.00
Obj	60.74	58.49	75.00	78.95
ObjCls	37.66	52.73	55.00	66.67
Ours	85.90	89.47	81.48	95.00

Table 4 Attack success rate in ablation analysis

Method	ASR (%)				
	5 km/h	20 km/h	40 km/h	60 km/h	Mean
Without motion blur model	80.77	65.45	71.43	80.00	74.41
With motion blur model	85.90	89.47	81.48	95.00	87.96

motion blur characteristics but also provides precise and effective guidance for the development of adversarial example generation strategies in the next section.

4.3 Ablation Analysis

To validate the necessity of incorporating the motion blur model in the proposed method, this section conducts an ablation experiment. The experiment compares adversarial examples generated with and without the dynamic driving scenario model, with the results presented in Table 4.

Table 4 clearly demonstrates that incorporating the motion blur model significantly enhances the attack success rate at various speeds, yielding an average increase of 13.55%. Remarkably, at a high-speed setting of 60 km/h, the attack success rate increased by 15%. These findings suggest that the integration of the motion blur model improves the effectiveness of adversarial examples in real-world driving scenarios, further confirming the model's ability to elevate attack success rates.

5 Conclusion

This paper presents a novel adversarial example generation method guided by an image motion blur model, specifically designed to address challenges in high-speed scenarios. Unlike existing methods that assume clear capture conditions, this

approach accounts for motion blur effects, enabling more realistic and robust adversarial examples under dynamic driving conditions. First, a motion blur model for dynamic driving scenario is established based on vehicle kinematics, revealing the non-uniform distribution characteristics of motion blur parameters in images. The proposed method then incorporates motion blur according to the scene information and pedestrian position, thereby enhancing the attack success rate in high-speed driving conditions. Experimental validation, including real-world tests and ablation studies, verifies that the proposed method attains a higher attack success rate at various vehicle speeds, with particularly notable improvements observed under high-speed conditions. This study focuses on motion blur during constant speed scenarios, highlighting the impact of vehicle speed on adversarial effectiveness. However, it does not address motion blur from turning or rapid speed changes, which are common in real-world driving. Future research should explore these dynamic conditions to enhance the understanding of motion blur effects on adversarial example generation.

Acknowledgements Project is supported by the Natural Science Foundation of Chongqing [CSTB2023NSCQ-MSX0063]

References

1. Zhang Y et al (2023) A small target pedestrian detection model based on autonomous driving. J Adv Transp 2023(1):5349965
2. Khan AH, Nawaz MS, Dengel A (2023) Localized semantic feature mixers for efficient pedestrian detection in autonomous driving, pp. 5476–5485
3. Wang N et al (2023) Does physical adversarial example really matter to autonomous driving? Towards system-level effect of adversarial object evasion attack. In: IEEE/CVF international conference on computer vision. IEEE, Paris, France
4. Thys S, Van Ranst W, Goedem TE (2019) Fooling automated surveillance cameras: adversarial patches to attack person detection. In: IEEE/CVF conference on computer vision and pattern recognition workshops
5. Macas M, Wu C, Fuertes W (2024) Adversarial examples: a survey of attacks and defenses in deep learning-enabled cybersecurity systems. Expert Syst Appl 238:122223
6. Gunasekaran M et al. (2023) Adversarial attacks against neural networks using projected gradient descent with line search algorithm. IEEE, pp 1–6
7. Goodfellow IJ, Shlens J, Szegedy C (2015) Explaining and harnessing adversarial examples. In: International conference on learning representations. San Diego, CA, USA OpenReview.net
8. Kurakin A, Goodfellow IJ, Bengio S (2018) Adversarial examples in the physical world. In: Artificial intelligence safety and security. Chapman and Hall/CRC, pp 99–112
9. Madry A et al (2018) Towards deep learning models resistant to adversarial attacks. in international conference on learning representations. Vancouver, BC, Canada OpenReview.net
10. Dong M, Xu C (2023) Adversarial robustness via random projection filters, pp 4077–4086
11. Saleh HM, Marouane H, Fakhfakh A (2024) Stochastic gradient descent intrusions detection for wireless sensor network attack detection system using machine learning. IEEE Access
12. Hu Z et al (2022) Adversarial texture for fooling person detectors in the physical world. In: IEEE/CVF conference on computer vision and pattern recognition
13. Koren N Color management and color science. http://www.normankoren.com/color_management.html

14. Sharif M, Bhagavatula S, Bauer L, Reiter MK (2016) Accessorize to a crime: Real and stealthy attacks on state-of-the-art face recognition. In: ACM sigsac conference on computer and communications security, pp 1528–1540

Open Access This chapter is licensed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License (http://creativecommons.org/licenses/by-nc-nd/4.0/), which permits any noncommercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if you modified the licensed material. You do not have permission under this license to share adapted material derived from this chapter or parts of it.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.



Research on Performance of Traction Winch Hydraulic System in Lifting Condition Based on AMESim



Zehao Chen, Junbo Ge, Jingying Xu, Jingqiao Liu, and Rui Guo

Abstract By constructing the simulation model of the hydraulic system of traction winch, this study focuses on the influence of cable friction, motor moment of inertia and motor displacement on the lifting performance of the winch. The results show that the moment of inertia and motor displacement have great influence on the winch speed, while the cable friction has relatively little influence on the lifting speed. By changing these key variables, the hoist performance and system stability can be significantly affected. The research results have certain guiding significance for the optimization of traction winch system.

Keywords Traction winch · AMESim · Hydraulic system · Two-point motor · Stability

1 Introduction

The traction winch is mainly composed of power device, winch, supporting box and control system [1]. In the process of operation, it is combined with the cable storage winch and the cable banding system to form a complete towing system. In addition, the traction winch is also divided into two-wheel single-drive structure and two-wheel dual-drive structure [2, 3]. Due to the hydraulic overload protection, the hydraulic winch has strong safety and reliability, and has been applied in industrial development because of its easy operation and smooth speed regulation [4]. Aiming at the no-load jitter of conventional hydraulic winch, Ju Chen proposed three anti-jitter measures through simulation analysis of electro-hydraulic control system. After the

Z. Chen (✉) · J. Ge · J. Xu · J. Liu
Marine Design and Research Institute of China, Shanghai, China
e-mail: 1010848115@qq.com

R. Guo
Yanshan University Lifting Machinery Key Technology National Key Laboratory, Qinhuangdao, Hebei, China

© The Author(s) 2026

H. Saman et al. (eds.), *The 9th International Conference on Advances in Construction Machinery and Vehicle Engineering*, Lecture Notes in Mechanical Engineering,
https://doi.org/10.1007/978-981-96-5527-4_25

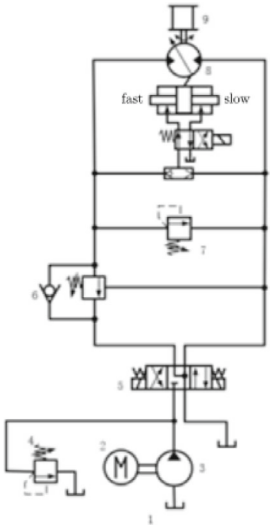
third scheme is actually adopted, the jitter problem is solved, which has a reference value for improving the stability of the hydraulic winch system [5].

Domestic and foreign research has been carried out on the hydraulic system of winch. Chen Jiahuan established a simulation model of the hydraulic system of Marine towing winch, simulated the effect of flow delay loading on the "reverse towing" phenomenon, and proposed an optimization scheme. After optimization, the maximum backtow speed and backtow distance of 630 kg load are reduced by 28% and 35% respectively. When 450 kg load is released, the system basically stops "backward dragging" and speed fluctuation is reduced [6]. Li designed the composition scheme of the lifting winch for ship transportation, calculated the hydraulic system of the 400 kN lifting winch, determined drum parameters, and calculated its operating characteristics, involving lifting, lowering and speed regulation processes [7]. Zhou et al. designed the structure of hydraulic winch with 3D software and introduced its working principle. ANSYS software is used to conduct three-dimensional modeling and stress deformation analysis of the drum, and it is verified that the mechanical structure strength and stiffness of the hydraulic winch drum meet the requirements [8]. Meng et al. use Amesim modeling to analyze the hydraulic viscous winch, and the results show that: (1) the disc brake is affected by the accumulator; (2) The viscous brake is affected by the throttle valve; (3) Good braking performance. This study provides theoretical support for winch design [9]. Yue et al. analyzed the structural characteristics, control loop, velocity solution and hydraulic motor drive form of the winch hydraulic system, and established a mathematical model system, which laid a foundation for the establishment of the system simulation model [10]. Meng et al. analyzed the hydraulic system of the water-viscous winch by simulation, and found that the accumulator affected disc brake pressure and flow rate, the throttle valve aperture affected the hydraulic viscous brake pressure and oil flow rate, and the deceleration during braking was small, so the system met the requirements and provided theoretical support for the optimization design.

The research object of this paper is the hydraulic system of traction winch. The simulation model is built based on AMESim to realize the simulation analysis of the hydraulic system of traction winch, and the influence of parameters such as moment of inertia, cable length and friction force on the performance of traction winch is analyzed, which has certain reference value for the optimization of its hydraulic system.

2 Hydraulic System Design

In this paper, the traction winch is a double reel hydraulic winch, which is composed of hydraulic motor, reducer, reel, brake device (composed of brake belt components and hydraulic cylinder), clutch device, spindle, hydraulic valve, common base, etc. With the continuous progress of technology, the inertia of the load required by the traction winch is also increasing, and the requirement for the accurate control of the



1- fuel tank 2- Motor 3- hydraulic pump 4- relief valve 5- electromagnetic directional valve 6- balance valve 7- relief valve 8- hydraulic motor 9- drum

Fig. 1 Schematic diagram of hydraulic system of traction winch

traction winch is also increasing. Therefore, it is necessary to analyze the influence of various parameters of traction winch on its performance.

The main body of the traction hydraulic winch is a motor, a hydraulic pump group, a control valve group and a mechanical device, and the hydraulic system schematic diagram is shown in Fig. 1. The power drive mode of the hydraulic system of the traction winch is driven by the motor 2 hydraulic pump 3. The reverse direction of the three-position four-way electromagnetic reversing valve realizes the positive and negative rotating work of the traction winch. When the three-position four-way solenoid valve is in the left position, the reel is turning, the reel is pulling, and the weight is lowered; When the three-position four-way solenoid valve is in the right position, the reel is reversed, the reel is pulled up, and the weight is lifted; When the three-position four-way solenoid valve is in the middle position, the system flow is unloaded directly through the relief valve.

3 Hydraulic System Modeling of Traction Winch Based on AMESim

AMESim can model complex systems in multiple disciplines and fields, and AMESim hydraulic system is modeled based on the hydraulic system schematic diagram of traction winch. Use the hydraulic component modules provided here,

such as the hydraulic pump model in the hydraulic library, the three-position four-way electromagnetic reversing valve model, the reducer model, the relief valve model, the motor model, and the traction mechanism model. The hydraulic component model is connected according to the oil circuit to form a traction winch simulation model, as shown in Fig. 2.

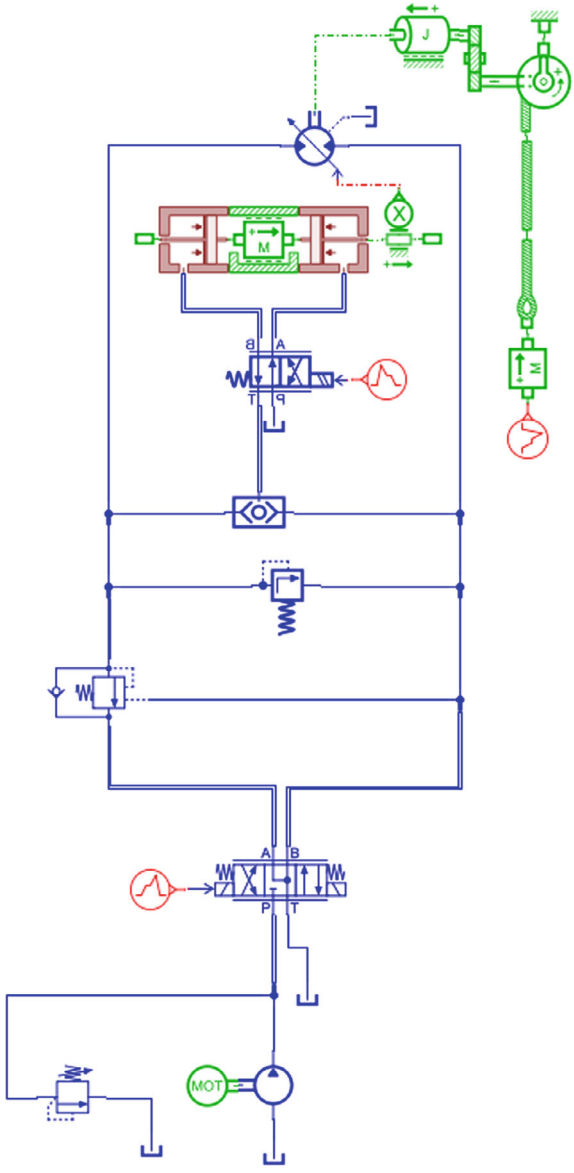


Fig. 2 AMESim simulation model of traction winch hydraulic system

The main simulation parameters of the traction winch model are shown in the following table:

4 Simulation Results and Analysis

Based on the built traction winch AMESim simulation model, the simulation is carried out in the running mode. The simulation parameters are shown in Table 1, and the simulation time is 0.4 s. From the analysis of Figs. 3 and 4, it can be seen that the winch system jitter is relatively stable during start-up, the stability time is 0.1 s, the maximum overshoot is 0.000837, and the sub-frequency of oscillation is small. When the hydraulic pump starts, the maximum pressure reaches 25 MPa (this is the set pressure of the system relief valve), and the vibration of the circuit is relatively severe.

Table 1 Main parameters of traction winch model simulation

Component parameter	Numerical value
Motor displacement mL/r	1848
System relief valve set pressure MPa	25
Reduction ratio	14.87
Lifting weight/kg	7000
Winch drum diameter/(mm)	550
Oil supply flow L/min	195

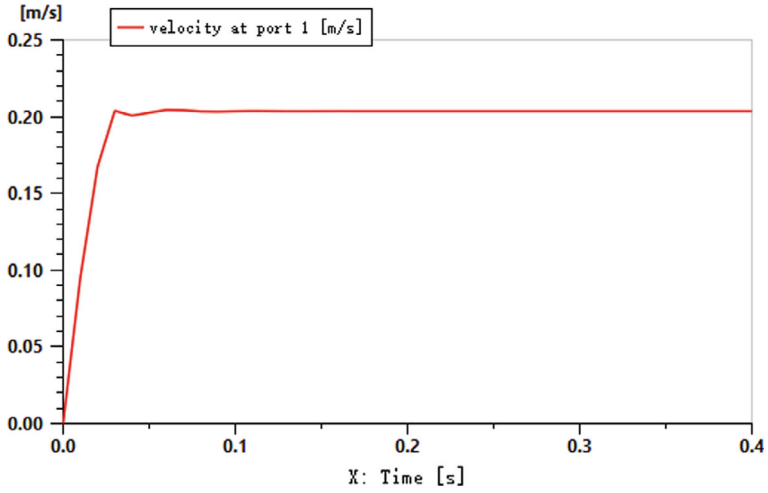


Fig. 3 Speed curve of heavy objects

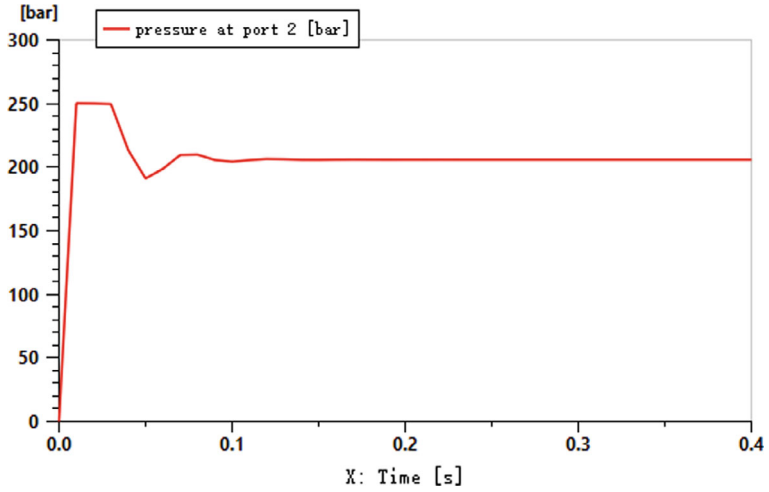


Fig. 4 Outlet pressure of hydraulic pump

The following is to analyze the influence of parameters such as characteristic flow rate, moment of inertia and motor displacement on the performance of traction winch at the maximum opening of the balance valve by changing the parameters of the simulation model.

(1) The influence of different cable friction on the traction winch system

Keep other parameters unchanged, change the cable friction for simulation, and simulate the speed of the weight.

When the friction force is 10^7 N,10^8 N, and 10^9 N respectively, the speed curve of the weight is shown in Fig. 5.The simulation results show that with the increase of cable friction, the dynamic response of weight speed slows down and the overshoot quantity increases. It can be seen that the increase of cable stiffness is not conducive to improving the stability and reliability of the system.

(2) Influence of characteristic flow on traction winch system at different maximum opening of balance valve

Keep other parameters unchanged, change the maximum opening of the balance valve to simulate the characteristic flow, and simulate the speed of the weight.

When the characteristic flow at the maximum opening of the balance valve is 50 L/min, 70 L/min, and 90 L/min respectively, the speed curve of the weight is shown in Fig. 6. According to the simulation results, with the increase of the characteristic flow rate at the maximum opening of the balance valve, although the dynamic response of the weight speed becomes faster, the overshot increases, the stability decreases obviously, and the fluctuation increases. It can be seen that the increase of characteristic flow when the maximum opening of the balance valve is not conducive to reducing the vibration of the traction winch and reducing the stability and reliability of the system.

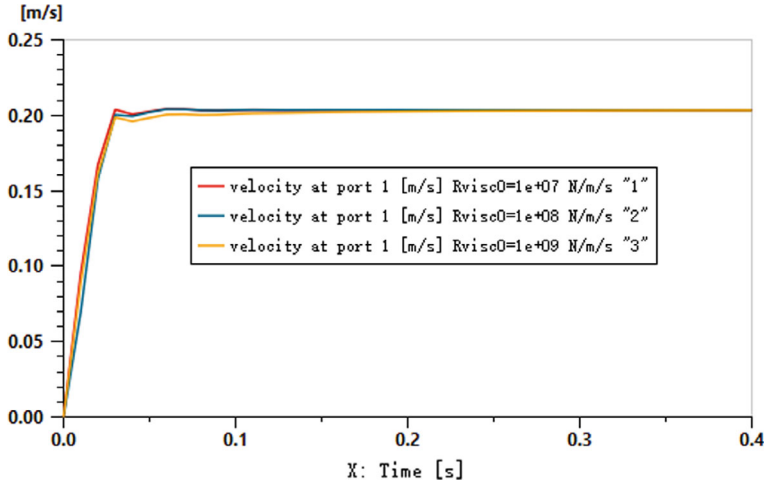


Fig. 5 Velocity curves of weights under different friction forces

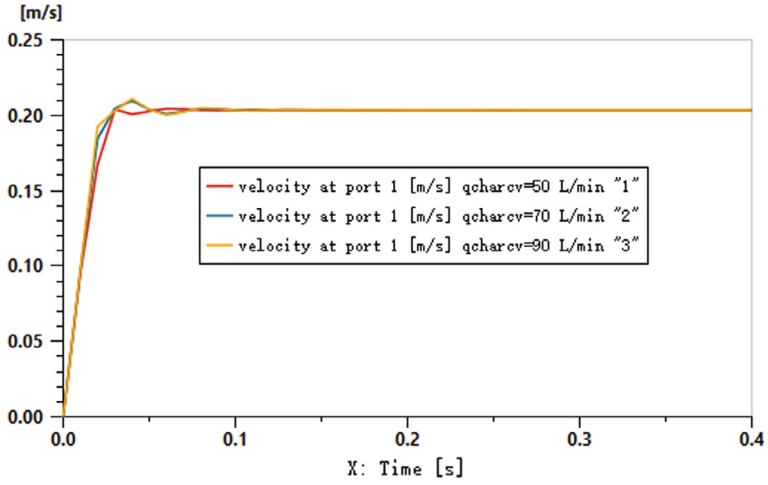


Fig. 6 Load speed curve under characteristic flow at different maximum opening of balance valve

(3) The influence of different moment of inertia on the traction winch system

Keep other parameters unchanged, change the moment of inertia for simulation, and simulate the speed of the weight.

The dynamic characteristics of the motor speed are influenced by the load and the moment of inertia of the motor converted to the drive shaft of the hydraulic motor. Reducing the natural frequency of the system can increase the corresponding speed of the motor speed.

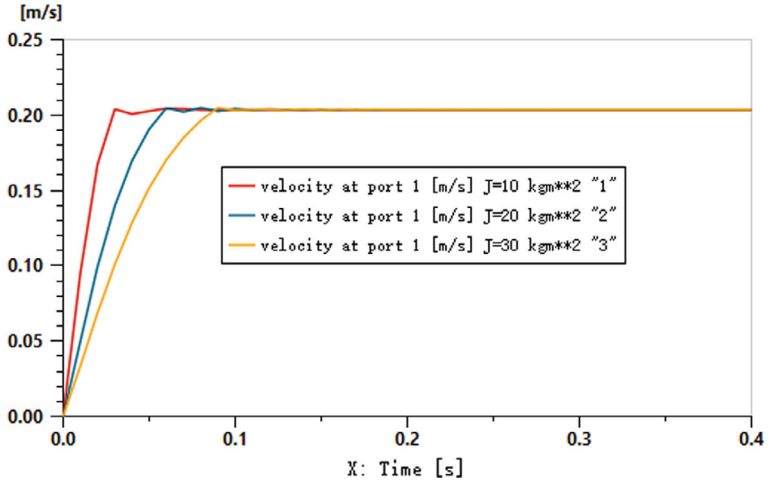


Fig. 7 Velocity curves of weights under different moment of inertia

When the moment of inertia is respectively 10 kg·m^2, 20 kg·m^2, and 30 kg·m^2, the speed curve of the weight is shown in Fig. 7.It can be seen from the simulation results that the increase of moment of inertia has a significant impact on the stability time of the system. When the moment of inertia is large enough, the number of oscillations decreases, which improves the stability of the system to a certain extent.

(4) Influence of different displacement on traction winch system

Keep other parameters unchanged, change the motor displacement for simulation, and simulate the speed of heavy objects.

When the motor displacement is 1848cc,924cc,616cc respectively, the speed curve of the weight is shown in Fig. 8. Larger displacement motors typically provide more power, so they can lift heavy objects faster under the same load and transmission conditions. With the increase of motor displacement, the response time of motor output shaft speed decreases, but when the motor displacement increases to a certain value, the fluctuation of motor output shaft increases and the stability value decreases.

(5) Whether the throttle valve has an impact on the traction winch system

As shown in Fig. 9, red and blue are respectively the curves of heavy weights in the hydraulic system with or without throttle valves. The simulation results show that the hydraulic system with throttle valve can provide better flow and pressure control than the system without throttle valve, so as to reduce overshoot and improve the stability and response performance of the system.

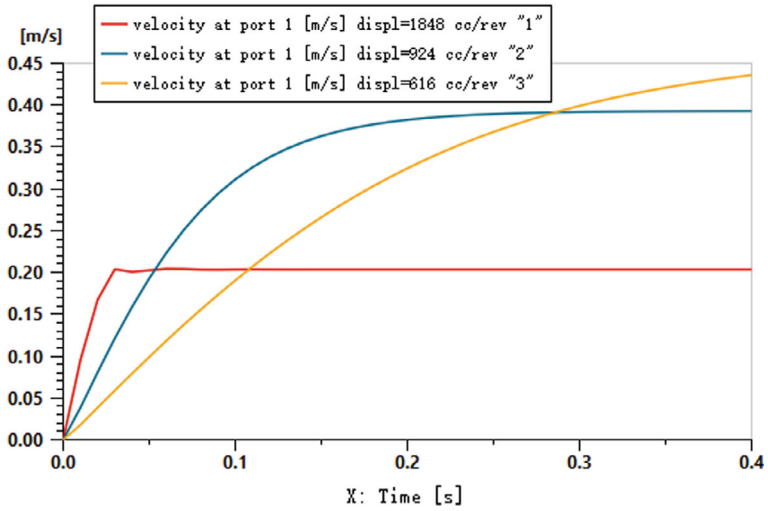


Fig. 8 Speed curves of weights under different motor displacement

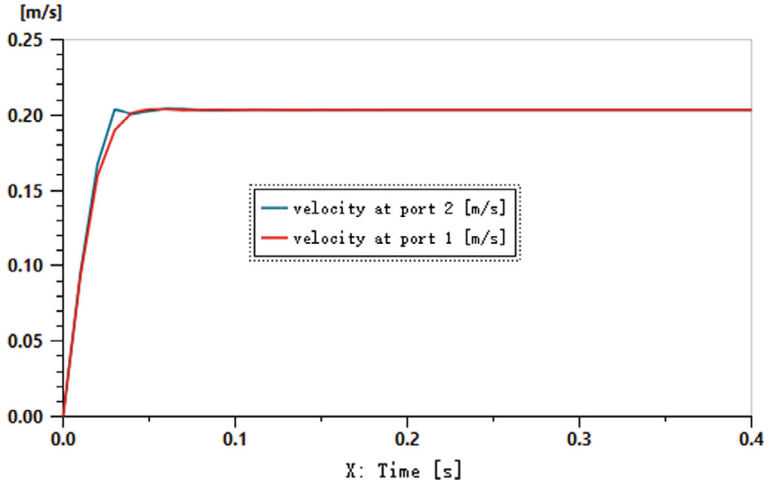


Fig. 9 Weight velocity curve with or without throttle

5 Conclusion

In this paper, the principle of the hydraulic system of traction winch is briefly described, and the hydraulic system of traction winch is simulated and analyzed based on AMESim, and the influence of different hydraulic components on the system under different parameters is compared.

Through the above analysis, the following conclusions are drawn: the increase of cable friction and the characteristic flow rate at the maximum opening of the balance valve will increase the overkill and reduce the stability, which is not conducive to the vibration of the traction winch, and reduce the stability and reliability of the system; The increase of moment of inertia has a significant effect on the stability time of the system. When the moment of inertia is large enough, the number of oscillation decreases, which improves the stability of the system to a certain extent. The increase of the motor displacement will reduce the load response time, and the hydraulic system after the throttle valve has better flow and pressure control, and the overshoot is smaller, which improves the system stability and response performance. The research and conclusion of this paper have certain guiding significance for the optimization of traction winch system.

References

1. Pengcheng Z, Qing Z, Wenhua Z et al (2022) Research on Tension Control of traction winch based on BP Fuzzy Neural Network PID Control. Mach Tools Hydraul 50(21):137–143
2. Ma H (2018) System design and research of constant tension control hydraulic winch. Zhejiang Normal University
3. Yu K, Wang X, Zhong C et al (2011) Design and research of double drum winch for laying deep water pipe. China Pet Mach 39(02):41–44
4. Zhang YS, Zeng Z-G, He Z et al (2009) Application of hydraulic transmission technology in hoisting winch. Mach Tool Hydraul 37(06):111–112
5. Ju C (2022) Treatment and analysis of jitter problem of a hydraulic winch system. Hydraul Pneum Seals 42(05):80–83
6. Chen J (2024) Modeling simulation and optimization of hydraulic system of marine towing winch. China Machinery (12):21–24+29
7. Li Z (2023) Design of hydraulic system for marine transport crane winch. Intern Combust Engine Accessories (11):30–32
8. Zhou X, Chen Z, Wu Y et al (2023) Structure design and finite element analysis of hydraulic winch. Hydraul Pneum Seals 43(01):65–68
9. Yue J-t, Li T, Zhou Y-l (2019) Modeling and analysis of winch hydraulic self-running system based on Ameim. In: DEStech transactions on computer science and engineering. Destech Publications, pp 21–25
10. Meng QR, Lin SF, Wang J (2013) Design and analysis of Hydro-viscous Winch Hydraulic System. In: Advanced designs and researches for manufacturing. Part 1.: Trans Tech Publications, pp 1295–1299

Open Access This chapter is licensed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License (http://creativecommons.org/licenses/by-nc-nd/4.0/), which permits any noncommercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if you modified the licensed material. You do not have permission under this license to share adapted material derived from this chapter or parts of it.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.



A Study on the Active Suspension System of a Wheel-Legged Excavator with Attitude Adjustment



Zeguang Hu, Hao Xiong, Dingxuan Zhao, and Haiwu Zheng

Abstract For the ET120 wheel-legged excavator model, a kinematic and dynamic model of the wheel-legged chassis was developed. This study explores a retrofit solution that employs an electro-hydraulic servo valve to control the lift cylinders of the wheel-legged chassis. Based on the original excavator configuration, the system was enhanced with functions such as active–passive mode switching, electro-hydraulic servo control, oil pressure maintenance, and fluid replenishment, along with the corresponding components. These upgrades enable the hydraulic system to deliver sufficient force and flow to support an active suspension function for the excavator. A PID controller was implemented to regulate both the excavator's attitude and the chassis ground clearance, thereby ensuring minimal variation in attitude angle when operating on uneven terrain.

Keywords Walking excavator · Multibody system · Electro-hydraulic servo · PID · Simulink

1 Introduction

In 2015, Siegwart [1] from the University of Zurich conducted calculations and component selection for the hydraulic pump, hydraulic motor, and planetary gears of the M545 wheel-legged excavator from Menzi Muck. This study provided estimates for the maximum speed, minimum output torque, and motor power.

In 2015, Zimmermann [2] at the University of Zurich conducted research on traction control for wheel-legged excavators to address issues of slippage when climbing steep or loose-soil slopes. The study aimed to improve traction behavior, thereby enhancing the excavator's climbing capability, control, and overall performance. The entire vehicle was modeled as a quasi-static system, neglecting forces due to acceleration, air resistance, and friction, and assuming single-point contact between each

Z. Hu · H. Xiong · D. Zhao (✉) · H. Zheng
Yanshan University, Qinhuangdao, Hebei, China
e-mail: zdx@ysu.edu.cn

© The Author(s) 2026

H. Saman et al. (eds.), *The 9th International Conference on Advances in Construction Machinery and Vehicle Engineering*, Lecture Notes in Mechanical Engineering,
https://doi.org/10.1007/978-981-96-5527-4_26

wheel and the ground. Based on this model, the study verified that a slip minimization strategy is the optimal approach for improving the excavator's traction performance.

In 2015, Hutter [3] at the University of Zurich studied ground reaction force optimization for the four legs of the M545 wheel-legged excavator. Initially, a single-cylinder force control strategy based on pressure feedback was developed on a test bench. Due to budget constraints, no force sensors were installed; instead, cylinder output force and friction were estimated based on piston position, velocity, and hydraulic pressure. Experiments revealed that the large hydraulic cylinders used on the M545 (cross-sectional area 7.85×10^{-3} m^2, $F_{max} = 200$ kN) exhibit significantly lower relative friction compared to the smaller cylinders on the HYQ hydraulic quadruped robot (cross-sectional area 2.01×10^{-4} m^2, $F_{max} = 3$ kN). The friction characteristics for the M545 cylinders could be modeled effectively using the Stribeck model, with friction effects accounting for approximately 1% of maximum drive force. The standard Duplomatic BLS6 proportional valve used in the M545 has 50% overlap at the zero-flow position, minimizing leakage when the valve is closed. To address overlap limitations, a double-valve configuration was implemented, allowing independent control of inflow and outflow for both cylinder chambers. On the single-cylinder test bench, force control without accounting for friction utilized a PI controller, showing bandwidths of 5 Hz for a single standard valve, 15 Hz for the double-valve setup, and 25 Hz for the MOOG valve. The double-valve configuration also demonstrated excellent zero-flow performance. Field tests on the M545's leg showed error control within 150 kg, with errors under 2%.

In 2017, Hutter [4] at the University of Zurich conducted research on compliance control for the legs of walking excavators. Viewing the walking excavator as a multi-rigid-body system, he found that, when the center of mass approaches the geometric center of contact points, nearly the entire load is borne by a single diagonal axis. This concentration of load can potentially damage terrain and wear down the excavator's suspension system due to high peak forces. Only by incorporating compliance in the structure, contact points, or actuation can all legs effectively support the machine, allowing forces to be autonomously distributed across each leg to achieve the intended movement trajectory. To implement adaptive contact force control—enabling precise regulation of ground reaction forces to modulate the excavator's dynamic behavior and overcome the limitations of position control—Hutter proposed a two-step approach. First, establishing whole-body force control requires force control for each individual hydraulic cylinder. On quadrupedal robots like humanoid robots and BigDog, joint force/torque sensors are employed for force control. However, instead of using additional force sensors, Hutter utilized hydraulic pressure sensors combined with Stribeck friction compensation to achieve single-cylinder force control. This allowed for direct joint force regulation without relying on cylinder position feedback, achieving a precision within 3% of the vehicle's weight. The second step involved developing a control algorithm that expresses the contact force between the robot and the ground as a function of the robot's leg joint states and chassis pose. This function is then optimized. Given the relatively low speed and acceleration of the excavator, the robot is simplified as a quasi-static system, with

virtual model control techniques used to combine virtual force with the Jacobian transpose method.

In 2019, Jelavic [5] from the University of Zurich conducted research on whole-body motion planning for the HEAP platform. By employing trajectory optimization (TO)-based planning algorithms, advanced tasks can be formulated as TO problems. Using the full degrees of freedom of HEAP, autonomous whole-body trajectory planning was achieved in simulations, enabling the system to perform tasks such as omnidirectional steering, on-the-spot turning, and crossing deep trenches and stairs with the assistance of the arm.

In 2020, Jelavic [6] from the University of Zurich conducted research on terrain-adaptive planning and control for complex movements. This study included a trajectory optimization-based planner that computes reference trajectories for the end effector and joints. A whole-body controller, based on hierarchical optimization methods, was utilized to track these reference trajectories, enabling terrain-adaptive control of the entire body. The controller was capable of calculating the desired output forces, positions, and velocities based on the performance of the hydraulic cylinders. During experiments with the HEAP platform, issuing high-level steering commands allowed the system to automatically plan paths according to the elevation information of small slopes ahead, utilizing the arm to stabilize the center of gravity and adaptively navigate uneven terrain to complete the commands.

In 2020, Egli [7] from the University of Zurich investigated the automation of the boom on the HEAP excavator. Utilizing data collected during the excavator's operation, a universal model was trained using neural networks, with hydraulic drive current as input and joint displacement as output, capturing the highly nonlinear dynamics of the system. A simulation model was developed by incorporating link lengths, and reinforcement learning (RL) was employed to train a neural network control strategy for end-effector position tracking. The control strategy, once trained in simulation, was directly implemented on the HEAP excavator, achieving accurate and stable position tracking. This approach demonstrated versatility, as it eliminated the need for detailed modeling of hydraulic components.

In 2021, Jelavic [8] from the University of Zurich focused on planning for the wheel-legged excavator using a combination of sampling and optimization techniques. This approach effectively avoided the pitfalls of local minima commonly encountered in trajectory optimization methods, thereby accelerating the convergence of the optimization process. The planning methodology expanded the roadmap to encompass the mapping of joint angles to the center of mass, allowing for real-time stability checks. The efficacy of this method was validated through simulation experiments, demonstrating its potential for enhancing the operational performance of the excavator in complex terrains.

In 2024, Hu [9] from Yanshan University proposed a new chassis dynamics modeling method for the ET120 walking excavator, focusing on the wheel-ground interaction forces represented as a spring-damping system with a spherical tire profile. Utilizing the principles of d'Alembert and virtual work, a comprehensive ten-degree-of-freedom dynamic model is developed that incorporates leg mass for the first time. The model addresses the position and orientation of the base through trajectory

planning, identifying the driving force of the four actuators and the vertical acceleration of the tire center as unknown variables. Notably, the vertical acceleration acts as an intermediate iterative parameter in the numerical calculations. The simulation framework is constructed using the MATLAB/Simulink/Simscape/Multibody toolbox, integrating crucial components such as the 6-DOF Joint and Spatial Contact Force modules. This configuration effectively simulates the six-degree-of-freedom motion and captures the interaction between the chassis and the ground. A comparative analysis of the simulation outcomes against the theoretical model results demonstrates the accuracy and reliability of the proposed dynamic model, contributing valuable insights into the performance dynamics of walking excavators in complex operational environments.

2 Chassis Mechanism and Hydraulic System of the Wheel-Legged Excavator

2.1 Chassis Leg Structure of the Excavator

As illustrated in Fig. 1, the chassis structure of the ET120 walking excavator comprises a base and four legs. Each leg is designed with three degrees of freedom—lifting, swinging, and lower-leg rotation—along with a hydraulically driven wheel. The front two legs are equipped with support claws at their ends, enhancing ground stability. This study concentrates exclusively on the motion control of the chassis during straight-line movement, deliberately excluding any steering dynamics. Consequently, the legs are simplified to possess a single lifting degree of freedom, while the remaining joints are treated as fixed.

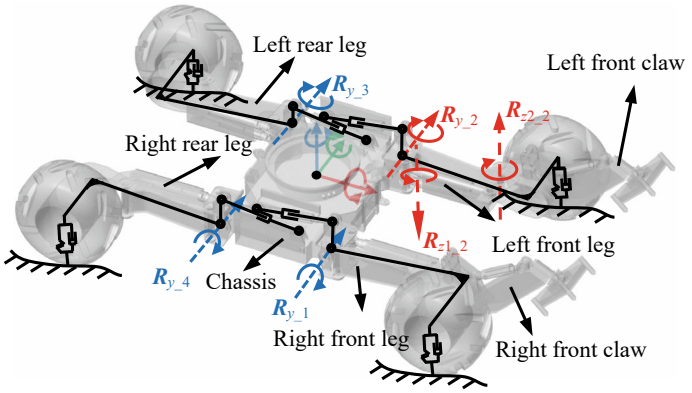


Fig. 1 Schematic diagram of the Chassis mechanism

The lifting function is actuated by a hydraulic cylinder, where one end of the cylinder is connected to the base hinge point b1 and the other end is linked to the leg hinge point b3. The leg and base are interconnected at hinge point b2. By modulating the extension and contraction of the lifting hydraulic cylinder, the legs facilitate rotational movement around hinge point b2, enabling the lifting function.

In this model, the tires are conceptualized as components of a spring-damping system with a spherical profile. Additionally, the differential speed capability of the wheel hub is acknowledged, assuming no relative sliding occurs between the tire and the ground in practical applications; thus, the friction force between the tire and the ground is considered negligible. Furthermore, the ground force acting on the tire is assumed to be directed vertically upward.

2.2 Kinematic Model of the Chassis

A body coordinate system $\{O_d\}$ is established at the center point directly beneath the rotating platform of the base. In this system, the X-axis is oriented horizontally toward the front of the base, the Z-axis points vertically upward, and the Y-axis follows the right-hand rule. For practical experiments, IMUs and other attitude sensors can be positioned within this coordinate framework. The center of mass of the base is denoted as O_{dz}, and $\{O_d\}$ and O_{dz} do not coincide, allowing for real-time updates to the mass parameters in the base's center of mass coordinate system O_{dz} as the excavator arm moves.

The leg's body coordinate system $\{O_{t_i}\}$ is established at the center of mass of each leg, with axis orientations as shown in Fig. 2. Here, the subscripts 1, 2, 3, and 4 refer to the right front leg, left front leg, left rear leg, and right rear leg, respectively. In the inverse kinematics and dynamics analysis of single-leg motion, the leg tilt angle θ_i is treated as a known variable, enabling the hydraulic cylinder drive parameters to be determined through inverse calculations.

In the base coordinate system $\{O_d\}$, as shown in Fig. 3, the coordinates ${}^d\boldsymbol{b}_{1_1}$ and ${}^d\boldsymbol{b}_{2_1}$ of the hinge points b_{1_1} and b_{2_1}, which are associated with the right front leg on the base, can be expressed as:

$$ {}^d\boldsymbol{b}_{1_1} = \left(-d_{1_x}\ -d_{1_y1}\ d_{1_z}\right)^{\mathrm{T}},\ {}^d\boldsymbol{b}_{2_1} = \left(d_{2_x}\ -d_{1_y1}\ d_{2_z}\right)^{\mathrm{T}} \tag{1} $$

The relationship between the length r_1 of the right front lifting hydraulic cylinder and the rotation angle θ_1 of the leg is given by:

$$ r_1 = \left(\left(d_{2_x} + d_3c\theta_1 + d_4s\theta_1 + d_{1_x}\right)^2 + \left(d_{2_z} - d_3s\theta_1 + d_4c\theta_1 - d_{1_z}\right)^2\right)^{\frac{1}{2}} \tag{2} $$

Differentiating both sides of Eq. (2) with respect to time provides the relationship between the hydraulic cylinder velocity ${}^d\boldsymbol{v}_{r1}$ and the leg rotational angular velocity ${}^d\boldsymbol{\omega}_1$:

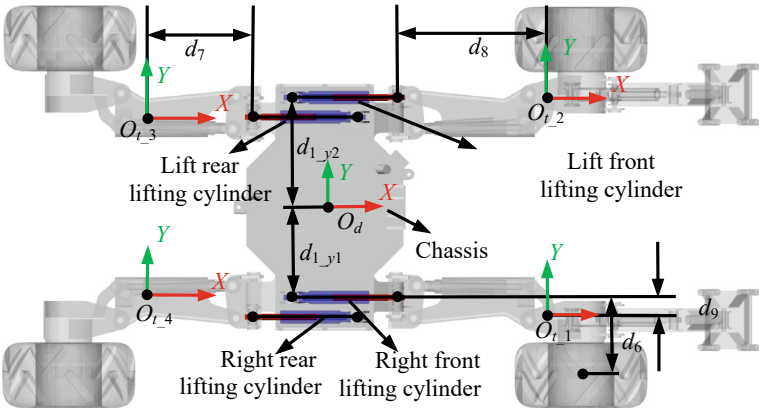


Fig. 2 Top view of the Chassis

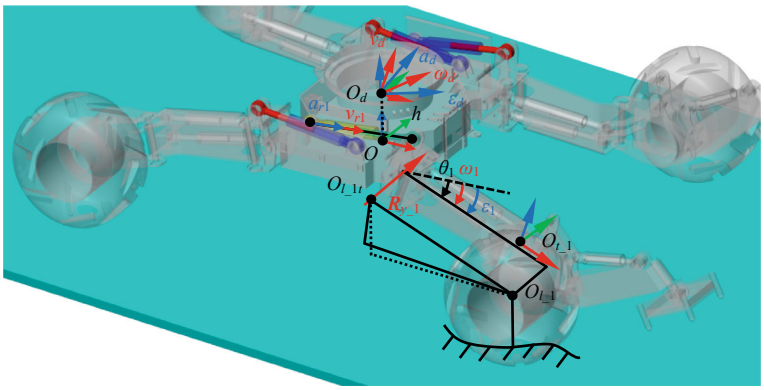


Fig. 3 Overall vehicle inverse kinematics schematic

$$ {}^{d}\boldsymbol{v}_{r1} = {}^{d}\boldsymbol{\omega}_1 \times {}^{d}_{t_1}\boldsymbol{R}^{d} \begin{pmatrix} d_3 \\ 0 \\ d_4 \end{pmatrix} = {}^{d}\hat{\boldsymbol{\omega}}_1 {}^{d}_{t_1}\boldsymbol{R}^{d} \begin{pmatrix} d_3 \\ 0 \\ d_4 \end{pmatrix} \tag{3} $$

Projecting the hydraulic cylinder velocity ${}^{d}\boldsymbol{v}_{r1}$ onto the axis of the hydraulic cylinder results in the scalar velocity v_{r1} of the hydraulic cylinder:

$$ v_{r1} = {}^{d}\boldsymbol{v}_{r1} \cdot {}^{d}\boldsymbol{\delta}_i \tag{4} $$

Given the chassis's lateral tilt angle α, pitch angle β, yaw angle γ, coordinates x_d and y_d of point O_d in coordinate system $\{O\}$, ground clearance height h, and the heights z_{l_1}, z_{l_2}, z_{l_3} and z_{l_4} of the centers of the four wheel hubs, the inverse kinematics solution for the entire vehicle's position can be determined. By using

geometric relationships, the attitude $O\,dz\boldsymbol{R}$, position $\boldsymbol{O}_{dz}$ in $\{O\}$, rotation angles θ_i of the four legs relative to the base, lengths r_i of the four lifting drive cylinders, position vectors $\boldsymbol{O}_{t_i}$, and attitudes $O_{t_i}\,\boldsymbol{R}$ of the centers of mass of the four legs in the world coordinate system $\{O\}$ are derived.

By substituting the solutions for θ_1, θ_2, θ_3, and θ_4 obtained from the equations, the lengths r_1, r_2, r_3, and r_4 of the lifting cylinders for each walking leg can be determined.

$$\begin{aligned} z_{l_1} &= h + (-c\alpha s\beta c\gamma + s\alpha s\gamma)\left(d_{2_x} + d_5 c\theta_1\right) \\ &\quad + (c\alpha s\beta s\gamma + s\alpha c\gamma)\left(-d_6 - d_{1_y1}\right) \\ &\quad + c\alpha c\beta\left(d_{2_z} - d_5 s\theta_1\right) \\ z_{l_i} &= r_l + q_i \end{aligned} \tag{5}$$

In the context of the $\{O\}$ coordinate system, the velocity component along the Z-axis for the point O_{l_1} is designated as the generalized velocity, denoted by $\dot{z}_{l_1}$. Conversely, the angular velocity, denoted as ω_1, is derived through the process of inversion:

$$\begin{aligned} \omega_1 = [&\dot{z}_{l_1} - h' - (-\omega_y c\beta c\gamma + \omega_x(s\alpha s\beta c\gamma + c\alpha s\gamma))(d_{2_x} + d_5 c\theta_1) \\ &- (\omega_y c\beta s\gamma + \omega_x(-s\alpha s\beta s\gamma + c\alpha c\gamma))(-d_6 - d_{1_y1}) \\ &- (-\omega_y s\beta - \omega_x s\alpha c\beta)(d_{2_z} - d_5 s\theta_1)] \\ &/\left[d_5 s\theta_1(c\alpha s\beta c\gamma - s\alpha s\gamma) - d_5 c\theta_1 c\alpha c\beta\right] \end{aligned} \tag{6}$$

2.3 Dynamic Model of the Chassis

In the analysis of the mechanical system under consideration, the force propelling the lifting cylinder is denoted as F_{rd_1}. The external force applied to the leg at the point of ground contact is represented by ${}^{d}\boldsymbol{F}_{d_1}$. The gravitational force influencing the leg is designated as ${}^{d}\boldsymbol{G}_{t_1}$, while the inertial force associated with the leg is denoted by ${}^{d}\boldsymbol{F}_{t_1}$. Additionally, the inertial torque generated by the leg is indicated by ${}^{d}\boldsymbol{T}_{t_1}$. These forces and torques can be mathematically articulated as follows:

$$\begin{aligned} {}^{d}\boldsymbol{G}_{t_1} &= m_{t_1}\boldsymbol{g},\ {}^{d}\boldsymbol{F}_{t_1} = -m_{t_1}{}^{d}\boldsymbol{a}_{t_1} \\ {}^{d}\boldsymbol{T}_{t_1} &= -{}^{d}\boldsymbol{I}_{t_1}{}^{d}\boldsymbol{\varepsilon}_1 - {}^{d}\boldsymbol{\omega}_1 \times \left({}^{d}\boldsymbol{I}_{t_1}{}^{d}\boldsymbol{\omega}_1\right) \end{aligned} \tag{7}$$

The derivation of the dynamic equations can be succinctly outlined as follows:

$$\begin{aligned} &{}^{d}\boldsymbol{F}_{d_1} \cdot {}^{d}\boldsymbol{\omega}_1 d_5 c\theta_1 + F_{rd_1} v_{r1} + {}^{d}\boldsymbol{G}_{t_1} \cdot {}^{d}\boldsymbol{\omega}_1 d_8 c\theta_1 \\ &\quad + {}^{d}\boldsymbol{F}_{t_1} \cdot {}^{d}\boldsymbol{v}_{t_1} + {}^{d}\boldsymbol{T}_{t_1} \cdot {}^{d}\boldsymbol{\omega}_1 = 0 \end{aligned} \tag{8}$$

The derivation of the inverse dynamics for the force F_{rd_1} exerted by the leg actuator can be articulated as follows:

$$\begin{aligned} F_{rd_1} = -\big({}^{d}\boldsymbol{F}_{d_1} \cdot {}^{d}\boldsymbol{\omega}_1 d_5 c\theta_1 + {}^{d}\boldsymbol{G}_{t_1} \cdot {}^{d}\boldsymbol{\omega}_1 d_8 c\theta_1 \\ + {}^{d}\boldsymbol{F}_{t_1} \cdot {}^{d}\boldsymbol{v}_{t_1} + {}^{d}\boldsymbol{T}_{t_1} \cdot {}^{d}\boldsymbol{\omega}_1\big)/v_{r1} \end{aligned} \tag{9}$$

In accordance with the principle of virtual work, the following relationship is established:

$$\begin{aligned} &\begin{bmatrix} F_{r1} \\ F_{r2} \\ F_{r3} \\ F_{r4} \\ F_{ld_1} \\ F_{ld_2} \\ F_{ld_3} \\ F_{ld_4} \end{bmatrix}^{\mathrm{T}} \begin{bmatrix} v_{r1} \\ v_{r2} \\ v_{r3} \\ v_{r4} \\ \dot{z}_{l_1} \\ \dot{z}_{l_2} \\ \dot{z}_{l_3} \\ \dot{z}_{l_4} \end{bmatrix} + \begin{bmatrix} \boldsymbol{G}_{dz} + \boldsymbol{F}_{dz} \\ \boldsymbol{T}_{dz} \end{bmatrix}^{\mathrm{T}} \begin{bmatrix} \boldsymbol{v}_{dz} \\ \boldsymbol{\omega}_{dz} \end{bmatrix} + \sum_{i=1}^{4} \begin{bmatrix} \boldsymbol{G}_{t_i} + \boldsymbol{F}_{t_i} \\ \boldsymbol{T}_{t_i} \end{bmatrix}^{\mathrm{T}} \begin{bmatrix} \boldsymbol{v}_{t_i} \\ \boldsymbol{\omega}_{t_i} \end{bmatrix} \\ &+ \begin{bmatrix} f_{ld_1} \\ f_{ld_2} \\ f_{ld_3} \\ f_{ld_4} \end{bmatrix}^{\mathrm{T}} \begin{bmatrix} v_{ld_1} \\ v_{ld_2} \\ v_{ld_3} \\ v_{ld_4} \end{bmatrix} = 0 \end{aligned} \tag{10}$$

In the context of mechanical systems, the forces and velocities described can be articulated with precision as follows: Let F_{ri} denote the driving force applied by the four actuating cylinders. The vertical support forces exerted by the four tires on the ground are represented by F_{ld_i}. The driving speed of the cylinders is given by v_{ri}, and the vertical velocity of the hub center points is denoted by $\dot{z}_{l_i}$. The gravitational force vector acting on the chassis is $\boldsymbol{G}_{dz}$, while the inertial force vector of the chassis is $\boldsymbol{F}_{dz}$ and the inertial torque vector is $\boldsymbol{T}_{dz}$. For the support legs, the gravitational force vectors are $\boldsymbol{G}_{t_i}$, the inertial force vectors are $\boldsymbol{F}_{t_i}$, and the inertial torque vectors are $\boldsymbol{T}_{t_i}$. The velocity vector of the chassis center of mass is $\boldsymbol{v}_{dz}$, and the angular velocity vector of the chassis is $\boldsymbol{\omega}_{dz}$. Similarly, the velocity vectors of the support leg center of mass are $\boldsymbol{v}_{t_i}$, and the angular velocity vectors of the support legs are $\boldsymbol{\omega}_{t_i}$.

2.4 Electro-hydraulic Servo System of the Chassis

In the original vehicle, the leg support of the wheel-leg chassis is controlled by the D1MW-UK series of electromagnetic Directional Control Valves, as shown in Fig. 4. When the supply voltage is 24 V, the P port is connected to the A port, and the B port is connected to the T port, causing the lifting cylinder to extend and the support leg to descend. When the supply voltage is − 24 V, the oil passage is reversed,

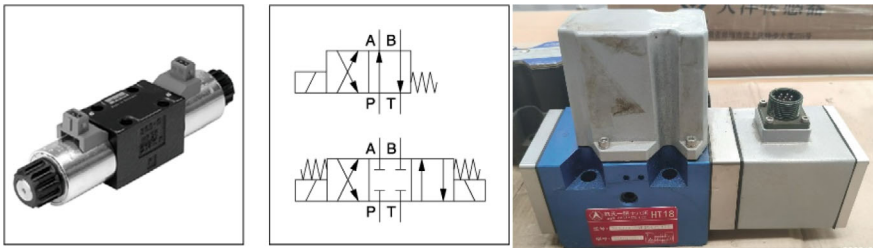


Fig. 4 D1MW-UK series solenoid directional valve and SFD234 electro-hydraulic flow servo valve

causing the lifting cylinder to retract and the support leg to ascend. When the supply voltage is 0, both the A and B ports are connected to the oil tank, and the hydraulic lock within the cylinder closes, preventing the support leg from descending, thus providing protection. This hydraulic system is simple and reliable but cannot control the flow rate of the hydraulic fluid, making it unsuitable for high-frequency servo control. It can only be manually adjusted, which is complex and fails to maximize the performance of the wheel-leg chassis. Therefore, an SFD electro-hydraulic servo valve has been added to the original vehicle. Its control signal is ± 10 mA, which is proportional to the valve opening. By controlling the valve opening, the flow rate in the system can be regulated. In conjunction with the magnetic hysteresis displacement sensor in the hydraulic cylinder, precise automatic control of the lifting cylinder's displacement can be achieved.

At the same time, to ensure the original vehicle's functionality, a two-position three-way solenoid directional control valve has been added. When switched to the active mode, high-pressure control oil pushes open the hydraulic control check valve in the active oil circuit, realizing the active function while locking the passive oil circuit. When switched to the passive mode, the hydraulic control check valve in the active oil circuit closes, realizing the passive function. At the same time, a 25 Mpa accumulator is added to match the 25 Mpa oil pressure of the original vehicle to stabilize the pressure and absorb shock.

3 PID-Based Chassis Attitude Control and Simulation

3.1 Simulink Simulation and Dynamic Model Validation of Wheel-Legged Excavator

Simulink, a widely utilized simulation and model-based design environment within MATLAB, enables engineers and researchers to construct, simulate, and analyze multi-domain dynamic systems with high precision. Originally developed for complex control systems and signal processing, Simulink provides a graphical interface that allows users to model system components using a library of pre-built blocks,

facilitating modular and hierarchical design. The platform supports continuous and discrete event simulations, making it suitable for applications ranging from mechanical and electrical systems to biological and economic models. Additionally, Simulink integrates seamlessly with MATLAB's scripting capabilities, allowing for extensive customization, parameter tuning, and post-processing analysis. Its real-time simulation support, combined with code generation tools, enables researchers to rapidly prototype and test control algorithms in hardware-in-the-loop configurations, thereby bridging the gap between simulation and practical implementation. The versatility and accuracy of Simulink have established it as a standard tool for system design, analysis, and optimization in both academic and industrial settings.

Simulink provides robust simulation capabilities, with the Simscape-Multibody toolbox offering a comprehensive set of components such as various joints, rigid body elements, flexible components, reference frames, and contact modules. It includes a dynamics engine that supports both open-chain and closed-chain configurations, enabling accurate and rapid simulation of multi-body systems as well as rigid-flexible hybrid systems. In this study, Simscape-Multibody is employed to simulate the wheel-leg excavator. The 3D models, carefully surveyed and developed, are imported into Simulink using the File Solid block, allowing for visual representation. Within the File Solid block, users can edit key properties of the rigid bodies, including mass, center of mass, inertia, and rotational inertia, which are critical for realistic dynamics.

To assemble the complete vehicle structure, various File Solid modules are interconnected using the Prismatic Joint or Revolute Joint blocks. Given that the excavator is free in the inertial frame, a 6-DOF Joint block is introduced between the reference coordinate system and the excavator, facilitating the representation of its six degrees of freedom. Furthermore, to model the interaction between the tires and the ground, a Spatial Contact Force module is incorporated, allowing for the configuration of parameters such as contact damping, contact stiffness, and friction forces. Finally, the input driving forces for the hydraulic cylinders are defined in the Prismatic Joint modules, enabling the overall vehicle simulation to run effectively. This comprehensive simulation setup facilitates the analysis of the wheel-leg excavator's dynamics, providing valuable insights into its operational capabilities and performance characteristics.

Simulink's Simscape-Fluids toolbox [10] offers a wide array of hydraulic components for simulating hydraulic systems, including various pumps, tanks, hydraulic fluids, valves, pipes, actuators, and modules for force transfer between Simscape-Fluids and Simscape-Multibody. In this study, the Reservoir (IL) module is utilized to represent the hydraulic tank, while the Pressure Source (IL) module serves as the pump, allowing the configuration of the system pressure. This pressure source is then connected to the 4-Way Directional Valve (IL) module, where users can define the functional relationship between the valve opening and the valve spool displacement signal. Subsequently, the valve is linked to a Double-Acting Actuator (IL) module through the Pipe (IL) module. Within the actuator configuration, parameters such as the areas of the rod and non-rod chambers and friction characteristics can be specified. This detailed setup ensures accurate representation of the hydraulic cylinder's behavior during operation. To facilitate interaction between the hydraulic

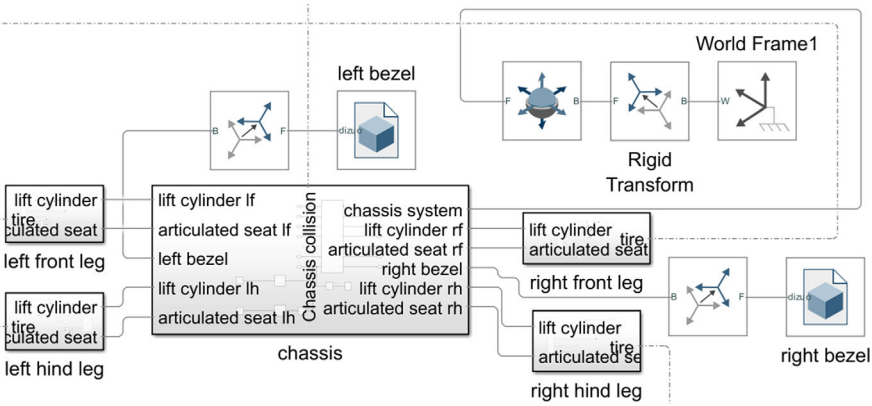


Fig. 5 ET120's SimMechanics blocks in Simulink

system and the multi-body dynamics, the Translational Multibody Interface module is employed. This module allows for the sharing of force and displacement data between the hydraulic cylinder in Simscape-Fluids and the rigid bodies in Simscape-Multibody, achieving a cohesive simulation of the hydraulic system alongside the multi-body dynamics system. Through this integrated approach, the simulation accurately models the behavior of hydraulic systems in conjunction with the mechanical aspects of the excavator, enhancing the overall analysis of system performance and efficiency. Figure 5 shows the complete simulation model of ET120, Fig. 6 shows the details of the wheel-ground module, Fig. 7 shows the six-degree-of-freedom joint module, Fig. 8 shows the rigid body module, Fig. 9 shows the mobile joint joint module, Fig. 10 shows the contact module, and Fig. 11 shows the drive module. Figure 12 shows the 4-Way Directional Valve, Fig. 13 shows the Double-Acting Actuator, and Fig. 14 shows the Simscape-Fluids Simulation Model.

After establishing the multi-body dynamics model and hydraulic system model of the legged excavator in Simulink, the kinematic and dynamic behaviors of the entire vehicle can be simulated and validated. Specifically, position drive functions are assigned to the four lifting cylinders as follows:

$$
\begin{aligned}
&y_\alpha = 0.06\sin(0.15(t-2.5)^3),\, y_\beta = 0.02\sin(0.3(t-2.5)^3),\\
&y_z = 0.04\sin(0.35(t-2.5)^3),\, l_1 = y_\alpha\\
&\quad + y_\beta + y_3,\, l_2 = -y_\alpha + y_\beta + y_z,\\
&l_3 = -y_\alpha - y_\beta + y_z,\, l_4 = y_\alpha - y_\beta + y_z
\end{aligned}
$$

As depicted in Fig. 15, the virtual power exerted by the driving force is quantified at a magnitude of 1e6 N·m/s, whereas the system's virtual power fluctuates within a range of ± 30 N·m/s. This indicates that throughout the simulation, the system adheres to the principle of virtual work at all instances, thereby validating the

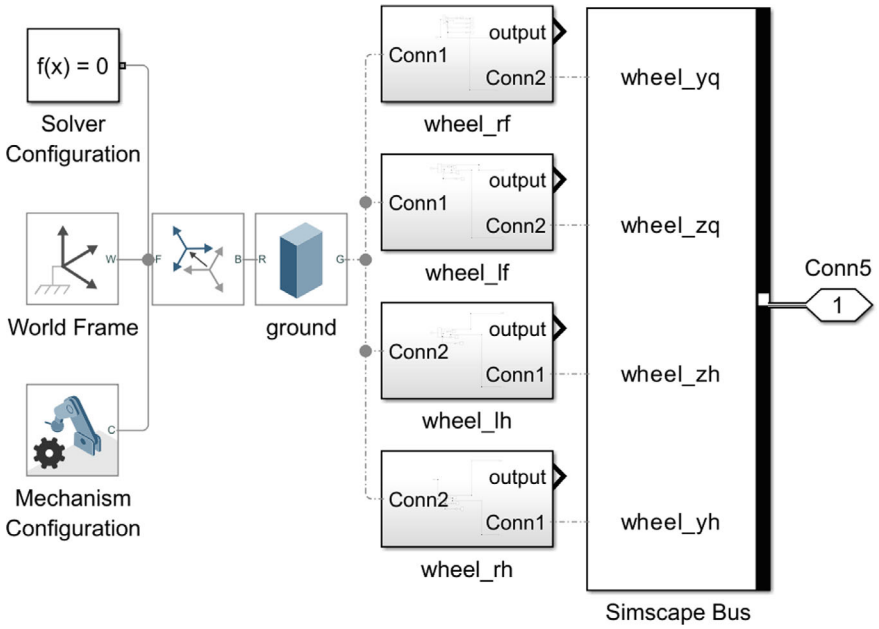


Fig. 6 Details for the wheel-ground block of the model

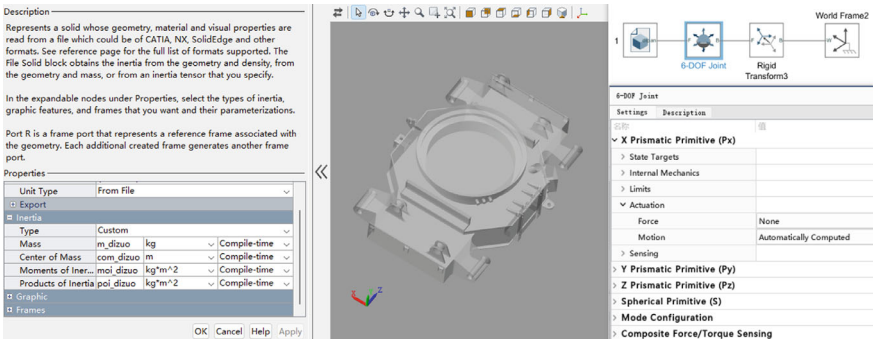


Fig. 7 Details for the 6-DOF Joint block of the model

compatibility and accuracy of the simulation system with the theoretical framework predicated on the principle of virtual work.

As can be seen from Fig. 16, during the movement, the displacement, velocity, acceleration and driving force of the lifting cylinder calculated by the formula are consistent with those in the simulation, which verifies the correctness of the theoretical formula.

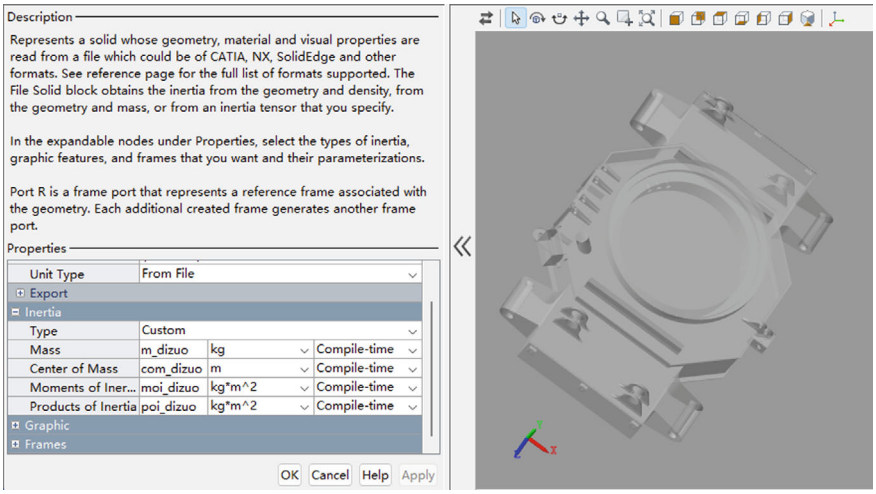


Fig. 8 File solid module

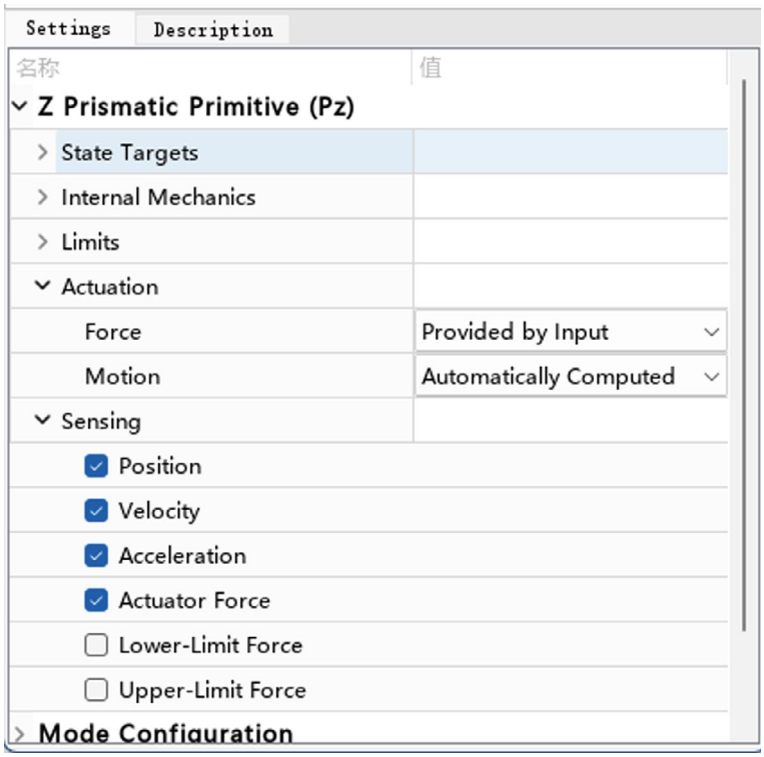


Fig. 9 Prismatic joint block

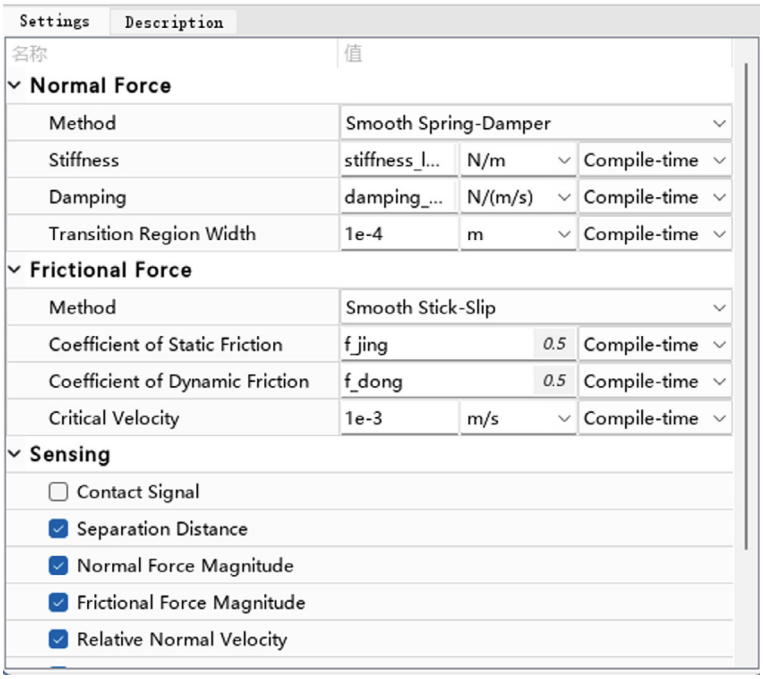


Fig. 10 Contact block

3.2 Position PID Control of a Single Leg

Based on the established multi-body dynamics and hydraulic system simulations, a PID controller was implemented to achieve position control of the lifting cylinder in a single leg [11]. The leg mass was set to 2000 kg, with hydraulic pressure at 25 MPa, an 8 mm pipe diameter, a 100 mm rod-side chamber diameter, and a 60 mm hydraulic rod diameter. The spool displacement was used as the control variable, while cylinder displacement served as feedback. By fine-tuning the proportional, integral, and derivative parameters of the PID controller, satisfactory simulation results were achieved. Figure 17 is a control flow chart.

Give the oil cylinder an ideal sinusoidal signal with an amplitude of 50 mm and a period of 4 s. Adjust the three parameters of PID. As shown in Fig. 18, the displacement of the oil cylinder can follow the sinusoidal signal well. The error is shown in Fig. 19. When the sinusoidal signal amplitude is 50 mm, the maximum error is 2 mm, which appears at the top and bottom of the signal curve. This is due to the nonlinearity when the displacement direction of the oil cylinder changes. The model control method can be used to solve this problem later. The valve core displacement curve is shown in Fig. 20.

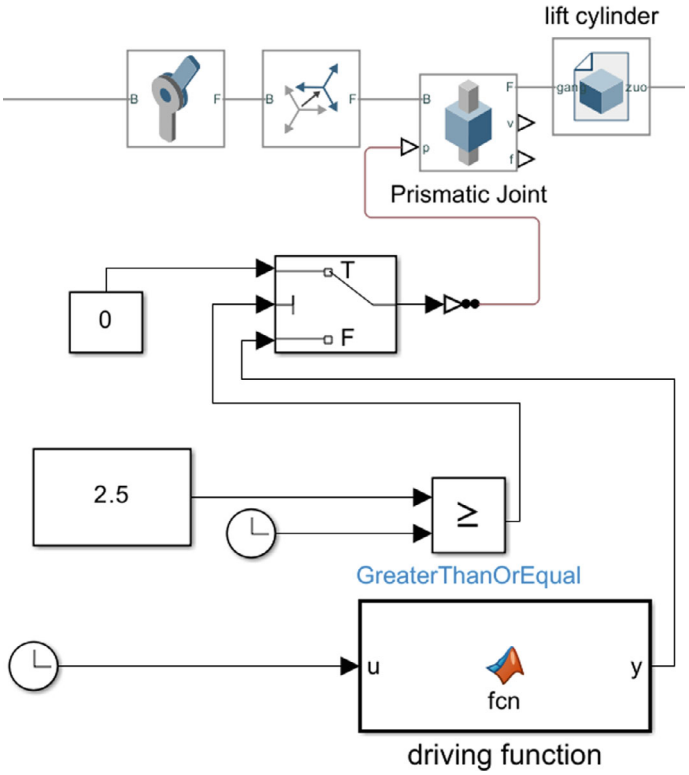


Fig. 11 Driver block

3.3 PID-Based Automatic Attitude Control of the Entire Vehicle

Excavators usually travel on rugged mountain roads. In order to minimize the spillage of materials in the bucket, the attitude angle of the chassis needs to be automatically adjusted, and the lifting cylinders of the four legs need to be controlled at the same time. Therefore, four points on the chassis are selected, and their coordinates in the chassis coordinate system are {1,− 1,0}, {1,1,0}, {− 1,1,0}, and {− 1,− 1,0} respectively. The distance from the fixed coordinate system of the chassis to the ground is set to 0.4 m, and the displacements of the lifting cylinders of the four legs corresponding to the four selected points are taken as the control variables. The Z coordinates of the four points in the world coordinate system are taken as the feedback variables. By fine-tuning the proportional, integral, and differential parameters of the PID controller, relatively satisfactory simulation results are obtained. Figure 21 is the PID control flow chart of the whole vehicle, Fig. 22 is the Vehicle posture control simulation in simulink, Fig. 23 is the simulation using PID control, and Fig. 24 is the simulation without PID control.

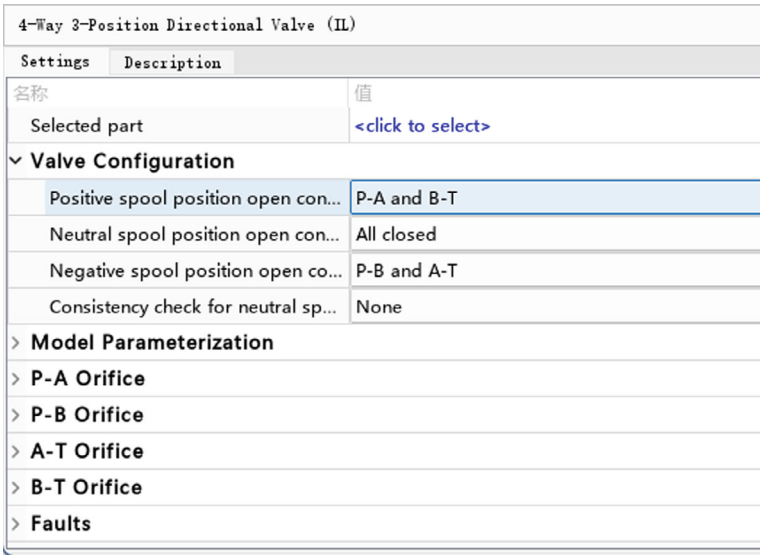


Fig. 12 4-Way directional valve and displacement-opening area curve

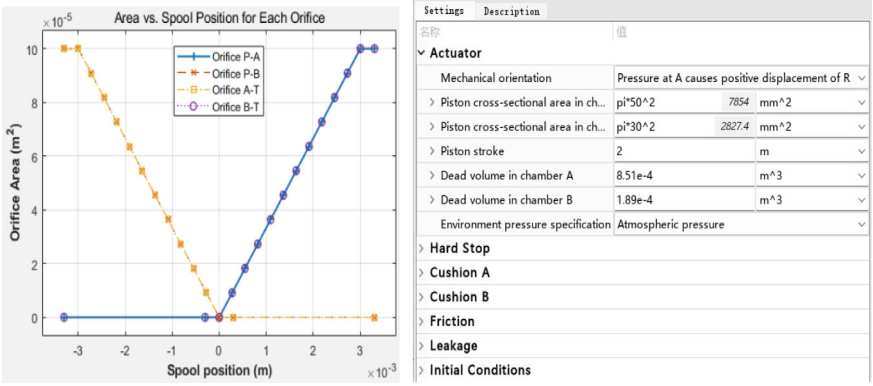


Fig. 13 Double-acting actuator (IL)

It can be seen from Figs. 25 and 26 that after the PID controller is used to perform closed-loop control of the chassis attitude, the three attitude angles are reduced, especially the pitch angle and roll angle. Therefore, the simulation proves that the PID-based chassis attitude control strategy is effective. However, it was found during the simulation that the four legs are not always close to the ground. This is a shortcoming of the position PID closed-loop control, which needs to be improved by adding laser radar and adding contact detection to the control method.

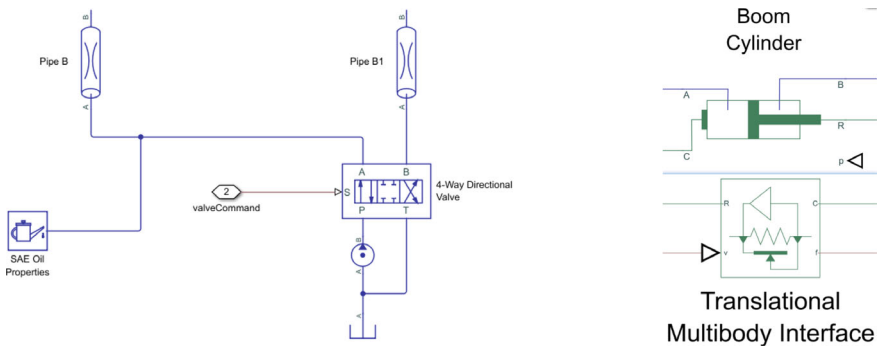


Fig. 14 Simscape-fluids simulation model

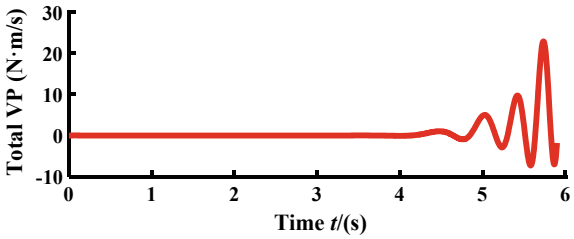


Fig. 15 System total virtual power

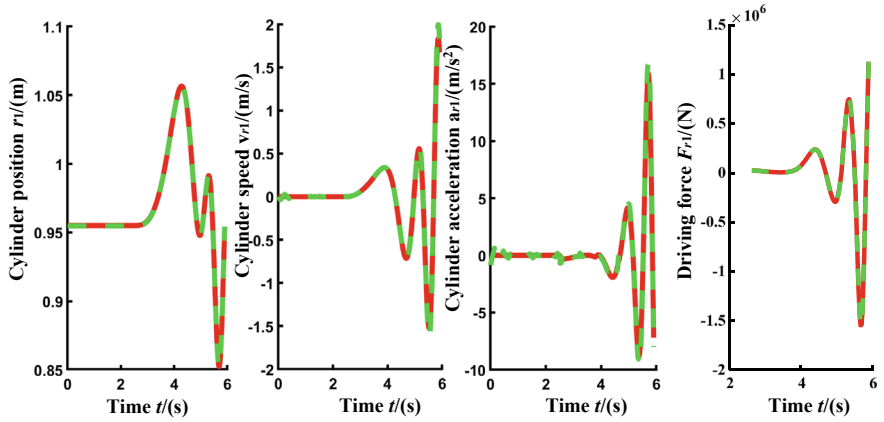


Fig. 16 Curve comparing simulation and theory

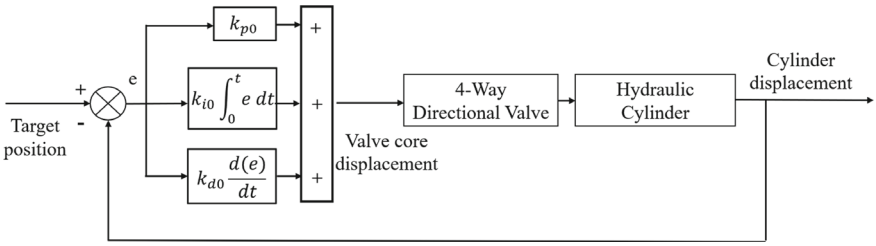


Fig. 17 Outrigger cylinder PID control principle diagram

Fig. 18 Cylinder displacement

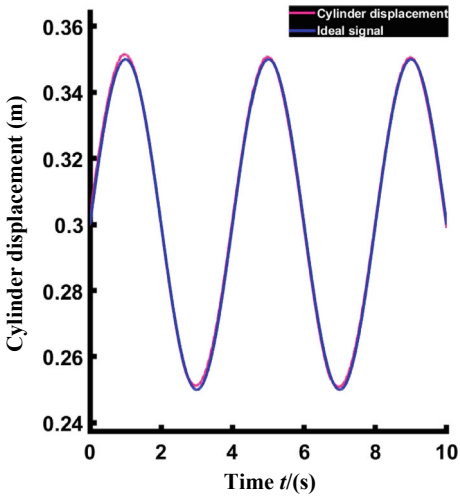
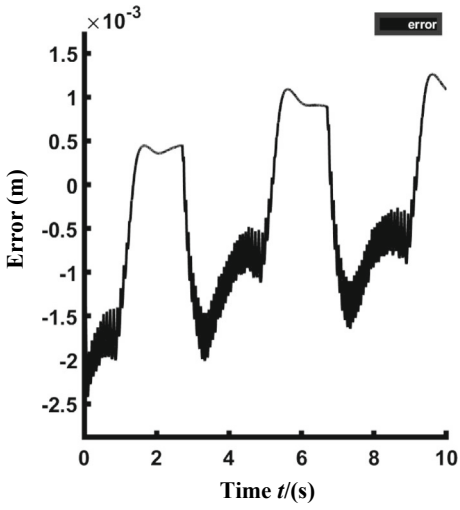


Fig. 19 Error



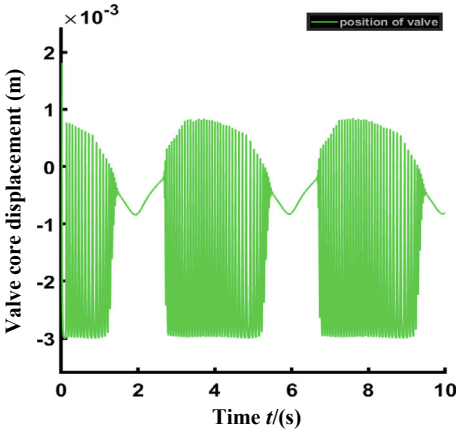


Fig. 20 Valve core displacement

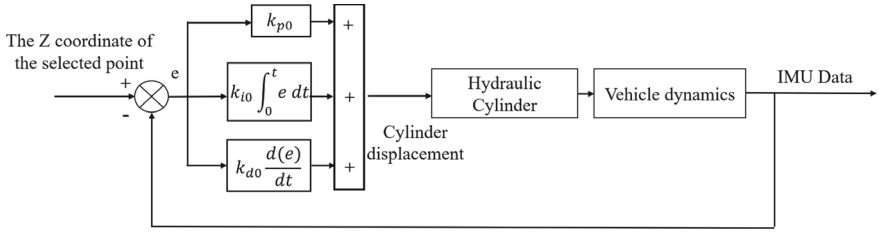


Fig. 21 Vehicle posture PID control principle diagram

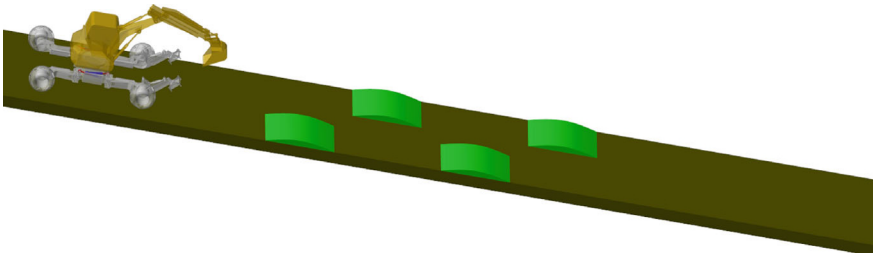


Fig. 22 Vehicle posture control simulation in simulink

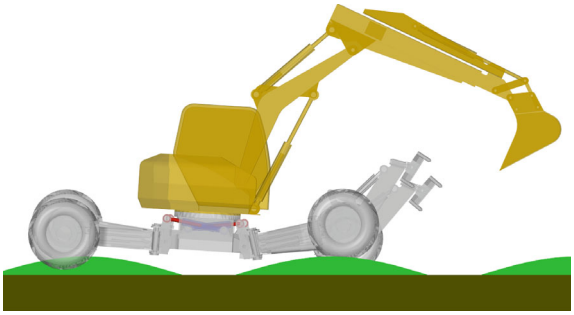


Fig. 23 Simulation using PID

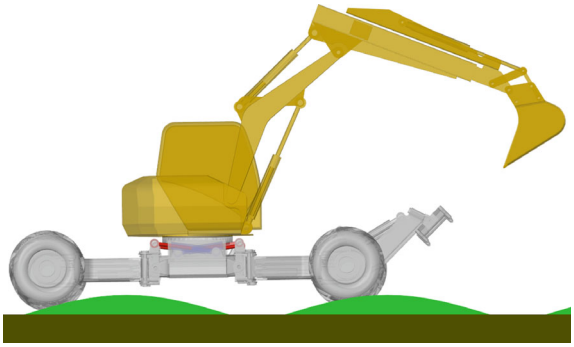


Fig. 24 Simulation without PID

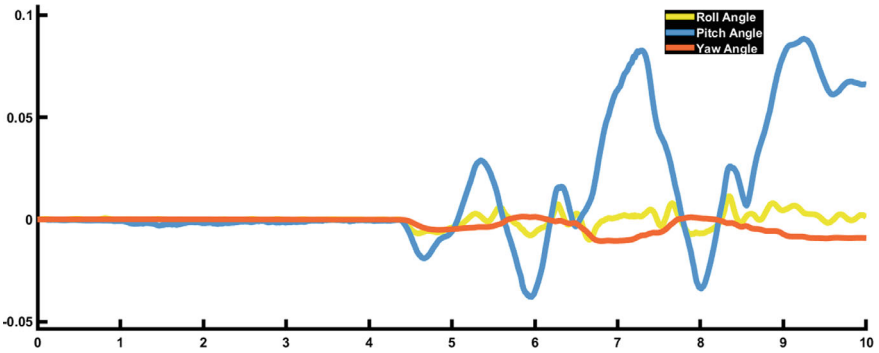


Fig. 25 Attitude angle using PID

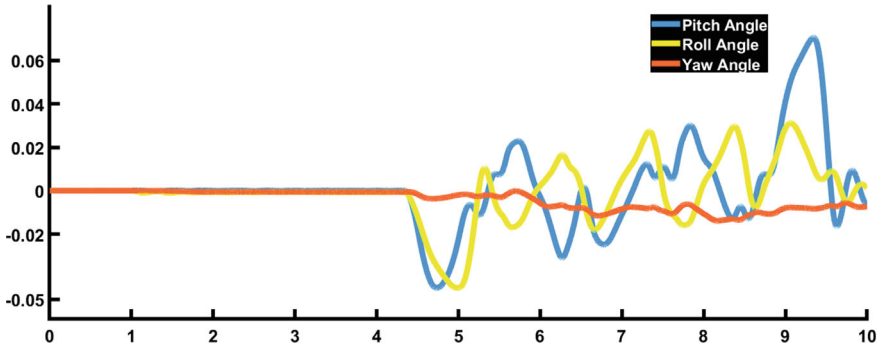


Fig. 26 Attitude angle without using PID

4 Conclusion

(1) Focusing on the ET120 walking excavator's chassis, a ten-degree-of-freedom dynamic model has been established due to the absence of a fixed rotational axis between the base and the inertial frame. This lack of a fixed axis, combined with the fact that ground-tire interactions allow for relative sliding (a unilateral constraint), results in inevitable yaw and translations along the X and Y axes, which occur alongside the mixed motion of pitch, roll, and elevation of the base. The model incorporates six degrees of freedom for the chassis itself and four additional rotational degrees of freedom for the support legs.

(2) To convert the vehicle's manual control to automatic control, an SFD234 electro-hydraulic flow servo valve was added to the existing hydraulic system, enabling precise flow control for the wheel-legged chassis joints. Additionally, a two-position three-way solenoid directional valve was incorporated to facilitate active–passive switching functionality. A 25 MPa accumulator was also installed to stabilize system pressure and reduce minor hydraulic shocks, enhancing overall system stability.

(3) In simulations of a hydraulic single-leg lifting cylinder, a PID controller was successfully implemented to maintain position control. By tuning PID parameters, the cylinder displacement closely tracked a sinusoidal input signal (50 mm amplitude, 4-s period), achieving a maximum tracking error of 2 mm. This error, concentrated at the peak and trough, highlights nonlinearity issues that could be addressed with model-based control. Extending the approach, a chassis attitude control strategy was applied to four legs to stabilize an excavator on uneven terrain. While effective in reducing pitch and roll angles, the simulation revealed that leg-ground contact was inconsistent, indicating that enhancements with laser radar and contact detection could further improve stability.

Acknowledgements The Joint Fund for Regional lnnovation and Development of the National Natural Science Foundation of China under Grant (U20A20332). Hebei Provincial Natural Science

Foundation Innovation Research Group Continuous Support Project: Basic Research on High-Performance Special-Purpose Transport Equipment (E2024203257).

References

1. Siegwart R (2015) Fast driving concepts for a walking excavator. ETH Zurich, Zurich, pp 5–10, 19–23
2. Simon Z (2015) Traction control concepts for a walking excavator. ETH Zurich, Zurich, pp 1–2, 16–18
3. Hutter M, Leemann P, Stevsic S et al (2015) Towards optimal force distribution for walking excavators. In: International conference on advanced robotics (ICAR), pp 295–301
4. Hutter M, Leemann P, Hottiger G et al (2016) Force control for active chassis balancing. Trans Mechatr 22(2):613–622
5. Jelavic E, Hutter M (2019) Whole-body motion planning for walking excavators. In: IEEE/RSJ international conference on intelligent robots and systems (IROS), pp 2292–2299
6. Jelavic E, Berdou Y, Jud D et al (2020) Terrain-adaptive planning and control of complex motions for walking excavators. In: IEEE/RSJ international conference on intelligent robots and systems (IROS), pp 2684–2691
7. Egli P, Hutter M (2020) Towards Rl-based hydraulic excavator automation. In: IEEE/RSJ international conference on intelligent robots and systems (IROS), pp 2692–2697
8. Jelavic E, Farshidian F, Hutter M (2021) Combined sampling and optimization based planning for legged-wheeled robots. In: IEEE international conference on robotics and automation (ICRA), pp 8366–8372
9. Hu Z, Zhao D (2024) Dynamics and simulation of walking excavator chassis based on the virtual work principle. Multibody Syst Dyn 1–24
10. Tiwari D, Mittal G (2024) Modelling of a hydraulically actuated four bar loading dock lift. In: AIP conference proceedings, pp 2978–2981
11. Zhang Z, Zhang J, Luo N (2022) A study on trajectory control system of hydraulic excavators based on multi-domain physical model. Int J Fluid Power 395–410

Open Access This chapter is licensed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License (http://creativecommons.org/licenses/by-nc-nd/4.0/), which permits any noncommercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if you modified the licensed material. You do not have permission under this license to share adapted material derived from this chapter or parts of it.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.



Construction and Application of Crawler Crane Digital Twin System



Juntao Liu, Bing Li, Yang Zhang, Yulan Wei, Tongxi Zhu, and Yuanqi Hu

Abstract Crawler crane is a key equipment in large-scale infrastructure. In order to ensure operation safety, digital twin technology can be applied to crawler crane to monitor the working state of crawler crane in real time. Firstly, the physical model and virtual model of crawler crane were built, and the surface reduction optimization of the model was carried out by the edge collapse algorithm, so as to construct the reduced-order model. Secondly, the working condition data of the crawler crane is collected in real time through the pose sensor. Finally, the digital twin system was tested. The results show that by collecting the real-time working condition data of the experimental prototype and sending it to the crawled crane digital twin system, the virtual model can be driven to do the corresponding movement, and the real-time monitoring of the working attitude and data can be realized.

Keywords Crawler crane · Digital twin system · Condition monitoring

1 Introduction

At present, the digital twin system has entered a variety of fields, such as manufacturing, energy, construction, etc., which provides new ideas and methods for the optimal design, operation management, and fault diagnosis of entities [1]. In manufacturing, digital twin systems can be used to simulate product design and manufacturing processes, improving production efficiency and product quality [2]. In the

J. Liu · B. Li (✉) · Y. Wei · T. Zhu
School of Engineering, Huzhou University, Huzhous, China
e-mail: bingli@zjhu.edu.cn

Y. Zhang
Zhejiang Sany Equipment Co Ltd, Huzhou, China

Y. Hu
School of Information Engineering, Huzhou University, Huzhou, China

© The Author(s) 2026

H. Saman et al. (eds.), *The 9th International Conference on Advances in Construction Machinery and Vehicle Engineering*, Lecture Notes in Mechanical Engineering,
https://doi.org/10.1007/978-981-96-5527-4_27

field of energy, digital twin systems can be used for simulation analysis and optimized operation of power systems to improve energy efficiency and reliability [3]. With the continuous development of computer technology and sensor technology, the digital twin system has gradually matured and become an important technology to help the design of reality, which provides new ideas and methods for the design, operation management, and fault diagnosis of entities [4–6]. In recent years, digital twin systems have also been developed in the field of construction machinery. Wang Yijun established the virtual simulation model of bridge crane based on digital twin technology, and realized the functions of equipment operation status monitoring, fault diagnosis, maintenance and repair plan recommendation [7]. The constructed digital twin model for safe operation of the cable machine is stated to monitor hidden dangers and faults, and realize the safety monitoring, parametric control and visual display of the operation process of the cable machine [8]. Li Hui used the proposed variable universe fuzzy PID control method in crane digital twin system, and the system achieved the goal of virtual and real similarity of digital twin technology, and had good 3D visualization effect [9]. Deng Xu put forward the hydraulic gelled crane force and stability prediction monitoring system, through the finite element model and BP neural network prediction model to analyze the crane structure force, experimental results show that the system is feasible [10].

2 Construction of Crawler Crane Digital Twin Model and Environment Model

The crawler crane is primarily operated through key components such as the crawler chassis, slewing platform, and boom to execute intricate tasks including simultaneous lifting and movement. The crawler chassis serves as a foundational support for the entire mechanical structure. Positioned on the crawler chassis, the slewing platform provides support for the boom, hoist, and other essential equipment while enabling seamless 360° rotation through its integrated slewing mechanism. Employing a truss structure design, the crane arm maximizes material strength utilization, reduces dead weight, and enhances stiffness.

In the crawler crane digital twin system, the physical entity and virtual model serve as the foundational elements and primary data carriers. The twin mapping between the virtual model and the physical entity is achieved by creating a digital twin model for the crawler crane. Based on the specified functional requirements for the digital twin system, this chapter details the development of a physical model test prototype, the construction and optimization of the virtual model, and the establishment of a digital twin simulation environment within Unity 3D.

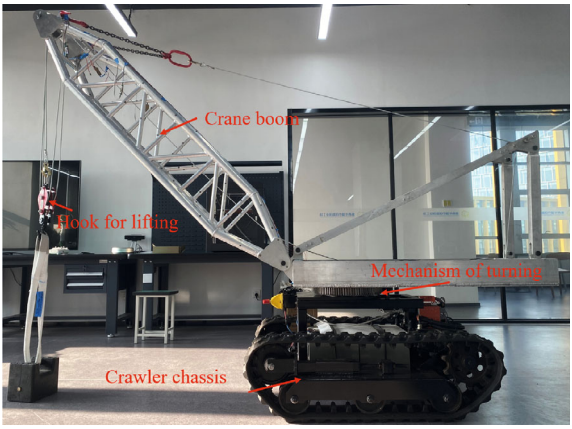


Fig. 1 Experimental prototype machine

2.1 Physical Entity Model Construction

The physical entity is the model basic object of the crawler crane digital twin system, which describes the physical information of the whole crawler crane entity. However, the actual crawler crane has a complex and large mechanical structure, which is difficult to assemble on site and requires a large number of personnel cooperation, and has high requirements for the operator's operation, and the actual operation safety risk is high.

Therefore, in order to facilitate the research and development of crawler crane digital twin system, the real crawler crane simulation model is used as the experimental object, which simplifies the mechanical structure of the solid crawler crane and has little relevance to the research, and reduces the modeling complexity of the virtual 3D model. The basic motion structure of the crawler crane includes the walking mechanism, the rotary mechanism, the boom variable amplitude mechanism and the lifting mechanism. The specific experimental prototype and constituent mechanisms are shown in Fig. 1.

2.2 Virtual Model Building and Optimization

The construction of crawler crane virtual model is an important part of the whole development of digital twin system, and the virtual model is the carrier of physical crawler crane twin data. To establish the visualization of digital twin, it is necessary to build a 3D model that is consistent with the geometric and physical size, color, shape and material of the crawler crane in the real work site. The established model can truly reflect the origin, motion relationship, subordinate relationship and relative relationship of each component structure of the crawler crane digital twin system. You have to have a visual twin as well as a structural twin.

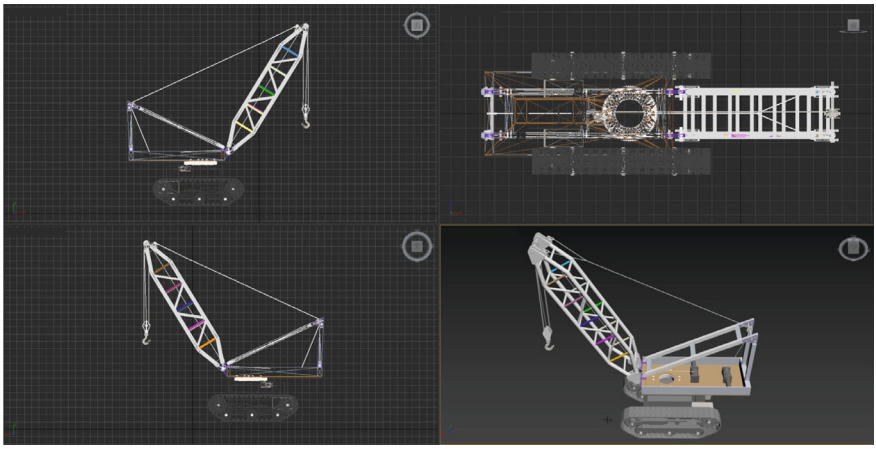


Fig. 2 Render the rendering

Rendering is used to give the crawler crane model with materials, maps and textures. Finally, the FBX format supported by unity3D is exported, and the specific rendering effect is shown in Fig. 2.

The established crawler crane model usually has too many part models and a large number of triangular faces for each part, which leads to a large memory occupancy, thus lengthening the model loading time and reducing the running speed of the whole digital twin system. Therefore, while not damaging the overall look and feel of the crawler crane model, consider reducing the number of parts in the model, by simplifying the geometry, reducing the number of facets or removing unnecessary details to reduce the complexity of the model. Some optimization tools and algorithms can help remove redundant information to reduce the memory footprint of the model.

Edge collapse is a method of simplifying 3D models, aiming to reduce the complexity of the model and preserve as much detail as possible. The key idea is to reduce the number of triangular facets by removing some edges and related vertices, thereby simplifying the model. The steps of the algorithm consist of merging two vertices of an edge into a single vertex, which minimally changes the original model shape.

In order to preserve the shape, a quadratic error measure is introduced, that is, the new vertex should be as close as possible to the plane where the triangles around the first two old vertices are merged, and the effective simplification of the model is achieved by minimizing the sum of squared distances of all planes. The formula is as follows.

$$E = \sum_i (ax_i + by_i + cz_i + d)^2 \tag{1}$$

where, a, b, c, d are the coefficients of the plane equation $ax + by + cz + d = 0$, (x_i, y_i, z_i) are the coordinates of the new vertex, and i is the index of the adjacent triangle.

Model optimization software was used to simplify the mesh of each structural module in the crawler crane model. Prior to optimization, the model contained 208,424 vertices and 422,310 triangular facets. After optimization, these were reduced to 192,216 vertices and 389,904 facets, achieving an overall optimization rate of 7.8% in vertex reduction and triangle count.

3 Construction of Crawler Crane Motion Simulation System

When constructing the digital twin system of crawler crane, in order to ensure the virtual model in the virtual scene to achieve the visualization effect of "virtual to real", it is necessary to make the motion mode and attitude of each motion mechanism consistent with the entity, and realize the synchronous simulation. To this end, the physical entity of the crawled crane and the digital twin model need to realize the mapping and interaction through data exchange, and the sensor data on the physical entity are sent to the digital twin system through the wireless transmission method with flexibility, easy scalability and strong anti-interference. WIFI transmission has the characteristics of low power consumption, low cost, short distance reliable transmission and easy to expand to multiple devices. It becomes an ideal choice for digital twin systems.

3.1 Digital Twin System Script Run Process Analysis

Script is an important component to control the behavior transformation of moving objects in virtual environment. It is a code file written in C# or JavaScript, which can be attached to the moving objects to realize the functions of responding to input, controlling motion, handling collision and so on.

Various components on the moving object can be accessed and manipulated through scripts, such as Transform, Renderer, Collider, etc., and the Get Component method can be used to obtain the reference of the corresponding component for operation. Scripts contain a set of lifecycle methods that are called at different stages of the script's lifecycle, allowing you to execute specific code when objects are created, updated, destroyed, and so on. The script running flow chart of the crawler crane motion simulation system is roughly shown in Fig. 3.

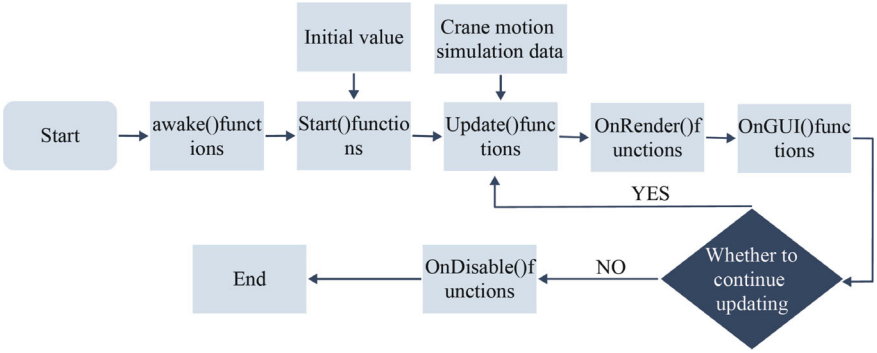


Fig. 3 Flowchart of script execution

3.2 Scene Resource Integration and System Testing

By sorting out the resources in the resource management, the pre-built scene model and the crawler crane model are dragged into the scene, and the position of each model is adjusted to build the virtual operating environment of the crawler crane. Adjust the relative position and Angle between the camera and the crawler crane model in the scene window, set the projection mode to perspective, adjust the size of the field of view and the distance of the clipping plane, and keep the crawler crane model in the real center of the camera perspective. Finally, the simulation platform built in Unity 3D is shown in Fig. 4.

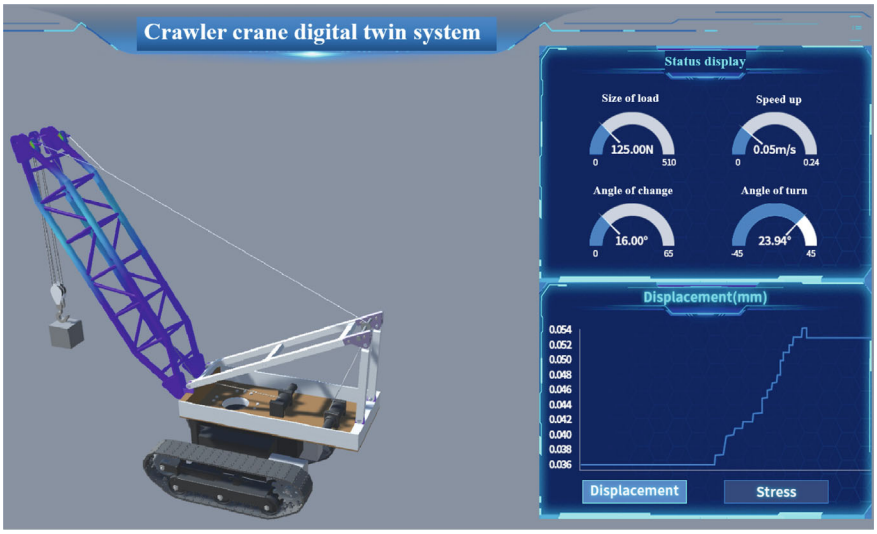


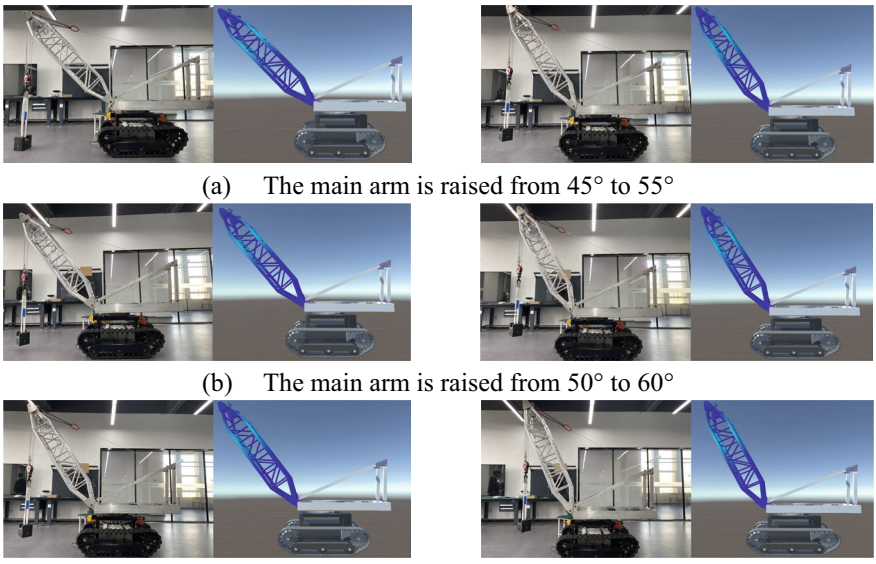
Fig. 4 Crawler crane simulation platform

The virtual model of the Siamese platform moves synchronously when the movement of the variable amplitude mechanism of the crawler crane is tested. The system response delay during the movement of the crawler crane's main arm was tested. The timestamp of the attitude sensor and the beginning of data transmission was obtained. After the data transmission and processing, the virtual model was controlled to move, and the timestamp was recorded again. The specific data are shown in Table 1.

For the crawler crane's main arm, three sets of system response delay records are made in the process of movement. In general, the delay of the system is less than or equal to 5 ms, indicating that the whole digital twin system has good real-time performance. Through the comparison of all data, it can be obtained that the slowest signal transmission time delay is 1.43 ms, and the fastest can reach 0.88 ms. The twin model in the digital twin system can map the motion attitude of the crawler crane prototype in the physical world in real time. Figure 5 is the correspondence diagram between the physical model and the main arm Angle of the twin model.

Table 1 System response delay test

Mechanism of motion	Angle before movement/°	Angle after motion/°	Signal acquisition frequency/Hz	Signal transmission time delay/ms	
				The slowest	The fastest
The main arm is raised	45	55	200	0.96	1.23
	50	60	200	0.91	1.03
	55	65	200	0.88	1.43



(a) The main arm is raised from 45° to 55°

(b) The main arm is raised from 50° to 60°

(c) The main arm is raised from 55° to 65°

Fig. 5 Correspondence diagram between the physical model and the main arm angle of the twin model

4 Conclusion

In order to ensure the operation safety of the crawler crane and monitor the posture of the crane in real time, this paper first builds the physical model and digital twin model of the crawler crane, and uses the edge collapse algorithm to optimize the surface reduction of the model, so as to construct the reduced-order model with an optimization rate of 7.8%. Secondly, the working condition data of the crawler crane is collected in real time through the pose sensor. Finally, the digital twin system was tested. The results show that by collecting the real-time working condition data of the experimental prototype and sending it to the crawled crane digital twin system, the virtual model can be driven to do the corresponding motion, and the real-time monitoring of the working attitude and data can be realized. The signal delay transmission is up to 0.88 ms. Finally, the digital twin system was tested. The results show that by collecting the real-time working condition data of the experimental prototype and sending it to the crawled crane digital twin system, the virtual model can be driven to do the corresponding movement, and the real-time monitoring of the working attitude and data can be realized.

Acknowledgements Thanks to the Huzhou Public Welfare Application Research Project (2022GZ13), the Zhejiang Provincial "Jianbing" "Lingyan" Research and Development Program of China (2022C01132), and the Postgraduate Research and Innovation Project of Huzhou University Project (2024KYCX56).

References

1. LuJ, Xia L, Zhang H et al (2022) Research and application of manufacturing enterprise digital twin ecosystem. Comput Integr Manuf Syst 28(8):2273–2290
2. Tao F, Zhang H, Qi Q et al (2021) Theory of digital twin modeling and its application. Comput Integr Manuf Syst27(1):1–15
3. Zhao P,Pu T, Wang X et al (2022) Key technologies and perspectives of power internet of things facing with digital twins of the energy internet. Proc CSEE 42(2):447–458
4. Zhang F, Ge S (2023) Construction method and evolution mechanism of mine digital twins. J China Coal Soc 48(1):510–522
5. Deng J, Wei S, Shi X et al (2021) Digital twin-based distribution management system. Comput Integr Manuf Syst 27(2):585–604
6. Chenyuan Z, Fei T (2021) Evaluation index system for digital twin model. Comput Integr Manuf Syst 27(8):2171–2186
7. Wang Y, Chen X, Zhang Z et al (2024) Based on the number of twin bridge crane intelligent health management system design. J Mach Des9(S2):61–65
8. Lu S, Wang J et al (2012) Construction method of cable crane safe operation digital twin model based on Unity3D. J Saf Sci Technol 20(01):154–159
9. Li H, Hui Y, Wang G et al (2024) Research on autonomous operation system of bridge crane based on digital twin. J Mech Electr Eng41(04):691–700
10. Deng X, Zhang A (2024) Research and application of force and stability monitoring system for hydraulic gane crane. Mod Manuf Eng(07):135–143

Open Access This chapter is licensed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License (http://creativecommons.org/licenses/by-nc-nd/4.0/), which permits any noncommercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if you modified the licensed material. You do not have permission under this license to share adapted material derived from this chapter or parts of it.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.



Gearbox Fault Diagnosis Method Based on Model Transfer and Multi-sensor Self-learning Weighted Fusion



Yu Zhou, Long Chen, and Dongguang Zhang

Abstract In response to the problem that single sensor data is easily affected and interfered with by the environment, making it difficult to collect accurate gearbox fault signals, resulting in low fault recognition accuracy, this paper proposes a multi-sensor self-learning weighted fusion method. This method not only effectively improves the accuracy of fault diagnosis but also has good interpretability. Firstly, the signals of different sensors are transformed into time–frequency images through sliding window and continuous wavelet transform to improve the signal to noise ratio; In order to fully extract fault features from the time–frequency images, a feature extraction module is constructed using model transfer method for feature extraction; Then, the self-learning weighted fusion module amplifies the fault features in the feature matrix while suppressing interference features to improve the accuracy of gearbox fault diagnosis; Finally, the classification is carried out through the classification module. The average accuracy of this method under different operating conditions on the gearbox dataset of Southeast University and the WT planetary gearbox dataset of Beijing University of Technology is higher than 99.9%.

Keywords Residual network · Model transfer · Self-learning weighted fusion · Multi-sensor · Gearbox · Interpretability

1 Introduction

As a key mechanical transmission component, the gearbox plays a crucial role in the operation of the system. However, gearboxes are prone to failure in harsh working environments, which not only affects their own performance but may also cause abnormalities in other components of the mechanical system, potentially threatening

Y. Zhou · L. Chen (✉) · D. Zhang
College of Mechanical Engineering, Taiyuan University of Technology, Taiyuan, China
e-mail: chenlong@tyut.edu.cn

L. Chen
China Coal Science and Technology Group Taiyuan Research Institute Co., Ltd., Taiyuan, China

© The Author(s) 2026

H. Saman et al. (eds.), *The 9th International Conference on Advances in Construction Machinery and Vehicle Engineering*, Lecture Notes in Mechanical Engineering,
https://doi.org/10.1007/978-981-96-5527-4_28

the stability of the entire system. Therefore, fault diagnosis of gearboxes is crucial [1].

In recent years, driven by artificial intelligence technology, deep learning has also been widely used for fault diagnosis [2]. For example: Shang et al. proposed an interpretable end-to-end network for machine fault diagnosis by embedding continuous wavelet transform into deep learning models [3]. Han et al. achieved more accurate fault diagnosis of wind turbines with only 30 labeled samples based on an improved generative adversarial network [4]. Qian et al. proposed a relationship transfer diagnosis framework based on generative adversarial networks to solve the domain confusion problem in practical diagnostic scenarios, which can reduce the distribution difference between the source domain and the unknown target domain, for rotating machinery fault diagnosis [5]. Zhang et al. proposed a signal enhanced semi supervised learning scheme based on generative adversarial networks for bearing fault diagnosis [6].

The above studies are all single sensor fault diagnosis methods. Due to the poor stability of the signals collected by a single sensor, multiple sensors need to be applied in practical engineering to detect the working status of the gearbox. In addition, compared to a single sensor, multiple sensors can fully reflect the fault situation of the gearbox, making their detection results more accurate [7]. Therefore, multi-source sensor fusion technology for machine fault diagnosis is currently one of the research hotspots. For example: Chen et al. proposed a method for wind turbine fault diagnosis that utilizes attention mechanism to fuse frequency domain features of multiple sensors. Although this method has high accuracy, the fusion method is relatively complex [8]. Wang et al. fused multiple vibration sensor signals into feature rich multi-channel images for rotating machinery fault diagnosis. Although this method can fully utilize the original data to overcome feature information loss, it lacks interpretability [9]. Jinag et al. proposed a multi-sensor decision level fusion method based on Dempster Shafe evidence theory for mechanical equipment fault diagnosis. Although this method has good interpretability, the information loss after fusion is severe [10]. Wang et al. proposed a newly designed multi signal Glaman angle difference field method that combines sensor signals from different positions into a single image for permanent magnet synchronous motor fault diagnosis. Although this method has high accuracy, it lacks interpretability [11].

Based on the above issues, this paper proposes a gearbox fault diagnosis method based on model transfer and multi-sensor self-learning weight fusion. The innovation of this method is as follows:

(1) In order to fully extract the fault features from the time–frequency graph, a feature extraction module was built using model transfer method.
(2) A interpretable self-learning weighted fusion method has been proposed that can amplify fault features in the feature matrix while suppressing interference features.

2 Proposed Method

The gearbox fault diagnosis model based on model transfer and multi-sensor self-learning weight fusion proposed in this article consists of a signal processing module (SPM), a feature extraction module (EMF), a self-learning weighted fusion module (SLWFM), and a classification module (CM). Its overall structure is shown in Fig. 1.

Firstly, use a sliding window to divide the vibration signals collected by different vibration sensors into fixed length time-domain signals, and then convert them into time–frequency pictures through the signal processing module. Then input the time–frequency pictures of different sensors into different feature extraction modules for feature extraction. After feature extraction, use the self-learning weighted fusion module to fuse the feature matrices of different sensors to amplify fault features and suppress interference features. Finally, use the classification module for classification and recognition.

2.1 Signal Processing Module

In order to improve the signal-to-noise ratio, this paper chooses continuous wavelet transform as the signal processing method.

Continuous wavelet transform is achieved through the integration of wavelet basis functions and objective functions, and its mathematical definition is as follows [12]:

$$W_{a,b} = \frac{1}{\sqrt{|a|}} \int x(t)\Phi^*\left(\frac{t-b}{a}\right)dt \tag{1}$$

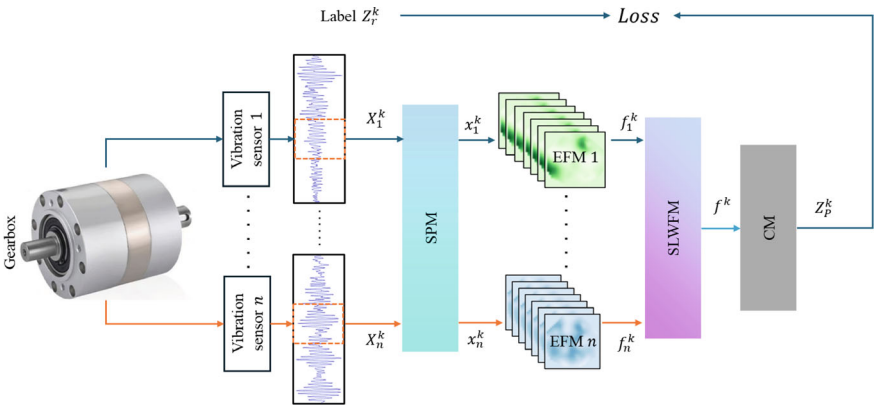


Fig. 1 The framework of the proposed method

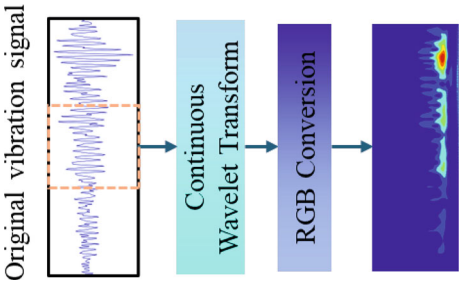


Fig. 2 Signal processing flow

where a is the scale parameter inversely proportional to frequency, b is the translation parameter, $x(t)$ is the one-dimensional time-domain signal, $\Phi^*\left(\frac{t-b}{a}\right)$ is the wavelet basis function, and $W_{a,b}$ is the wavelet coefficient.

In this article, the CMOR wavelet basis function is selected to perform wavelet transform on the vibration signal collected by the sensor, and then the RGB transform is used to convert it into a time–frequency pictures, as shown in Fig. 2. The process is described mathematically as follows:

$$x_i^k = RGB\left(W_{a,b}\left(X_i^k\right)\right) = S\left(X_i^k\right) \tag{2}$$

where X_i^k is the k-th sample of the i-th sensor, x_i^k is the k-th time–frequency diagram of the i-th sensor, $RGB(.)$ is the mathematical expression for RGB transformation, and $S(.)$ is the mathematical expression for the signal processing module.

2.2 *Feature Extraction Module*

In order to accurately extract fault features from time–frequency maps, the feature extraction module is built by transferring the Resnet18 model. The specific method of transfer is shown in Fig. 3. From Fig. 3, it can be seen that the feature extraction module consists of convolutional layers and pooling layers, and its mathematical expression can be described as:

$$f_i^k = F_i\left(x_i^k\right) \tag{3}$$

where $F_i(.)$ is the mathematical expression function of the i-th EFM, and f_i^k is the k-th feature matrix extracted by the i-th EFM.

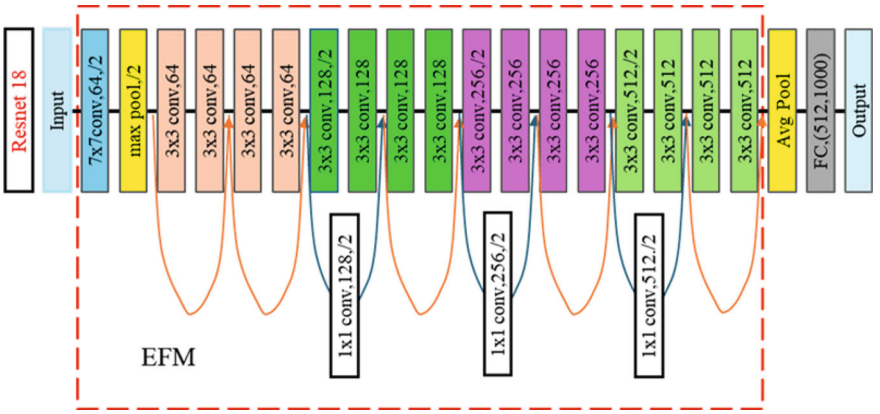


Fig. 3 Detailed structure of feature extraction module

2.3 Self-learning Weighted Fusion Module

Due to the feature matrix extracted by the feature extraction module from the original vibration sensor signal, in most cases, its fault features are not obvious, making it difficult for the classifier to accurately classify; Therefore, this article proposes a multi-sensor self-learning weighted fusion method that can amplify the fault features of the feature matrix while suppressing interference features, enabling the classifier to classify more accurately. The principle is as follows:

When the fault features of the feature matrix extracted are not obvious (as shown in Example 1 of Fig. 4), it is difficult for the classifier to accurately classify them; If a simple addition method is used for fusion (as shown in Example 2 of Fig. 4), the fault features are still not obvious after fusion, so it is still difficult for the classifier to accurately classify them; However, when multiplying the two feature matrices by a weight coefficient and then adding them together (as shown in Example 3 of Fig. 4), the fault features become more apparent, and the classifier can accurately classify them.

Based on the inspiration from the above situation, in order to find a fusion weight that amplifies the fault features as much as possible, this paper proposes a multi-sensor self-learning weighted fusion method, whose mathematical expression is as follows:

$$f^k = \sum_{i=1}^{n} a_i f_i^k a_i \in [0,1] \text{and} \sum_{i=1}^{n} a_i = 1 \tag{4}$$

where f^k is the k-th feature matrix after fusion, a_i is the fusion weight of the i-th feature matrix that can be trained together with the model, and n is the number of sensors.

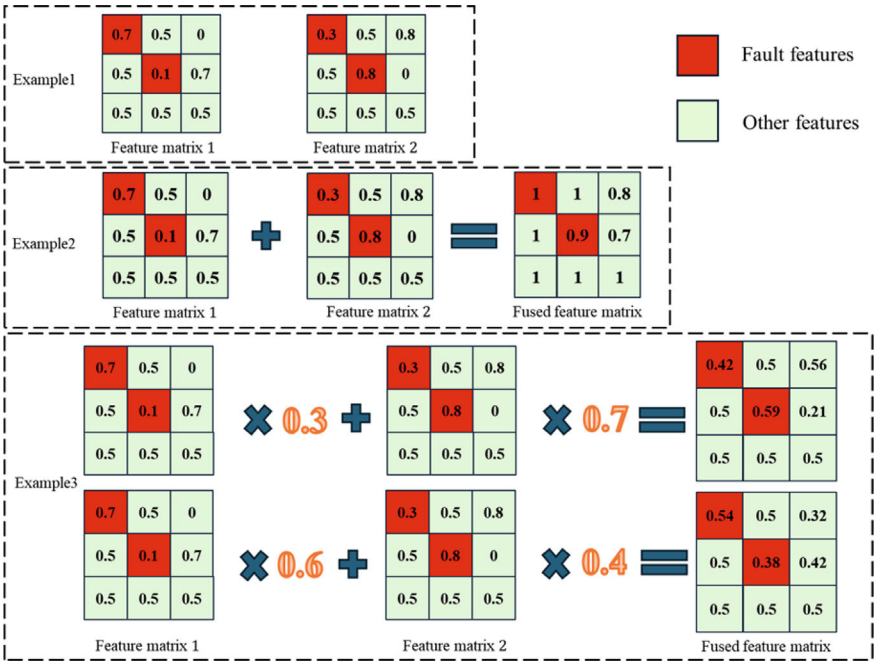


Fig. 4 The impact effect of weighted fusion

Self-learning weighted fusion is mainly reflected in giving an initial value of $a_1 a_2 \ldots a_n$, and then continuously adjusting it as the model is trained. When the loss function of the model reaches convergence, the weights $a_1 a_2 \ldots a_n$ can amplify the fault features as much as possible, making it easier for the classifier to accurately classify.

2.4 Classification Module

The classification module consists of three parts: adaptive average pooling layer, dimensionality reduction layer, and fully connected layer. The specific process is shown in Fig. 5, and the specific parameters are shown in Table 1.

The adaptive average pooling layer aims to further reduce the size of the fused feature matrix to reduce computational complexity. The process can be described as [13]:

$$p_i^{l+1} = \underset{(j-1)w+1 \le t \le jw}{avg} \left\{ q_i^l(t) \right\} \tag{5}$$

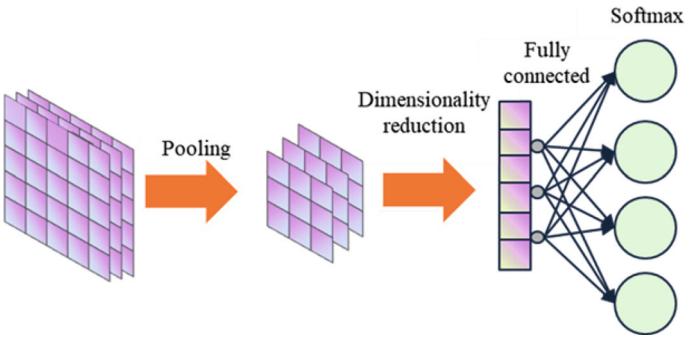


Fig. 5 Detailed structure of classification module

Table 1 Classification module parameters

Layer	Output size
Input	$8 \times 8 \times 512$
Adaptive average pooling	$2 \times 2 \times 512$
Dimensionality reduction	2048
Fully connected	5

where $q_i^l(t)$ is the value of the t-th neuron in the i-th frame of layer l; w is the width of the pooling area; p_i^{l+1} is the corresponding value of the neuron in the $l+1$ layer of the pooling operation.

The fully connected layer aims to convert one-dimensional feature vectors into feature vectors with a length equal to the number of categories through matrix multiplication, and normalize the output feature vectors. The process can be described as:

$$\begin{bmatrix} z_1 \\ \vdots \\ z_s \end{bmatrix}^T = Softmax\left(\begin{bmatrix} x_1 \\ \vdots \\ x_k \end{bmatrix}^T \begin{bmatrix} w_{11} & \dots & w_{1s} \\ \vdots & \ddots & \vdots \\ w_{k1} & \dots & w_{ks} \end{bmatrix} + \begin{bmatrix} b_1 \\ \vdots \\ b_s \end{bmatrix}^T \right) = \begin{bmatrix} \frac{e^{(\sum_{j=1}^{k} x_j w_{j1}) + b_1}}{\sum_{i=1}^{s} e^{(\sum_{j=1}^{k} x_j w_{ji}) + b_i}} \\ \vdots \\ \frac{e^{(\sum_{j=1}^{k} x_j w_{js}) + b_s}}{\sum_{i=1}^{s} e^{(\sum_{j=1}^{k} x_j w_{ji}) + b_i}} \end{bmatrix}^T \tag{6}$$

where $\begin{bmatrix} x_1 & \dots & x_k \end{bmatrix}$ is the input vector, $\begin{bmatrix} w_{11} & \dots & w_{1s} \\ \vdots & \ddots & \vdots \\ w_{k1} & \dots & w_{ks} \end{bmatrix}$ is the weight matrix of the fully connected layer, $\begin{bmatrix} b_1 & \dots & b_s \end{bmatrix}$ is the bias term, Softmax(.) is the normalization function, $\begin{bmatrix} z_1 & \dots & z_s \end{bmatrix}$ is the predicted label, and s must be equal to the number of categories to be classified.

The entire classification module can be described as:

$$Z_p^k = Softmax\left(D\left(avg\left(f^k\right)\right).w + b\right) = C\left(f^k\right) \tag{7}$$

where $D(.)$ is the mathematical expression of the dimensionality reduction layer, and Z_p^k is the k-th predicted label.

2.5 Loss Function

In this article, the Cross entropy loss function is used to calculate the classification loss of predicted labels and true labels. Since the number of sensors n in this article is 2, the loss function can be rewritten as:

$$Loss = -\frac{1}{N}\sum_{k=1}^{N} Z_r^k \ln\left(C\left(\sum_{i=1}^{n} a_i F_i\left(S_i\left(X_i^k\right)\right)\right)\right) \tag{8}$$

where Z_r^k is the true label of the k-th sample, N is the number of samples.

Since the number of sensors n in this article is 2, the loss function can be rewritten as:

$$Loss = -\frac{1}{N}\sum_{k=1}^{N} Z_r^k \ln\left(C\left(a_1 F_1\left(S_1\left(X_1^k\right)\right) + a_2 F_2\left(S_2\left(X_2^k\right)\right)\right)\right) = -\frac{1}{N}\sum_{k=1}^{N} Z_r^k \ln\left(Z_p^k\right) \tag{9}$$

where X_1^k and X_2^k are the k-th samples collected by the sensor in the X and Y directions, respectively.

Finally, by continuously optimizing the *Loss* through gradient descent algorithm, the model training can be achieved. After the model is trained, the vibration signals collected by the sensors can be directly input into the model to predict the fault type of the gearbox.

3 Experiment

In order to verify the effectiveness of the multi-sensor self-learning weighted fusion (MSSLWF) method proposed in this paper, it was validated on the Southeast University gearbox dataset [14] and the Beijing University of Technology WT planetary gearbox dataset [15], and three control groups (single sensor X, single sensor Y, and multi sensor non self-learning weight fusion (MSNSLWF) were set up for comparison.

3.1 Southeast University Gearbox Dataset

The Southeast University gearbox dataset includes two sub datasets: bearing data and gear data, with a sampling frequency of 1024 Hz. The gear dataset is divided into two working conditions: 20 Hz–0 V and 30 Hz–2 V. Each working condition has 5 types of faults (Health, Surface, Root, Miss, Chipped). The Excel file for each fault type has 8 columns, with columns 1–8 being the motor vibration signal, planetary gearbox X, Y, Z direction vibration signal, motor torque signal, and parallel gearbox X, Y, Z direction vibration signal (where columns 2 and 3 are the original vibration signals X and Y used for testing). This dataset is derived from the Drive System Dynamics Simulator (DDS). The collected driving system dynamics simulator is shown in Fig. 6.

Firstly, the original vibration signal of the X-direction sensor is transformed into a time–frequency images with 512 data points, and then divided into a training set and a testing set. The training set accounts for 80%, with a total of 8000 images, 1600 images for each fault type, and 20% for the testing set, with a total of 2000 images and 400 images for each fault type; The original vibration signal of the Y-direction sensor is also divided into a training set and a testing set using the same method. During training, use the Adam optimizer, set the Batchsize to 100, set the learning rate to 0.001, and iterate 50 times. The test results under two operating conditions of 20 Hz–0 V and 30 Hz–2 V are shown in Figs. 7 and 8, respectively.

From point and line charts and bar charts in Figs. 7a and 8a, it can be seen that the multi-sensor self-learning weighted fusion method has the highest fault recognition accuracy compared to the multi-sensor non self-learning weighted fusion method, single sensor X, and single sensor Y under two different operating conditions; From the confusion matrices in Figs. 7b and 8b, it can be seen that the highest accuracy is 100% under both operating conditions of 20 Hz–0 V and 30 Hz–2 V.

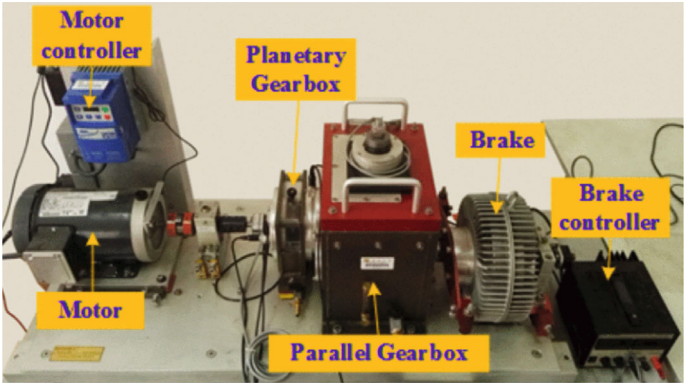


Fig. 6 Drive system dynamics simulator at southeast university

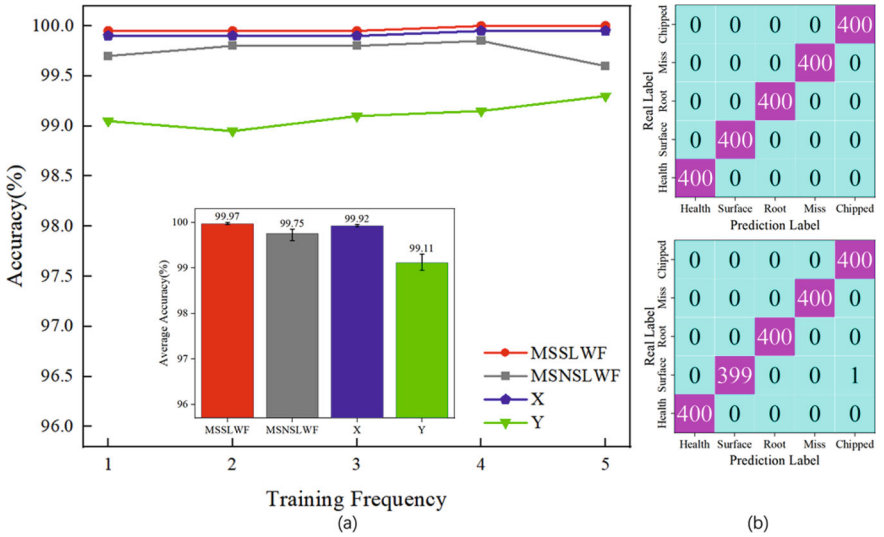


Fig. 7 Test results of Southeast University gearbox dataset under 20 Hz–0 V operating conditions, **a** Accuracy and average accuracy of 5 training sessions using different methods, **b** Confusion matrix of the highest and lowest accuracy in 5 training sessions using MSSLWF method

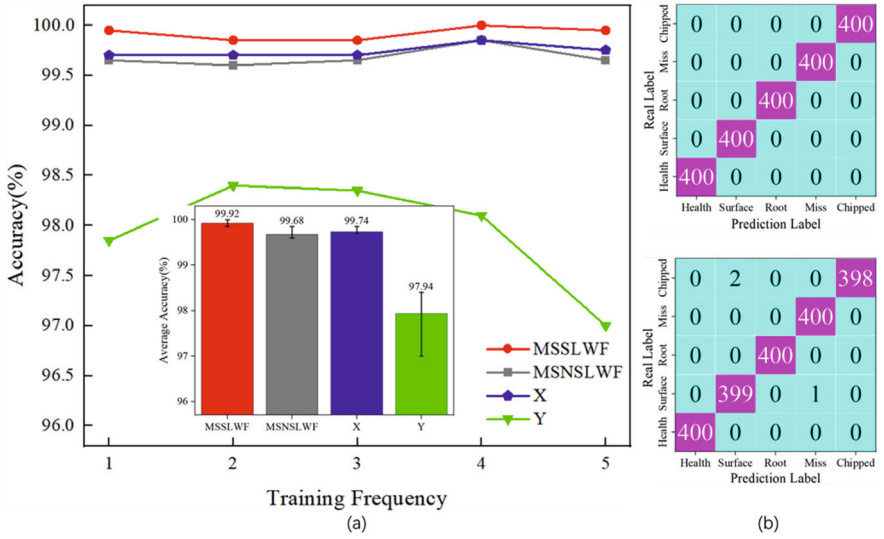


Fig. 8 Test results of Southeast University gearbox dataset under 30 Hz–2 V operating conditions, **a** accuracy and average accuracy of 5 training sessions using different methods, **b** confusion matrix of the highest and lowest accuracy in 5 training sessions using MSSLWF method

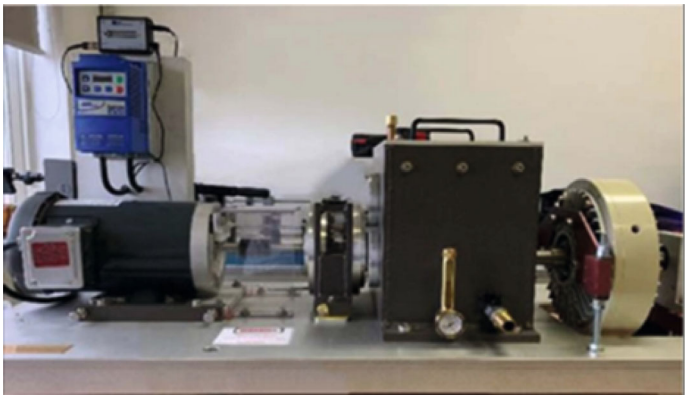


Fig. 9 The driving system dynamics simulator of Beijing institute of technology

3.2 Beijing University of Technology WT Planetary Gearbox Dataset

The sampling frequency of the WT planetary gearbox dataset at Beijing University of Science and Technology is 48,000 Hz, with 8 working conditions including 20, 25, 30, 35, 40, 45, 50, 55 Hz. Each working condition has 5 types of faults (Healthy condition (H), gear with a broken tooth (B), wear gear (W), crack occurs in the root (R), missing one tooth (M)), and the mat file for each fault type has 4 columns. Columns 1 to 3 are the vibration signals of the X and Y direction sensors of the planetary gearbox and the motor speed signals (where columns 1 and 2 are the original vibration signals X and Y used for testing). This dataset was also collected from the Drive System Dynamics Simulator (DDS), as shown in Fig. 9.

Due to the high sampling frequency of the dataset, a downsampling method was first used to reduce the frequency to 4800 Hz, and then the vibration signals of the X and Y direction sensors after downsampling were converted into time–frequency images with 600 data points. The dataset partitioning method and hyperparameter settings during training are the same as above. The test results at 20 Hz and 30 Hz are shown in Figs. 10 and 11, respectively.

From point and line charts and bar charts in Figs. 10a and 11a, it can be seen that the multi-sensor self-learning weighted fusion method has the highest fault recognition accuracy compared to the multi-sensor non self-learning weighted fusion method, single sensor X, and single sensor Y under two different operating conditions; From the confusion matrices in Figs. 10b and 11b, it can be seen that the highest accuracy is 100% under both operating conditions of 20 Hz–0 V and 30 Hz–2 V.

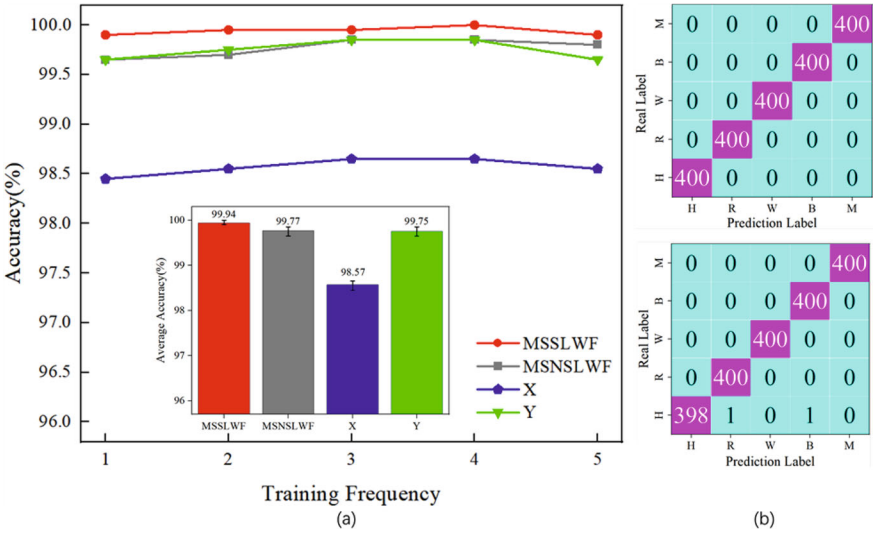


Fig. 10 Test results of the WT planetary gearbox dataset from Beijing University of Technology under 20 Hz operating conditions, **a** accuracy and average accuracy of different methods trained 5 times, **b** confusion matrix of the highest and lowest accuracy of the MSSLWF method trained 5 times

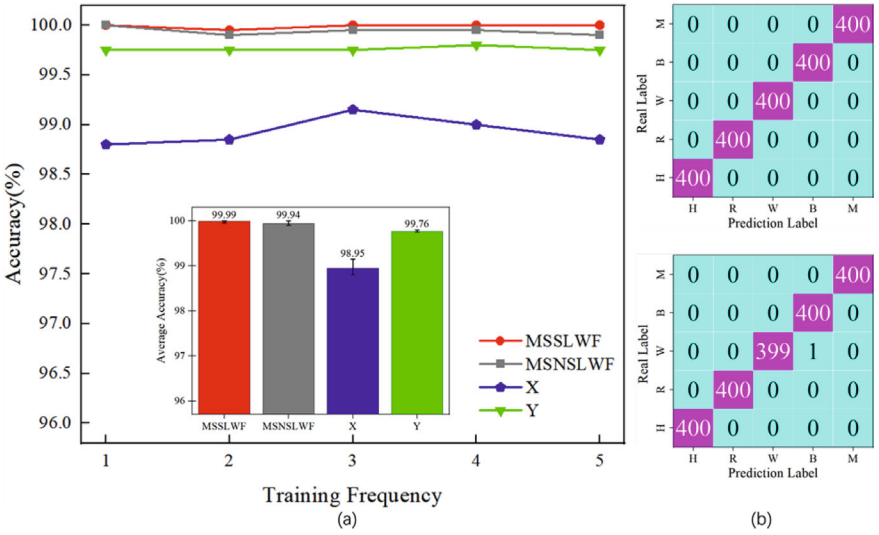


Fig. 11 Test results of the WT planetary gearbox dataset from Beijing University of Technology under 30 Hz operating conditions, **a** accuracy and average accuracy of different methods trained 5 times, **b** confusion matrix of the highest and lowest accuracy of the MSSLWF method trained 5 times

4 Conclusions

In order to solve the problem of low accuracy in gearbox fault diagnosis using a single sensor, this paper proposes a gearbox fault diagnosis method based on model transfer and multi-sensor self-learning weight fusion. Through experimental analysis, the following conclusions are drawn:

(1) The feature extraction module built using model transfer can fully extract fault features from images.
(2) The multi-sensor self-learning weight fusion method proposed in this article can effectively amplify the fault features in the feature matrix while suppressing interference features to improve the accuracy of fault diagnosis, and has interpretability.
(3) The test results on the Southeast University gearbox dataset and the Beijing University of Technology WT planetary gearbox dataset show that the multi-sensor self-learning weight fusion method proposed in this paper has the highest accuracy compared to the single sensor X, single sensor Y, and multi-sensor non self-learning weight fusion method, with an average accuracy of over 99.9% and the highest accuracy of 100%. Therefore, the method proposed in this paper is effective.

References

1. He C, Yasenjiang J, Lv L, Xu L, Lan Z (2024) Gearbox fault diagnosis based on MSCNN-LSTM-CBAM-SE. Sensors 24:4682. https://doi.org/10.3390/s24144682
2. Xu Q, Jiang H, Zhang X, Li J, Chen L (2023) Multiscale convolutional neural network based on channel space attention for gearbox compound fault diagnosis. Sensors 23:3827. https://doi.org/10.3390/s23083827
3. Shang Z, Zhao Z, Yan R (2023) Denoising fault-aware wavelet network: a signal processing informed neural network for fault diagnosis. Chin J Mech Eng 36:9
4. Han T, Xie W, Pei Z (2023) Semi-supervised adversarial discriminative learning approach for intelligent fault diagnosis of wind turbine. Inf Sci 648:119496. ISSN 0020-0255
5. Qian Q, Zhou J, Qin Y (2023) Relationship transfer domain generalization network for rotating machinery fault diagnosis under different working conditions. IEEE Trans Ind Inf 19(9):9898–9908. https://doi.org/10.1109/TII.2022.3232842
6. Zhang T, Chen J, Pan T, Zhou Z (2020) Towards Intelligent fault diagnosis under small sample condition via a signals augmented semi-supervised learning framework. In: 2020 IEEE 18th international conference on industrial informatics (INDIN), Warwick, United Kingdom, pp 669–672. https://doi.org/10.1109/INDIN45582.2020.9442224
7. Guo J, He Q, Zhen D, Gu F, Ball AD (2024) Multiscale cyclic frequency demodulation-based feature fusion framework for multi-sensor driven gearbox intelligent fault detection. Knowl Based Syst 283:111203. ISSN 0950-7051, https://doi.org/10.1016/j.knosys.2023.111203
8. Chen Y, Zheng X (2023) Fault diagnosis of wind turbine based on multi-signal CNN-GRU model. Proc Inst Mech Eng Part A J Power Energy 237(5):1113–1124. https://doi.org/10.1177/09576509231151482

9. Wang H, Zhu H, Li H (2023) A rotating machinery fault diagnosis method based on multi-sensor fusion and ECA-CNN. IEEE Access 11:106443–106455. https://doi.org/10.1109/ACCESS.2023.3320065
10. Jiang D, Wang Z (2023) Research on mechanical equipment fault diagnosis method based on deep learning and information fusion. Sensors 23:6999
11. Wang J, Ma J, Meng D, Zhao X, Zhang K (2023) Fault diagnosis of PMSMs based on image features of multi-sensor fusion. Sensors 23:8592
12. Teta A, Korich B, Bakria D et al (2024) Fault detection and diagnosis of grid-connected photovoltaic systems using energy valley optimizer based lightweight CNN and wavelet transform. Sci Rep 14:18907. https://doi.org/10.1038/s41598-024-69890-7
13. Geng Y, Su S, Zhang T, Zhu Z (2023) A novel multi-scale one-dimensional convolutional neural network for intelligent fault diagnosis of centrifugal pumps. J Mar Sci Eng 11:2278. https://doi.org/10.3390/jmse11122278
14. Shao S, McAleer S, Yan R, Baldi P (2019) Highly accurate machine fault diagnosis using deep transfer learning. IEEE Trans Ind Inf 15(4):2446–2455. https://doi.org/10.1109/TII.2018.2864759
15. Liu D, Cui L, Cheng W (2024) A review on deep learning in planetary gearbox health state recognition: methods, applications, and dataset publication. Meas Sci Technol 35:012002

Open Access This chapter is licensed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License (http://creativecommons.org/licenses/by-nc-nd/4.0/), which permits any noncommercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if you modified the licensed material. You do not have permission under this license to share adapted material derived from this chapter or parts of it.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.



Research on Automatic Control System of Sprinkler Truck Filling Robot Arm in Mining Area



Yandong Ma, Deli Liu, Kun Ma, Meng Zhang, and Zhong Wang

Abstract To solve the vibration problem of the existing intelligent water station robotic arm operation in open-pit coal mines and avoid safety issues caused by vibration, the motion characteristics of a four degree of freedom robotic arm are studied. Using geometric methods to solve its kinematic model, and then simulating a single hydraulic system in Matlab. Finally, the coordinates of the sprinkler truck were provided through the Beidou system. Based on the given system model, the actual displacement and angle of each degree of freedom were obtained. The system automatically controlled the water inlet to reach the designated position, achieving the automatic control function of the water injection robotic arm.

Keywords Water injection robotic arm · Trajectory planning · Trajectory tracking

1 Introduction

The dust and dirt generated during surface stripping, blasting, shoveling, transportation, unloading, and other processes in mining operations cause significant air pollution. In the dust reduction operation of open-pit mines, there are various types of dust removal equipment, among which the mining sprinkler truck is the most widely used and concentrated [1]. At the same time, coal mine fires are frequent, and prevention and control are of utmost importance. The fire water cannon equipped on the sprinkler truck is a relatively effective firefighting method [2]. At present, the intelligent water adding station used in open-pit coal mines has a large size. When the water adding mechanism mechanical arm is positioned, it generates vibration during startup and braking, resulting in inaccurate and unstable positioning, reducing work efficiency, damaging its own structure, and reducing its service life. To improve the

Y. Ma · D. Liu
Jarud Banner Zhahanaoer Coal Industry Co., Ltd., Inner Mongolia, China

K. Ma (✉) · M. Zhang · Z. Wang
School of Vehicle and Energy, Yanshan University, Qinhuangdao, Hebei, China
e-mail: 495157649@qq.com

© The Author(s) 2026

H. Saman et al. (eds.), *The 9th International Conference on Advances in Construction Machinery and Vehicle Engineering*, Lecture Notes in Mechanical Engineering,
https://doi.org/10.1007/978-981-96-5527-4_29

motion accuracy and stability of robotic arms, many scholars at home and abroad have studied their motion characteristics [3].

Dong has designed a four axis palletizing robot for production line palletizing operations to meet the needs of cargo handling Perform kinematic analysis on the robot using the D-H motion method [4]. MIN proposed a passive yaw pitch structure (CBM) consisting of springs and metal wires. Through geometric analysis and experiments, it has been demonstrated that the proposed CBM can effectively compensate for the gravitational torque caused by the weight and payload of the robot [5]. Lu proposed a method for position control of a robotic arm based on a linear quadratic regulator (LQR), established a dynamic model of the robotic arm, derived the state equation using the dynamic model, and added fuzzy algorithm to traditional LQR control to improve the adaptability of the control system [6]. References [7, 8] respectively applied polynomial interpolation and high-order polynomial methods to study the trajectory planning and smooth control problems of multi degree of freedom robotic arms. In summary, there are few research results on solving mechanical vibration that are applicable to mine water injection robotic arms.

In view of this, this article takes a four degree of freedom robotic arm as the research object. Based on the parking position of the sprinkler truck, GPS is used to input its coordinates. In the planning process, geometric methods are applied to solve the inverse solutions of its various joints. Then, based on the obtained results, PID control method is used to compare the closed-loop trajectory tracking control of the example valve, in order to provide research ideas and data support for more accurate and stable operation of the water station robotic arm.

2 Kinematic Model of Robotic Arm

The water adding mechanism of the automatic water injection station is a 4-degree-of-freedom robotic arm, which is composed of four rotating pairs. Its power supply comes from the hydraulic system, which includes two hydraulic cylinders and two hydraulic motors. The model is shown in Fig. 1.

The trajectory of the water outlet is determined by the combined motion of the boom and the water outlet pipe. To achieve the expected trajectory control of the water outlet pose, it is necessary to establish the motion relationship between the water outlet and each joint, and solve its kinematic model. The commonly used robot kinematics solving methods include D-H method, geometric method, traversal iteration method, spinor theory solving method, etc. [9] each of which has its own applicable characteristics. This article uses geometric methods to solve its inverse kinematics model.

Among them, the water adding robotic arm is 3.8 m above the ground, and the hydraulic motor of the water outlet pipe and the hydraulic motor of the boom can achieve a rotation operation of − 180° ~ 180°. The hydraulic cylinder of the water outlet pipe and the hydraulic cylinder of the boom can extend and retract to push the water outlet pipe and the boom to achieve a pitch operation of − 13° ~ 18°. When

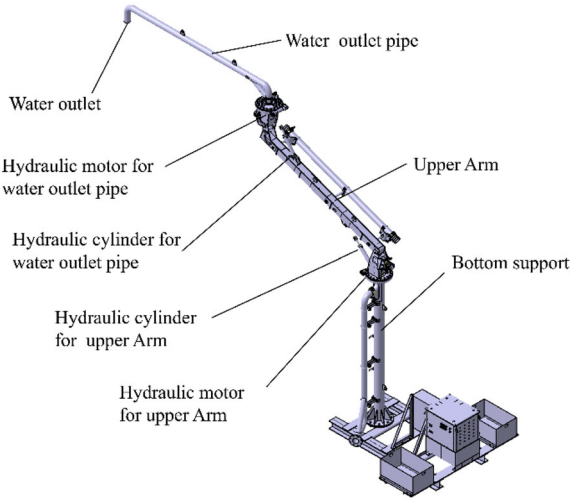


Fig. 1 Three dimensional diagram of water injection robotic arm

adding water, ensure that the water outlet is perpendicular to the water adding port of the sprinkler truck.

Firstly, simplify the water adding robotic arm by solving for the joint angles, and then solve for the control variables such as hydraulic cylinder stroke. Establish a spatial coordinate system, where points are the coordinates of the water outlet, line segments are the projections of the water outlet pipe and the boom on a plane, and line segments are the projections of the water outlet pipe and the boom on a plane. The green square area is the workspace of the water adding robotic arm, and counterclockwise is defined as the positive direction, as shown in Fig. 2.

Using the cosine theorem, it can be concluded that:

$$cos\gamma_2 = -cos(\pi - \gamma_2) = \frac{x^2 + y^2 - (OB')^2 - (A''B')^2}{2OB' \times A''B'} \tag{1}$$

$$\varphi = arccos\frac{(OB')^2 + (x^2 + y^2) - (A''B')^2}{2OB'\sqrt{x^2 + y^2}}, \varphi \in [0, 90°) \tag{2}$$

$$B'O = \sqrt{L_B^2 - (z - AA')^2} \tag{3}$$

$$BO = B'Ocos\gamma_1 \tag{4}$$

$$\beta_1 = arctan\frac{z - AA'}{\sqrt{BO^2 - (z - AA')^2}} \tag{5}$$

$$\beta_2 = \frac{\pi}{2} - \beta_1 \tag{6}$$

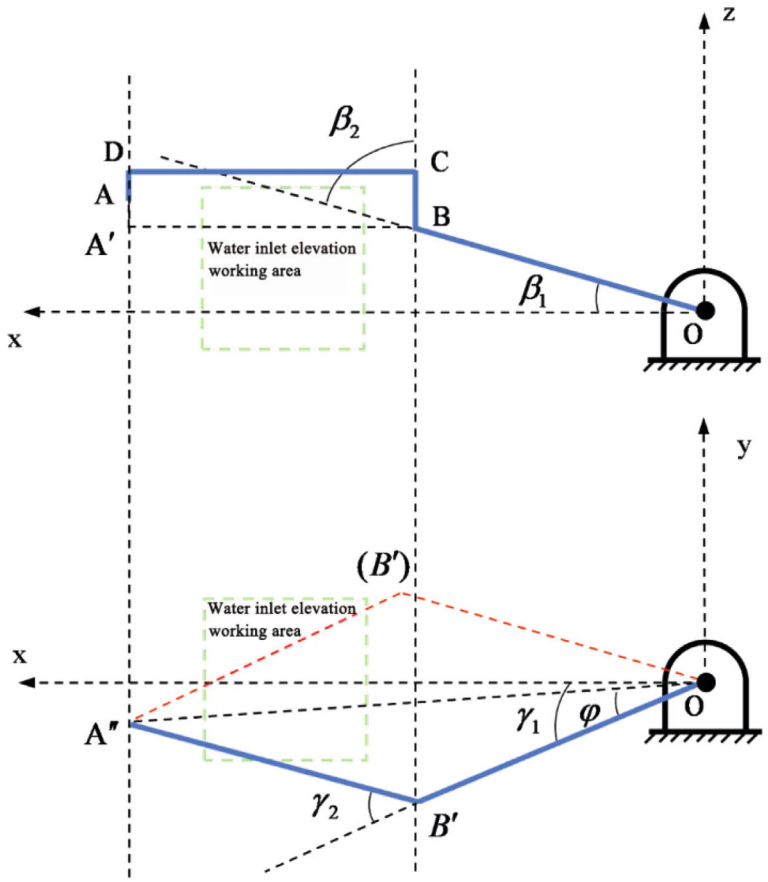


Fig. 2 Schematic diagram of robotic arm

where $L_B = 3505$ mm is the length of upper arm.

The schematic diagram of the installation position of the hydraulic cylinder is shown in Fig. 3, and the two hydraulic cylinders have the same model. ED and FG are both composed of the extension of the push rod and the combined size of the cylinder body and hinge. The cylinder body size is 500 mm and the hinge size is 40 mm.

The partial two-dimensional diagram of the water injection robotic arm is shown in Fig. 4.

Using the cosine theorem to solve for the angle corresponding to the stroke of the hydraulic cylinder.

$$\angle BOF = arccos\left(\frac{L_B^2 + OF^2 - BF^2}{2 \cdot L_B \cdot OF}\right) \tag{7}$$

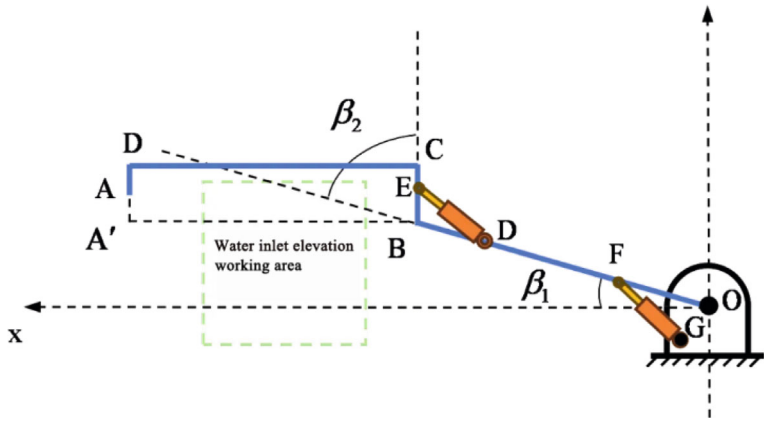


Fig. 3 Installation position diagram of hydraulic cylinder

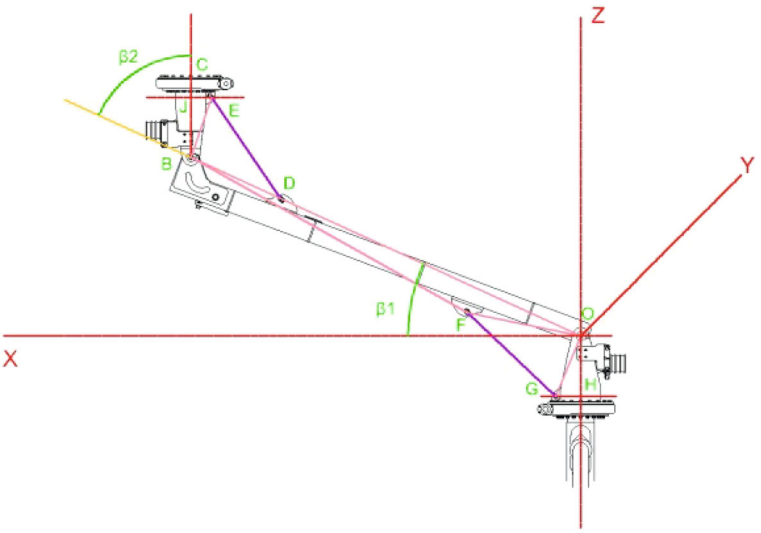


Fig. 4 Partial 2D image of water adding robotic arm

$$\angle GOH = arccos\left(\frac{GO^2 + OH^2 - GH^2}{2 \cdot GO \cdot OH}\right) \tag{8}$$

$$\angle FOG = \beta_1 + \frac{\pi}{2} - \angle BOF - \angle GOH \tag{9}$$

$$\angle JBE = arccos\left(\frac{JB^2 + BE^2 - JE^2}{2 \cdot JB \cdot BE}\right) \tag{10}$$

$$\angle DBO = arccos\left(\frac{DB^2 + L_B^2 - DO^2}{2 \cdot DB \cdot L_B}\right) \tag{11}$$

$$\angle EBD = \pi - \beta_2 - \angle JBE - \angle DBO \tag{12}$$

By combining the above formulas, the kinematic equation of the robotic arm can be established

$$\begin{cases} \gamma_2 = arccos\left(\dfrac{x^2 + y^2 - L_B^2 + (z - AA')^2 - (A''B')^2}{2\sqrt{L_B^2 - (z - AA')^2} \cdot A''B'}\right) \\ \gamma_1 = \begin{cases} arctan\dfrac{y}{x} + \gamma_2, -180° < \gamma_2 < 0°, 0° \le \varphi < 90° \\ arctan\dfrac{y}{x} - \gamma_2, 0 < \gamma_2 < 180°, 0° \le \varphi < 90° \end{cases} \\ \beta_1 = arctan\dfrac{z - AA'}{\sqrt{cos^2\gamma_1 \cdot \left[L_B^2 - (z - AA')^2\right] - (z - AA')^2}} \\ \Delta_B = \sqrt{FO^2 + OG^2 - 2FO \cdot OG \cdot cos(\beta_1 + \dfrac{\pi}{2} - \angle BOF - \angle GOH)} \\ \qquad - 500\,\text{mm} - 40\,\text{mm} \\ \Delta_{out} = \sqrt{EB^2 + BD^2 - 2EB \cdot BD \cdot cos(\beta_1 + \dfrac{\pi}{2} - \angle JBE - \angle DBO)} \\ \qquad - 500\,\text{mm} - 40\,\text{mm} \end{cases} \tag{13}$$

In the formula, γ_1, γ_2 are the rotation angles of the boom and the water outlet pipe, Δ_B and Δ_{out} are the strokes of the boom hydraulic cylinder and the water outlet pipe hydraulic cylinder, respectively. The size of γ_1 is related to the value of γ_2, and the positive or negative value of the value γ_2 will affect the final posture of the water adding robotic arm. When $\gamma_2 < 0$, the posture of the water adding robotic arm is shown by the blue line in Fig. 2; When $\gamma_2 > 0$, the posture of the water adding robotic arm is shown by the red dashed line in Fig. 2. The numerical values of each constant in the above formula are shown in Table 1.

Table 1 Constant value

Name	Value	Name	Value	Name	Value
OG	532	FO	941 mm	EB	515 mm
AA'	643 mm	$\angle BOF$	13°	BD	807 mm
$A''B'$	4100 mm	$\angle GOH$	22°	$\angle JBE$	19°

3 Algorithm Design for Water Injection Robot Arm

When the water injection robotic arm is working, the coordinates of the water inlet of the sprinkler truck are first obtained through the positioning antenna installed on the sprinkler truck (x, y), and the Z-axis coordinate $z = H_2 - 596.4 - H - h$ is obtained through parameter calculation. Among them, H_2 represents the height value of the vehicle, and different vehicle models correspond to different parameters. H represents the distance from the installation plane of the robotic arm to the coordinate origin, and h represents the distance from the installation plane of the water injection robotic arm to the parking plane of the sprinkler truck; After obtaining the three-axis coordinates of the water inlet of the sprinkler truck, the algorithm determines whether the water inlet meets the workspace requirements. If it does not meet the requirements, the position of the sprinkler truck water inlet needs to be adjusted to ensure that it is located in the ideal workspace. Subsequently, the algorithm calculates the pitch and rotation angles γ_1, Δ_B, γ_2, and Δ_{out} of the upper and lower arms. Then, send the four values to the hydraulic cylinder, output the corresponding expansion and contraction amount as well as the hydraulic motor, output the corresponding rotation angle, and reach the water injection condition. The robotic arm begins to inject water into the sprinkler truck. Finally, after the water injection is completed, the robotic arm returns to its initial position. The algorithm flowchart is shown in Fig. 5.

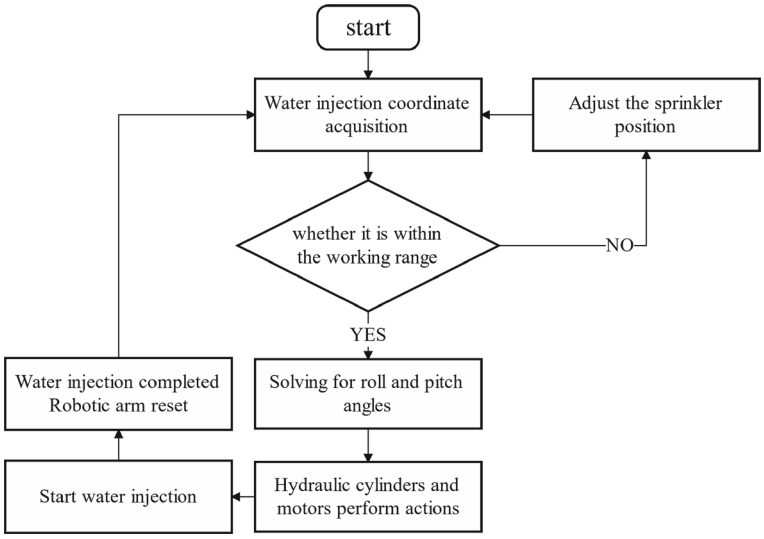


Fig. 5 Algorithm flowchart of water injection robotic arm

4 Trajectory Tracking Control of Hydraulic System

In construction machinery, both electro-hydraulic proportional valves and electro-hydraulic servo valves can accurately control the displacement of hydraulic cylinders and hydraulic motors. Among them, electro-hydraulic servo valves have high precision and fast response, and are used in high-precision machine tools, advanced flight control systems, and other situations where control precision and response speed are extremely high. Although electro-hydraulic proportional valves have slower response than electro-hydraulic servo valves, they have low cost and strong anti-fouling ability. After adding sensors, closed-loop control of hydraulic actuators can be achieved, this can meet the control requirements of general water injection robotic arms. Therefore, in order to achieve automatic control of the water injection robotic arm, an electro-hydraulic proportional valve is selected to replace the manual control valve of the hydraulic cylinder of each component of the control arm.

According to the kinematic model of the water injection robotic arm, the motion trajectory of the water outlet is determined by the displacement of the hydraulic cylinder and the rotation angle of the hydraulic motor. Among them, the displacement control system of the hydraulic cylinder and the control system of the rotation angle of the hydraulic motor both use electro-hydraulic proportional valves, and the system structures are shown in Fig. 6. The stroke of the hydraulic cylinder and the rotation angle of the hydraulic motor can be fed back through displacement sensors, forming a position closed loop with the reference displacement signal calculated by inverse solution, and then driving the robotic arm to drive the water outlet to the desired position.

Considering the electromagnetic proportional valve as a second-order component [10], the transfer function can be expressed as:

$$G_1(s) = \frac{K_q}{\frac{s^2}{\omega_v^2} + \frac{2\delta_v s}{\omega_v} + 1} \tag{14}$$

In the formula, K_q represents the flow gain of the electromagnetic proportional valve; ω_v represents the phase bandwidth of the electromagnetic proportional valve; δ_v represents the damping ratio of the proportional valve, with a range of values from 0.5 to 0.7.

When the load can be ignored, a hydraulic cylinder with very small elastic load can be regarded as a combination of integral and second-order components. After

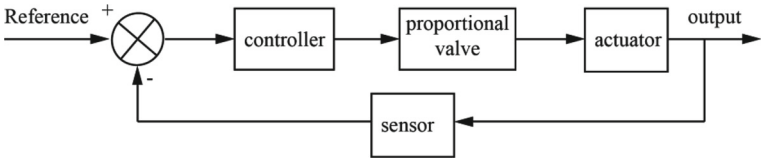


Fig. 6 Closed-loop control logic diagram of water injection robotic arm

deduction and calculation, the transfer function of the hydraulic power mechanism can be expressed as:

$$G_2(s) = \frac{1/A_h}{s\left(\frac{s^2}{\omega_h^2} + \frac{2\delta_h s}{\omega_h} + 1\right)} \tag{15}$$

In the formula, A_h represents the effective working area of the hydraulic cylinder. For asymmetric hydraulic cylinders, A_1 or A_2 can be taken according to the direction of motion; δ_h represents the damping ratio of the hydraulic cylinder load system, with a range of values from 0.1 to 0.2; ω_h represents the natural frequency of the hydraulic cylinder load link.

The relationship between the rotation angle of the valve controlled hydraulic motor and the displacement of the solenoid proportional valve core is:

$$G_3(s) = \frac{\theta(s)}{X_v(s)} = \frac{K_v/D_m}{s\left(\frac{s^2}{\omega_m^2} + \frac{2\delta_m s}{\omega_m} + 1\right)} \tag{16}$$

In the formula, ω_m is the undammed natural frequency of the hydraulic system; K_v is the flow gain of the solenoid valve; δ_m is the damping ratio of the hydraulic system; D_m is the displacement of the hydraulic motor; $\theta(s)$ is the angle of rotation of the hydraulic motor; $X_v(s)$ is the displacement of the valve core.

Using the above transfer function, a simulation environment was built in Matlab based on the actual selection parameters, and a PID controller was used to calibrate and compensate the system. The simulation results are shown in Fig. 7.

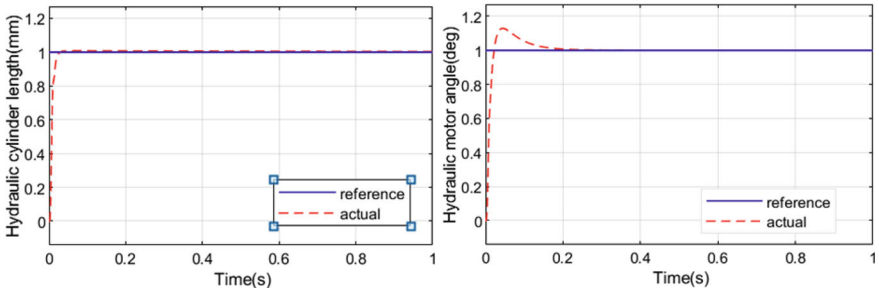


Fig. 7 Hydraulic system simulation curve

5 Trajectory Tracking Control of Hydraulic System

To achieve the fixed-point water injection operation of the water injection robotic arm, an automatic water injection system based on CAN bus is designed, as shown in Fig. 8, mainly composed of an electro-hydraulic position proportional control system and a Beidou guidance control system.

According to Fig. 8, the electro-hydraulic position proportional control system uses the CAN bus as the main network, with each node connected to the rotation angle sensor of the water outlet pipe and the boom, as well as the displacement sensor of the hydraulic cylinder, providing feedback values for achieving closed-loop control; Install a Beidou antenna on both sides of the bottom support, and use RTK positioning technology to achieve precise positioning of the water outlet for the sprinkler truck's water inlet.

The fixed-point water injection operation process of the water injection robotic arm is shown in Fig. 10: Firstly, the water injection nozzle of the sprinkler truck and the water outlet of the robotic arm are positioned. This step is operated by the Beidou guidance control system, which records the target coordinate position of the water injection nozzle and the initial coordinate position of the water outlet; Secondly, the control system accurately moves the water outlet of the robotic arm to the target position where water needs to be injected, and feedback the position information through displacement sensors to form a closed-loop control; Then turn on the water injection system, and according to the set parameters such as flow rate and pressure, let the water be injected into the water inlet of the sprinkler truck through the mechanical arm, and monitor the water injection situation in real time, including changes in water level and pressure, to ensure stable water injection process; Finally, turn off the water injection system, move the robotic arm back to its initial position, and wait for the next operation.

The water injection process was tested to evaluate the closed-loop control effectiveness of two hydraulic cylinders and two hydraulic motors, as well as the tracking performance of the water outlet. As depicted in Fig. 9, the water outlet positioning capability remains stable and accurate, meeting the design expectations, thus confirming the feasibility of the technical scheme.

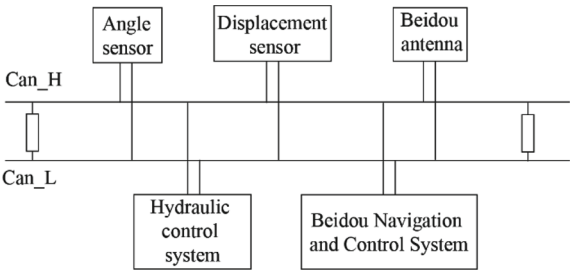
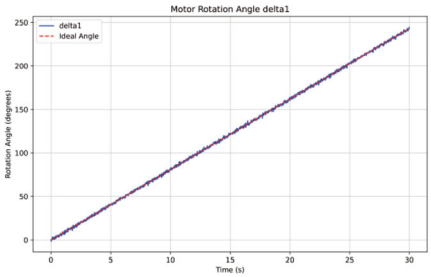
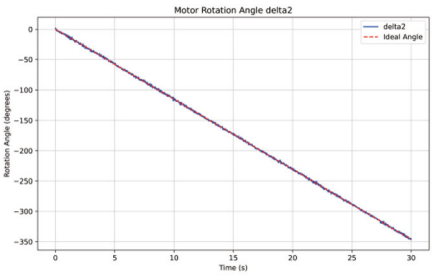


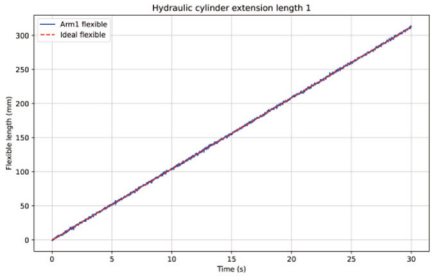
Fig. 8 Distributed arm automatic water injection system based on CAN bus



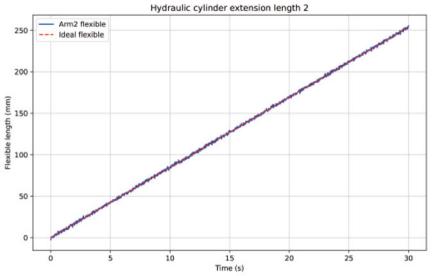
(a)Motor Rotation Angle delta 1



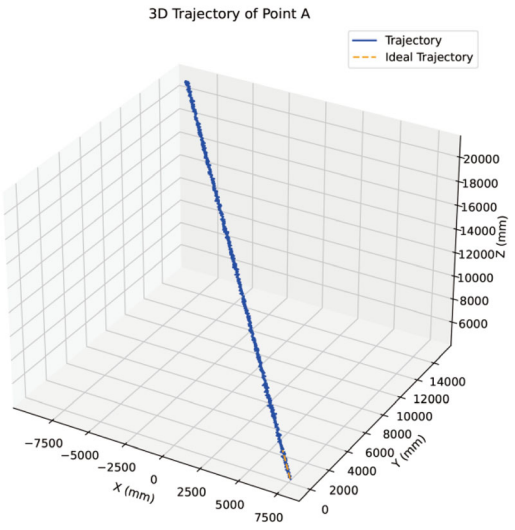
(b) Motor Rotation Angle delta 2



(c) Hydraulic cylinder extension length 1



(d) Hydraulic cylinder extension length 2



(e) 3D Trajectory of Point A

Fig. 9 Position tracking curve



(a) Initial position (b) Inject water after precise positioning of the control system

Fig. 10 Actual homework process

6 Conclusion

Based on the four degree of freedom robotic arm of the mine water injection station, the inverse kinematics model is derived using geometric methods, and the operating range is determined based on the mechanical structure. The extension length and rotation angle of the hydraulic cylinder and motor are derived based on the position of the sprinkler truck.

Based on the inverse calculation results, plan the automatic process and provide a method for determining whether the water inlet meets the water injection position. Simulate hydraulic systems in Maltab and the tracking performance of the hydraulic cylinder, hydraulic motor, and water outlet during the water injection process has been examined, and the efficacy of the technical approach has been confirmed through the empirical results.

The designed system is programmed in C language and applied in practical systems. Experimental results show that the designed automatic water injection robotic arm meets the requirements of actual working conditions and can achieve automatic water addition function.

References

1. Li G, Zhou Z, Hu J (2023) Study progress and prospect analysis of mine dust control technology based on knowledge graph in China. Metal Mine 7:28–39
2. Akshatha M, Kavyashree MK (2020) Survey paper on various methods of automating the water system using dust sensor to suppress the dust in mining. Int J Eng Res Technol 9(5):1163–1168
3. Manjunath A, Kanthraju KM (2022) An efficient automatic water sprinkler design for operate in dust environment. Trends Sci 19(14):4982–4982
4. Dong H, Wu W (2020) Mechanism design and dynamics simulation of a new four-axis palletizing robot. Packag Eng 41(3):176–181
5. Min JK, Kim DW, Song JB (2020) A wall-mounted robot arm equipped with a 4-DOF yaw-pitch-yaw-pitch counterbalance mechanism. IEEE Rob Autom Lett 5(3):3768–3774
6. Lu E, Yang XF, Li W (2016) Tip position control of single flexible manipulators based on LQR with the Mamdani model. J Vibro Eng 18(6):3695–3708

7. Sun S, Qi X, Su L et al (2011) Trajectory planning of 5-DOF manipulator based on polynomial method. Mechatronics 17(11):33–37
8. Parika PA, Trivedi RR, Joshi KD (2021) Trajectory planning of a 5 DOF feeding serial manipulator using 6 order polynomial method. J Phys Conf Ser 1921(1):19–20
9. Guo J, Liang Y, Zhang X et al (2023) Kinematic analysis and calculation of hexapod robot based on gait planning. Mach Tool Hydraul 51(5):66–73
10. Wang Z, Jiao X, Cheng M (2018) Trajectory tracking control with disturbance attenuation for hull derusting manipulators. China Mech Eng 29(6):731–737

Open Access This chapter is licensed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License (http://creativecommons.org/licenses/by-nc-nd/4.0/), which permits any noncommercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if you modified the licensed material. You do not have permission under this license to share adapted material derived from this chapter or parts of it.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.



Temperature and Thermal Deformation Analysis of Key Components of Wheel Hub Electric Wheel Under Combined Braking Conditions



Kaikun Pei, Lijun Zhang, Dejian Meng, Wei Tian, and Zhuang Zhang

Abstract Under the background of the current global energy transition and environmental protection, the development of new energy vehicles has become a global hot spot. As one of the key technologies of new energy vehicles, the electric wheel drive system has gradually become the key to the development of electric vehicles because of its high integration and high-power density. In this paper, the temperature rise and thermal deformation of the key components of the electric wheel under the compound braking condition are studied. Firstly, the geometric modeling and parameter setting of the electric wheel are carried out, and then the temperature field and thermal deformation under the joint action of friction brake and hub motor are simulated based on ABAQUS finite element analysis software. The results show that the temperature rise of the brake disc decreases under composite braking, and the temperature rise of the motor is mainly caused by the braking loss of the motor itself, in which the temperature rise of the permanent magnet is the highest. In terms of deformation, the deformation of the brake disc is the most significant, while the deformation of the motor part is relatively small and can be ignored.

Keywords Combined braking conditions · ABAQUS · Temperature field · Deformation field · Finite element analysis

1 Introduction

With the proposal of the development direction of automobile electrification, the automobile industry has gradually transformed from fuel vehicles to electric vehicles [1]. Hub motor, the core component of electric vehicle drive system, has become the focus of current research because of its compact structure and high energy conversion efficiency [2–4]. However, due to the particularity of the working environment of the wheel motor, there are installation space limitations and high power density and other

K. Pei · L. Zhang · D. Meng · W. Tian (✉) · Z. Zhang
School of Automotive Studies, Tongji University, Shanghai, China
e-mail: 971013314@qq.com

© The Author(s) 2026

H. Saman et al. (eds.), *The 9th International Conference on Advances in Construction Machinery and Vehicle Engineering*, Lecture Notes in Mechanical Engineering,
https://doi.org/10.1007/978-981-96-5527-4_30

characteristics lead to its large heat load, how to effectively control the temperature of the motor has become one of the key challenges in the development of technology [5].

The hub electric wheel will brake system, transmission system and drive system are highly integrated, which causes the key problem is the temperature rise problem. Compound braking, the vehicle through mechanical braking and electric braking at the same time, to achieve parking or deceleration, in this process, not only the heat generated by the mechanical brake, but also the heat generated by the electric brake. At present, it has attracted the attention of some scholars [6]. Li et al. studied the thermal behavior of axial flux hub motor by finite element analysis model, and verified the accuracy of the model combined with experimental data. In addition, they also conducted thermal stress analysis on the structure of the motor to ensure the safety and reliability of the design [7]. Shao analyzed the vibration and noise problems of the wheel motor, and verified the validity of his analytical model by comparing the temperature field simulation with the experimental data. In addition, he also studied the possibility of reducing vibration and noise through optimal design [8]. Zhang used multi-physics simulation software to optimize the structure of the wheel motor to improve its performance under high load conditions. In addition, he also analyzed the influence of different heat dissipation schemes on the motor temperature rise and mechanical properties [9]. Huang analyzed the temperature field of the wheel motor through the finite element model, and analyzed the heat dissipation effect of the motor in combination with the thermal network method, providing a theoretical basis for optimizing the motor design [10]. Zhu adopted a multi-physical field coupling model to study the interaction of electromagnetic field, temperature field and mechanical field of wheel motor, providing a systematic perspective for understanding the motor's behavior in actual operation. In addition, they also studied the possibility of reducing losses and temperature rise through structural optimization [11]. Shi designed a lightweight low-speed hub motor, and verified the rationality of its design through experiments, and analyzed its operating efficiency and economy through performance tests. In addition, some scholars proposed relevant methods for thermal management of electric wheels [12–16]. Xu used CFD and finite element analysis (FEA) to conduct initial thermal tests on the entire duty cycle of aircraft taxiing operation, and predicted the results used to verify the motor design [17]. Wu proposed an energy management strategy, which effectively reduced the consumption of electric energy [18]. Shi established the simulation model of permanent magnet electric wheel and analyzed the temperature distribution and heat source [19].

The existing research shows that as the core component of electric vehicle, the performance optimization of hub motor is very important to improve the overall performance and safety of electric vehicle, and many scholars have carried out research on this, but the interaction between mechanical heat and motor heat under compound braking conditions is rarely studied. Therefore, in this paper, the temperature rise and thermal deformation of the key components of the electric wheel are studied under the compound braking condition. Firstly, the geometric modeling and parameter setting of the electric wheel are carried out. Then, based on ABAQUS finite

element analysis software, the temperature field and thermal deformation under the joint action of the friction brake and the wheel motor are simulated and analyzed.

2 Loss Analysis

The loss of the motor is not only related to the magnetic field inside the motor, but also has a close relationship with the material properties of the various parts of the motor, the assembly relationship and the heat dissipation conditions. According to previous studies, the motor loss is generally divided into iron core loss, winding loss, permanent magnet loss, etc. The calculation method of each part is as follows.

2.1 Iron Core Loss

At present, the iron loss separation theory proposed by Bertotti et al. is mainly used for the calculation of iron core loss. The iron loss separation theory is mainly classified and calculated according to the different causes of the loss, and finally the calculation results of each part can be added together to calculate the total core loss. The specific formula is shown as follows.

$$P_{Fe} = P_e + P_{ex} + P_h \tag{1}$$

where P_{Fe} is the total core loss (W/kg), P_e eddy current loss (W/kg), P_{ex} additional loss (W/kg), and P_h hysteresis loss (W/kg).

The specific calculation formula of the loss of these three parts is as follows:

$$P_e = \frac{\sigma\delta^2}{12\rho}\frac{1}{T}\int_0^T \left(\frac{dB(t)}{d(t)}\right)^2 dt \tag{2}$$

$$P_{ex} = \frac{\sqrt{\sigma GVS}}{\rho}\frac{1}{T}\int_0^T \left|\frac{dB(t)}{d(t)}\right|^{1.5} dt \tag{3}$$

$$P_h = k_h f B_m^{\alpha} \tag{4}$$

where f is the alternating current frequency (Hz), B_m is the amplitude of the magnetic field density (T), B is the magnetic field density (T), σ is the conductivity of the stator core silicon steel sheet (S/m), δ is the stator core silicon steel sheet thickness (m), ρ is the stator core silicon steel sheet density (kg/m^3), S is the stator core silicon steel sheet cross-sectional area (m^2), T is the electromagnetic field change period (s)

inside the motor, V and G are the parameters related to the magnetic properties of the stator iron core silicon steel sheet.

It can be seen from the above formula that the loss value changes with the change of the magnetic field density of the electromagnetic field. The variation law of the magnetic field density is generally sinusoidal waveform or other forms of waveform. The following are specific expressions of sinusoidal and general waveforms.

When the magnetic field density of the electromagnetic field changes in a sinusoidal waveform:

$$P_{Fe} = k_e f^2 B_m^2 + k_{ex} f^{1.5} B_m^{1.5} + k_h f B_m^{\alpha} \tag{5}$$

where k_e is the stator core eddy current loss coefficient (W/kg), k_{ex} is the stator core additional loss coefficient (W/kg), k_h is the stator core hysteresis loss coefficient (W/kg).

When the electromagnetic field magnetic density changes in accordance with the general waveform, the magnetic field density is decomposed into two components to represent:

$$P_{Fe} = k_e f^2 \int_0^T \left[\left(\frac{dB_n(\theta)}{d(\theta)} \right)^2 + \left(\frac{dB_t(\theta)}{d(\theta)} \right)^2 \right] d\theta + k_{ex} f^{1.5} \int_0^T \left[\left(\frac{dB_n(\theta)}{d(\theta)} \right)^{1.5} \right] + d\theta + k_h f B_m^{\alpha} \tag{6}$$

where θ is the motor rotation Angle (rad).

2.2 Copper Loss in Winding

Since the motor winding is a good conductor of electricity, when the current passes through the stator copper winding, it can be known by Joule's law that heat will be generated, and the heat generated at this time will be the loss of copper loss, and the winding copper loss is mainly divided into two parts, basic copper consumption and additional copper consumption.

According to Joule's law, the basic copper loss is equal to the resistance of the winding multiplied by the second power of the winding's current. For the permanent magnet synchronous motor in this paper, the current of the three-phase winding is the same, so the calculation formula of the basic winding copper consumption is:

$$P_{cu} = 3I^2 \mathrm{R} \tag{7}$$

where P_{cu} is the motor winding copper consumption (W), I is the phase current of the permanent magnet synchronous motor (A), R is the winding resistance of the permanent magnet synchronous motor (Ω).

2.3 Permanent Magnet Loss

When the permanent magnet synchronous motor is running, because the stator current supply mode is PWM, it will produce the time harmonics and space harmonics of the current, the open slot of the stator core will make the permeability of the air gap change, and the magnetomotive force of the copper wire winding is not a complete sinusoidal curve, which will cause eddy current loss in the permanent magnet. In the total loss, the loss of space harmonics is relatively small, and the loss of time harmonics is relatively large. The specific calculation formula is as follows:

$$P_{evm} = \int \frac{|J|^2}{2\sigma} dv \tag{8}$$

where vm is the volume of the permanent magnet equivalent to A ring (m^3), σ is the conductivity of the permanent magnet (S/m), and J is the current density in the permanent magnet (A/m^2).

3 Parameters

3.1 Brake Force Distribution

Compound braking, which can greatly reduce the consumption of electric energy and is in line with a strategic task of sustainable development in our country, is one of the most important technologies in the field of electric vehicles. In general, single motor braking cannot meet the requirements of all braking conditions, and it is necessary to cooperate with friction braking to meet the requirements of braking strength and the reasonable match of energy recovery to the greatest extent.

First, according to the braking regulations required by ECE-R13, the standard of front and rear axle braking force distribution under different braking intensities is calculated, and then based on the electric braking performance curve, the maximum braking intensity of the electric brake is calculated. Based on this, the situation of braking force distribution under emergency braking is comprehensively determined. The basic dimensions of the entire vehicle are shown in Table 1.

The research object of this paper is the electric vehicle driven by the front axle. On the premise of meeting the braking strength requirements, the braking should occur in the front axle as far as possible, that is, the braking force distribution coefficient $= 1$. β for M1 type vehicles, according to ECE-R13 braking regulations, when $0.1 < Z < 0.61$, the braking force distribution coefficient of the front and rear axle should meet the following requirements.

$$\beta \geq \frac{b + zh_g}{L} \tag{9}$$

Table 1 Basic dimensions of the vehicle

Basic dimensions	Parameter values
Wheelbase L (mm)	2700
Height of center of mass (mm)h_g	610 (no load); 560 (full load)
Distance from center of mass to front axle a (mm)	1188 (no load); 1323 (full load)
Distance from center of mass to rear axis b (mm)	1512 (no load); 1377 (full load)
Brake force distribution coefficient β	0.5
Unloaded axle load distribution (front/rear) (kg)	816/739
Full load distribution (front/rear) (kg)	1016/989

$$\beta \geq 1 - \frac{(z + 0.07)(a - zh_g)}{0.85zL} \tag{10}$$

$$\beta \geq \frac{(\mathrm{z} + 0.07)(\mathrm{b} + \mathrm{z}h_g)}{0.85zL} \tag{11}$$

where β is the braking force distribution coefficient, A is the distance between the center of mass and the front axle (mm), B is the distance between the center of mass and the rear axle (mm), Z is the braking strength, h_g is the height of the center of mass (mm), L is the distance between the front and rear axle (mm).

When the braking force distribution coefficient = 1, the height of the center of mass and the distance between the center of mass and the front and rear axles a and b are substituted into the above formula, the braking force can be completely provided by the front axle when $\mathrm{Z} \leq 0.11$ or $\mathrm{Z} \geq 1.57$, and the other ranges are braking jointly by the front and rear axles. βh_g for the sake of safety, $\mathrm{Z} = 0.11$ is taken as the threshold value of electromechanical decoupling in the case of high intensity braking. In addition to considering the decoupling threshold under the ECE-R13 braking regulations, it is also necessary to consider the maximum braking force that can be generated by the motor-wheel permanent magnet synchronous motor. The braking characteristic curve of the motor is shown in Fig. 1.

It can be obtained from the figure that the maximum braking torque that the motor can produce during braking is 260Nm. When the braking force is completely

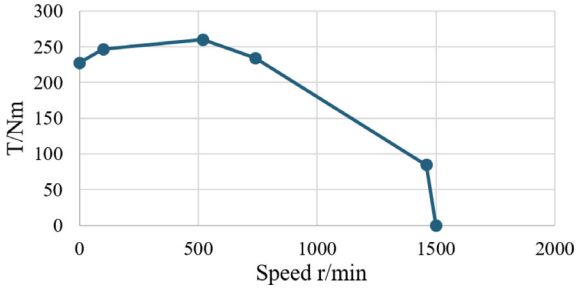


Fig. 1 Braking characteristic curve of the motor

Table 2 Compound braking force distribution

Braking Strength	Motor brake	Mechanical braking
Z < 0.1	Yes	No
0.1 < Z < 0.7	Yes	Yes
Z < 0.7	No	Yes

provided by the motor, according to the formula:

$$Z_{max} = (T_{max} \cdot i_0 \cdot \eta)/\mathrm{mg} \cdot \mathrm{r} \tag{12}$$

where Z_{max} is the maximum braking strength, T_{max} is the maximum braking moment (Nm), i_0 is the main deceleration ratio, where 1 is the direct drive, the value is 1, η is the transmission efficiency, here is 0.99, M is the total mass (kg), r is the braking force radius (mm).

The parameters of the electric vehicle and the maximum braking force of the motor can be obtained in the no-load and full load under the two conditions of the braking strength of 0.17, 0.14. Considering the requirements of braking laws and regulations and the limit of the maximum braking force of the motor, Z = 0.1 is selected as the boundary between single motor braking and compound braking. That is, when the braking strength is less than 0.1, the single motor is used for braking; when the braking strength is greater than 0.1, compound braking or single mechanical braking. When z > 0.7, for safety and battery and ultracapacitor life considerations, do not take the motor for braking. Therefore, the braking force distribution of composite braking is shown as Table 2.

In this study, the working condition is an emergency braking condition, with a deceleration rate of 6.12 m/s^2. Therefore, the braking intensity is 0.63. Under this braking intensity, the braking force is provided by both the front and rear axles, and it is a composite braking condition. The braking force of the electric brake varies with Fig. 1.

3.2 Convective Heat Transfer Coefficient

The main materials of electric wheel parts in this paper include: stator core material silicon steel sheet DW315, permanent magnet pole material NdFeB (NdFEB), brake disc solid gray cast iron, hub motor shell material aluminum alloy, hub motor outer rotor material silicon steel, hub motor coil winding material Cu (copper). The convective heat dissipation coefficient of these surfaces can be calculated by the related theory of heat transfer.

3.3 Heat Flux

Heat Flux of Mechanical Braking Under composite braking, the change of total braking torque over time is consistent with that of single friction braking. Combined with the braking characteristic curve of the motor, the change law of friction braking torque over time can be obtained under composite braking, as shown in Fig. 2.

The calculation method of the total heat flux of the electric wheel in the braking process is as follows.

$$I \bullet \frac{a}{R_c} = M = 2 \cdot P \cdot S_P \cdot \mu \cdot R_d \tag{13}$$

where I is the moment of inertia of 1/4 electric vehicle (kg m^2), A is the car deceleration (m/s^2), R_c is the tire equivalent radius (mm), M is 1/4 electric vehicle braking torque (Nm), P is the brake pressure Mpa, S_P is the piston area (mm^2), μ is the brake disc friction pair friction coefficient, R_d is the effective radius of the brake disc (mm).

Thus, the time history change of frictional brake heat flux under combined braking can be obtained, and the calculation results are shown in Fig. 3.

Motor Brake Heat Flux In this paper, only the loss of the stator core, the loss of the rotor core, the winding loss, the loss of the permanent magnet, etc., do not consider the mechanical loss and other additional losses. Mechanical loss and other additional damage are selected according to 10% of the total loss, and the heat flux of each part is shown in Fig. 4.

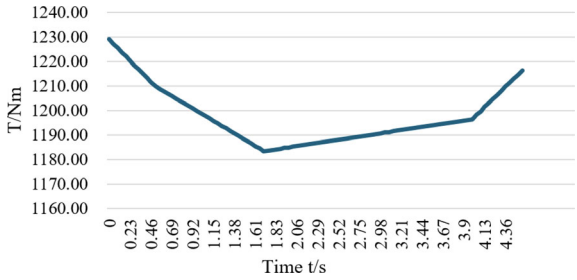


Fig. 2 Time history of friction braking torque under composite braking

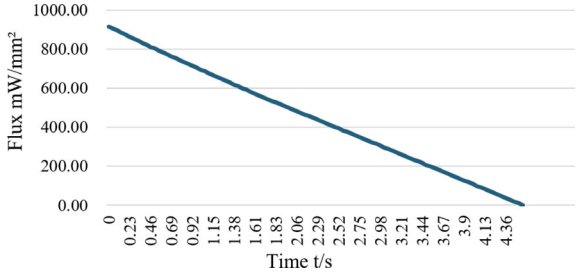


Fig. 3 Heat flux of composite brake-friction brake

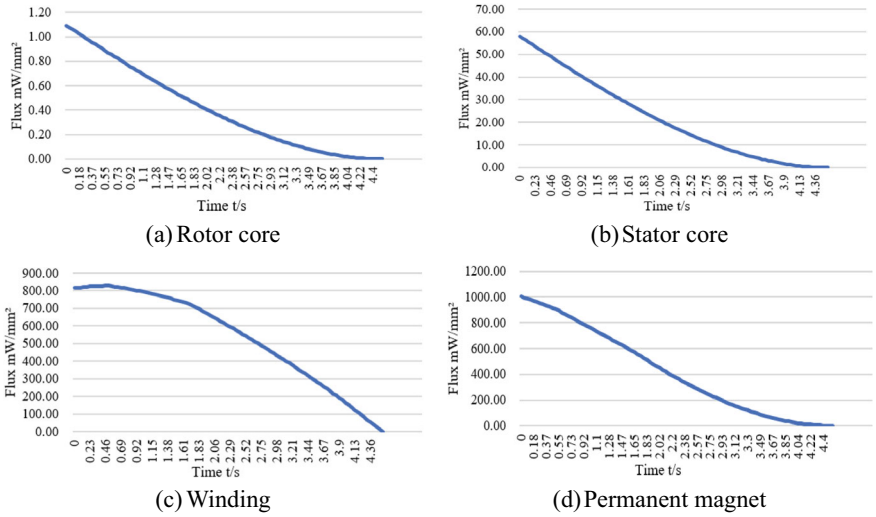


Fig. 4 Composite brake—electric brake heat flux

4 Model Results and Analysis

The following will analyze the simulation results of composite braking conditions from two aspects of temperature and thermal deformation.

4.1 Temperature Field Analysis

Figure 5 shows the temperature field distribution cloud diagram of the electric wheel model under the combined braking condition. The key parts of the electric wheel, such as brake disc, motor shell, motor rotor, motor permanent magnet, winding and stator core, are analyzed in combination with the cloud image.

Figure 6 is the temperature field distribution cloud diagram of key components of the electric wheel, from which it can be seen that the temperature of the disc surface of the brake disc is relatively high, and the temperature from the disc surface to the disc surface gradually decreases. The maximum temperature of the disc reaches 62.3 °C, which is relatively high, and the disc surface temperature increases first and then decreases, mainly due to the relatively large heat at the beginning of the braking stage. As the braking process of rotational speed continues, the heat dissipation speed of the disc surface is greater than the heating speed, resulting in a downward trend in the second half of the brake. The inner circumference of the shell near the end of the brake disc has the highest temperature rise, but the temperature rise is slightly lower than that of the same position when the single brake disc is stopped, which is nearly 0.4 degrees Celsius. It can be seen that this part is mainly affected by the brake disc. The

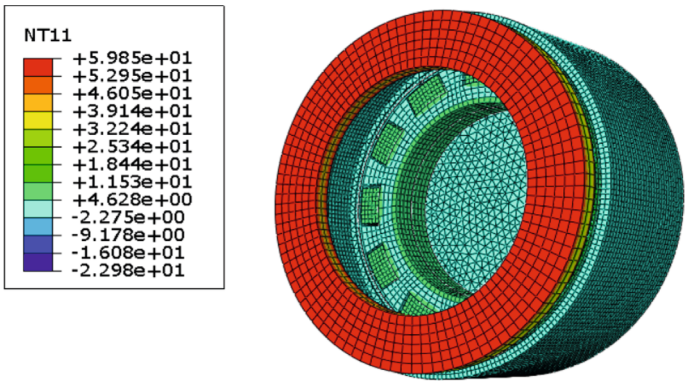


Fig. 5 The whole temperature field distribution cloud map of electric wheel

temperature of the part connected with the rotor and the brake disc is relatively high, and the temperature decreases rapidly along the axis, and because the braking time is relatively short, the temperature rise of the rotor is not very high. The temperature rise of the middle part of the permanent magnet is higher, and the temperature rise from the middle to both sides will gradually be lower. In the composite braking, the highest temperature of the middle part of the permanent magnet can reach 62.7 °C, and the temperature distribution is relatively concentrated. The temperature of the middle part of the winding is higher, the two sides are lower, and the highest temperature rise can reach 16.5 °C, but as the part with the largest power loss, the temperature rise is not very high, mainly because the thermal conductivity of the winding is 50 times that of the permanent magnet, and the volume of the winding is much larger than the permanent magnet, and there is sufficient place to accommodate the heat generated by the loss. The stator temperature rise is relatively high where the stator groove part and the water jacket part, the temperature rise of the stator core water jacket is higher, the temperature rise here is not because of the loss of the stator core and the temperature rise, but because the convection heat transfer intensity here is relatively high.

4.2 Thermal Deformation Analysis

Figure 7 shows the overall thermal deformation distribution of the composite brake electric wheel. It can be seen from the thermal deformation cloud map that the largest part of the deformation is the brake disc, and the rotor and the housing have different degrees of deformation. The deformation of each component will be analysed separately below.

Figure 8 shows the thermal deformation distribution of key components of the electric wheel, from which it can be seen that the brake disc produces different

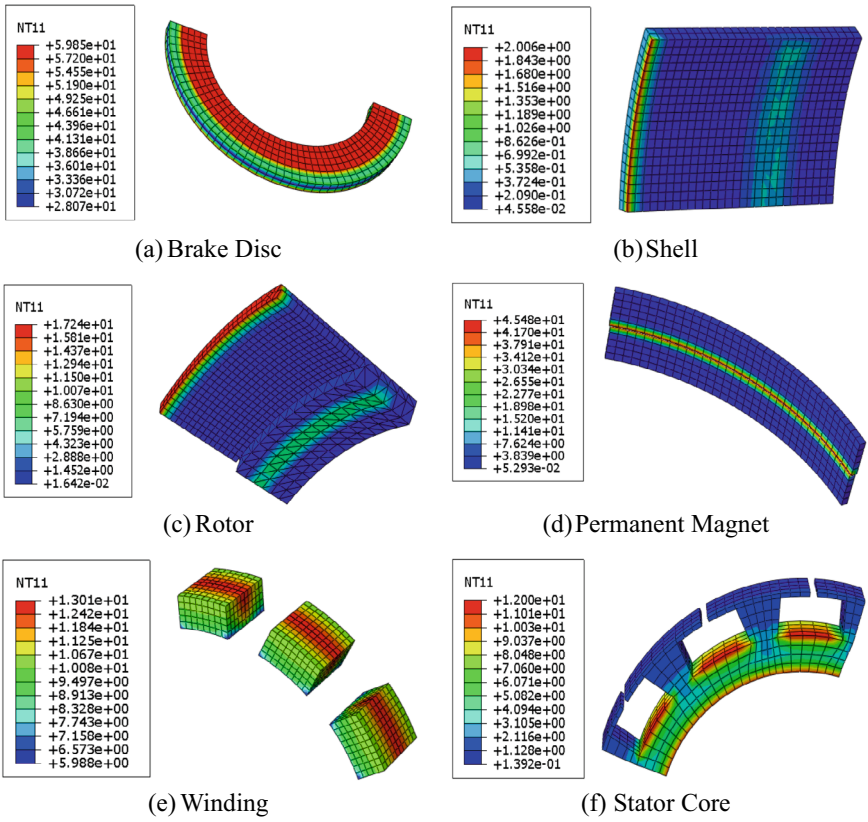


Fig. 6 Cloud picture of temperature field distribution of key components of electric wheel

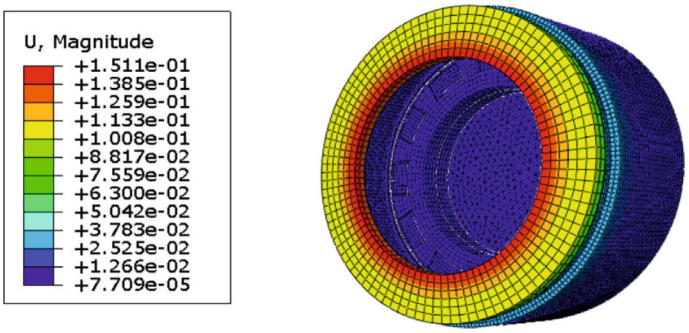
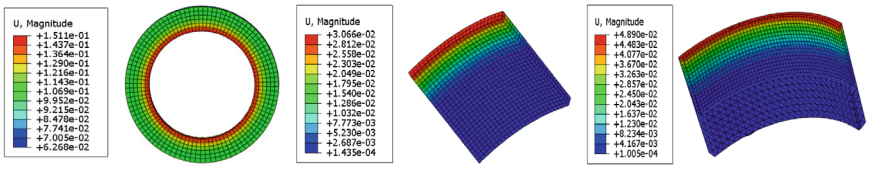


Fig. 7 Cloud image of thermal deformation of electric wheel



(a) Brake Disc (b) Shell (c) Rotor

Fig. 8 Cloud image of thermal deformation of key components of electric wheel

deformations at different radial radii after composite braking. The smaller the radius, the larger the shape variable, the maximum deformation on the outer surface of the brake disc is 0.15 mm, and the maximum deformation on the inner surface of the brake disc is 0.12 mm. The main reason is that part of the heat on the inner surface of the brake disc is transferred to the outer rotor, and the heat on the inner surface is less than the heat on the outer surface, so the maximum shape variable on the outer surface of the brake disc is greater than the maximum shape variable on the outer surface, and the maximum deformation occurs at the minimum radius of the brake disc. The deformation of the inner surface of the shell gradually decreases, infinitely tends to 0 mm, the maximum deformation occurs in the shell near the end face of the brake disc, the main reason is that the heat generated by the friction of the brake disc is transmitted to the end face of the shell, resulting in the shell end face deformation is larger, the maximum deformation is 0.03 mm. The deformation of the inner surface of the rotor gradually decreases, infinite tends to 0 mm, and the maximum deformation occurs at the end face of the rotor connected with the brake disc. The main reason is that the heat generated by the friction of the brake disc is transmitted to the end face of the rotor, resulting in a larger deformation of the end face of the rotor, and the maximum deformation is 0.048 mm. Permanent magnet, winding, stator iron core and brake disc, rotor and other components compared with the deformation is relatively small, can be ignored.

5 Conclusion

In this paper, firstly, the heat flux density of friction braking and motor braking loaded to brake disc, rotor, stator, winding, permanent magnet and other components is calculated respectively. Then, the simulation analysis of composite braking conditions is done, and the conclusion is as follows: in composite braking, the highest point of temperature rise is still the brake disc surface, but the temperature rise is reduced. The highest temperature rise of part of the motor permanent magnet is followed by the winding. The temperature rise of the rotor and stator core due to the motor braking is relatively low and can be ignored. In terms of deformation,

the maximum deformation occurs at the minimum radius of the brake disc, and the maximum deformation of the motor part is the rotor shell, and the stator core and other components of the permanent magnet winding are relatively small and can be ignored.

Acknowledgements Project is supported by the Natural Science Foundation of Chongqing [CSTB2023NSCQ-MSX0063]

References

1. Zhang G, Wang S, Wang J et al Electric passenger car wheel hub motor cooling technology overview. Mech Electr Eng 1–9. http://kns.cnki.net/kcms/detail/44.1522.TH.20240926.1500.008.html. Accessed 6 Nov 2024
2. Wenjie S, Junping F, Jingbo Z et al (2024) Research review on multi-physical coupling characteristics of wheel hub motors for electric vehicles. J Changzhou Inst Technol 37(04):25–31
3. (2024) J CSEE44(07):2871–2886. https://doi.org/10.13334/J.0258-8013.PCSE.222954
4. Guan T, Liu D, He Y (2024) Permanent magnet wheel motor technology development review. J Electrotech 33(02)6:378–396. https://doi.org/10.19595/j.carolcarrollnki.1000-6753 (The TCES. 221656)
5. Yawei L, Chengye L, Jingbo Z et al (2023) Review on vertical vibration coupling and active control of vehicle driven by wheel hub motor. J Changzhou Inst Technol 36(03):9–13
6. Guo J (2023) Research on matching design and control of distributed steering system of electric vehicle with hub motor. Shandong University of Science and Technology
7. Li T, Zhang Y, Liang Y et al (2021) Thermal analysis and research of axial flux hub motor. Chin J Eng Thermophys 42(06):1561–1568
8. Shao H (2021) The electric car wheel motor vibration and temperature analysis. East China Jiaotong University. https://doi.org/10.27147/,dcnki.Ghdju.2021.000545
9. Zhang L (2022) Switched reluctance rotor hub motor design and magnetic thermal coupling simulation. Hebei University of Science and Technology. https://doi.org/10.27107/,dcnki.Ghbku.2022.000302
10. Huang J (2019) The electric car wheel motor temperature field and thermal analysis. East China Jiaotong University. The https://doi.org/10.27147/,dcnki.Ghdju.2019.000050
11. Zhu J (2023) Research on multi-physical field analysis and loss suppression of integrated wheel motor with drive and brake. Yantai University
12. Ušakovs I, Mishkinis D, Galkin IA et al (2023) Experimental thermal characterization of the in-wheel electric motor with loop heat pipe thermal management system. Case Stud Therm Eng 47:103069
13. Raj JAPS, Asirvatham LG, Angeline AA et al (2024) Thermal management strategies and power ratings of electric vehicle motors. Renew Sustain Energy Rev 189:113874
14. Garud KS, Lee MY (2024) Thermal management characteristics of electric vehicle driving motor with oil spray cooling based on spray locations and oil types. Appl Therm Eng 248:123234
15. Chakraborty S, El Baghdadi M, Jaguemont J et al (2021) Parameterized cloud-connected electro-thermal modelling of a battery electric vehicle. In: 2021 IEEE vehicle power and propulsion conference (VPPC). IEEE, pp 1–8
16. Aloeyi E F, Ali N, Wang Q (2022) A review of in-wheel motors for electric vehicle propulsion. In: 2022 IEEE transportation electrification conference and expo, Asia-Pacific (ITEC Asia-Pacific). IEEE, pp 1–6

17. Xu Z, Galea M, Tighe C et al (2014) Mechanical and thermal management design of a motor for an aircraft wheel actuator. In: 2014 17th international conference on electrical machines and systems (ICEMS). IEEE, pp 3268–3273
18. Wu Y, Li J, Lei Z et al (2021) Optimal energy management for a dual-motor all-wheel drive electric vehicle considering battery temperature. In: 2021 IEEE vehicle power and propulsion conference (VPPC). IEEE, pp 1–6
19. Shi H, Deng Z, Zhang B et al (2022) Thermal-force coupling analysis of permanent magnet electrodynamic wheel system for maglev car. IEEE Trans Magn 59(1):1–9

Open Access This chapter is licensed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License (http://creativecommons.org/licenses/by-nc-nd/4.0/), which permits any noncommercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if you modified the licensed material. You do not have permission under this license to share adapted material derived from this chapter or parts of it.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.



Gap Detection Scheme of Safety Mechanism of Three Gorges Ship Lift Based on Sensing Detection



Weibi Tang and Qingyun Liang

Abstract According to the structural characteristics and operation mode of the safety mechanism of the Three Gorges ship lift, the influence factors of the clearance of the thread pair of the safety mechanism are analyzed. The safety mechanism is affected by the installation and manufacturing error of the equipment, the clearance of the mechanical transmission system, the structural deformation and deformation caused by the external load and the wear of the operating parts of the equipment. The maximum change value of the theoretical clearance can reach ± 49 mm, and the maximum value in actual operation does not exceed ± 15 mm. Based on the sensing detection technology, the clearance detection scheme of the safety mechanism of the Three Gorges ship lift is proposed. The scheme can realize the real-time detection and early warning of the clearance of the safety mechanism, and avoid the excessive load of the clearance detection device, the accuracy of device requirements were meets through the gap test, and the scheme has guiding significance for engineering practice.

Keywords Three Gorges ship lift · Security mechanism · Thread pair · Gap detection · Sensing detection

1 Introduction

The thread pair clearance of the safety mechanism of the Three Gorges ship lift is an important parameter affecting the safe operation of the ship lift. The size of the clearance needs to be moderate, which is too large or too small will affect the normal operation of the ship lift. When the ship chamber has travel deviation during the up/down operation, the gap of the screw pair of the safety mechanism will become smaller until the gap disappears, and the safety mechanism will be jammed. The transmission mechanism will rotate under the action of external load, and the torque

W. Tang (✉) · Q. Liang
Three Gorges Navigation Authority, Yichang, Hubei, China
e-mail: 274941548@qq.com

© The Author(s) 2026

H. Saman et al. (eds.), *The 9th International Conference on Advances in Construction Machinery and Vehicle Engineering*, Lecture Notes in Mechanical Engineering,
https://doi.org/10.1007/978-981-96-5527-4_31

transmitted by the safety clutch will rise sharply, causing major equipment failure. Therefore, the real-time detection of the clearance of the thread pair of the safety mechanism is the premise to ensure the safe and stable operation of the ship lift.

In the field of mechanical testing, based on the sensing detection technology, domestic scholars had proposed the detection methods of gear tooth side clearance and disc clearance in paper mill. This method has the characteristics of simple operation, high measurement accuracy, large adjustable range and real-time detection [1, 2]. Based on machine vision technology, Gan [3] adopted image contrast enhancement method and edge extraction method to better realize the installation clearance detection of motor rotating shaft. Sun [4] improved the high-precision sub-edge detection algorithm and applied it to the boundary extraction of tool setting image, which reduced the error caused by image positioning and improved the tool setting accuracy after coordinate transformation, and completed the visual detection of milling cutter wear. In addition, some scholars [5, 6] adopted the principle of electromagnetic induction and line structured light measurement to establish the gap measurement model or complete the design of the gap measurement device, which has high accuracy and engineering application value. In foreign countries, Lins [7] applied machine vision technology to the monitoring of tool wear during precision machining. Smith [8] used ultrasonic technology to design a new fuel plate gap probe (CGP) system for plate gap measurement. It can be seen that sensing detection technology and machine vision which are widely used in the field of mechanical detection. In this paper, the safety mechanism of the Three Gorges ship lift is taken as the research object. According to its structural characteristics, the sensing detection technology is used to propose a safety mechanism thread pair clearance detection scheme to meet the requirements of simple, reliable and real-time detection in the clearance detection process.

2 Structural Characteristics of Safety Mechanism

The ‘long nut-short screw’ type safety mechanism of the Three Gorges ship lift is the safety guarantee mechanism of the ship lift (as shown in Fig. 1). The safety mechanism is coupled with the adjacent driving mechanism through the mechanical shaft system, and the two are lifted and lowered synchronously. During the operation of the safety mechanism on the upper/lower chamber, the rotating screw lifts and drops synchronously with the small gear of the driving mechanism, and the screw teeth do not contact with the thread surface of the nut column. There is a 60 mm gap between the two in the upper and lower directions during installation.

As shown in Fig. 2, the safety mechanism is mainly composed of rotating screw, support rod, upper guide frame, lower guide frame, universal shaft and so on. The rotating screw is installed between the upper/lower guide frame, and the upper/lower guide frame slides up and down along the nut column to guide the positioning. The rotating screw rotates in a spiral motion relative to the nut column.

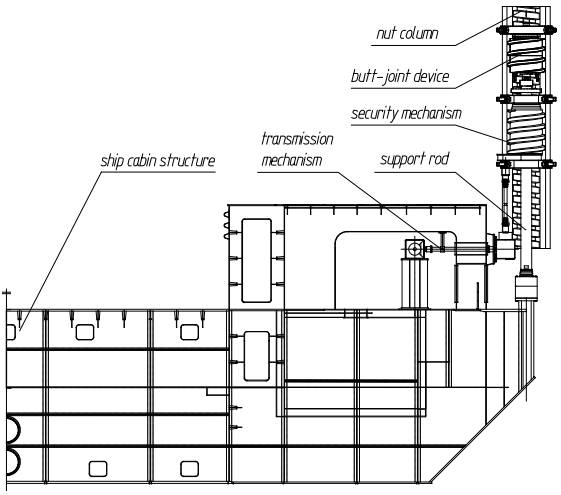


Fig. 1 General layout of security institutions

The safety mechanism has the following characteristics: (1) When the water leakage, shipwreck, overload or the ship chamber is submerged by the water in the ship chamber during the docking of the lower lock head, the full balance state of the ship chamber is destroyed, and the gap of the screw pair of the safety mechanism becomes smaller until the gap disappears completely. The unbalanced load of the ship chamber is transmitted to the nut column through the screw, and then the nut column is transmitted to the concrete structure of the tower column through the embedded parts; (2) when the ship cabin needs to be emptied for maintenance, the safety mechanism acts as the locking support of the ship cabin and plays a safety protection role; (3) the occurrence of stroke deviation during the lifting/downtime of the ship chamber will cause the clearance of the thread pair of the safety mechanism to become smaller until the clearance disappears, so that the safety mechanism is stuck, causing major equipment failure.

3 Influencing Factors of Thread Pair Clearance of Safety Mechanism

According to Fig. 3, the influencing factors of the thread pair clearance are analyzed as follows.

(1) Manufacturing installation error. Manufacturing installation errors mainly include transmission ratio error and screw, nut tooth size manufacturing error, etc., as shown in Table 1.

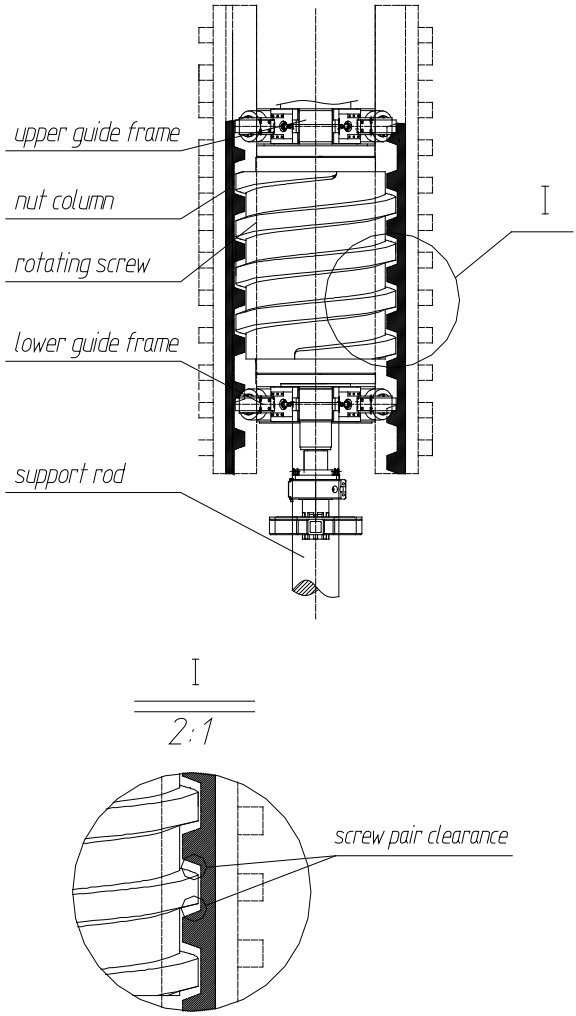


Fig. 2 Details of security institutions

(2) Structural deformation and displacement caused by external load. The influencing factors mainly include the transverse and longitudinal displacement of the ship chamber relative to the nut column and the elastic deformation of the ship chamber, as shown in Table 2.
(3) Wear caused by operation. Wear factors are shown in Table 3.

Comprehensive analysis of the above factors affecting the clearance of each thread pair, according to the principle of superposition of the results of each factor in the same direction, the maximum clearance change value is Eq. (1).

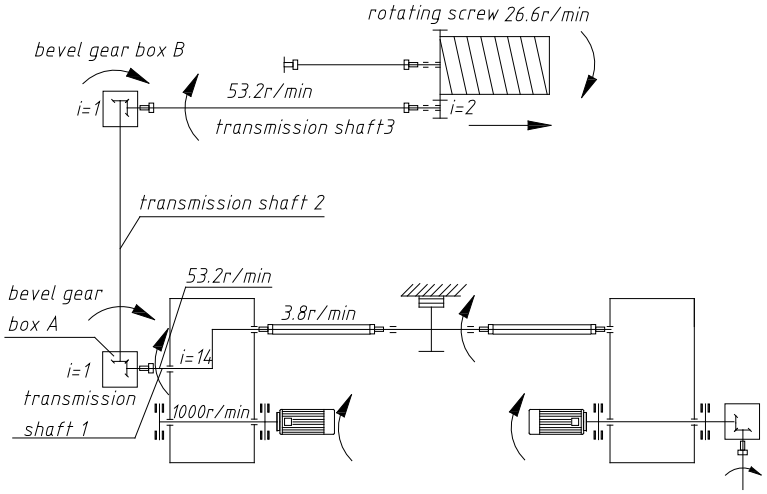


Fig. 3 Transmission diagram of safety mechanism

Table 1 Influence factors of manufacturing error on clearance

Serial number	The main factors affecting the thread pair clearance	Code	Change of gap (mm)
1	The transmission ratio error between the gear shaft of the driving mechanism and the screw shaft of the safety mechanism	Δt1	0 ~ 1.13
2	Screw, nut tooth size manufacturing error	Δt2	− 2.4 ~ 0
3	Nut column thread pitch manufacturing error	Δt3	± 3
4	The gap between the guide wheel and the guide rail surface of the screw guide frame	Δt4	± 0.3
5	Cumulative error of rack pitch (including gear/rack backlash)	Δt5	± 5.6
6	The relative error of the height position of the adjacent rack and nut column at the same elevation	Δt6	± 2
7	The relative error of height position between two pieces in the same nut column	Δt7	± 0.5
8	The installation error of the screw relative to the height of the nut column	Δt8	± 0.2
9	The influence of the adjustment gap between the two nut columns	Δt9	± 0.5
10	Influence of bearing radial clearance in pinion bracket mechanism	Δt10	± 0.87 ~ ± 1.90

Table 2 Influence factors of structural deformation and displacement on clearance

Serial number	The main factors affecting the thread pair clearance		Code	Change of gap (mm)
11	The influence of the elastic deformation of the ship chamber on the clearance of the thread pair under external load		Δt11	− 2.7 ~ 6.14
12	The transverse displacement of the ship chamber relative to the rack is 110 mm	Longitudinal and transverse displacement of ship chamber	Δt12	1.98
13	The longitudinal displacement of the chamber relative to the nut column is 95 mm		Δt13	− 1.42
14	The transverse displacement of the ship chamber relative to the nut column is 110 mm		Δt14	− 1.90
15	The ship chamber rotates around the transverse axis		Δt15	0
16	The ship chamber rotates around the vertical axis		Δt16	0
17	The ship chamber rotates around the longitudinal axis		Δt17	0
18	The influence of elastic deformation of pinion bracket mechanism on the clearance of thread pair		Δt18	0–10.4
19	The influence of the elastic deformation of the pin shaft at the hinged joint on the clearance of the thread pair in the gear bracket mechanism		Δt19	± 0.45
20	The influence of elastic torsion of pinion-universal coupling-main reducer-brake-bevel gear box shafting		Δt20	± (2.72 ~ 7.73)
21	The influence of different temperature changes between nut column and rack		Δt21	± (2.4 ~ 3.7)
22	Elastoplastic displacement of rack and pinion in vertical direction		Δt22	− 0.12 ~ 0.19

Table 3 Influence factors of running wear on clearance

Serial number	The main factors affecting the thread pair clearance		Code	Change of gap (mm)
23	Gear rack tooth thickness wear	The screw position can be adjusted through regular maintenance	Δt23	+ 5
24	Bearing wear in pinion bracket mechanism		Δt24	+ 1.03
25	The guide wheel and guide rail of the screw guide frame are worn out		Δt25	0

$$\Delta t = \sum_{i=1}^{n} \Delta t_i \tag{1}$$

In Eq. (1), Δt is the maximum clearance variation value of the thread pair, Δt_i is the influencing factors value of each thread pair. According to the calculation $\Delta t \approx 49$ mm, in the actual situation, the influence of various factors will be partly offset by each other. In addition, the actual accuracy is less than the theoretical accuracy range in the process of equipment manufacturing and installation, and the design value of the thread pair clearance of the safety mechanism is ± 60 mm, which can meet the clearance requirements. Since the trial navigation in 2016 year, the operation practice of the Three Gorges ship lift shows that the maximum variation of the thread clearance of the safety mechanism does not exceed ± 15 mm in the normal operation condition and the whole lift stroke of the ship chamber.

4 Gap Detection Scheme

According to the analysis results of the thread pair clearance of the safety mechanism, the clearance detection scheme of the safety mechanism of the Three Gorges ship lift is proposed.

4.1 Gap Detection Device

As shown in Fig. 4, the gap detection device is fixed in the thread extension section of the safety rotating screw, and its thread angle and thread tooth angle are the same as those of the safety rotating screw, and its thread thickness is larger than the safety rotating screw, so as to ensure that the gap detection device first contacts with the nut column thread. The nut column is composed of two tile structures. Each tile structure has a thread in the range of a°. There are grooves between the two pieces, and each

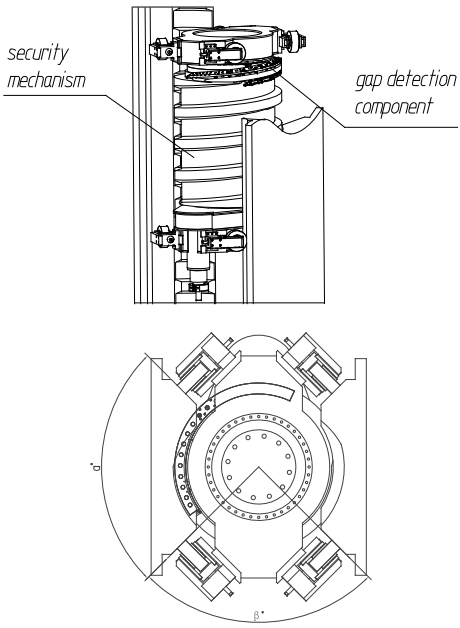


Fig. 4 Installation diagram of gap detection device

groove has a vacancy of $\beta = 180 - a°$. The arc length range of the gap detection device is greater than $\beta°$, which ensures that the safety screw has overlapping parts when it rotates to any angle relative to the nut column, and realizes the real-time detection of the gap of the safety mechanism.

As shown in Fig. 5, the gap detection device includes steel ball roller, upper detection block, fixed block, guide device, lower detection block, detection module, upper guide roller a, lower guide roller a, lower guide roller b, connecting seat and upper guide roller b. Multiple sets of steel ball rollers are installed in the threaded holes of the upper detection block and the lower detection block respectively. Multiple sets of guide devices are connected to the upper detection block and the fixed block, the lower detection block and the fixed block respectively to ensure that the upper detection block and the lower detection block move up and down along the fixed block respectively.

In the guide device, one end of the guide rod and the lower detection block are fixedly connected by screw two, and the other end is equipped with a support pad and a screw one. The spring is installed on the guide rod and is located between the fixed block and the lower detection block. The steel balls are assembled in the spherical groove of the roller seat. The upper guide roller a and the lower guide roller a are respectively installed on the inclined end surfaces of the upper detection block and the lower detection block. The upper guide roller b and the lower guide roller b are respectively installed on the other end surfaces of the upper detection block and the lower detection block, which are used to guide the detection block when it just enters the nut column and reduce the friction resistance.

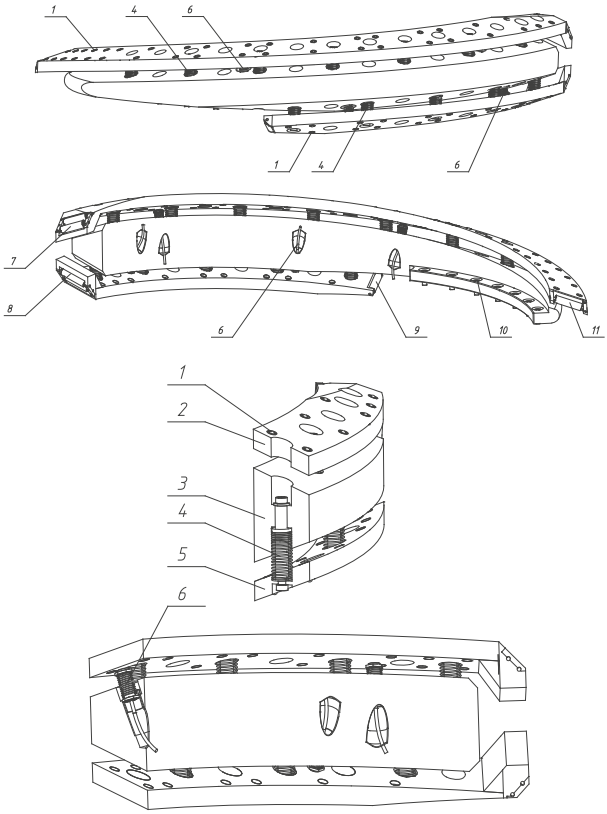


Fig. 5 Structure diagram of gap detection device. 1—Steel ball roller, 2—upper detection block, 3—fixed block, 4—guiding device, 5—lower detection block, 6—detecting module,7—top tumbler a, 8—bottom tumbler a, 9—bottom tumbler b, 10—connecting socket, 11 top tumbler b

As shown in Fig. 6, there are four groups of detection modules, two of which are used to detect the gap between the fixed block and the upper detection block, and the other two are used to detect the gap between the fixed block and the lower detection block. The detection module includes a positioning sleeve, an upper lock nut, a proximity switch, a spring two, and a lower lock nut. The proximity switch is installed in the positioning sleeve, and is limited by the upper lock nut and the lower lock nut. The spring two is installed on the proximity switch and is located between the lower lock nut and the step hole of the positioning sleeve.

As shown in Fig. 7, taking the gap between the detection fixed block and the lower detection block as an example, the end face of the proximity switch and the lower detection block maintain sufficient spacing n1, which is in an untriggered state. When the current detection block begins to contact the tooth surface of the nut column and continues to compress the lower detection block, the proximity switch spacing changes from n1 to n2 (n1 > n2), and the safety screw stops rotating,

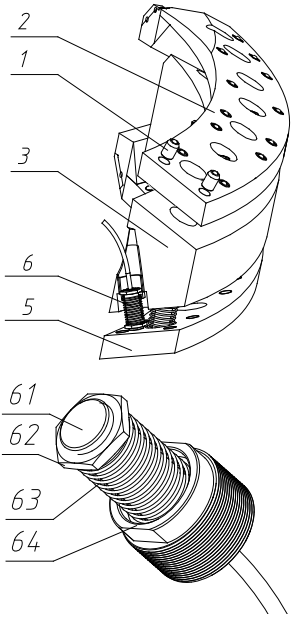


Fig. 6 Local diagram of gap detection device. 61—proximity switch, 62—upper/lower locking nuts, 63—spring, 64—set collar

suggesting a small limit fault of the gap, where n2 is the proximity switch trigger spacing; when the current detection block and the nut column tooth surface continue to contact and compress, after the proximity switch spacing becomes 0, the spring two is compressed, and the proximity switch is retracted back into the positioning sleeve; when the current detection block is in full contact with the fixed block, the proximity switch is in the maximum retraction state, so that the extrusion force on the proximity switch is equivalent to the spring two compression force, so as to ensure that the proximity switch is not crushed. When the gap between the safety screw and the nut column returns to normal, the lower detection block and the fixed block are synchronously out of contact and the initial spacing is restored, and the proximity switch returns to the initial position under the spring two force.

4.2 Gap Detection Method

(1) Device installation. The fixing is fixedly connected to the upper end of the safety screw through the connecting seat to ensure the continuous transition of the gap detection component and the safety screw thread without angle and displacement deviation. In the initial state, the upper detection block/lower detection block and the fixed block are installed to maintain a fixed gap under the action of the guiding device, and the gap between the upper detection block/lower detection block

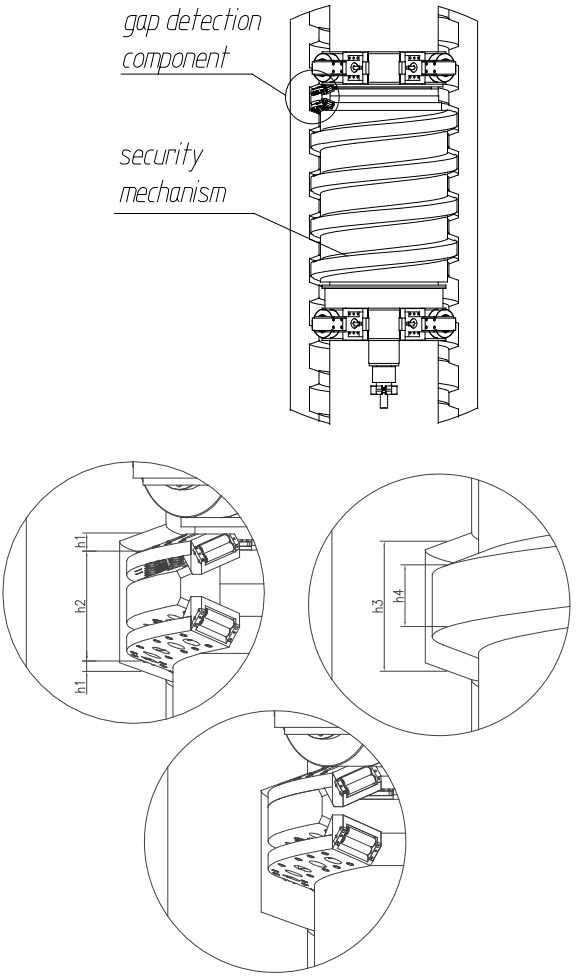


Fig. 7 Gap detection and safety protection diagram

and the nut column is h1; the distance between the end face of the proximity switch and the upper detection block/lower detection block is n1.

(2) Gap detection. During the up/down process of the ship lift, the safety screw rotates synchronously in the nut column. When the gap between the safety screw and the nut column changes, the upper detection block or the lower detection block in the gap detection component begins to contact with the nut column thread. When the upper detection block or the lower detection block continues to contact and compress the nut column thread surface for a certain distance, the proximity switch sends a message, the safety screw stops running, and the system stops. When the ship stops, the safety screw moves up and down relative to the nut column, resulting in a change in the gap. The gap detection between the safety screw and the nut column is transmitted through the proximity switch.

(3) Security protection. In the initial natural state, the total thickness of the screw teeth composed of the upper detection block and the lower detection block in the gap detection component is h2, the gap between the nut column (1) screw teeth is h3, h3 = h2 + 2 * h1, and the thickness of the safety screw teeth is h4, where h2 is greater than h4. When the upper detection block/lower detection block is in full contact with the fixed block, the safety screw thread contacts with the nut stud thread synchronously and bears the load, which can protect the clearance detection assembly. When the gap between the safety screw teeth changes, the proximity switch is sent, and the proximity switch can be retracted in the positioning sleeve, which can protect the proximity switch from being crushed.

4.3 Gap Detection Test

After the installation and debugging of the gap detection device is completed, the detection accuracy test of the gap detection device is carried out by setting the initial distance between the adjustment proximity switch and the upper detection block/lower detection block and the initial gap between the fixed block and the upper detection block/lower detection block to demonstrate its safety and reliability. First, the cabin is parked at a fixed elevation, and the coupling between the universal shaft assembly and the lower guide frame is separated. The pinion shaft in the lower guide frame is turned. The pinion drives the large gear to rotate, and the large gear is connected with the safety mechanism to make the safety mechanism rotate relative to the nut column. The gap test results after multiple tests are shown in Table 4.

In Table 4, by setting the transmitting gap of the detection device under different working conditions and comparing with the measured gap, the detection error of the gap detection device is less than 1 mm. It can be seen that the detection accuracy of the gap detection device meets the requirements, and the device is stable and reliable.

Table 4 Influence factors of manufacturing error on clearance

Gap number	Set the signalling gap/mm	Measured gap of safety mechanism/mm	Detection error/mm
1	45	44	1
2	40	40	0
3	35	34	1
4	30	31	1
5	25	25	0

5 Conclusions

a. The maximum theoretical value of the influence of manufacturing and installation errors, structural deformation caused by external loads, displacement and wear caused by operation on the thread pair clearance of the safety mechanism can reach ± 49 mm, and the maximum variation in actual operation does not exceed ± 15 mm.
b. The gap monitoring device has overlapping parts when it rotates to any angle relative to the nut column, so as to realize the real-time detection of the gap of the safety mechanism. When the upper detection block/lower detection block in the device is in full contact with the fixed block, the safety screw thread and the nut column thread are in synchronous contact and bear the load, which can avoid the excessive load on the gap detection device.
c. By adjusting the initial distance between the proximity switch and the upper detection block/lower detection block and the initial gap between the fixed block and the upper detection block/lower detection block, the detection and transmission of the safety mechanism under different minimum gaps can be realized, which is convenient for equipment installation and debugging, through the test of the gap detection device, the detection error values are less than 1 mm, and can adapt to the gap detection of the safety mechanism under different working conditions.

References

1. Jiang W, Wang Y, Luo J et al (2022) A gear tooth side clearance adjustment device and measurement method. Mech Des 39(S2):103–107
2. Xing Y, Dong J, Liu Y et al (2020) Research progress of real-time on-line detection technology for disc mill gap. China Paper 39(12):75–80
3. Sun Q, Li Y (2022) Research on the detection method of tool clearance based on machine vision. J Changchun Univ Sci Technol 45(05):53–60
4. Gan Y (2020) Research on installation gap detection method in complex environment based on machine vision. Hunan University of Science and Technology
5. Jing Y, Wang S, Feng W et al (2023) Research on suspension gap detection method based on electromagnet additional coil. Chin J Electr Eng 43(17):6807–6816
6. Zhu Y, Zhuo X, Huang J et al (2023) Automatic detection of riveting gaps in aircraft skins based on line structured light. Manuf Autom 45(10):38–44
7. Lins RG, de Araujo PRM, Corazzim M (2020) In-process machine vision monitoring of tool wear for cyber-physical production systems. Rob Comput Integr Manuf 61:101859–101876
8. Smith JA, Jesse CJ, Hanson WA et al (2023) Channel gap measurements of irradiated plate fuel and comparison with post-irradiation plate thickness. Nucl Eng Technol 55(6):2195–2205

Open Access This chapter is licensed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License (http://creativecommons.org/licenses/by-nc-nd/4.0/), which permits any noncommercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if you modified the licensed material. You do not have permission under this license to share adapted material derived from this chapter or parts of it.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.



Mechanical Engineering and Dynamics Analysis Research

An Integrated Approach for Tolerance Synthesis of a Spiral Bevel Gearbox



Siyuan Hu, Yujie Huang, and Wenhui Zeng

Abstract This paper proposes a systematic and integrated methodology for the tolerance synthesis of spiral bevel gearbox. In this methodology, for the complex structure of this gearbox, dimensional tolerances and geometric tolerances are modeled by the screw theory. Parallel chains produced by the fits of gear shaft and holes are solved by a new integrated method based on localization tolerancing with contact influence method and screw theory. Leverage effect of geometrical structures on assembly deviation is considered to enhance the precision based on unified Jacobian-Torsor model. At last, bottleneck tolerances are identified for tolerance optimization by contribution analysis and sensitivity analysis. By means of this approach, tolerance synthesis results of spiral bevel gearbox are more accurate and comprehensive for practical application.

Keyword Spiral bevel gearbox · Parallel chains · Unified Jacobian-Torsor model · Screw theory

1 Introduction

Due to high contact ratio, excellent load bearing capacity, smooth operation and low noise, spiral bevel gears are commonly used in power transmission systems with intersecting or crossed axes, such as transmission devices of aircraft, ships, helicopters and other machines, To ensure high precision assembly while minimizing costs and compensating for inevitable errors, it is essential to focus on the tolerance synthesis of the spiral bevel gearbox.

In recent decades, a lot of researches about the gears are the tooth contact analysis such as the ones by Litvin [1] and Simon [2, 3]. However, few of them involve the tolerance analysis of gears. Classical tolerance analysis models, such as T-Map model

S. Hu · Y. Huang · W. Zeng (✉)
School of Mechanical and Electrical Engineering, Wuhan University of Technology, Wuhan, People's Republic of China
e-mail: wenhuizeng@whut.edu.cn

© The Author(s) 2026

H. Saman et al. (eds.), *The 9th International Conference on Advances in Construction Machinery and Vehicle Engineering*, Lecture Notes in Mechanical Engineering,
https://doi.org/10.1007/978-981-96-5527-4_32

[4], Matrix model [5], Unified Jacobian-Torsor model [6], DLM [7, 8], and SDT-DP model [9], can be used to calculate assembly deviations. These models focus on the research hotspots of the 3-D (three dimensional) tolerance analysis [10–13]. For the specific mechanical structure, Desrochers [5] proposed a matrix approach applied to a simplified gear pump assembly, while Asante [14] introduced a constraint-based tolerance analysis for multi-operation single setup and multi-operation multi-setup part–fixture assembly. However, few of these approaches can be applied to the spiral bevel gearbox because of the complexity of 3-D dimension chains [15–17].

In a spiral bevel gearbox, part deviations have an impact on assembly deviations, which may cause transmission error, tooth contact misalignment, meshing interference, clearance, etc. Functional requirements should be satisfied by limiting the related deviations to ensure a high assembly quality. Moreover, the parallel chains produced by the fits of gear shaft and holes in the 3-D dimension chains of spiral bevel gearbox could not be solved by the proposed approach in [18]. So far, few researchers focused their attention on the solution of the parallel chains [19]. Some existing methods may be applied to solve the parallel chains [20], for example CLIC (localization tolerancing with contact influence) method [21, 22], this method is applied to define the functional specifications of parts and the displacements of contact points [23]. However, the calculation result of this method only accounts for the position deviation of translation parameters (u, v, w) without considering the orientation deviation of rotation parameters (α, β, δ).

This paper aims to introduce a systematic approach on the worst-case tolerance analysis of spiral bevel gearbox. CLIC method and screw theory will be combined to develop a new approach to solve parallel chains with calculating the six degrees of freedom: translation (u, v, w) and rotation (α, β, δ). Considering the leverage effect of geometrical structures, the torsor model is used for tolerance representation and the Jacobian matrix for tolerance propagation. Using this approach, the final assembly deviation can be calculated based on the Unified Jacobian-Torsor model, and bottleneck tolerances can be found for tolerance optimization by contribution analysis and sensitivity analysis.

The remainder of this paper is organized as follows: the detailed process to construct tolerance analysis model of spiral bevel gearbox is discussed in Sect. 2; in Sect. 3 an integrated approach is proposed to solve the parallel chains; in Sect. 4 the industrial application of spiral bevel gearbox is shown; finally, the systematic approach is summarized and conclusion is presented in Sect. 5.

2 Tolerance Analysis of Spiral Bevel Gearbox

2.1 Functional Requirement of the Gearbox

In this paper, Unified Jacobian-Torsor model is used to calculate the final accumulation deviation of the gearbox, with a detailed description introduced by Desrochers in [6]. A spiral bevel gearbox mainly consists of gear shafting components, pinion shafting components and housing. Figure 1 shows the assembly drawing of spiral bevel gearbox. The assembly quality of spiral bevel gearbox affects the distribution of gear meshing, contact path and vibration, which in turn affects its mechanical characteristics and operational lifespan. The challenge in assembly is to decrease the shaft angle deviation f_Σ and the axis intersection point deviation f_a and improve its assembly precision [24]. The assembly requirements of f_Σ and f_a should be fulfilled that:

$$(f_\Sigma)_{\max} \leq f_\Sigma; \quad (f_a)_{\max} \leq f_a \tag{1}$$

2.2 Assembly Graph of the Gearbox

Input and output shafts are mounted with clearance on the gearbox housing as shown in Fig. 1. Due to the interference fit between the shafts and bearings, this aspect is not discussed in this paper. Actually, the fit of the shaft and holes are very common in many mechanisms; the input shaft or output shaft is fitted with two holes on the gearbox housing. As shown in Fig. 2, it is the assembly connection graph of gearbox with parallel chains. This is presented as an example to illustrate the parallel chains produced by the fit of the shaft and holes.

As shown in Fig. 2, there are two internal pairs IFE8, IFE9 and five kinematic pairs CFE1, CFE2, CFE3, CFE4 and CFE7. The kinematic pairs CFE1, CFE2, CFE3 and CFE4 are produced by the fit of shaft and holes. The FR1 and FR2 are the radial deviations of gear1 and gear 2, respectively. The input shaft and output shaft are each fitted with clearance in the respective holes. It is evident that the chains of 1.1 $\rightarrow$ 1.3 $\rightarrow$ 4.1 $\rightarrow$ 4.3 and the chains of 1.1 $\rightarrow$ 1.5 $\rightarrow$ 4.2 $\rightarrow$ 4.3 are parallel; and the chains of 1.2 $\rightarrow$ 1.4 $\rightarrow$ 5.1 $\rightarrow$ 5.3 and the chains of 1.2 $\rightarrow$ 1.6 $\rightarrow$ 5.2 $\rightarrow$ 5.3 are parallel. The two parallel chains are produced by the fit between shaft and holes. However, ignoring the parallel chains would lead to a decrease in the accuracy and reliability of the Unified Jacobian-Torsor model. Typically, simplification has been used to avoid the parallel chains, but it could lead large errors. In this section, an integrated approach based on CLIC method and screw theory is proposed to address the parallel chains produced by the fit between shaft and holes.

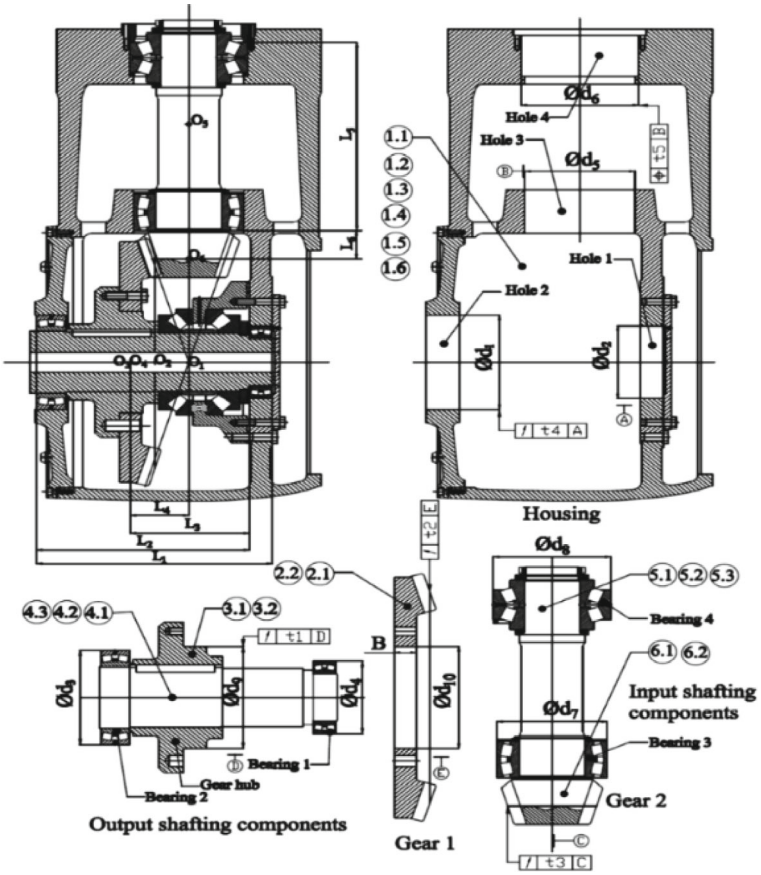


Fig. 1 Assembly drawing of spiral bevel gearbox

3 The Integrated Approach to Address the Parallel Chains

The key to solving the parallel chains produced by the fit of shaft and holes lies in obtaining the torsor parameters of the shaft; in other words we need to calculate the torsor parameters of an arbitrary point in the nominal axis of shaft. In this section, the integrated approach to solve the parallel chains is proposed.

3.1 Description of a Cylindrical Feature

Firstly, for a cylindrical feature, the position deviation of the cylinder must be defined using its axis. As illustrated in Fig. 3, a cylindrical feature of axis AB is defined by the two endpoints A and B, with the center point O aligned along AB. For any point

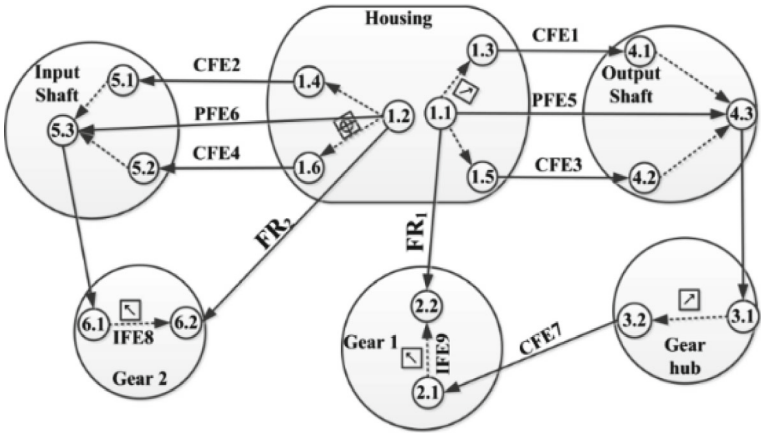


Fig. 2 Assembly graph of kinematic chains

M on the axis AB of the cylinder, the following expressions are hold:

$$\begin{cases} x_{A1} - x_A = \delta(y_A - y_M) + x_{M1} \\ z_{A1} - z_A = \alpha(y_A - y_M) + z_{M1} \\ x_{B1} - x_B = \delta(y_B - y_M) + x_{M1} \\ z_{B1} - z_B = \alpha(y_B - y_M) + z_{M1} \end{cases} \quad \left(-\frac{t_d}{L} \le \alpha \le \frac{t_d}{L}, -\frac{t_d}{L} \le \delta \le \frac{t_d}{L}\right) \tag{2}$$

where $(x_A,\ y_A,\ z_A)$, $(x_B,\ y_B,\ z_B)$ and $(x_M,\ y_M,\ z_M)$ are the coordinates of point A, point B and point M on the nominal axis AB, respectively. $(x_{A1},\ z_{A1})$, $(x_{B1},\ z_{B1})$ and $(x_{M1},\ z_{M1})$ are the x and z coordinates of point A, point B and point P on the real axis, respectively. Because point A and point B are on the nominal axis, the values of x_A, x_B, z_A and z_B are equal to 0.

By expressing (2) in matrix form and rearranging it, the small displacement torsor parameters of point M can be derived as [25]:

$$\begin{bmatrix} \alpha \\ \delta \\ x_{M1} \\ z_{M1} \end{bmatrix} = \begin{bmatrix} 0 & (y_A - y_M) & 1 & 0 \\ (y_A - y_M) & 0 & 0 & 1 \\ 0 & (y_B - y_M) & 1 & 0 \\ (y_B - y_M) & 0 & 0 & 1 \end{bmatrix}^{-1} \begin{bmatrix} x_{A1} \\ z_{A1} \\ x_{B1} \\ z_{B1} \end{bmatrix} \tag{3}$$

Besides, for a cylindrical feature, the small displacement torsor v and β are meaningless because they do not affect the position variation according to the open degrees of freedom, therefore v and β are set as 0.

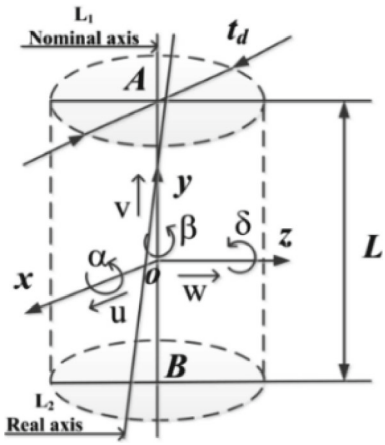


Fig. 3 Torsor model for cylindrical feature

3.2 Solution to Obtain the Torsor Parameters of Arbitrary Point M

According to (3), α, δ, x_{P1} and z_{P1} of point M can be obtained:

$$\begin{cases} \alpha = \frac{z_{A1}-z_{B1}}{y_A-y_B} \\ \delta = \frac{x_{A1}-x_{B1}}{y_A-y_B} \end{cases} \tag{4}$$

$$\begin{cases} x_{M1} = x_{A1} - \frac{x_{A1}-x_{B1}}{y_A-y_B} \cdot (y_A - y_M) \\ z_{M1} = z_{A1} - \frac{x_{A1}-x_{B1}}{y_A-y_B} \cdot (y_A - y_M) \end{cases} \tag{5}$$

In view of (4) and (5), it is obvious that if the y coordinate of points A, B and M and values of $(x_{A1},\ z_{A1})$ and $(x_{B1},\ z_{B1})$ are known, $\left[x_{M1}\ z_{M1}\ \alpha\ \delta\right]^T$ of point M can be calculated.

As shown in Fig. 4, when the shaft is fitted with clearance in the holes, point P_1 of hole1 is in contact with point P_{1s} and point P_2 of hole 2 is in contact with point P_{2s} of the shaft. Points A and B lie on the nominal axis at the centers of the shaft's end surfaces. Points A_{SA} and B_{SA} lie on the actual axis corresponding to the nominal points A and B. Δd_1 and Δd_2 are the biggest displacement between the shaft and holes. The biggest displacement of point A in direction f_1 is given by the vector $\overrightarrow{AA_{SA}} \cdot \overrightarrow{f_1}$ that can be described as:

$$\Delta d_1 = \overrightarrow{AA_{SA}} \cdot \overrightarrow{f_1} \tag{6}$$

Similarly, the biggest displacement of point B in direction f_2 is given by the vector $\overrightarrow{BB_{SA}} \cdot \overrightarrow{f_2}$ that can be described as:

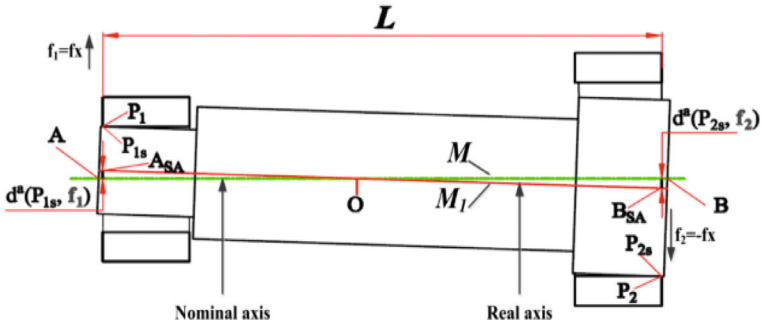


Fig. 4 Displacements of analysis points A and B in the opposite direction

$$\mathrm{d}_2 = \overrightarrow{\mathrm{BB}_{\mathrm{SA}}} \cdot \vec{f_2} \tag{7}$$

The biggest deviations of points A and B in the directions f_x and f_z are the same, so they can be obtained that: $|x_{A1}| = |z_{A1}| = \Delta \mathrm{d}_1$, $|x_{B1}| = |z_{B1}| = \mathrm{d}_2$. For worst-case tolerance analysis, we only need to obtain the biggest absolute values of x_{M1}, and z_{M1}, α and δ. If we want to obtain the biggest absolute value of α and δ, , in view of (4), the sign of x_{A1} and x_{B1} are opposite. As shown in Fig. 4, the offset directions of point A and point B are opposite, and assume that $x_{A1} = \Delta \mathrm{d}_1$ and $x_{B1} = -\Delta \mathrm{d}_2$. . According to (4), the biggest absolute value of δ can be obtained:

$$|\delta|_{max} = \delta = \frac{\Delta \mathrm{d}_1 + \Delta \mathrm{d}_2}{L} \tag{8}$$

Similarly, with $z_{A1} = \Delta \mathrm{d}_1$ and $z_{B1} = -\Delta \mathrm{d}_2$, the biggest absolute value of α can be obtained:

$$|\alpha|_{max} = \alpha = \frac{\Delta \mathrm{d}_1 + \Delta \mathrm{d}_2}{L} \tag{9}$$

Assume that $y_A - y_M = \mathrm{AM}$. According to the (4), (8) and (9), the absolute values of the corresponding parameters x_{M1} and z_{M1} can be calculated:

$$\begin{cases} |x_{M1}| = \left| \Delta \mathrm{d}_1 - \frac{\Delta \mathrm{d}_1 + \Delta \mathrm{d}_2}{L} \cdot \mathrm{AM} \right| \\ |z_{M1}| = \left| \Delta \mathrm{d}_1 - \frac{\Delta \mathrm{d}_1 + \Delta \mathrm{d}_2}{L} \cdot \mathrm{AM} \right| \end{cases} \tag{10}$$

According to the calculation results of (8), (9) and (10), with the biggest orientation deviations α and δ the torsor parameters of point M can be written as:

$$T_M = \begin{bmatrix} \pm\left|\Delta d_1 - \frac{\Delta d_1+\Delta d_2}{L} \cdot AM\right| \\ 0 \\ \pm\left|\Delta d_1 - \frac{\Delta d_1+\Delta d_2}{L} \cdot AM\right| \\ \pm\frac{\Delta d_1+\Delta d_2}{L} \\ 0 \\ \pm\frac{\Delta d_1+\Delta d_2}{L} \end{bmatrix} \tag{11}$$

where AM represents the distance between point A and point M. However, If we want to obtain the biggest absolute value of x_{M1} and z_{M1}, in view of (4), the sign of x_{A1} and x_{B1} are the same. As shown in Fig. 5, the offset directions of point A and point B are the same. Assume that $x_{A1} = z_{A1} = \Delta d_1$ and $x_{B1} = z_{B1} = \Delta d_2$. . According to (4), the biggest absolute value of x_{M1} and z_{M1} can be calculated:

$$\begin{cases} |x_{M1}|_{\max} = \left|\Delta d_1 - \frac{\Delta d_1-\Delta d_2}{L} \cdot AM\right| \\ |z_{M1}|_{\max} = \left|\Delta d_1 - \frac{\Delta d_1-\Delta d_2}{L} \cdot AM\right| \end{cases} \tag{12}$$

Simultaneously, the absolute values of the corresponding parameters α and δ can be calculated:

$$\begin{cases} |\alpha| = \left|\frac{\Delta d_1-\Delta d_2}{L}\right| \\ |\delta| = \left|\frac{\Delta d_1-\Delta d_2}{L}\right| \end{cases} \tag{13}$$

According to the calculation results of (15) and (16), with the biggest position deviations $|x_{M1}|$ and $|z_{M1}|$ the torsor parameters of point M can be written as:

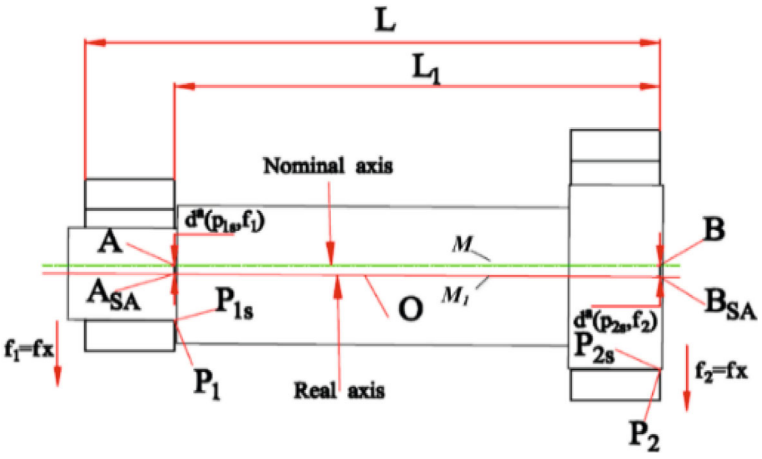


Fig. 5 Displacements of analysis points A and B in the same direction

$$T_M = \begin{bmatrix} \pm\left|\Delta d_1 - \frac{\Delta d_1 - \Delta d_2}{L} \cdot AM\right| \\ 0 \\ \pm\left|\Delta d_1 - \frac{\Delta d_1 - \Delta d_2}{L} \cdot AM\right| \\ \pm\left|\frac{\Delta d_1 - \Delta d_2}{L}\right| \\ 0 \\ \pm\left|\frac{\Delta d_1 - \Delta d_2}{L}\right| \end{bmatrix} \tag{14}$$

The final torsor parameters of point M can be expressed as:

$$T_M = \begin{bmatrix} \pm\left|\Delta d_1 - \frac{\Delta d_1 + \Delta d_2}{L} \cdot AM\right| \\ 0 \\ \pm\left|\Delta d_1 - \frac{\Delta d_1 + \Delta d_2}{L} \cdot AM\right| \\ \pm\frac{\Delta d_1 + \Delta d_2}{L} \\ 0 \\ \pm\frac{\Delta d_1 + \Delta d_2}{L} \end{bmatrix} \text{or} \begin{bmatrix} \pm\left|\Delta d_1 - \frac{\Delta d_1 - \Delta d_2}{L} \cdot AM\right| \\ 0 \\ \pm\left|\Delta d_1 - \frac{\Delta d_1 - \Delta d_2}{L} \cdot AM\right| \\ \pm\left|\frac{\Delta d_1 - \Delta d_2}{L}\right| \\ 0 \\ \pm\left|\frac{\Delta d_1 - \Delta d_2}{L}\right| \end{bmatrix} \tag{15}$$

where L is the distance between the two contact points P_1 and P_2, AM is the distance between point A and point M. It should be noted that if the local reference frame of the functional requirement is located between the two contact points A and B, Eq. (14) should be selected to compute the maximum position deviation, alternatively, Eq. (11) should be used for the maximum orientation deviation. Additionally, if the reference system of the functional requirement is positioned along the extension line of contact points A and B, Eq. (14) must always be chosen to calculate the maximum orientation deviation of the functional requirement.

4 Industrial Application

4.1 Problem Formulation

The challenge of this project is to decrease the shaft angle deviation f_Σ and the axis intersection point deviation f_a of spiral bevel gearbox and improve its assembly precision. The parameters of the spiral bevel gear pair are provided in Table 1. The dimensions and tolerances of this assembly are listed in Table 2. The assembly requirements shaft angle deviation f_Σ and axis intersection point deviation f_a should be fulfilled that:

$$f_\Sigma \leq 0.0015\,\text{rad}; \quad f_a \leq 0.700\,\text{mm} \tag{16}$$

Table 1 Parameters of spiral bevel gears pair

Gear	Number of teeth Z	Normal module m_n	Transverse module m_t	Reference cone angle δ	Facewidth b (mm)
Gear 1	56	17.0499	24.11	74.055°	220
Gear 2	16	17.0499	24.11	15.945°	220

Table 2 Tolerances of the assembly

Dimension	Values (mm)	Dimension	Values (mm)	Geometric tolerances	Values (mm)
d_1	520 (+ 0.07,0)	d_6	533.4 (+ 0.07,0)	t_1	0.02
d_2	400 (+ 0.089,0)	d_7	500 (0,− 0.035)	t_2	0.048
d_3	520 (0,− 0.04)	d_8	533.4 (+ 0.025,0)	t_3	0.04
d_4	520 (0,− 0.036)	d_9	560 (± 0.022)	t_4	0.05
d_5	500 (+ 0.063,0)	d_{10}	560 (+ 0.07,0)	t_5	0.1

4.2 Tolerance Analysis of Spiral Bevel Gearbox

For output shafting components, PFE5 is set to replace the parallel chains of 1.1 → 1.3 → 4.1 → 4.3 and 1.1 → 1.5 → 4.2 → 4.3 shown in Fig. 2. The point A lies on the nominal axis of bearing 2, while point B lies on the nominal axis of bearing 1. According to (6) and (7), the biggest displacement of point A and point B of output shaft can be obtained:

$$\Delta d_{1out} = \frac{d_1 - d_3 + t_4}{2} = 0.08\,\text{mm} \tag{17}$$

$$\Delta d_{2out} = \frac{d_2 - d_4}{2} = 0.0625\,\text{mm} \tag{18}$$

The location of the gear 1 is between points A and B along the nominal axis of the output shaft. According to the criterion of the integrated approach, the orientation deviation of gear 1 is biggest when the offset directions of the contact points P_1 and P_2 are opposite as shown in Fig. 4. Point M corresponds to point O_3 on the nominal axis of the shaft as shown in Fig. 2, and AM represents the distance between point O_3 and the right end surface of bearing 1.

However, the position deviation of gear 1 is biggest when the offset directions of the contact points P_1 and P_2 are aligned as shown in Fig. 5. AM represents the distance between point O_3 and left end surface of bearing 1, and it is equal to L_3.

For input shafting components, PFE6 is set to replace the parallel chains of 1.2 → 1.4 → 5.1 → 5.3 and 1.2 → 1.6 → 5.2 → 5.3 as shown in Fig. 3. For input shaft, point A lies on the nominal axis of bearing 3, while point B lies on the nominal axis of bearing 4, according to (6) and (7), the biggest displacement of point A and point B on the nominal axis of output shaft can be obtained:

$$\Delta d_{1in} = \frac{d_5 - d_7}{2} = 0.085\,\text{mm} \tag{19}$$

$$\Delta d_{2in} = \frac{d_6 - d_8 + t_5}{2} = 0.049\,\text{mm} \tag{20}$$

The location of the gear 2 is on the extension line of the nominal axis of the input shaft. According to the criterion of the integrated approach, the position deviation and orientation deviation of gear will be biggest when the offset directions of the contact points are the opposite as shown in Fig. 5. In order to simplify the torsor parameters of PFE6, we can choose a point M on the nominal axis of input shaft as point O5 to let $\left|\Delta d_1 - \frac{\Delta d_1 + \Delta d_2}{L} \cdot AM\right|$ be 0.

As the tolerances shown in Table 3, for the input shafting components, as the dimensions shown in Table 4, AM can be calculated and it is 655.9 mm. Then values of l_{56} can be calculated by:

$$l_{56} = O_6O_5 = 1034 - AM + 155.67 = 533.77\,\text{mm} \tag{21}$$

The l_{32} is the distance between the points O_3 and O_2, and it can be obtained by:

$$l_{32} = L_4 - R_1 = L_4 - \left(\frac{Z_1 \cdot m_t}{2 \cdot \sin\delta_1} - \frac{b}{2}\right)\cos\delta_1$$

Table 3 Dimensions of the assembly

Dimensions	$L1$	$L2$	$L3$	$L4$	$L5$	$L6$	B
Values (mm)	1069	965	540	265	1034	155.67	100

Table 4 Torsor parameters of PFE5, PFE6, CFE7, IFE8 and IFE9

PFE5	[[±0.006], [0,0], [±0.006], [±0.000133], [0,0], [±0.000133]]^T (when AM = 644 mm) [[±0.07], [0,0], [±0.07], [±0.000018], [0,0], [±0.000018]]^T (when AM = 540 mm)
PFE6	[[0,0], [0,0], [0,0], [±0.000130], [0,0], [±0.000130]]^T
CFE7	[±0.056,0, ±0.056, ±0.00112,0, ±0.00112]^T
IFE8	[±0.02,0, ±0.02, ±0.00018,0, ±0.00018]^T
IFE9	[±0.024,0, ±0.024, ±0.00022,0, ±0.00022]^T

$$= 265 - 162.65 = 102.35\,\text{mm} \tag{22}$$

The point O_4 coincides with point O_3; then l_{42} is also 102.35 mm. For output shafting components, when calculating the maximum orientation deviation of gear 1, point M is the point O_3 which is coincided with the center point O_4 of gear hub as shown in Fig. 2, and AM is equal to $(L_1 - L_2 + L_3)$ and it is 644 mm. When calculating the maximum position deviation of gear 1, AM is equal to L_3 and it is 540 mm. The torsor parameters of PFE5, PFE6, CFE7, IFE8 and IFE9 can be calculated and they are shown in Table 4, and the final calculated assembly deviations of the gear 1 and gear 2 are shown in Table 5 according to the Unified Jacobian-Torsor model.

In this paper, the theoretical shaft angle of the spiral bevel gear box is 90°. After obtaining the deviations of gear 1 and gear 2, the shaft angle deviation f_Σ and axis intersection point deviation f_a can be calculated according to the approach proposed by Zeng [24], and the function $\max(f_\Sigma)$ about the tolerances can be obtained that:

$$\begin{aligned}\max(f_\Sigma) = {} & 0.000468((d_1 - d_3) + t_4 + (d_2 - d_4)) \\ & + 0.004545(t_2 + t_3) + 0.01((d_{10} - d_9) + t_1) \\ & + 0.000484((d_6 - d_8) + t_5 + (d_5 - d_7))\end{aligned} \tag{23}$$

The maximum axis intersection point deviation f_a can be derived as:

$$\begin{aligned}\max(f_a) = {} & 0.358((d_1 - d_3) + t_4) + 0.142(d_2 - d_4) \\ & + 3.15((d_{10} - d_9) + t_1) + 1.239t_2 \\ & + 0.533((d_6 - d_8) + t_5 + (d_5 - d_7)) + 3.088t_3\end{aligned} \tag{24}$$

Substituting the tolerances values shown in Table 2 into (23) and (24), the final f_Σ and f_a can be calculated and they are 0.00178 rad and 0.754 mm, respectively.

Table 5 Deviations of gear 1 and gear 2

Gear 1				Gear 2	
$d_{r1-omax}$ (mm)	$\alpha_{r1-omax}$ (rad)	$d_{r1-pmax}$ (mm)	$\alpha_{r1-pmax}$ (rad)	d_{r2-max} (mm)	α_{r2-max} (rad)
0.214	0.00147	0.267	0.00136	0.089	0.00031

The maximum radial deviation and orientation deviation of gear 1 is ± 0.266 mm and ± 0.00147 rad, respectively.
The maximum radial deviation and orientation deviation of gear 2 is ± 0.089 mm and ± 0.0003 rad, respectively.

4.3 Contribution Analysis and Sensitivity Analysis (Bottleneck Tolerance Determination)

By evaluating the contribution of functional elements (component ring) on the assembly functional requirements, the bottleneck tolerances can be determined and optimized. To some extent, the sensitivity coefficient can reflect the leverage effect of geometrical structures on the deviation. The contributions analysis and sensitivity analysis of deviations on the shaft angle deviation f_Σ and axis intersection point deviation f_a are shown in Table 6.

From Table 6, it can be seen that the sensitivity coefficients of $(d_{10} - d_9 + t_1)$ on the deviations f_Σ and f_a both are the largest with 0.01 and 3.15, respectively, and its contributions on the deviations f_Σ and f_a also both the largest with 62.82% and 46.81%, respectively. Therefore, for the functional requirements: f_Σ and f_a, $(d_{10} - d_9 + t_1)$ is the bottleneck deviation. The detailed constituent elements of deviation $(d_{10} - d_9 + t_1)$ are shown in Table 7; the contributions of fit tolerance of d_{10} and d_9 on the deviations f_Σ and f_a are 51.60% and 38.45%, respectively. So for the functional requirements: f_Σ and f_a, d_{10} and d_9 are the bottleneck tolerances.

Table 6 Contributions and sensitivity coefficients of the tolerances on f_Σ and f_a

Deviations	Values (mm)	Sensitivity coefficients of deviations		Contributions of deviations	
		On f_Σ	On f_a	On f_Σ (%)	On f_a (%)
$d_1 - d_3 + t_4$	0.16	0.000468	0.358	4.20	7.60
$d_2 - d_4$	0.125	0.000468	0.142	3.28	2.36
$d_{10} - d_9 + t_1$	0.112	0.01	3.15	62.82	46.81
t_2	0.048	0.004545	1.239	12.24	7..89
$d_6 - d_8 + t_5 + d_5 - d_7$	0.268	0.000484	0.533	7.27	18.95
t_3	0.04	0.004545	3.088	10.20	16.39

Table 7 The detail constituent elements of deviation $(d_{10} - d_9 + t_1)$

Deviations	Fit	Tolerance	Deviation values		Contribution on f_Σ (%)	Contribution on f_a (%)
d_{10}-$d_9 + t_1$	$\frac{H7}{js6}$	d_{10}	(+ 0.07,0)	0.092	51.60	38.45
		d_9	± 0.022			
	–	t_1	0.02		11.22	8.36

Table 8 Values of max (f_Σ) and max (f_a) in different fit tolerances conditions (mm)

Tolerance	I	II	III	IV
Fit t^1_{7d} t^1_{6d}	H7/js6 $\frac{(+0.07,0)}{(\pm 0.022)}$	H6/js6 $\frac{(+0.044,0)}{(\pm 0.022)}$	H7/js5 $\frac{(+0.07,0)}{(\pm 0.016)}$	H6/js5 $\frac{(+0.044,0)}{(\pm 0.016)}$
$(d_{10} - d_9 + t1)$	0.112	0.086	0.106	0.08
Max (f_Σ)	0.00178	0.00152	0.00172	0.00146
Max (f_a)	0.754	0.672	0.735	0.653

4.4 Tolerance Optimization

When handling with the tolerance optimization and allocation, the following basic criteria should be considered: (1) product performance requirements; (2) manufacturing process capability; (3) the fit type not changed. In this application case, the main bottleneck fit tolerance of d_{10} and d_9 should be optimized to improve assembly precision. Based on the fit tolerance standard applied in the industrial company, the fit tolerance H7/js6 of d_{10} 560 (+ 0.07, 0) and d_9 560 (± 0.022) can be replaced by tolerance fit H6/js6, H7/js5 or H6/js5. Thus, four conditions can be obtained and condition I is the original tolerances.

Table 8 gives out the fit tolerances and corresponding values of f_Σ and f_a in four conditions. It can be found that condition IV most effectively satisfies the functional requirements. As a result, by enhancing the fit tolerances $(H7/js6)d_{10}(+0.07, 0)/d_9(\pm 0.022)$ to $(H6/js5)$ $d_{10}(+0.044, 0)/d_9(\pm 0.016)$, the final assembly quality can satisfy the functional requirements.

5 Conclusion

This paper presents a tolerance synthesis methodology for the spiral bevel gearbox, where dimensional and geometric tolerances of parts are considered and described by the screw theory. In the process of the assembly deviation accumulation and propagation, two main problems are devoted: the first one is to address the parallel chains produced by the fit of shaft and holes and the second one is to take into account the leverage effect of geometrical structures on assembly deviation. For the first problem, the deviations of contact points are calculated when the shaft is fitted in the holes; then the obtained deviations are utilized to compute the torsor parameters of the shaft based on screw theory. For the second problem, the leverage effect of geometrical structure of parts for the assembly requirement is remarkable and cannot be neglected; then the unified Jacobian-Torsor model is used to solve the deviation accumulation problem in the whole transmission system.

By means of this method, the final calculated results can directly reflect the relationship between the parts deviations and the assembly requirements by calculating the contributions. Meanwhile the sensitivity coefficients which are related with leverage effect of geometrical structures can be directly obtained as the polynomial coefficients of the function relation between the parts deviations and functional requirements. The obtained explicit function formula between the assembly requirements and tolerances of parts can be convenient and effective for the tolerance optimization of mechanism. Moreover the final analysis results are more accurate. The industrial application of the all-direction propeller's spiral bevel gearbox is given to illustrate the proposed method.

Acknowledgements This research was founded by the National Natural Science Foundation of China (Grant No.52205575), the National Key Research and Development Program of China (Grant No. 2023YFE0203200), the Key Research and Development Program of Hubei Province (Grant No. 2023BCB085) and the Dawning Program of Wuhan Municipal Bureau of Science and Technology (Grant No. 2023020201020326), all supports are gratefully acknowledged.

References

1. Litvin FL, Fuentes A, Hayasaka K (2006) Design, manufacture, stress analysis, and experimental tests of low-noise high endurance spiral bevel gears. Mech Mach Theory 41(1):83–118
2. Simon V (2007) Computer simulation of tooth contact analysis of mismatched spiral bevel gears. Mech Mach Theory 42(3):365–381
3. Simon V (2008) Influence of tooth errors and misalignments on tooth contact in spiral bevel gears. Mech Mach Theory 43(10):1253–1267
4. Davidson JK, Mujezinovic A, Shah JJ (2002) A new mathematical model for geometric tolerances as applied to round faces. J Mech Des 124(4):609–622
5. Desrochers A, Rivière A (1997) A matrix approach to the representation of tolerance zones and clearances. Int J Adv Manuf Technol 13(9):630–636
6. Desrochers A, Ghie W, Laperriere L (2003) Application of a unified jacobian—torsor model for tolerance analysis. J Comput Inf Sci Eng 3(1):2–14
7. Chase KW, Gao J, Magleby SP (1995) General 2-D tolerance analysis of mechanical assemblies with small kinematic adjustments. J Des Manuf 5:263–274
8. Gao J, Chase KW, Magleby SP (1998) Generalized 3-D tolerance analysis of mechanical assemblies with small kinematic adjustments. IIE Trans 30(4):367–377
9. Li H, Zhu H, Li P, He F (2014) Tolerance analysis of mechanical assemblies based on small displacement torsor and deviation propagation theories. Int J Adv Manuf Technol 72(1–4):89–99
10. Guo J, Li B, Liu Z, Hong J, Wu X (2016) Integration of geometric variation and part deformation into variation propagation of 3-D assemblies. Int J Prod Res 54(19):5708–5721
11. Corrado A, Polini W, Moroni G, Petrò S (2016) 3D Tolerance analysis with manufacturing signature and operating conditions. Proc CIRP 43:130–135
12. Ramos Barbero B, Pérez Azcona J, González Pérez J (2015) A tolerance analysis and optimization methodology. The combined use of 3D CAT, a dimensional hierarchization matrix and an optimization algorithm. Int J Adv Manuf Technol 81(1–4):371–385
13. Jin S, Chen H, Li Z, Lai X (2015) A small displacement torsor model for 3D tolerance analysis of conical structures. Proc Inst Mech Eng C J Mech Eng Sci 229(14):2514–2523

14. Asante JN (2013) A constraint-based tolerance analysis in a multi-operation single setup and multi-operation multi-setup part–fixture assembly. Int J Adv Manuf Technol 68(5–8):1001–1014
15. Kyung M-H, Sacks E (2003) Nonlinear kinematic tolerance analysis of planar mechanical systems. Comput Aided Des 35(10):901–911
16. Sacks E, Joskowicz L (1998) Parametric kinematic tolerance analysis of general planar systems. Comput Aided Des 30(9):707–714
17. Jin S, Ding S, Li Z, Yang F, Ma X (2018) Point-based solution using Jacobian-Torsor theory into partial parallel chains for revolving components assembly. J Manuf Syst 46:46–58
18. Bruyère J, Dantan J-Y, Bigot R, Martin P (2007) Statistical tolerance analysis of bevel gear by tooth contact analysis and Monte Carlo simulation. Mech Mach Theory 42(10):1326–1351
19. Chaosheng S, Shuaitao Z, Siyuan L, Cheng X, Caichao Z (2024) Influence of assembly errors on meshing characteristics of spiral bevel gear and compensation approach of misalignment. Proc Inst Mech Eng C J Mech Eng Sci 238(10):4743–4756
20. Chen H, Jin S, Li Z, Lai X (2015) A solution of partial parallel connections for the unified Jacobian-Torsor model. Mech Mach Theory 91:39–49
21. Anselmetti B (2010) Part optimization and tolerances synthesis. Int J Adv Manuf Technol 48(9–12):1221–1237
22. Anselmetti B (2008) Cotation fonctionnelle tridimensionnelle et statistique. Hermes Sciences, Paris
23. Chavanne R, Anselmetti B (2012) Functional tolerancing: virtual material condition on complex junctions. Comput Ind 63(3):210–221
24. Zeng W, Rao Y, Wang P, Yi W (2017) A solution of worst-case tolerance analysis for partial parallel chains based on the Unified Jacobian-Torsor model. Precis Eng 47:276–291
25. Asante JN (2009) A small displacement torsor model for tolerance analysis in a workpiece-fixture assembly. Proc Inst Mech Eng Part B J Eng Manuf 223(8):1005–1020

Open Access This chapter is licensed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License (http://creativecommons.org/licenses/by-nc-nd/4.0/), which permits any noncommercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if you modified the licensed material. You do not have permission under this license to share adapted material derived from this chapter or parts of it.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.



Multi-body Dynamics Analysis of All-Terrain Crane Steering Member System



Ling Wang, Xin Wang, Shuai Song, Xifeng Zhang, and Chenghao Wang

Abstract In design stage, the ultimate load of crane steering rod system, as input of finite element analysis to check the structural strength, is almost obtained from spot turn case, but there exists members bending or fracture. Taking a certain type of all-terrain crane as an example, the crane dynamics model under different levels of road surface is established based on MATLAB and ADAMS. The experimental data are fitted to obtain a simplified model of the "magic formula", which was used as the input condition of the dynamic model to measure the dynamic load of the steering tie rods under two kinds of road surfaces. The load curves show that the ultimate load of the steering tie rods under driving steering condition is larger than that under in-situ steering condition, therefore, the ultimate load under driving steering condition should be used as the input for steering rod system strength checking.

Keywords Crane steering · Tie rods · Ultimate load · Magic formulae

1 Introduction

In recent years, the domestic demand for large all-terrain cranes is growing. As one of the key parts of the steering system of all-terrain cranes, the strength of the tie rod has an important impact on driving safety of the vehicle. Affected by self-weight and road conditions, all-terrain cranes are traveling under poor working conditions, the phenomenon of tie rod bending or even fracture occurs from time to time.

At present, universities and research institutes at home and abroad have carried out certain research on the load of crane steering tie rods, but mainly focused on the in-situ steering condition [1–3] and the research on the load of steering tie rod under driving state is still to be in-depth.

L. Wang · X. Wang (✉) · C. Wang
Dalian University of Technology, Dalian, China
e-mail: wangxbd21@163.com

S. Song · X. Zhang
Taiyuan Heavy Industry Co., Ltd., Taiyuan, China

© The Author(s) 2026

H. Saman et al. (eds.), *The 9th International Conference on Advances in Construction Machinery and Vehicle Engineering*, Lecture Notes in Mechanical Engineering,
https://doi.org/10.1007/978-981-96-5527-4_33

Therefore, the paper is based on MATLAB/Simulink to establish a pavement spectral model, which is imported into ADAMS to construct a virtual pavement, and establish a crane virtual prototype model to simulate the steering of the vehicle in the driving process. After that, the experimental data are fitted to obtain a set of equations that can express the change of steering torque with the turning angle, and this is used as the dynamic input to obtain the dynamic load of the steering tie rods in the driving process.

2 Steering Torque

2.1 Components of Steering Torque

The tire in-situ steering resistance moment consists mainly of the following components.

(1) Friction moment at the main pin

The tyre generates a large frictional resistance moment during rotation around the main pin. The frictional moment at the main pin is mainly determined by the size of the tyre load.

(2) The return moment caused by the inner inclination angle of the main pin

Under the action of the tire's own load, When the wheels turn, it will generate a return moment M_z around the main pin to return to the position automatically.

(3) Frictional resistance moment between the tire and the ground

When a vehicle is steered in place, the tyres first torsion occurs, followed by frictional resistance moments M_s as they slip against the road surface. From the above analysis, we can assume that the static in-situ steering resistance moment M of the tire is calculated as,

$$M = M_f + M_z + M_s \tag{1}$$

2.2 Calculation of Maximum Steering Resistance Moment

At present, empirical and semi-empirical formulas are mainly used in calculating the in-situ steering torque of tyres [4], which cannot represent the dynamic change of the steering tie rod load. Therefore, the empirical and semi-empirical formulas have great limitations in practical application. Currently the most commonly used equation for calculating the maximum in-situ steering resistance moment is Eq. (2).

Table 1 Michelin 14.00R25 tyre in-situ steering maximum moment of resistance

Payloads (10^4 N)	3	4	5	6	7
Maximum in-situ steering torque (10^6 N mm)	1.647	2.455	3.367	4.324	5.267

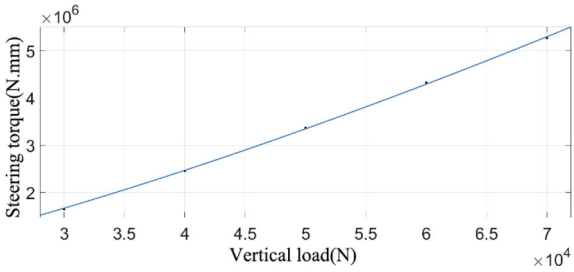


Fig. 1 Fitted plot of maximum steering torque for in-situ steering

$$M = \frac{\mu}{3}\sqrt{\frac{F_z^3}{p}} \tag{2}$$

$$M = a.\mu\frac{F_z^b}{P^c} \tag{3}$$

In order to improve the calculation accuracy, the general form of semi-empirical formula is used to fit the maximum steering resisting moment test data (Table 1).

The general form of the semi-empirical formula is Eq. (3), where a, b, c are the parameter to be fitted. The fitting of the experimental data in MATLAB is shown in Fig. 1.

From the above figure, it can be seen that the general form of the semi-empirical formula can be used to fit the maximum steering resisting moment test data better, and the values of the fitted parameters are shown in Table 3.

2.3 Fitting the 'Magic Formula'

The 'magic formula' tyre model is proposed by Professor Pacejka [5]. Based on the tyre in-situ steering moment of resistance test data [6, 7], Wang Yunchao of Jimei University simplified the above equation, and the simplified tyre in-situ steering model magic formula is:

$$M = M_{\max}\sin[C\arctan(E\alpha + B)] \tag{4}$$

Substituting Eq. (3) into the above equation as,

$$M(F_z, \alpha, P, \mu) = a.\mu\frac{F_z^b}{P^c}\sin[C\arctan(E\alpha + B)] \tag{5}$$

Table 2 In-situ steering torque variation with cornering angle test data (6 t tyre load)

Tyre turning angle (degree)	0	2.5	5	7.5	10	12.5	15	17.5	20	22.5	25
In-situ steering torque (10^6 N mm)	0.4	1.75	2.8	3.4	3.82	4.2	4.45	4.63	4.77	4.87	5

Table 3 Fitting parameters

Fit parameter	*a*	b	c	B	C	E
Numerical value	0.4075	0.3285	1.545	0.1235	0.5005	0.1369

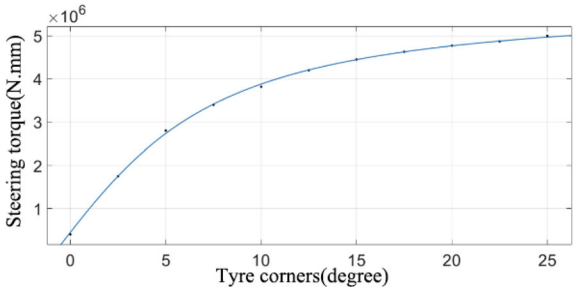


Fig. 2 Steering resistance moment fitted to tyre angle

where B, C and E are the parameters to be fitted to extract the in-situ steering resistance moment of the tyres of the heavy multi-axle steering vehicle in a paper, the test data of the variation of the in-situ steering resistance moment with the angle of rotation for a vertical tyre load of 6 t are shown in Table 2.

The test data are fitted according to Eq. (5) and the fitted curve is shown in Fig. 2, from which it can be seen that Eq. (5) can fit the test data better. The fitting parameters in this equation are shown in Table 3.

3 In-Situ Steering Condition Tie Rod Force Analysis

In order to study the force on the steering tie rods of all-terrain cranes in the steering process, parametric modelling is used to establish a rigid body model of the crane steering system dynamics in ADAMS View as shown in Fig. 3. Steering tie rod force and the size of the corner is positively correlated, due to the 1st axle corner is the largest, so the 1st axle steering tie rod force is the largest, this paper will follow up with the 1st axle as an example of the steering tie rod force analysis.

The resisting moment is applied to the left and right tyres of the axle to simulate the resistance of the tyres during rotation, and the resisting moment function is given in Eq. (6).

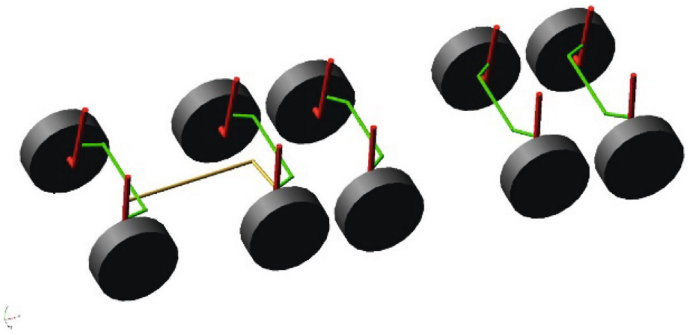


Fig. 3 Steering system dynamics model diagram

$$\begin{cases} M(\alpha_z) = a.\mu \dfrac{F_z^b}{P^c} \sin[C \arctan(E\alpha_z + B)] \\ M(\alpha_y) = a.\mu \dfrac{F_z^b}{P^c} \sin[C \arctan(E\alpha_y + B)] \end{cases} \tag{6}$$

$$\begin{cases} M_{d1} = (F_c + F_c\prime)m_1 \\ F_c = \dfrac{\pi D_c^2}{4} P_1 \\ F_c\prime = \dfrac{\pi (D_c^2 - d_c^2)}{4} P_1 \end{cases} \tag{7}$$

When the crane is steering, the driving torque needs to be greater than the resistance torque. The left and right side cylinders are rod cavity and rodless cavity, due to the booster cylinder work oil pressure is equal but the force area is different, which leads to different driving torque on both sides, the proportionality between the two can be calculated by Eq. (7).

In this paper, take the left turn as an example, according to the actual size, the ratio is taken as 39/64. At this time, the left side is a rodless cavity, the right side is a rod cavity, then the right side of the driving torque function is,

$$M_{dy} = \frac{39}{103}[M(\alpha_z) + M(\alpha_y)] \tag{8}$$

3.1 In-Situ Road Steering Conditions

The vertical load of a single tyre under road steering condition is 6.5 t, the tyre air pressure is 0.9 MPa, and the COF is 0.8, then the left and right side tyre resisting moment functions are,

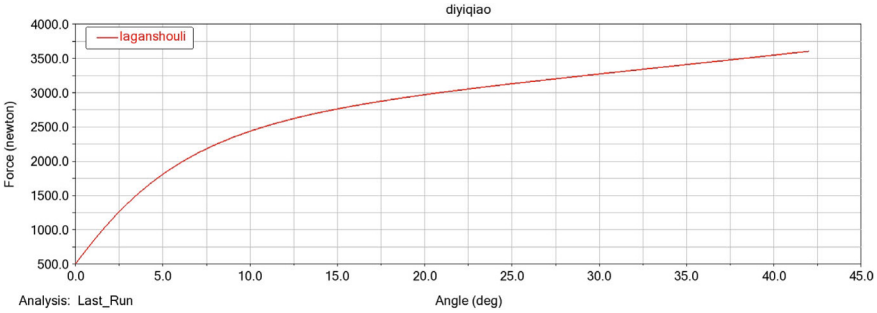


Fig. 4 Steering tie rod force diagram

$$\begin{cases} M(\alpha_z) = 0.4075 \times 0.8 \times \dfrac{63700^{0.3285}}{0.9^{1.545}} \\ \qquad \sin[0.5005 \times \arctan(0.1369 \times \alpha_z + 0.1235)] \\ M(\alpha_y) = 0.4075 \times 0.8 \times \dfrac{63700^{0.3285}}{0.9^{1.545}} \\ \qquad \sin[0.5005 \times \arctan(0.1369 \times \alpha_y + 0.1235)] \end{cases} \tag{9}$$

The right cylinder drive force function is Eq. (8). The rotary drive STEP (time, 0, 0d, 5.42d) is applied to the left suspension cylinder [8]. The force on the steering crossbar is shown in Fig. 4. (Horizontal coordinate is left tyre angle).

3.2 Off-Road In-Situ Steering Conditions

Under the off-road steering condition, the vertical load of a single tyre is 20 t, the tyre air pressure is 0.9 MPa, and the COF is 0.8. Similarly, by adding the corresponding drive and moment of resistance functions, the force on the steering crossbar is shown in Fig. 5.

3.3 Tie Rod Force Analysis

As can be seen from Fig. 5, the maximum force on the steering tie rod under the off-road in-situ steering condition is close to 20,000 N. Steering tie rod in 0°–20° corner of the tension increases faster, in 20°–42° corner of the driving moment and steering resistance moment changes are small, due to the gradual decrease of the force arm, the pulling force of the steering tie rod still increases slowly.

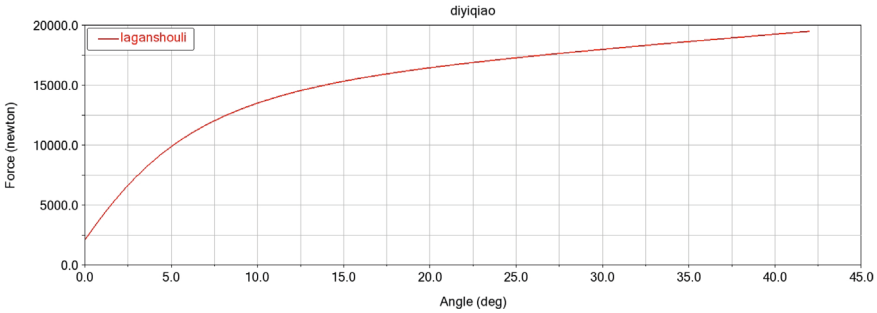


Fig. 5 Steering tie rod force diagram

4 Tie Rod Force Analysis Under Driving and Steering Conditions

4.1 Pavement Spectrum Establishment

In this paper, MATLAB/Simulink as a platform to establish a white noise random time domain pavement spectral model [9]. Selecting the class B road as the road condition and the class E road as the non-road condition. The parameters corresponding to the two road conditions are shown in Table 4, from which the road surface spectra under the two conditions are obtained as shown in Fig. 6, and the distance between the two roads is set to be equal, As the Class B road travelling speed is four times as fast as the Class E, the travelling time is 1/4 of the Class E.

Table 4 Vehicle speeds and pavement gains for on-road and off-road conditions

Road classification	Speed (km/h)	K1	K2
B	60	0.0205	1.047
E	15	0.0824	0.264

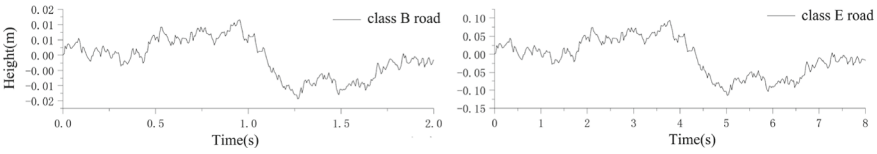


Fig. 6 Pavement spectra for on-road and off-road conditions

4.2 Tyre Force Analysis

Import the pavement spectrum file into ADAMS to create SPLINE. Set up a moving vice between the pavement and the earth of the system, and add the driving function AKISPL (time, 0, SPLINE, 0) to this moving vice to simulate the bumpy pavement.

Set the contact force between the tyre and the road surface, set the solver as a 'collision' model, and open the Coulomb friction model. The contact parameters between tyre and road surface are shown in Table 5.

(1) Tyre contact force for road conditions

Under road conditions, a single tyre carries 6.5 t, a mass of 6.5 t is applied to the tyre and a rotating sub and rotary drive are applied to make the vehicle dynamics model travelling on the road at 60 km/h, and the contact force between the tyre and the road surface in the direction perpendicular to the road surface is measured as shown in Fig. 7.

(2) Off-road tyre contact force

Off-road conditions under a single axle load 40 t, at this time a single tyre load 20 t, the same in the tyre to apply 20 t mass and apply a rotary drive so that the model can be in the off-road conditions at a speed of 15 km/h, measure the tyre and the ground in the vertical direction of the road surface contact force as shown in Fig. 8.

Table 5 Tyre-road contact parameters

Parametric	Stiffness	Force exponent	Damping	Penetration depth	SCOF	DCOF
Numbers	1E5	2.2	1.5E+5	1E−3	0.8	0.5

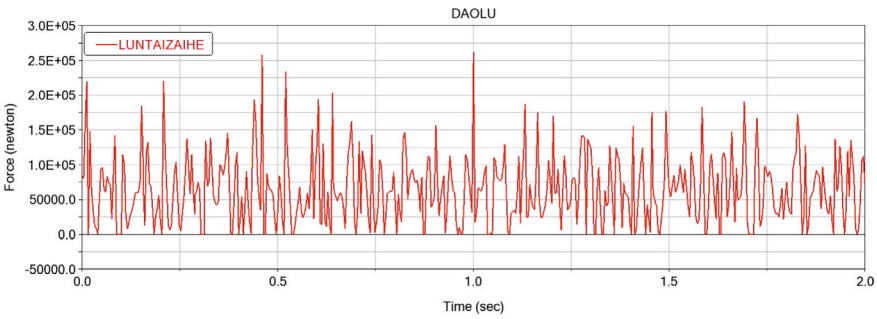


Fig. 7 Force diagram for road conditions

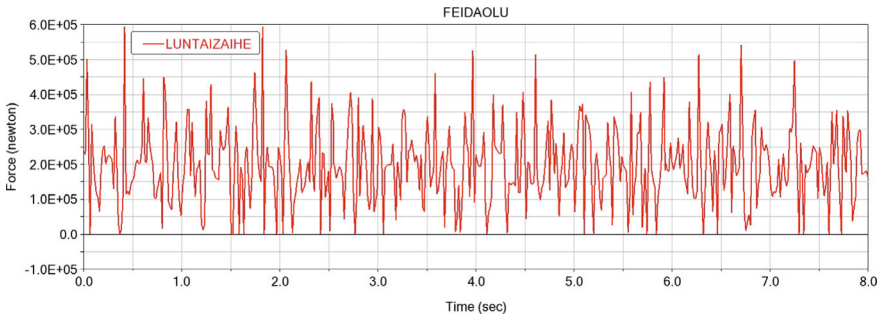


Fig. 8 Off-road force diagram

4.3 Steering Tie Rod Force Analysis

Tested and theoretically verified [10], when the vehicle is steering in the driving state, the friction resistance moment between the tyre and the ground is generally ½–1/3 times of that of the in-situ steering, in order to prevent the calculated tie-rod load from being small, according to the research results of Zhang Xiaojiang et al., this paper chooses the friction resistance moment in driving to be 1/2 times of that of the in-situ steering.

(1) Steering tie rod force under road conditions

The load on a single tyre under road conditions is imported into the vehicle dynamics model for calculation. Then the left and right side resistance moment functions are,

$$\begin{cases} M(\alpha_z) = 0.5 \times 0.4075 \times 0.8 \times \dfrac{ABS(AKISPL(time, 0, SPLINE, 0))^{0.3285}}{0.9^{1.545}} \\ \qquad \sin[0.5005 \times \arctan(0.1369 \times \alpha_z + 0.1235)] \\ M(\alpha_y) = 0.5 \times 0.4075 \times 0.8 \times \dfrac{ABS(AKISPL(time, 0, SPLINE, 0))^{0.3285}}{0.9^{1.545}} \\ \qquad \sin[0.5005 \times \arctan(0.1369 \times \alpha_y + 0.1235)] \end{cases} \tag{10}$$

The right cylinder drive force function is Eq. (8). This gives the dynamic steering resistance moment during driving. The steering tie rod forces were measured after applying rotational actuation to the left suspension cylinder as shown in Fig. 9.

(2) Off-road steering tie rod force

The single tyre force curve under off-road conditions was imported into the kinematic model for tie rod force analysis, and the corresponding driving and resisting moment functions were applied, and the measured steering tie rod forces are shown in Fig. 10.

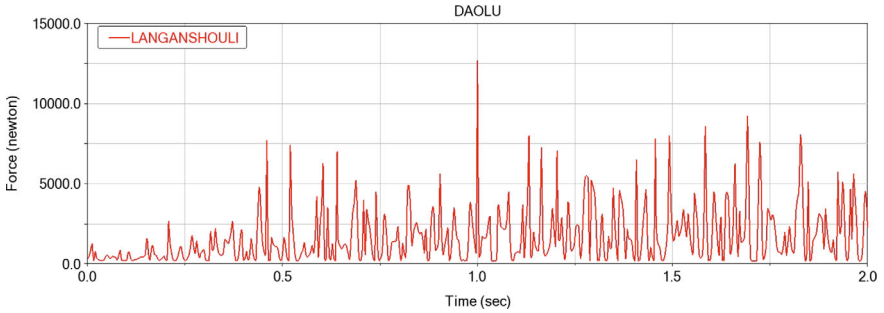


Fig. 9 Force diagram for road conditions

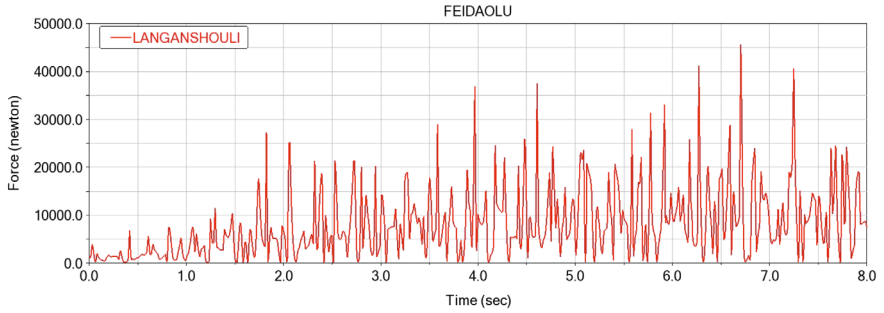


Fig. 10 Off-road force diagram

(3) Tie rod force analysis

From the above two calculations, it can be seen that the steering tie-rod load under both conditions shows an increasing trend. The maximum tension force on the steering tie rods in the steering process is 45,000 N, which is larger than that in the in-situ steering process, which is 20,000 N.

5 Conclusion

In order to obtain the force curve of the steering tie rods of the all-terrain crane when driving and steering, this paper creates B and E level road spectra based on the MATLAB/simlink platform, and then imports them into ADAMS to set up the contact between the tyre and the road surface to obtain the tyre load of the all-terrain crane when driving.

Secondly, the 'magic formula' is used to fit the experimental data, and the steering torque expression formula is obtained, and the vertical load of the tyre and the change of the tyre angle are introduced into this formula in ADAMS to obtain the left and

right tyre torque curves when driving and steering, and then the dynamic analysis is used to obtain the dynamic load curves of the steering tie rods when driving and steering. The dynamic load curve of the steering tie rod during steering is then obtained through dynamic analysis. The result obtained is as follow:

The ultimate load of the steering tie rod is related to factors such as tyre angle and vertical load, while the road surface grade and driving speed will have a greater impact on the vertical load of the tyre. Therefore, when determining the ultimate load, it is necessary to consider both in-situ and driving steering conditions, and select the one with larger ultimate load as the finite element input of the steering tie rod for static strength analysis.

Acknowledgements This research was supported by Taiyuan key core technology research unveiled list of marshal projects.

References

1. Wenjia W (2019) Optimised design of steering mechanism of a six-axle all-terrain crane. Hunan University, Changsha
2. Xie M, Xin Q, Long J et al (2023) Fatigue life analysis of fillet weld with root gap in bogie tie rod seat under road traffic load. 2:36–40
3. Topac et al (2020) Analysis of the failure modes and design improvement of an eccentrically loaded connecting rod for a double front axle steering linkage prototype. Eng Fail Anal 1:102–106
4. Gao J, Qi X (2023) Optimal design of steering trapezoid considering multiple performances of the whole vehicle. 18:114–126
5. Pacejka HB (2012) Tire and vehicle dynamics, 3rd edn. Elsevier Ltd., Great Britain
6. Du H, He Y, Yang Y et al (2021) Study on tie rod force characteristics in electro-hydraulic power steering system for heavy vehicle 235(2–3):564–579
7. Luca F, Manzoni S, Cigada A et al (2022) Automatic detection of real damage in operating tie-rods. Sensors 22(04):124–132
8. Xie M, Xin Q, Long J et al (2023) Fatigue life analysis of fillet weld with root gap in bogie tie rod seat under road traffic load. Eng Fail Anal 5:175–180
9. Verma P, Vaishnav D, Nandan A et al (2024) Study on failure of a tie rod for two-wheeler automotive gear mechanism. Can Metall Quart 12:64–72
10. Fenglin Y, Wenjun M, Jie Z et al (2020) The relationship between eccentric structure and super-lift device of all-terrain crane based on the overall stability. J Mechan Sci Technol 17(6–8):118–131

Open Access This chapter is licensed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License (http://creativecommons.org/licenses/by-nc-nd/4.0/), which permits any noncommercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if you modified the licensed material. You do not have permission under this license to share adapted material derived from this chapter or parts of it.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.



Estimation of Coefficient of Road Adhesion Based on Stacked Generalisation and Bayesian Optimisation



Guan Zhou and Chenxi Gao

Abstract Coefficient of road adhesion is an important parameter affecting vehicle handling stability and safety. This paper designs a method for estimation of coefficient of road adhesion based on stacked generalisation and bayesian optimisation. Firstly, by changing the road adhesion coefficient and vehicle speed, the estimation simulation test is designed. Secondly, the long and short-term memory network and convolutional neural network model are selected as the base learner to initially estimate the coefficient of road adhesion, and the preliminary estimation results are later utilized to train the meta-learner to reduce the estimation error. Finally, optimizing meta-learner hyperparameter models using Bayesian algorithms. The validity of the model proposed in this paper is verified by simulation tests.

Keywords Coefficient of road adhesion · Stacked generalization · Convolutional neural networks · Hyperparameter optimization

1 Introduction

Vehicles are accelerating in the direction [1] of electrification, intelligence, connectivity and sharing, and active safety systems are becoming more and more popular. The coefficient of road adhesion μ [2] is an important factor in determining the effectiveness of active safety systems. When the vehicle is traveling, the coefficient is affected by many factors such as humidity, temperature, and material of the road surface. Moreover, it is currently impossible to accurately determine μ using onboard sensors. For automotive active safety systems, the rapid and accurate estimation of μ t is extremely critical to the effectiveness of the system. Accurate estimation [3] can effectively reduce the occurrence of tailgating, skidding and other accidents, and improve the stability of the vehicle and driving safety.

G. Zhou (✉) · C. Gao
College of Energy and Power Engineering, Nanjing University of Aeronautics and Astronautics, Nanjing, China
e-mail: nuaagcx@nuaa.edu.cn

© The Author(s) 2026

H. Saman et al. (eds.), *The 9th International Conference on Advances in Construction Machinery and Vehicle Engineering*, Lecture Notes in Mechanical Engineering,
https://doi.org/10.1007/978-981-96-5527-4_34

A lot of scholars from universities and research institutes have conducted many approaches on estimating μ. Mainly divided into [4] experimental method and model method categories. The experimental method [5] typically utilizes sensors to obtain ratings related to μ. Dogan [6] reduced the estimation cost in the estimation process by using sound signals to estimate μ and classified them into four categories. However, the cost of such methods [7] still suffers from the problem of high estimated costs, and there are certain requirements for the spatial arrangement of the vehicle, and the operating status of the sensors [8] and other components is not stable in harsh environments.

The model-based estimation approach [9] requires a simplified vehicle dynamics model that contains two main parts of the model, the vehicle and the tires. The estimation cost can be significantly reduced compared to the first type of estimation methods. Gustafsson [10] vachieves the estimation by fitting the slopes of the curves and obtaining a function curve related to μ. Other scholars consider the influence of the tire model and combine the state observer to achieve the estimation. Wang et al. [11] choose the feed-forward neural network to achieve the online estimation of multi-module.

The estimation method of coefficient of road adhesion proposed in this paper takes into account the estimation error of the model performance and the influence of hyperparameters on the model performance, integrates the long and short-term memory network model and the convolutional neural network model in the estimation model, which in turn reduces the estimation error that exists in the estimation process of a single model, and then selects the meta-learners hyperparameters through the Bayesian optimization algorithm, which reduces the workload and randomness of the manual selection and improves the estimation accuracy of the model. The estimation accuracy of the model is improved.

2 Establishment of Dataset

Since this estimation model is an offline estimation model, it is necessary to establish a dataset to train the estimator, and the dataset should be the coefficient of road adhesion and the parameters related to it, i.e., the dynamics parameters of the vehicle driving process. In this paper, the longitudinal force, lateral force, vertical force, the longitudinal slip rate and the tire side deflection angle of each wheel of the vehicle, the vehicle speed, and acceleration of the vehicle are selected to be collected.

The reasonable establishment of the data set has a great influence on the estimation effect of the regression model, and the actual process is affected by the road conditions and the driver's operation, and the driving status of the vehicle is also very different, so the data set should be established to cover as many vehicle driving conditions as possible. Therefore, in this paper, the simulation is collected by changing the driving trajectory, speed, and road conditions of the vehicle.

3 Estimator of Stacked Generalization Model Based on Bayesian Optimization

3.1 Stacked Generalization Estimation Module

Long Short—Term Memory (LSTM) model is a special Recurrent Neural Network (RNN) architecture. It is improved on the basis of traditional RNN, and can effectively solve the gradient vanishing and gradient explosion problems that tend to occur in RNN when dealing with long sequence data, so as to better capture the long-term dependencies in sequence data. The basic unit structure is shown in Fig. 1.

The output vector h_t of the LSTM network model at time t is defined as:

$$h_t = o_t * \tanh(c_t) \tag{1}$$

$$\begin{cases} o_t = \sigma(W_o x_t + U_o h_{t-1} + b_o) \\ c_t = f_t c_{t-1} + i_t c_i \\ f_t = \sigma\left(W_f x_t + U_f h_{t-1} + b_f\right) \\ i_t = \sigma(W_i x_t + U_i h_{t-1} + b_i) \\ c_i = \tanh(W_c x_t + U_c h_{t-1}) \end{cases} \tag{2}$$

where, i_t, o_t, f_t are time vector functions respectively; σ is the activation function; x_t is the input vector at time t; c_t is the cell state at time t; W_o, W_f, W_i, W_c are weight matrixs; b_o, b_f, b_i are offsets of each gate.

Convolutional Neural Network (CNN) is a deep learning model widely used in image recognition, video analysis, natural language processing and other fields. It is capable of automatically extracting features from data through structures such as convolutional layers, pooling layers and fully connected layers, and has powerful

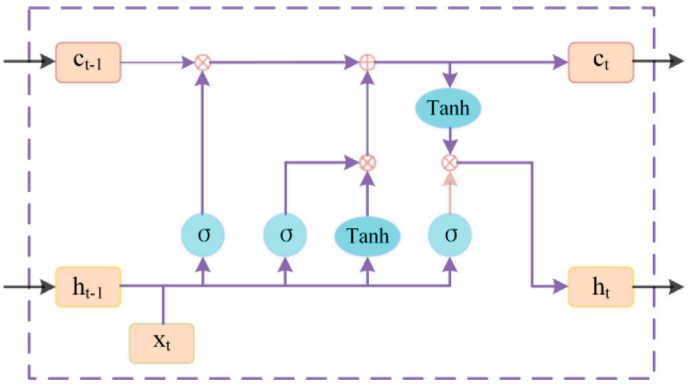


Fig. 1 Basic unit structure

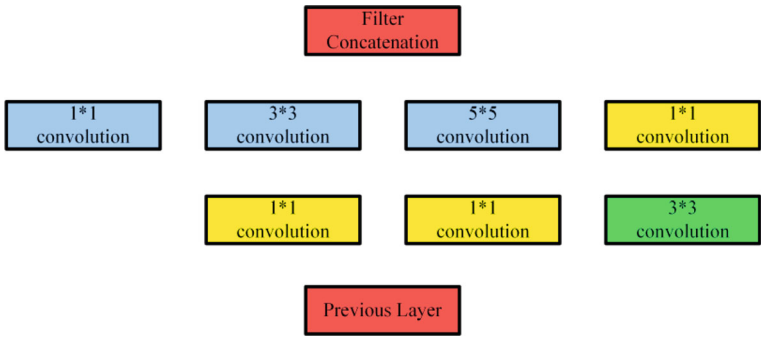


Fig. 2 Principle structure of a CNN network model

feature learning and pattern recognition capabilities. The principle structure of a CNN network model is illustrated in Fig. 2.

3.2 μ Estimator Based on Bayesian Optimized Stacked Generalization Model

When building and training the estimator, the function and estimator architecture chosen can have unpredictable effects on the training results. Therefore, it is extremely critical to carefully select and finalize the hyperparameter combinations. In stacked generalized estimation models, the network hyperparameters of the base learner are usually roughly adjusted with the aim of avoiding overfitting of the model, while those of the meta-learner are finely adjusted. In view of this, this paper proposes a Bayesian algorithm to optimize the hyperparameter combinations of the meta-learners in the model, which can target the best hyperparameter combinations within the given hyperparameter range. This not only improves the efficiency of manual adjustment, but also enhances the estimation accuracy of the model. Structure diagram of the optimized estimation model is shown in Fig. 3.

Bayesian optimization is an effective method for hyperparameter tuning, which guides the search process by constructing an agent model of the objective function to find the optimal hyperparameter combination. Its optimization process is defining the objective function, defining the search space, initialization, constructing the agent model, replacing the hyperparameter combinations, repeating the above steps and outputting the final hyperparameter combinations.

In this study, we meticulously screened out the hyperparameters that have a key impact on the performance of the meta-learner, including the optimizer type, the number of nodes in each layer of the neural network, and the discard rate. After determining the value interval of each hyperparameter, a hyperparameter domain is constructed, and the specific parameter ranges are detailed in Table 1.

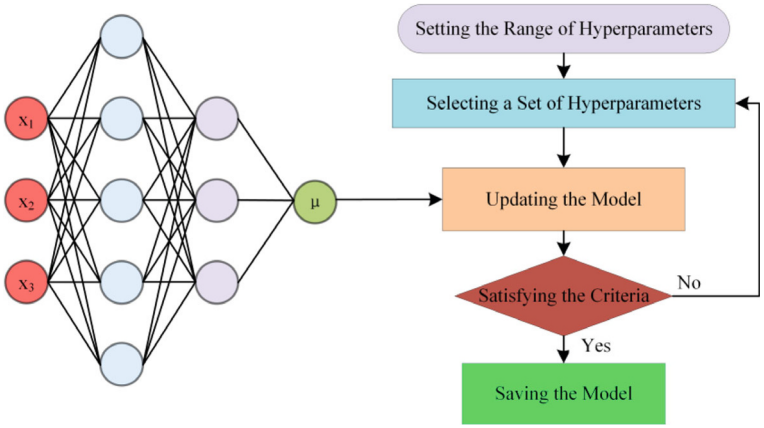


Fig. 3 Structure diagram of the optimized estimation model

Table 1 Hyperparameter domain space and optimal hyperparameter combination

	Type	Hyperparameter	Range	Combination
1	Trainning	Optimizer	SGD, Rmsprop	SGD
2	Trainning	Leraning rate	[0.0002, 0.002]	0.0011
3	Trainning	Batch size	64, 128, 256	256
4	Structual	Dense1	64, 128, 256	64
5	Structual	Dense2	32, 64, 128	32
6	Structual	Dense3	16, 32, 64	16
7	Trainning	Dropout	[0.0, 0.25]	0.07

Through the optimization of the meta-learner by the Bayesian optimization algorithm, the final estimation model consists of four layers, where the number of nodes in some of the layers are 64, 32, and 16, respectively. The training volume of each batch of data is set to 256, the dropout rate is 0.07, and the learning rate of the optimizer is 0.0011. Afterwards, the data are normalized, so as to map the features of the data onto a uniform scale and to reduce the data's impact on the training of the model. The data is then normalized so as to map the features of the data, unify the scale and reduce the influence of the data on the model training process and model accuracy.

4 Simulation Verification

4.1 Verification Test of the Fixed Speed and the Fixed μ

In this experiment, the specific conditions of the fixed speed and fixed validation experiments are as follows: there are two cases of fixed speed, namely low and high speed, which are set as 30 and 100 km/h. The fixed validation experiments are also divided into two categories, which are low coefficient of road adhesion and high coefficient of road adhesion, and the corresponding values are set as 0.3 and 0.8, and the results of the simulation experiments are shown in Figs. 4, 5, 6 and 7. results obtained from this simulation experiment are shown in Figs. 4, 5, 6 and 7.

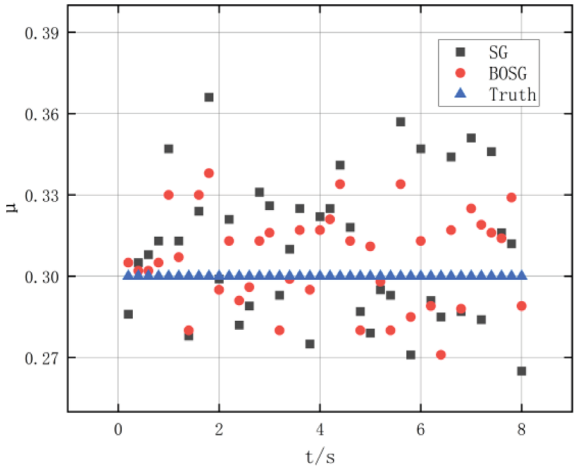


Fig. 4 Simulation results at a vehicle speed of 30 km/h and μ of 0.3

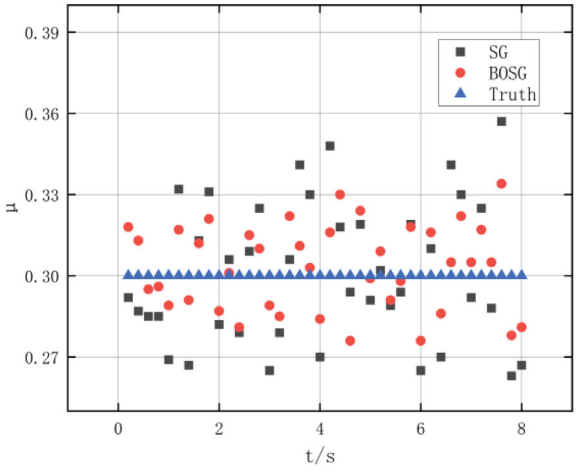


Fig. 5 Simulation results at a vehicle speed of 100 km/h and μ of 0.3

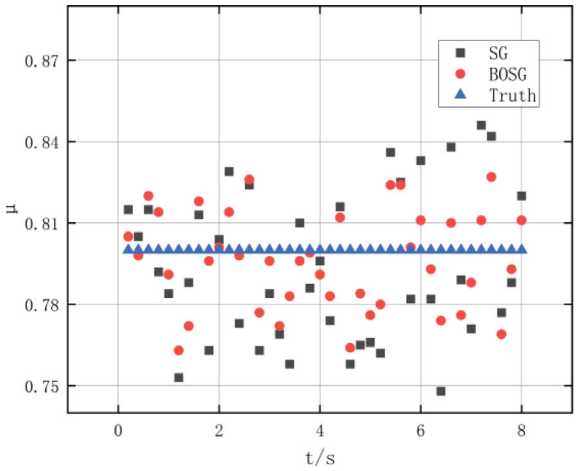


Fig. 6 Simulation results at a vehicle speed of 30 km/h and μ of 0.8

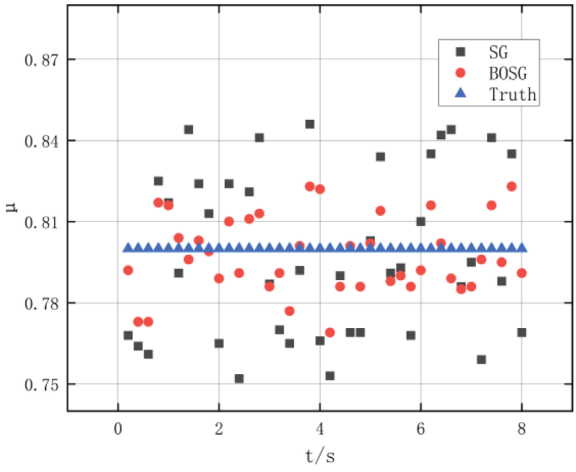


Fig. 7 Simulation results at a vehicle speed of 100 km/h and μ of 0.8

From Figs. 4 and 5, it can be seen that under the road condition of low coefficient of road adhesion of 0.3, the road adhesion coefficient estimator designed in this chapter can realize accurate estimation of road adhesion coefficient, and the error range is kept within 0.05, and accurate estimation can be realized under the driving conditions of both low and high vehicle speeds. The estimation effect of the optimized stacked generalization model is significantly lower than that of the unoptimized estimator, which verifies the effectiveness of the Bayesian optimization algorithm. As can be seen from Figs. 6 and 7, under the road conditions with high coefficient of road adhesion of 0.8, the estimator of coefficient of road adhesion can also realize the accurate estimation of coefficient of road adhesion, and the error range is still kept within 0.05. The estimation performance of the estimator is not affected regardless of high or low speed conditions, which verifies the robustness of the estimator and

proves that it can be adapted to different working conditions and maintain good estimation performance.

4.2 *Verification Test of Variable Speed and the Fixed μ*

In the validation test of the fixed coefficient of road adhesion, the conditions for variable speed are as follows: there are two cases of variable speed, one is accelerating from 30 to 100 km/h, and the other is decelerating from 100 to 30 km/h, in which the fixed value is set to be 0.5. The results of this simulation are presented in Figs. 8 and 9.

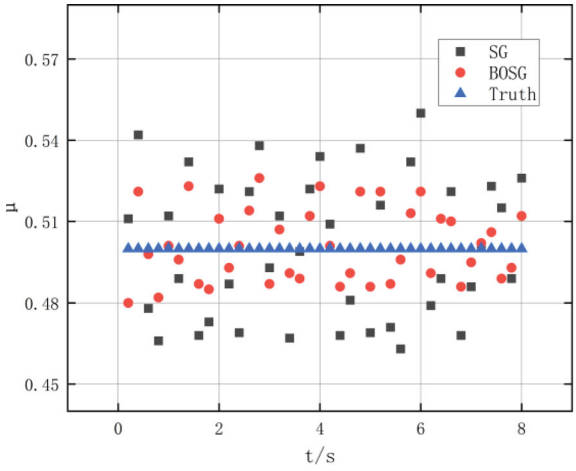


Fig. 8 Simulation results of vehicle speeds (30–100 km/h) and μ of 0.5

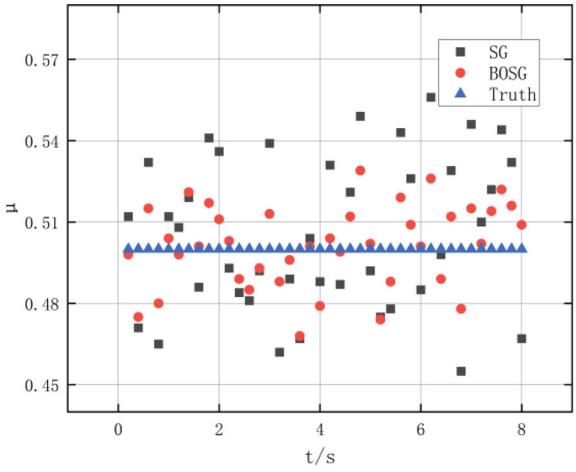


Fig. 9 Simulation results of vehicle speed (100–30 km/h) and μ of 0.5

As can be seen from Figs. 8 and 9, both models achieve a certain degree of accuracy in estimating the coefficient of road adhesion under the conditions of varying vehicle speed and fixed coefficient of road adhesion. The change in vehicle speed has little effect on the estimation performance of the models, which further validates the robustness of the models. However, the optimized superposition generalization model performs significantly better than the unoptimized model, and the fluctuation range of the estimation results in the simulation is smaller than that of the unoptimized model.

In the comparative analysis, the mean absolute error of the SG model is reduced by 22.7% on average and the mean square error is reduced by 24.1% compared to the estimation results of the two base learner models, LSTM and CNN, which confirms the feasibility of the stacked generalization model applied to the estimation. Further, the Bayesian-optimized model significantly reduces the MAE value by 30.5% and the MSE value by 45.9% compared to the SG model without fine parameter tuning, which fully proves the feasibility and high efficiency of the Bayesian optimization algorithm for the optimization of the SG model. The experimental results show that the BOSG model proposed in this paper exhibits high accuracy and excellent estimation ability under different conditions, and has strong estimation stability.

5 Conclusion

In this paper, a neural network regression model is used to estimate the coefficient of road adhesion, and for the problem of the strong or weak position of the correlation between vehicle dynamics parameters and the coefficient of road adhesion when the data set is established, the initial data set is screened for features through the F-test, so as to improve the training speed of the estimation model and reduce the risk of overfitting. To address the estimation error of a single model, two models, LSTM and CNN, are selected as the base learner, and then the base learner estimation results are input into the meta-learner for training by using stacked generalization, which improves the accuracy of the estimation model, and the hyper-parameters of the meta-learner MLP are regulated by combining with the Bayesian optimization algorithm. Through three types of simulation tests, it is verified that the designed BOSG estimation model has high estimation accuracy under the working conditions of different coefficients of road adhesion and different vehicle speeds, and the estimation model also has strong robustness.

The proposed attachment coefficient estimation method in this paper uses a model combination of CNN and LSTM, and the estimation performance of more model combinations can be tested subsequently to further improve the pavement attachment coefficient estimation accuracy.

References

1. Peng H, Chen X (2021) Active safety control of X-by-wire electric vehicles: a survey. SAE Int J Veh Dyn Stabil NVH 6:115–33
2. Guo H, Yin Z, Cao D et al (2019) A review of estimation for vehicle tire-road interactions toward automated driving. IEEE Trans Syst Man Cybern 49(1):14–30
3. Ma B, Lv C, Liu Y et al (2018) Estimation of road adhesion coefficient based on tire aligning torque distribution. J Dyn Syst Meas Control 140(5):05–10
4. Hu J, Rakheja S, Zhang Y (2020) Real-time estimation of tire–road friction coefficient based on lateral vehicle dynamics. Proc IMechE Part D: J Automob Eng 234(10–11):2444–2457
5. Xia X, Xiong L, Sun K et al (2016) Estimation of maximum road friction coefficient based on Lyapunov method. Int J Autom Technol 17(6):991–1002
6. Dogan D (2017) Road-types classification using audio signal processing and SVM method. In: 25th signal processing and communications applications conference (SIU), Antalya, Turkey, 15–18, pp 1–4
7. Song Y, Shu H, Chen X et al (2019) Direct-yaw-moment control of four-wheel-drive electrical vehicle based on lateral tyre–road forces and side slip angle observer. IET Intell Transp Syst 13(2):303–312
8. Li L, Yang K, Jia G et al (2015) Comprehensive tire–road friction coefficient estimation based on signal fusion method under complex maneuvering operations. Mech Syst Signal Process 56–57:259–276
9. Lin M, Yoon J, Kim B (2015) A study and analysis for calculating the brake-application time of AEB systems considering the gradient. Int J Control Autom 8(6):283–292
10. Gustafsson F (1998) Estimation and change detection of tire – road friction using the wheel slip. IEEE Control Syst Mag 18(4):42–49
11. Wang G, Li S, Feng G et al (2023) Road adhesion coefficient estimation by multi-sensors with LM-MMSOFNN algorithm. Adv Mech Eng 6(15):12

Open Access This chapter is licensed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License (http://creativecommons.org/licenses/by-nc-nd/4.0/), which permits any noncommercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if you modified the licensed material. You do not have permission under this license to share adapted material derived from this chapter or parts of it.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.



Bifurcation and Chaos Properties of Magnetic Gear Two-Speed Transmission System



Junliang Du, Dawei Liu, and Tingzhi Ren

Abstract In this paper, a magnetic gear two-speed transmission system is proposed. The system has the advantages of isolated vibration and overload protection. A nonlinear dynamics model including nonlinear magnetic coupling torque, torque ripple, time-varying stiffness of mechanical gears, tooth-side clearance, and integrated transmission error of magnetic gears is established. The dimensionless vibrational differential equations of the system are derived, and the bifurcation and chaos characteristics of the system are analyzed by using the bifurcation diagram and the Poincaré cross-section diagram. The results show that the system has multi-frequency narrowband bifurcation intervals accompanied by amplitude jumps. The inverse doubling period bifurcation, the inverse Hough bifurcation, which restores stability to the system, can be observed in this interval. Mechanical gears in the system are more prone to chaotic vibration than magnetic gears, but by improving the intermediate shaft damping can effectively reduce the multi-frequency jumping phenomenon of the system and avoid chaotic vibration of the system.

Keywords Magnetic-mechanical compound vibration · Bifurcation · Chaos · Two-speed transmission · Nonlinear dynamics

1 Introduction

Magnetic gears have received extensive attention from research scholars for their non-contact transmission and automatic overload protection capability [1, 2]. In particular, the electric drive system equipped with magnetic field modulation gear is capable of effectively reducing vibration and reducing the difficulty of system maintenance [3, 4]. Wong et al. employed a Halbach pole arrangement to obtain magnetic gears with torque densities close to those of mechanical gears, with torque densities approaching 300 Nm/L [5]. This offers a structural solution to facilitate

J. Du · D. Liu (✉) · T. Ren
School of Mechanical Engineering, Yanshan University, Qinhuangdao, Hebei, P. R. China
e-mail: liudw@ysu.edu.cn

© The Author(s) 2026

H. Saman et al. (eds.), *The 9th International Conference on Advances in Construction Machinery and Vehicle Engineering*, Lecture Notes in Mechanical Engineering,
https://doi.org/10.1007/978-981-96-5527-4_35

the implementation of magnetic gears in variable speed drive systems. For instance, magnetic gears are integrated with electric motors as a power shunt element for hybrid vehicles or a hub drive element for purely electric vehicles [6, 7]. However, this will greatly increase the design and manufacturing difficult, and is not conducive to compatibility with existing equipment.

In contrast, the use of magnetic gears to replace the planetary gears of the transmission system is more practical [8]. However, in comparison to mechanical gears, magnetic gears exhibit transmission characteristics that are non-contact, weakly damped, and low-stiffness. This results in suboptimal dynamic characteristics of the drivetrain with magnetic-containing gears [9]. It is therefore of particular importance to analyze the vibration mechanism of magnetic gearing systems and to implement improvements to enhance their dynamic response performance.

In the case of systems comprising solely magnetic gearing, the vibration characteristics may be evaluated through the utilization of multiscale methodologies and dynamic simulations [10, 11]. Furthermore, the incorporation of damping windings into magnetic gears has been demonstrated to be an efficacious method for attenuating the oscillation of magnetic gears and enhancing the response speed [12]. But the involvement of both magnetic and mechanical gears in a transmission system gives rise to a complex magnetic mechanical interaction, which may result in chaotic vibration phenomena. It is necessary to investigate the dynamic characteristics of composite transmission systems that incorporate both magnetic and mechanical gears.

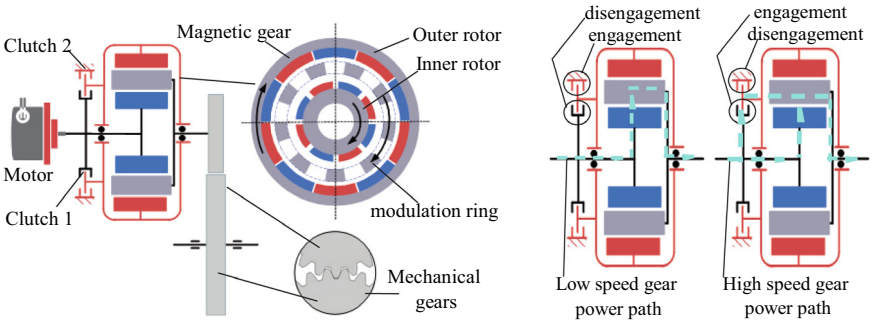
This paper puts forth a proposal for the implementation of a magnetic two-speed transmission system in electric vehicles. The system is designed to serve two primary functions: the isolation of vibration and the safeguarding of the motor. Subsequently, analyzing the bifurcation and chaos characteristics of the magnetic-mechanical composite transmission system. By studying its bifurcation characteristics and chaotic state, identifying the chaotic interval and controlling it can help improve the vibration characteristics of the system and reduce the occurrence of instability.

2 Magnetic Gear Two Speed Transmission System and Its Dynamic Model

2.1 Magnetic Gear Two Speed Transmission System

The magnetic gear two-speed transmission configuration and the power transmission path of the low/high speed gears is shown in Fig. 1. The system utilizes the non-contact differential characteristics of magnetic gears to achieve two speed ratios.

The kinematic relationship of the three components inside the magnetic gear [13] satisfies Eq. (1).



(a) system configuration (b) power path

Fig. 1 Magnetic gear two-speed transmission system configuration

$$\begin{cases} N_s = p_1 + p_2 \\ p_1\omega_1 + p_2\omega_2 = N_s\omega_3 \end{cases} \tag{1}$$

In Eq. (1), N_s is the number of iron blocks of the modulation ring. p_1 and p_2 are the number of magnetic pole pairs of the inner and outer rotors, respectively. ω_1, ω_2, ω_3 are the rotation speeds of the inner and outer rotors and the modulation ring, respectively.

The magnetic coupling torque in the air gap satisfies Eq. (2).

$$\begin{cases} T_i = \dfrac{p_i}{N_s}\left(T_{\max}\sin(p_2\theta_2 + p_1\theta_1 - N_s\theta_3) - T_{ripple}\sin(U_1\omega_1 t)\right), i = 1, 2 \\ U_1 = 2p_1N_s \end{cases} \tag{2}$$

T_i is the magnetic coupling torque in the air gap, T_{max} is the maximum coupling torque. T_{ripple} is the max ripple torque, and U_1 is the torque pulsation factor.

To enhance the dynamic response of the magnetic gear, a damping winding can be added, whose damping force can be expressed as a function of the relative rotational speed [12].

$$\begin{cases} T_{mc} = c_{MG}(p_2\omega_2 + p_1\omega_1 - N_s\omega_3) \\ c_{MG} = 2\xi\sqrt{K_{MG}J_{MG}} \end{cases} \tag{3}$$

where T_{mc} is the damping force, J_{MG} is the equivalent moment of inertia k_{MG} is the equivalent stiffness. ξ is the damping ratio.

In low gear, the inner and outer rotors of the magnetic gear move relative to each other, causing rippling torque in the air gap. In high gear, there's no relative motion. Thus, the analysis should focus on low gear to understand the underlying mechanisms.

2.2 System Dynamics Model and Its Differential Equations

Figure 2 shows the torsional dynamics model of the magnetic gear two-speed transmission system. This model only considers pure torsional vibration and has been simplified using the lumped parameter method.

The torsional stiffness of the clutch k_2 and its damping c_2, the stiffness k_3 and damping c_3 of the intermediate drive shaft as well as the mechanical gear inter tooth clearance b_0, time-varying stiffness k, and the combined transmission error e are considered in Fig. 2. Θ_1, $\Theta_{2,}\Theta_{3,}\Theta_{R1,}\Theta_{R2}$ are the relative rotation angles of each component, T_f is the load torque, which is balanced with the input torque T_M. R_1 and R_2 are the radius of mechanical gears.

The relative rotation angles, and their normalized variables are shown in Eq. 4.

$$\begin{cases} q_1 = p_1\theta_1 + p_2\theta_2 - N\theta_3 \\ q_2 = \theta_2 \\ q_3 = \theta_3 - \theta_R \\ q_4 = R_1\theta_{R1} - R_2\theta_{R2} - e \end{cases}, \begin{cases} \omega_n = \sqrt{\left(\frac{p_1}{J_1}\frac{p_1}{N_s} - \frac{p_2}{J_2}\frac{p_2}{N_s} + \frac{N_s^2}{J_3 N_s}\right)T_{\max}} \\ q_i = b_c\delta_i,\ i = 1, 2, 3, 4 \end{cases} \tag{4}$$

In Eq. 4, q_i ($i = 1, 2, 3, 4$) is the relative displacement variable, δ_i is the dimensionless displacement, b_c is the normalized length scale, and ω_n is the time scale coefficient, which is taken as the time scale $\tau = \omega_n t$.

The normalized differential equation with respect to the relative rotating angle is obtained as shown in Eq. (5).

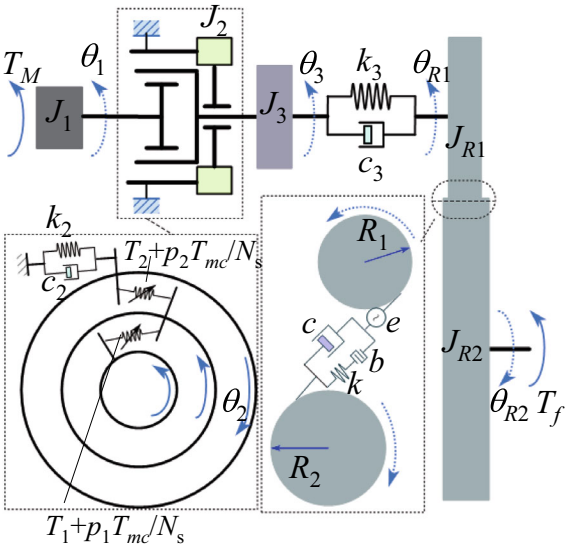


Fig. 2 Dynamics model of the magnetic gear two-speed transmission system

$$\begin{cases} \ddot{\delta}_1 + \dfrac{k_{MG}}{\omega_n^2} T_{kc1} - \dfrac{p_2}{J_2} T_{kc2} - \dfrac{N}{J_3} T_{kc3} - \dfrac{p_1}{J_1 \omega_n^2} T_M = -\dfrac{k_{MG} T_{ripple}}{\omega_n^2} \sin\left(U_1 \dfrac{\omega_1}{\omega_n} \tau \right) \\ \ddot{\delta}_2 + \dfrac{1}{J_2} T_{kc2} - \dfrac{p_2}{N_s J_2 \omega_n^2} T_{kc1} = \dfrac{p_2 T_{ripple}}{N_s J_2 \omega_n^2} \sin\left(U_1 \dfrac{\omega_1}{\omega_n} \tau \right) \\ \ddot{\delta}_3 + \left(\dfrac{1}{J_3} + \dfrac{1}{J_{R1}} \right) T_{kc3} - \dfrac{R_1}{J_{R1}} \left(\dfrac{k b_c}{\omega_n^2} \cdot f(\delta_4) + \dfrac{c b_c}{\omega_n} \dot{\delta}_4 \right) - \dfrac{T_{kc1}}{J_3 \omega_n^2} \\ \qquad = \dfrac{T_{ripple}}{\omega_n^2 J_3} \sin\left(U_1 \dfrac{\omega_1}{\omega_n} \tau \right) \\ \ddot{\delta}_4 + \left(\dfrac{R_1^2}{J_{R1}} + \dfrac{R_2^2}{J_{R2}} \right) \left(\dfrac{k}{\omega_n^2} \cdot f(\delta_4) + \dfrac{c}{\omega_n} \dot{\delta}_4 \right) - \dfrac{R_1}{J_{R1}} \left(\dfrac{k_3}{b_c \omega_n^2} \delta_3 + \dfrac{c_3}{b_c \omega_n} \dot{\delta}_3 \right) \\ \qquad - \dfrac{R_2}{J_{R2} b_c \omega_n^2} T_f = -\ddot{e}(\tau) \end{cases} \tag{5}$$

In Eq. (5), the following equation holds.

$$\begin{cases} T_{kc1} = T_{\max} \sin(\delta_1) + c_{MG} \omega_n \dot{\delta}_1, T_{kc2} = \left(\dfrac{k_2}{\omega_n^2} \delta_2 + \dfrac{c_2}{\omega_n} \dot{\delta}_2 \right), \\ \qquad T_{kc3} = \left(\dfrac{k_3}{\omega_n^2} \delta_3 + \dfrac{c_3}{\omega_n} \dot{\delta}_3 \right) \\ \begin{cases} \ddot{e}(\tau) = e_{10} \left(\dfrac{\omega_1}{\omega_n} \dfrac{p_1}{N} z_1 \right)^2 \cos\left(\dfrac{\omega_1}{\omega_n} \dfrac{p_1}{N} z \tau \right), f(\delta) = \begin{cases} \delta - b/b_c, \delta \geq b/b_c \\ 0, |\delta| < b/b_c \\ \delta + b/b_c, \delta \leq -b/b_c \end{cases} \\ e_\Delta = e_{10} b_c \end{cases} \end{cases} \tag{6}$$

Among them, f is a nonlinear function of tooth flank clearance,

The time-varying meshing stiffness k of the gear pair as well as the combined transmission error e can be expressed as a trigonometric function with respect to the meshing frequency.

$$\begin{cases} k(\theta) = k_m + \sum_{h=1}^{n} k_\Delta \sin(\omega_h \theta + \varphi_h) \\ e(\theta) = \sum_{h=1}^{n} e_\Delta \sin(\omega_h \theta + \varphi_h) \end{cases} \tag{7}$$

In Eq. (7), ω_h is the frequency of gear mesh.

The inter tooth damping c and the stiffness damping c_3 of the transmission shaft are calculated using Eq. (8).

$$\begin{cases} c = 2\xi \sqrt{k_m J_1 J_2 R_1^2 R_2^2 / (J_1 R_1^2 + J_2 R_2^2)} \\ c_3 = 2\xi_3 \sqrt{k_3 / (1/J_3 + 1/J_{R1})} \end{cases} \tag{8}$$

3 System Solution and Its Bifurcation and Chaos Characteristics

3.1 Solution and Bifurcation Analysis of Nonlinear Dynamic Response of the System

Take dimensionless displacement δ, dimensionless velocity $\dot{\delta}$ as state variables. The fourth-order Runge–Kutta method program for simulation calculations is written in MATLAB R2021a. Of these, the time step is 10^{-4}. The simulation durations are all taken as 2000 to ensure that the system is able to reach steady state. The initial simulation value is set to 0. The structural parameters are shown in the Table 1.

Taking the position of the amplitude maximum as the initial phase, the bifurcation map using stroboscopic method can reflect the amplitude-frequency characteristics and bifurcation characteristics at the same time. Calculate with T_M of 35 Nm and obtain the result shown in Fig. 3.

From Fig. 3, the magnetic gear relative rotation angle δ_1 and the mechanical gear relative rotation angle δ_3 have similar amplitude-frequency variation characteristics. The amplitude reaches its maximum value near the excitation frequency of 0.62, which is known to be the resonant frequency of the system. And the bifurcation diagram depicts the multi-valued bifurcation property. $\delta_1, \delta_3, \delta_4$ with amplitude jumps around frequencies of about 0.62, 1.37, 1.8 and 2.14. And, as the amplitude jumps, bifurcated intervals appear. This reflects the interaction between variable stiffness magnetic gears and mechanical gears containing backlash. Subsequently, the system re-enters periodic motion via an inverse doubling period bifurcation or an inverse Hough bifurcation. Typical inverse doubling period bifurcations and inverse Hough bifurcations occur at frequencies of approximately 1.83 and 2.25, respectively.

The Poincaré cross-section of the magnetic gear at varying frequencies is illustrated in Fig. 4.

Table 1 Dynamics system parameters

Parameter	Value	Parameter	Value
Magnetic pole p_1:p_2:N_s	4:5:9	Ripple torque	4 Nm
Maximum magnetic coupling torque T_{max}	100 Nm	Average mesh stiffness k_m	3.3×10^8 Nm
Gear module m	2	Stiffness fluctuation	24%
Number of gear teeth z_1, z_2	21, 107	Tooth flank clearance	0.07 mm
Comprehensive transmission error e	0.05 mm	Clutch torsional stiffness k_2	2000 Nm/rad
Damping ratio of transmission shaft ζ_3	0.03	Transmission shaft torsional stiffness k_3	5800 Nm/rad
Gear mesh damping ratio ζ	0.1	Clutch torsional damping c_2	5 Nms/rad

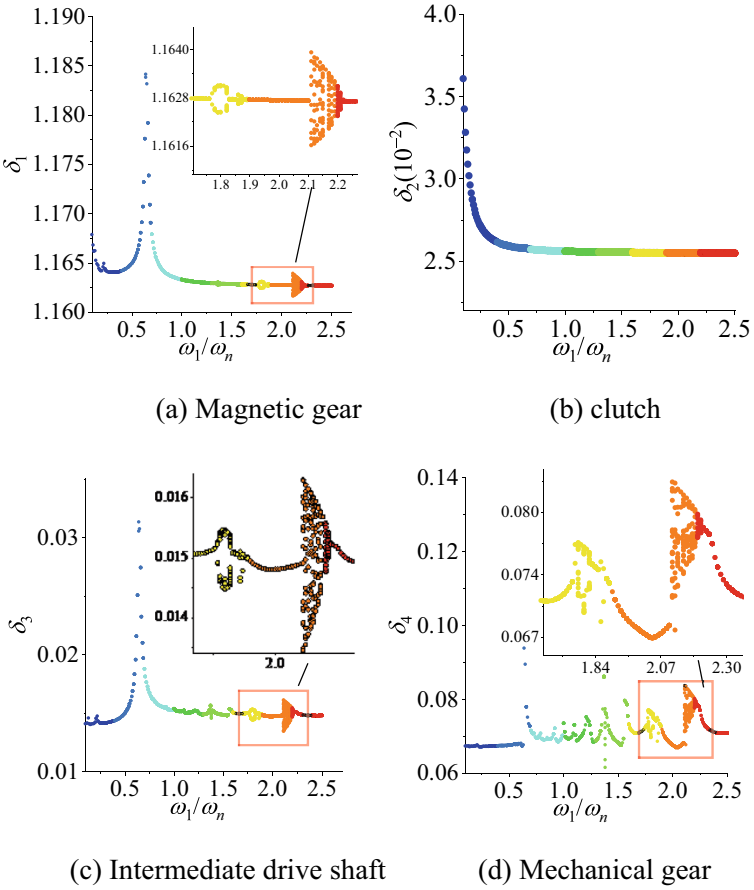


Fig. 3 System dimensionless displacement bifurcation diagram

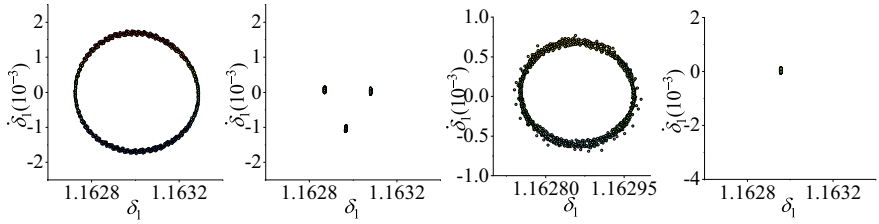


Fig. 4 Poincaré cross-section at typical frequencies

In Fig. 4, the bifurcation intervals in the proximity of 1.8 and 2.14 manifest quasi-periodic motion, and the emergence of chaotic behavior is not discernible. At frequencies of 1.86 and 2.35, the system then exhibits multiplicative periodic motion and periodic motion, respectively. This finding is consistent with the results of the bifurcation diagram analysis, which indicate the presence of multiple vibration mode switching during the increase in system speed.

3.2 *The Influence of Damping on System Bifurcation and Chaos Characteristics*

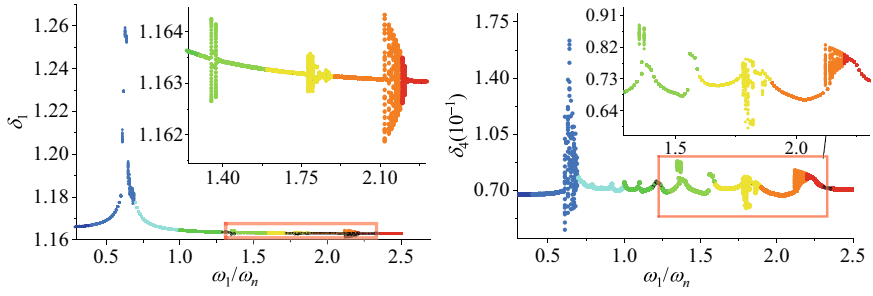
It is of particular importance to consider the effect of damping in order to assess the potential risk of multi-peak resonance instability in this system. When the damping ratio of the mechanical gear meshing is 0.06, the resulting bifurcation diagram is illustrated in Fig. 5.

A comparison of Fig. 3 with the present case reveals that a reduction in the damping of the mechanical gear results in the magnetic gear producing a new case of oscillation bifurcation at a frequency of 1.39. Mechanical gears have a significantly wider potential chaos interval around a frequency of 0.64.

Figure 6 illustrates the bifurcation characteristics of the mechanical and magnetic gears when the intermediate axis damping is 0.01.

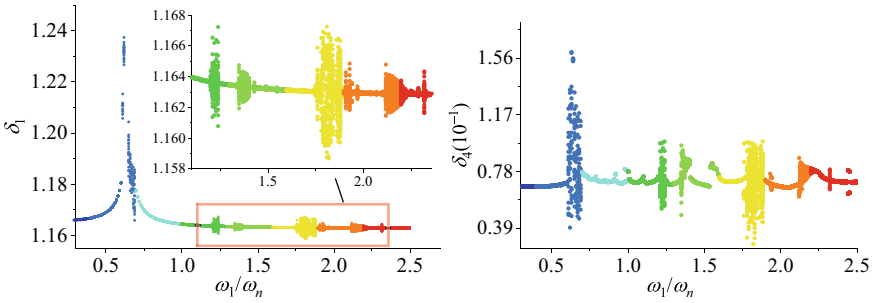
In Fig. 6, a reduction in the damping ratio of the intermediate drive shaft results in the bifurcation of the relative displacements of the mechanical and magnetic gears at frequencies of 0.63, 1.25, 1.37, 1.8, and 2.21. In comparison to magnetic gears, mechanical gears exhibit a more pronounced amplitude jump.

The reason for this phenomenon is the enhanced torsional vibration interaction between the mechanical and magnetic gears due to reduced drive shaft damping. This results in bifurcations occurring in close proximity to nearly every resonance frequency. Furthermore, this phenomenon gives rise to an enhanced propensity for chaotic behavior.



(a) Magnetic gear (b) Mechanical gear

Fig. 5 Poincaré cross-section at typical frequencies ($\zeta = 0.06$)

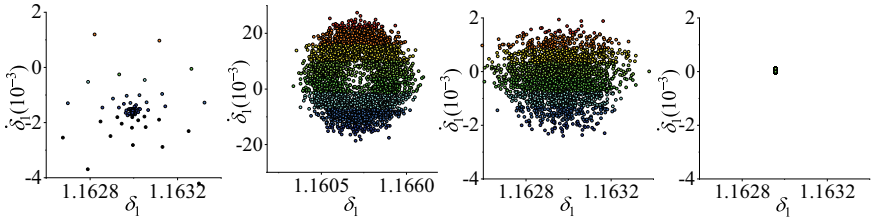


(a) Magnetic gear (b) Mechanical gear

Fig. 6 Poincaré cross-section at typical frequencies ($\zeta_3 = 0.01$)

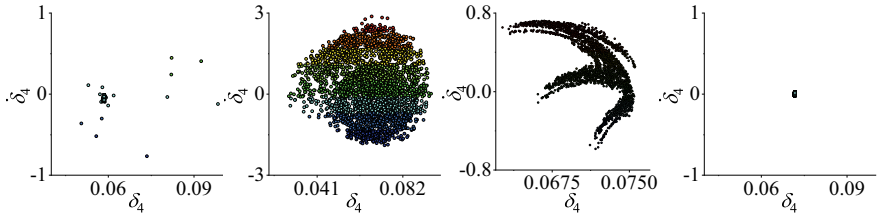
Figures 7 and 8 show how the magnetic and mechanical gears switch from stable to chaotic and back again when $\zeta_3 = 0.01$. It is established that in instances where damping is diminished, quasi-periodic motion will possess a sufficient frequency range to evolve into chaotic motion. At a frequency of 1.86, the cross-section exhibits irregular sheets but no fractal structure, indicating that the system has not yet entered a fully chaotic state. In Fig. 8, moreover, at a frequency of 1.86, the Poincaré cross-section exhibits a fractal structure, indicating that the mechanical gear has entered a state of fully chaotic vibration at this point.

In summary, it can be seen that the intermediate axis damping can significantly affect the vibration state of the system, and reducing the damping will lead to the system being more prone to instability. Moreover, mechanical gears are more prone to chaotic vibrations than magnetic gears.



(a) ω_1/ω_n=1.78 (b) ω_1/ω_n=1.82 (c) ω_1/ω_n=1.86 (d) ω_1/ω_n=1.9

Fig. 7 Poincaré cross-section at typical frequencies (magnetic gear, $\zeta_3 = 0.01$)



(a) $\omega_1/\ \omega_n$=1.78 (b) $\omega_1/\ \omega_n$=1.82 (c) $\omega_1/\ \omega_n$=1.86 (d) $\omega_1/\ \omega_n$=1.9

Fig. 8 Poincaré cross-section at typical frequencies (mechanical gear, $\zeta_3 = 0.01$)

4 Conclusion

In this paper, a magnetic gear two-speed transmission system using magnetic gears as variable speed transmission components is proposed. And analyzed the bifurcation and chaos characteristics. Results show that the system exhibits rich bifurcation characteristics under the combined effect of nonlinear magnetic coupling stiffness, time-varying meshing stiffness of gears, tooth-side clearance, and integrated transfer error. Include Multi-band bifurcation and amplitude jump. Reducing the damping leads to an increase in the frequency range of the system's chaotic intervals and an increase in the number of chaotic intervals. The effect of intermediate driveshaft damping is even more pronounced. Furthermore, mechanical gears are more likely to enter a chaotic state than magnetic gears.

It's important to note that ignoring efficiency doesn't affect identifying the system's vibration mechanism, but it does limit performance evaluation. In the future, we will improve the dynamics model by including multi-source excitation and friction factors to increase the accuracy of predictions about the system's behavior.

Acknowledgements This work was supported by the Hebei Province Graduate Innovation Ability Cultivation Fund project in China (grant number CXZZBS2023057).

References

1. Filippini M, Alotto P (2017) Coaxial magnetic gear design and optimization. IEEE Trans Ind Electron 64:9934–9942
2. Tzouganakis P, Gakos V, Kalligeros C, Tsolakis A, Spitas V (2023) Fast and efficient simulation of the dynamical response of coaxial magnetic gears through direct analytical torque modelling. Simul Model Pract Theory 123:102699
3. Jo I-H, Lee J, Lee H-W, Lee J-B, Lim J-H, Kim S-H, Park C-B (2023) A topology study for the application of magnetic geared motor as traction for urban railway vehicle. J Electr Eng Technol 18:4473–4479
4. Jing L, Tang W, Wang T, Ben T, Qu R (2022) Performance analysis of magnetically geared permanent magnet brushless motor for hybrid electric vehicles. IEEE Trans Transp Electrific 8:2874–2883

5. Wong HY, Baninajar H, Dechant BW, Southwick P, Bird JZ (2023) Experimentally testing a Halbach rotor coaxial magnetic gear with 279 Nm/L torque density. IEEE Trans Energy Convers 38:507–518
6. Chmelicek P, Calverley SD, Dragan RS, Atallah K (2019) Dual rotor magnetically geared power split device for hybrid electric vehicles. IEEE Trans Ind Applicat 55:1484–1494
7. Iwaki K, Ito K, Nakamura K (2023) A novel magnetic-geared motor combining magnetic gear and SR motor. IEEE Trans Magn 59:1–5
8. Du J, Liu D, Ren T (2024) Low-stiffness comfort shift characteristics of a permanent magnet two-speed transmission for low-speed EV
9. Park E-J, Kim C-S, Jung S-Y, Kim Y-J (2018) Analysis of torque characteristics according to non-uniform air gaps of coaxial magnetic gear. JMAG 23:675–679
10. Hao X, Yuan XM, Zhang HF, Zhang LJ (2014) Combination resonances of parametric vibration system of the field modulated magnetic gear. J Vibroeng
11. Chen CL, Tsai MC (2018) Kinematic and dynamic analysis of magnetic gear with dual-mechanical port using block diagrams. IEEE Trans Magn 54:1–5
12. Ma BH, Hou CC, Bao GQ (2016) Transient behavior of field modulated magnetic gears with damper windings. In: 2016 IEEE vehicle power and propulsion conference (VPPC), pp 1–6. IEEE, Hangzhou
13. Atallah K, Calverley SD, Howe D (2004) Design, analysis and realisation of a high-performance magnetic gear. IEE Proc Electr Power Appl 151:135

Open Access This chapter is licensed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License (http://creativecommons.org/licenses/by-nc-nd/4.0/), which permits any noncommercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if you modified the licensed material. You do not have permission under this license to share adapted material derived from this chapter or parts of it.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.



Force Analysis of All-Terrain Crane Lowering Arm and Suspension Cylinder



Ling Wang, Xin Wang, Shuai Song, and Xifeng Zhang

Abstract Due to the limited transverse bearing capacity of the suspension cylinder, a lower arm is introduced between the hub and the gearbox to distribute the transverse load on the suspension cylinder. To analyze the forces acting on this component and its impact on the suspension cylinder, force equations for the hinge points of the lower swing arm and suspension cylinder are derived based on the equilibrium of spatial forces and moments. These equations are validated by indirectly comparing them with the forces on the left and right tires obtained through FEA. The derived theoretical formula is further validated by calculating the stresses on the lower swing arm and comparing them with the simulation results. The comparison shows a good correlation between the theoretical and simulation results, demonstrating the applicability of the formula in engineering calculations.

Keywords Crane · Lower arm · Suspension cylinder · Force analysis

1 Introduction

In recent years, the rapid growth of industries like high-speed rail, wind power, petrochemicals, and metallurgy has increased the demand for large all-terrain cranes. However, these cranes often face challenging force conditions, particularly during braking and turning, due to factors such as self-weight and road conditions. The suspension cylinder, which plays a key role in load-bearing, has limited capacity to handle lateral forces [1–3], leading to the use of thrust rods to improve horizontal load capacity. Thrust rods not only help transfer traction [4] and braking forces but also assist in positioning the frame and axle [5], ensuring vehicle stability.

L. Wang · X. Wang (✉)
School of Mechanical Engineering, Dalian University of Technology, Dalian, China
e-mail: wangxbd21@163.com

S. Song · X. Zhang
Taiyuan Heavy Industry Co., Ltd., Taiyuan, China

© The Author(s) 2026

H. Saman et al. (eds.), *The 9th International Conference on Advances in Construction Machinery and Vehicle Engineering*, Lecture Notes in Mechanical Engineering,
https://doi.org/10.1007/978-981-96-5527-4_36

Research on thrust rods has mainly focused on structural strength and optimization [6]. For example, Wenku et al. [7] optimized the thrust rod's ball-hinge structure under extreme conditions. Other studies, such as those by Li and Feng [8], have developed models and force calculation formulas for thrust rods under various conditions.

While most studies focus on strength analysis, further exploration is needed on thrust rods in composite working conditions. This paper uses force and moment balance principles to conduct a spatial force analysis under side load, braking, and vertical forces, with simulations showing a strong match with theoretical calculations.

2 Theoretical Calculation

For ease of analysis, this paper focuses on the crane's single-axle suspension system, dividing the three-dimensional forces under combined conditions (side load, braking, and vertical force) into forces on the XZ and XY planes. Under the combined side load and vertical force conditions, the single-axle structure is subjected to forces in the XZ plane, while during braking, it experiences forces in the XY plane. The global coordinate system is shown in Fig. 1.

2.1 Side Loading and Vertical Force Working Condition Force Analysis

(1) Overall force analysis

Due to the presence of lateral roll acceleration, the vertical forces in the left and right tyres during side loading are greater and less than those during static loading, respectively, as shown in Fig. 1.

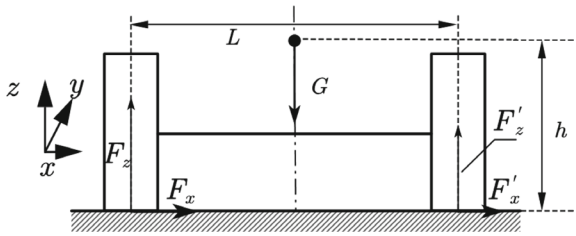


Fig. 1 Overall force diagram

$$\begin{cases} F_x + F_x' = ma_x \\ F_z + F_z' = G \\ F_x \big/ F_x' = F_z \big/ F_z' \\ ma_x h + F_z' L \big/ 2 - F_z L \big/ 2 = 0 \end{cases} \tag{1}$$

$$\begin{cases} F_b \cos\alpha = F_z \\ F_b \sin\alpha + F_{ox} = F_x \\ F_b' \cos\alpha' = F_z' \\ F_b' \sin\alpha' + F_x' = F_{ox}' \end{cases} \tag{2}$$

where G is the gravity of single axle, h is the height of the centre of gravity from the ground, L is the wheelbase, F_x is the lateral force generated by the left tyre by lateral tilt braking, F_z is the vertical force of the left wheel, F_x' is the lateral force of the right wheel caused by the lateral tilt braking, F_z' is the vertical force of the right wheel, μ is the tyre-ground attachment coefficient, which is obtained from Eq. (1):

$$\begin{cases} F_z = G \big/ 2 + ma_x h \big/ L \\ F_z' = G \big/ 2 - ma_x h \big/ L \end{cases} \tag{3}$$

$$\begin{cases} F_x = ma_x F_z \big/ (F_z + F_z') \\ F_x' = ma_x F_z' \big/ (F_z + F_z') \end{cases} \tag{4}$$

(2) Unilateral force analysis

After completing the analysis of the external forces on the axle, the axle is divided into the left and right parts of the force analysis is carried out separately as shown in Fig. 2, according to the relationship shown in the diagram can be obtained as Eq. (2). Where, F_b is the force at the connection between the frame and the cylinder, F_{ox} is the tension in the left lower swing arm, and F is the resultant force of F_x and F_z. The lower swing arm can be regarded as a two-force member, from Eq. (2):

$$\begin{cases} F_{ox} = F_x - F_z \tan\alpha \\ F_{ox}' = F_x' + F_z' \tan\alpha' \end{cases} \tag{5}$$

$$\begin{cases} F_{bx} = F_x - F_{ox} \\ F_{bz} = F_z \end{cases} \tag{6}$$

$$\begin{cases} F_{bx}' = F_{ox}' - F_x \\ F_{bz}' = F_z' \end{cases} \tag{7}$$

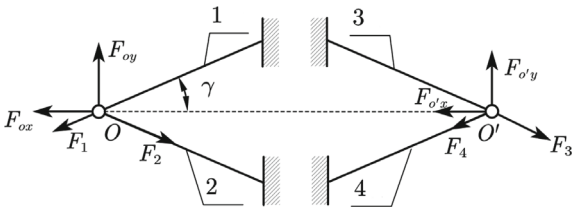


Fig. 2 Lower arm force diagram

2.2 *Braking Force Analysis*

Vehicle braking, longitudinal force on the left side of the single axle as shown in Fig. 3, connecting the cylinder and the frame hinge point B and the lower swing arm and the hub of the hinge point O to get the main pin axis as shown in Fig. 4, over the braking force point F_y of action for the above main pin axis of the plumbline, point D is the vertical foot. The braking force F_y is equivalent to point D. The resulting moment around the main pin axis is taken up by the steering linkage (part 4 in Fig. 5). The longitudinal forces on the right side are the same as those on the left side.

From the relationship shown in the graph, the vehicle brakes:

$$\begin{cases} F_y = ma_y / 2 \\ F_{oy} = F_y + F_{by} \\ aF_y = bF_{by} \end{cases} \tag{8}$$

$$\begin{cases} F_{oy} = (a+b)F_y / b \\ F_{oy} = aF_y / b \end{cases} \tag{9}$$

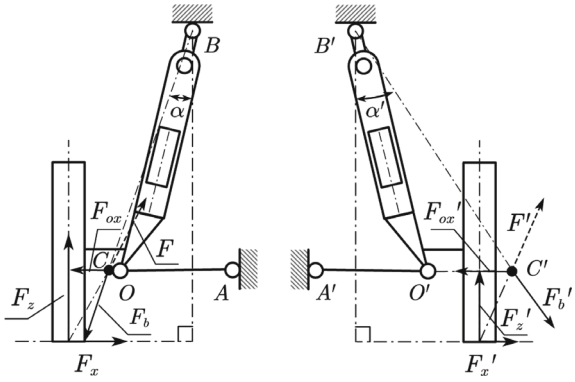


Fig. 3 Unilateral force diagram

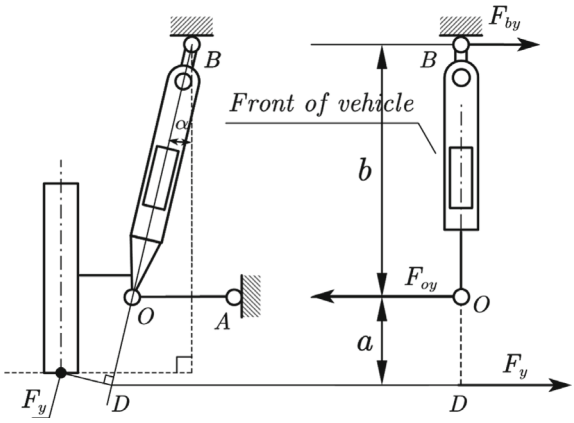


Fig. 4 Braking force of the left axle

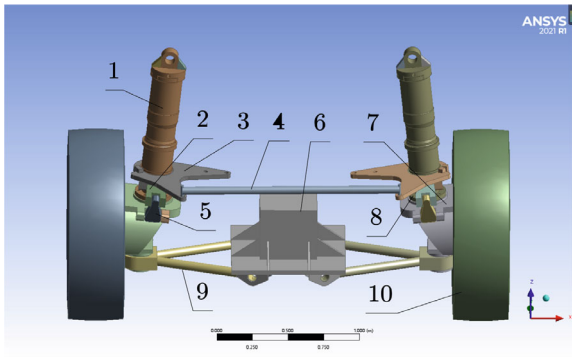


Fig. 5 Finite element model of a single bridge

2.3 *Extreme Working Condition Force Analysis*

In summary, when the vehicle is subjected to lateral acceleration a_x, braking acceleration a_y, and drooping acceleration a_z, the X-direction force F_{ox} and Y-direction force borne F_{oy} by the left and right side lowering arms are:

$$\begin{cases} F_{ox} = F_x - F_z \tan\alpha \\ F_{oy} = (a+b)F_y \big/ b \end{cases} \tag{10}$$

$$\begin{cases} F'_{ox} = F'_x + F'_z \tan\alpha' \\ F'_{oy} = F_{oy} \end{cases} \tag{11}$$

The component forces F'_{ox}, F'_{oy}, F'_{oz} in the X, Y, Z directions borne by the hinge points of the left and right side suspension cylinders are:

Table 1 Calculation of the required parameter values

Parameters	m/t	h/mm	L/mm	a/mm	b/mm	a_x/g	a_y/g	a_z/g	α	α'	γ
Numbers	13	2200	2580	440	1570	0.4	− 0.4	1	7°	19°	20°

$$\begin{cases} F_{bx} = F_x - F_o \\ F_{by} = aF_y \big/ b \\ F_{bz} = F_z \end{cases} \tag{12}$$

$$\begin{cases} F'_{bx} = F'_o - F'_x \\ F'_{by} = aF \\ F'_{bz} = F'_z \end{cases} \tag{13}$$

2.4 Force Analysis of the Lower Arm

When the vehicle is working in the extreme working condition, the lower swing arm is subjected to the acceleration in X and Y directions, and its force is shown in Fig. 2. According to the relationship shown in the diagram. Equation (14) can be obtained. The data required for the above calculations are shown in Table 1.

3 Static Analysis

3.1 Finite Element Modelling

The 3D model of the crane single bridge is built in SW and imported into Workbench as shown in Fig. 5. A tetrahedral unstructured mesh is used to divide the structure and the divided mesh diagram is shown in Fig. 6. The total number of mesh nodes is 840,058 and the total number of cells is 476,534. Subsequently, a local contact surface between the tyre and the ground is created in DM for the subsequent application of fixed support constraints on this contact surface.

3.2 Constraints and Load Application

A 13t mass point is applied at the centre of mass of the single bridge above the two suspension cylinders to simulate the dead weight of the bridge. Fixed supports were applied at the left tire-ground contact and the underside of the gearbox, restricting

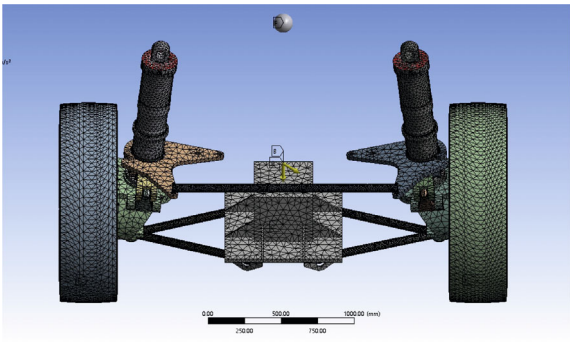


Fig. 6 Boundary conditions and mesh diagram

Table 2 Relative constraints between components

Contact part	Contact paraphernalia
1–8, 7–10	Bind
1–3, 2–3, 2–5, 5–7, 1–7	Rotary
3–4, 7–9	Spherical vice

all degrees of freedom. On the right tire, all degrees of freedom except X-direction displacement were constrained, allowing horizontal force transfer through the lower swing arm to the gearbox. Additionally, accelerations of 0.4 g in the X, − 0.4 g in the Y, and 1 g in the Z directions were applied to simulate extreme conditions. Figure 6 illustrates the external loads on the axle, while Table 2 details the component constraint relationships.

$$\begin{cases} F_1 = F_{oy}/2\sin\gamma - F_{ox}/2\cos\gamma \\ F_2 = F_{oy}/2\sin\gamma + F_{ox}/2\cos\gamma \\ F_3 = F_{o'y}/2\sin\gamma + F_{o'x}/2\cos\gamma \\ F_4 = F_{o'y}/2\sin\gamma - F_{o'x}/2\cos\gamma \end{cases} \tag{14}$$

3.3 Finite Element Calculation

The single bridge stress cloud is calculated as shown in Fig. 7, from which it can be seen that the maximum stress is 89 Mpa, which is located in the middle of the right suspension cylinder.

The stress cloud diagram of the lower swing arm is shown in Fig. 8, from which it can be seen that the stresses of 2 and 3 rods are larger, which is due to the acceleration vector sum of side loading and braking in the XY plane pointing to the lower right in Fig. 7, and this synthetic acceleration is mainly borne by the 2 and 3 rods.

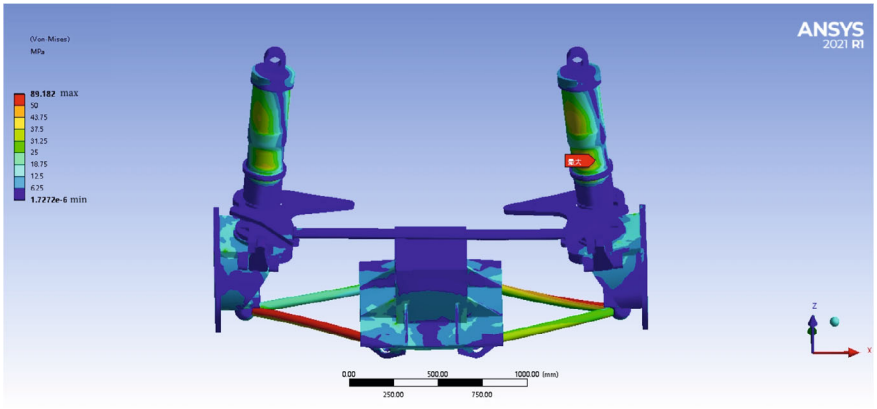


Fig. 7 Stress cloud (magnified 50 times)

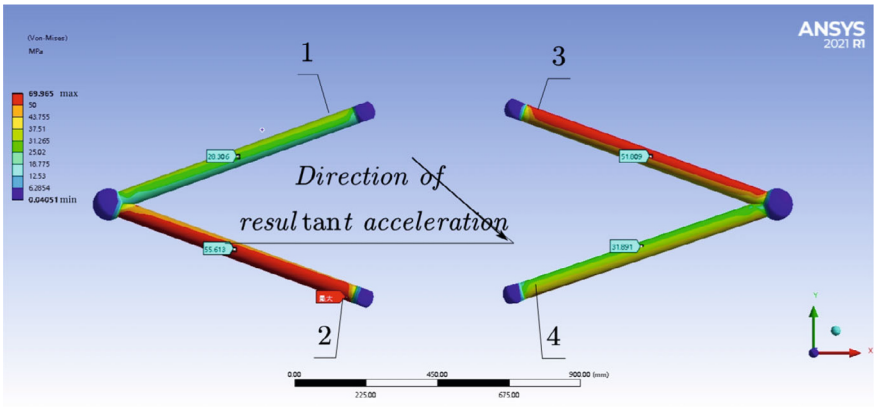


Fig. 8 Lower arm stress cloud

4 Comparative Analysis of Simulation and Theory

Compare the simulation and theoretical calculation of the left and right tyre X, Y, Z direction force indirectly compare the simulation and theory of the lower arm and suspension cylinder hinge point force, the specific data are shown in Table 3.

The X and Y forces of the lower arm and the X, Y and Z forces at the hinge point of the suspension cylinder are calculated as shown in Table 4, from which it can be seen that the lower arm greatly shares the X and Y forces borne by the suspension cylinder, which greatly improves the service life of the suspension cylinder.

Table 3 Simulated and theoretical forces on left and right tyres (unit: N)

Distributed	Simulation	Theoretical	Error (%)
Left-X	39,625	42,892	7.6
Left-Y	− 25,911	− 25,498	1.6
Left-Z	98,836	107,231	7.8
Right-X	11,371	8104	28.7
Right-Y	− 25,049	− 25,498	1.6
Right-Z	28,655	20,260	29.3

Table 4 Forces at the LCA and SC (unit: N)

Distributed	Left	Right
LCA-X	− 29,725	− 15,080
LCA-Y	32,644	32,644
SC-X	− 13,167	6976
SC-Y	− 7146	− 7146
SC-Z	− 107,231	− 20,260

5 Conclusion

To reduce the X and Y directional forces on the suspension cylinder under extreme conditions, a lower swing arm was added between the hub and the gearbox. Force equations for the swing arm and suspension cylinder under combined lateral, braking, and vertical forces were derived. A finite element model of the single axle was created, and the simulated tire forces were compared with theoretical values, confirming the accuracy of both the equations and the model. The stress analysis of the swing arm was further validated through comparison with theoretical results, and the X and Y directional forces on the suspension cylinder were verified. Due to limitations in the current conditions, this paper is temporarily unable to conduct experimental verification. This hypothesis will be validated during subsequent prototype testing. The key conclusions are:

(1) The combination of simulation and theoretical calculations confirmed the accuracy of the derived equations, providing a method for calculating the forces on the swing arm and suspension cylinder joints.
(2) The swing arm bears most of the X and Y directional forces under extreme conditions, significantly reducing stress on the suspension cylinder and extending its lifespan.

Acknowledgements This research was supported by Taiyuan key core technology research unveiled list of marshal projects.

References

1. Wang B, Peng J, Liang Y et al (2022) Design and characterisation of suspension mechanism of all-terrain mobile robot. J Mechan Eng 58(09):71–86
2. Özhan UD, Erdem A (2022) Design optimization of armored wheeled vehicle suspension lower control arm. Mater Test 64(7):932–944
3. Zhao Y, Zhao P (2018) Analysis of SUV control arm strength analysis of SUVs with rigidity and soft coupling characteristics. Mechan Des Manuf 08:80–83
4. Gao J, Qi X (2023) Optimal design of steering trapezoid considering multiple performances of the whole vehicle. 18:114–126
5. Liang GD, Cheng L (2022) Analysis and optimization of fatigue failure analysis and optimization of an front suspension of swing arm. Mechan Des Manuf 03:292–295
6. Zou XH, Gou LL, Fu L et al (2022) Optimization of vehicle dynamic performance multi-target optimization based on the entropy right TOPSIS method. J Weapon Equip Eng 43(07):301–308
7. Shi WK, Ke J, Wang Q et al (2013) V-type lock analysis of the V-type thrust rod and optimization of ball hinge structure. J Xi'an Jiaotong Univ 47(10):132–136
8. Feng YG, Shi KW, Ke J et al (2015) Analysis of commercial vehicle thrust rod load simulation. Appl Mech Mater 3744(713–715):159–163

Open Access This chapter is licensed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License (http://creativecommons.org/licenses/by-nc-nd/4.0/), which permits any noncommercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if you modified the licensed material. You do not have permission under this license to share adapted material derived from this chapter or parts of it.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.



Fluid–Structure Coupling Analysis of Corner Guide Vane in Large-Scale Low-Speed Wind Tunnel

Jiahang Li, Xuefei Wang, Rumin Teng, Linkai Zhou, and Zhenping Jiang

Abstract Taking the first corner guide vanes of a large low-speed wind tunnel as the research object. Fully considering the complexity of the structure of guide vane as well as the unevenness of wind load distribution, fluid–structure coupling calculation and analysis are conducted. The local flow field is used as a simplified representation of the full flow field for fluid–structure coupling calculations. The calculation results indicate that when the airflow passes through a guide vane, it is divided into two streams. The overall flow velocity initially increases rapidly and then gradually decreases. The flow velocity on the suction side is greater than that on the pressure side. The pressure on the windward side of the head skin is the highest. The airflow forms a positive pressure zone near the pressure side and a negative pressure zone near the suction side. The single guide vane is subjected to a load directed from the pressure side to the suction side. The deformation of the structure is maximum at the mid-span head skin position. The stress is maximum at the contact between the top bulkhead and the support. After verification, the structure meets the strength requirements under wind load.

Keywords Guide vane · Wind load · Fluid–structure coupling · Wind tunnel

J. Li · X. Wang · R. Teng (✉)
Dalian University of Technology, Dalian, China
e-mail: tengrumin@163.com

L. Zhou
Dalian Shipbuilding Industry Co., Ltd., Dalian, China

Z. Jiang
Dalian COSCO Shipping Heavy Industry Co., Ltd., Dalian, China

© The Author(s) 2026

H. Saman et al. (eds.), *The 9th International Conference on Advances in Construction Machinery and Vehicle Engineering*, Lecture Notes in Mechanical Engineering,
https://doi.org/10.1007/978-981-96-5527-4_37

1 Introduction

As an indispensable part of a return-flow wind tunnel, the guide vane in the corner section plays an important role in reducing the resistance loss and improving the uniformity of the flow field at the outlet in the corner section [1]. When the airflow passes through the corner, the flow direction changes suddenly. Due to the action of centrifugal force and the relatively large flow velocity, large pressure will be applied to structure of the guide vanes in the corner section. At the same time, the stiffness of the structure of a single guide vane is low, and the structure is only connected to the tunnel through supports at the upper and lower ends. It cannot be ignored that the influence of wind load on the guide vanes. Among the four corner sections, the airflow velocity in the first corner section downstream of the test section is the largest. Therefore, the guide vanes in the first corner section is selected as the research object.

The 4 m × 3 m large closed-loop aerodynamic and acoustic wind tunnel (BHAW) with low turbulence intensity and low background noise of Beijing University of Aeronautics and Astronautics is taken as the engineering application background. Liu [2] analyzed the influence of the installation angle of the corner deflector on the flow field under multiple expansion ratios by numerical simulation. In reference [3], the computer simulation show that the pressure loss of airflow is reduced by 62%, the corner vortex is eliminated and the uniformity of flow field is improved by using seven double-arc airfoil guide vanes. In Reference [4], the influence of solidity on guide vane on the aerodynamic loss of the corner of the wind tunnel was studied by using CFD tools and experimental verification, and it was determined that there was an optimal number of corner guide vane for a given wind tunnel. Zou et al. [5] used k-ξ turbulence model to analyze the flow field in the corner section of low-speed wind tunnel, and analyzed the influence of the spatial arrangement of guide vanes on the flow field quality with the resistance coefficient and fluctuating lift coefficient as the evaluation criteria.

A substantial amount of research on corner guide vane has almost invariably focused on the influence of factors such as the number, structural form, and installation angle of guide vanes on the quality of flow field. However, there is scant research on the effect of wind loads on wind tunnel guide vane. Therefore, the first corner guide vanes of a large low-speed return-flow wind tunnel is taken as the research object. Fully considering the complexity of the structure of biconvex airfoil guide vane and the unevenness of wind load distribution. The fluid–structure coupling analysis of the external air flow field and the structure of guide vane is carried out. The outer flow field distribution of the guide vane is obtained through numerical calculation of the three-dimensional flow field. The reliability of the structure of guide vane is verified through strength verification of the stress concentration parts of the structure, providing technical support for the safe operation of the guide vane.

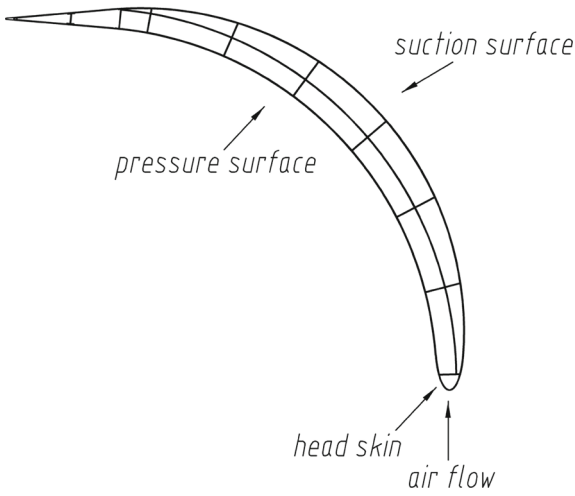


Fig. 1 Structure of guide vane

2 Computation Module

At the first corner, 22 individual guide vanes is arranged along the diagonal of the corner channel. The individual of the guide vane are skin welded structure, connected to the tunnel through five supports at the bottom and two supports at the top. The material of the support is Q355, and the material of the main load-bearing structure is Q235. As shown in Fig. 1, the inner curved surface of the guide vane with a shorter arc length is referred to as the pressure surface, and the outer curved surface with a longer arc length is called the suction surface. In order to improve computational efficiency, some details of structure are simplified while retaining the main load-bearing structures such as skin, bulkhead, and reinforcement plate.

3 Flow Field Analysis

The fluid–structure coupling analysis of the structure of guide vane under wind load is conducted. To enhance computational efficiency and reduce computational complexity, a local flow field of a single guide vane is employed for fluid–structure coupling calculation. In order to verify whether the local flow field can be used instead of the whole flow field, it is necessary to compare the surface load of the guide vanes in the two flow fields.

To ensure the accuracy of the calculation results, the inlet and outlet of the flow field need to be appropriately extended. The computational domain grid is shown in Fig. 2. The positive X direction is the inlet flow direction of the flow field, and the positive Y direction is the outlet flow direction of the flow field.

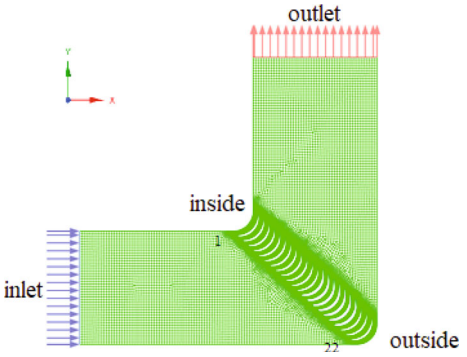


Fig. 2 Grid of full-flow computational domain

Steady-state fluid calculations is performed, and the k-epsilon model is selected as the turbulence model. The fluid domain material is air with the density of 1.225 kg/m^3 and the viscosity of 1.7894e−5 kg/m s. The velocity of inlet is set as 65 m/s and the outlet is set as the pressure outlet with a gauge pressure of 0 Pa. All other boundaries are set as non-slip walls to simulate solid boundaries. As shown in Fig. 2, the innermost guide vane is set as No.1 and the outermost guide vane is set as No.22, with the numbers of the guide vanes increasing in sequence. Extract As shown in Fig. 3, the force acting on the 22 guide vanes in the x and y directions is extracted.

From the positive and negative values of the force acting on the guide vanes, it can be observed that the direction of the force acting on each guide vane is from the inside of the corner to the outside of the corner. The wind load on the first guide vane is the highest, with 313.63 kN in the x-direction and − 347.33 kN in the y-direction. The wind load on the second guide vane drops sharply to nearly 200 kN in both directions. Then, it gradually increases from the second to the 21st guide vane. The wind load on the 21st guide vane is 263.20 kN in the x-direction and − 277.00 kN in the y-direction. The force on the 22nd guide vane suddenly drops to 136.69 kN in the x-direction and − 139.29 kN in the y-direction. Therefore, the local flow field of No.1 guide vane should be selected for fluid–structure coupling calculation.

To compare the difference between the local three-dimensional flow field and the full flow field, a calculation is conducted on the local outflow field of the first guide

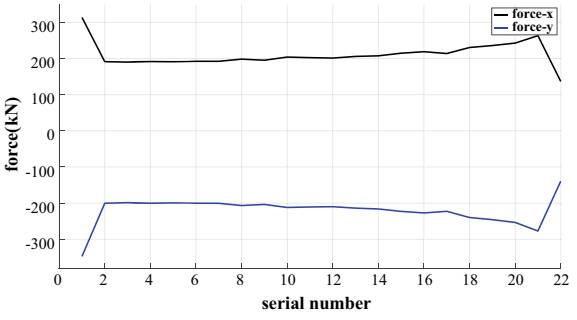


Fig. 3 Wind load situation of each guide vane at the first corner

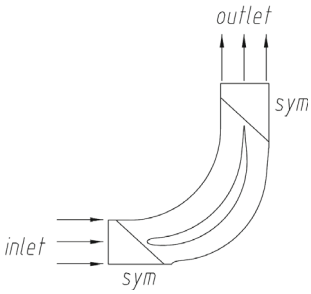


Fig. 4 Local flow field aerodynamic profile

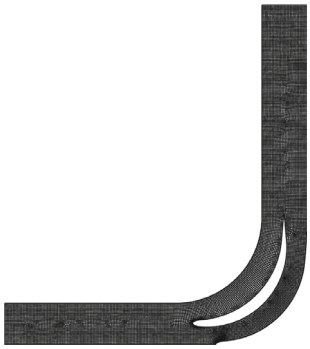


Fig. 5 Grid of local flow field

vane, and the wind load acting on it in the x and y directions are extracted. The calculation domain of local flow field is shown in Fig. 4, and the grid is shown in Fig. 5. The inlet and outlet of the actual calculation domain are slightly extended. The symmetry plane is used to simulate the fluid boundary.

In the three-dimensional flow field, the surface of the structure of guide vane is subjected to a force of 343.37 kN in the X direction and a force of − 379.47 kN in the Y direction, which is 8.5% different from the calculated results of the full-flow field. Therefore, it is reasonable to use the three-dimensional local flow field of the first guide vane for fluid–structure coupling calculation.

Figures 6 and 7 provides a three-dimensional airflow vector diagram and pressure distribution cloud map. From the picture, it can be seen that when the airflow passes through the guide vane, the windward side of the head skin blocks the airflow forward, leading to airflow obstruction and a rapid decrease in flow velocity and a sharp increase in pressure. Subsequently, the airflow splits into two streams along the surface of the guide vane. One stream flows on the pressure side of the guide vane, gradually slowing down and increasing in pressure. The other stream flows on the suction side, rapidly increasing in flow velocity and then gradually slowing down, the corresponding pressure decrease rapidly and then increases slowly. The pressure is highest on the windward side of the head skin of the guide vane. A positive pressure zone near the pressure surface and a negative pressure zone near the suction surface

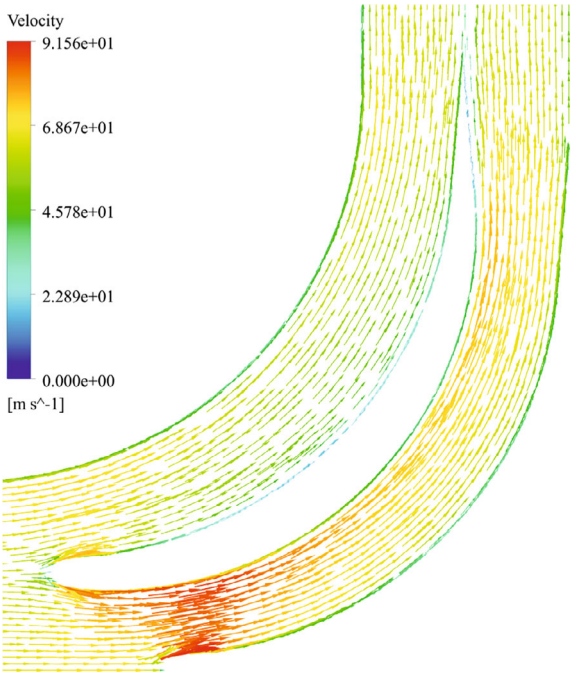


Fig. 6 Velocity vector distribution of local flow field

are formed, resulting in the guide vane being subjected to pressure from the pressure surface towards the suction surface.

4 Static Analysis of Guide Vane

Figure 8 provides the total displacement distribution of the structure of the first guide vane calculated through fluid–structure coupling. As can be seen from the picture, the maximum displacement of the structure is concentrated at the mid-span of the head skin, which is 4.75 mm. Combined with the pressure distribution on the surface of the guide vane, it can be concluded that the pressure in the windward area of the head skin is the main factor causing the maximum displacement at this location. Observing the overall deformation trend of the structure, it can be seen that the displacement at both ends of the structure is the smallest, and the displacement gradually increases from both ends towards the mid-span. There is a bending at the mid-span position, which is due to the support set on the upper and lower end faces of the structure of guide vane to limit its displacement, while the airflow applies a load from the pressure surface to the suction surface.

Figure 9 provides the stress distribution of the structure of guide vane obtained through fluid–structure coupling calculation. The maximum stress of the structure occurs at the position where the top end face of the guide vane contacts the support.

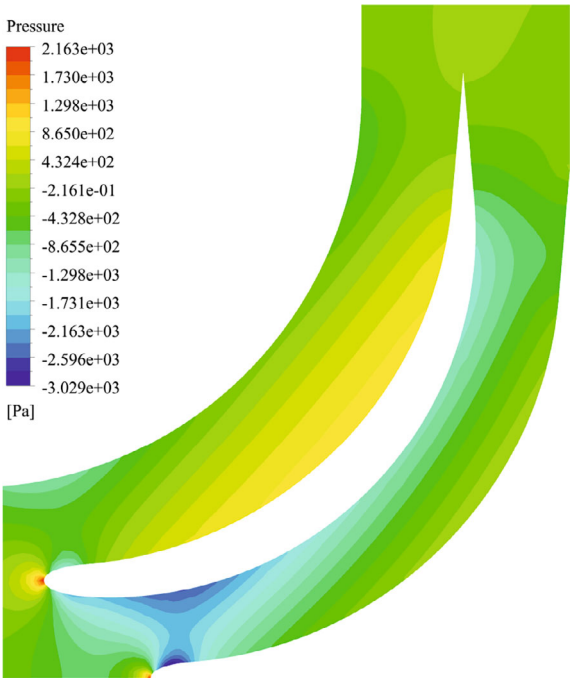


Fig. 7 Pressure distribution of local flow field

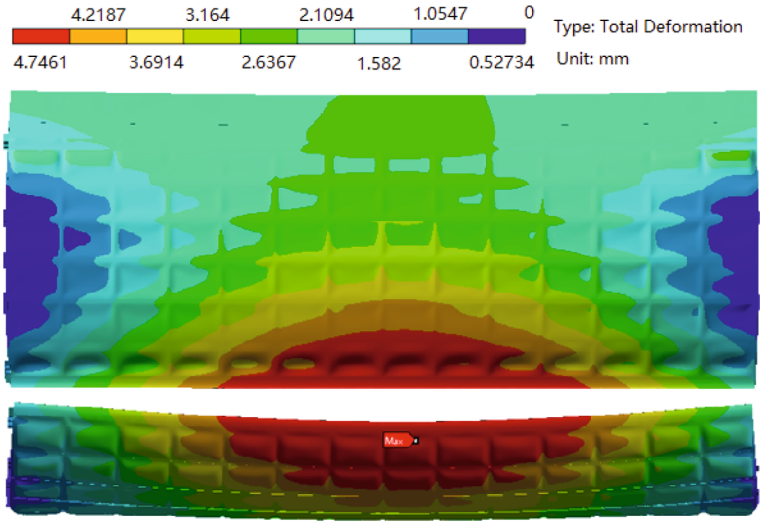


Fig. 8 Distribution of total displacement of guide vanes

As shown in Fig. 10 the maximum stress is 75.84 MPa. With a safety factor of 1.34, the allowable stresses for Q235 and Q355 are 175 and 265 MPa. Since the maximum stress of the structure is less than the allowable stress of the material, the structure meets the strength requirements under wind load.

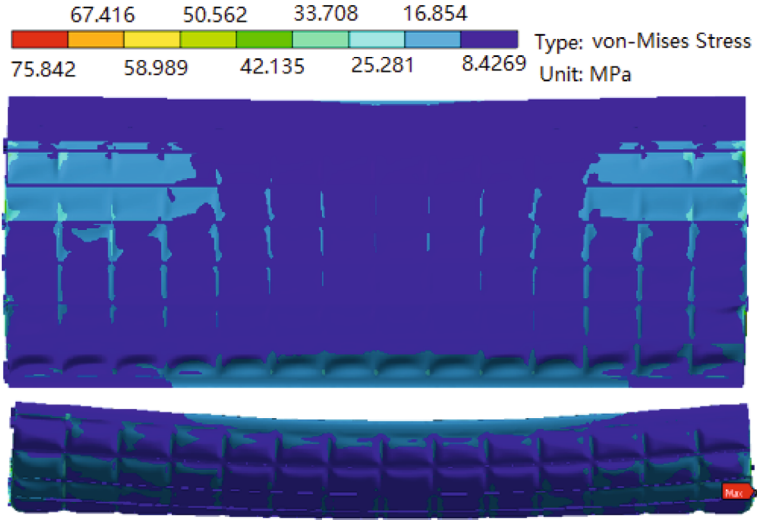
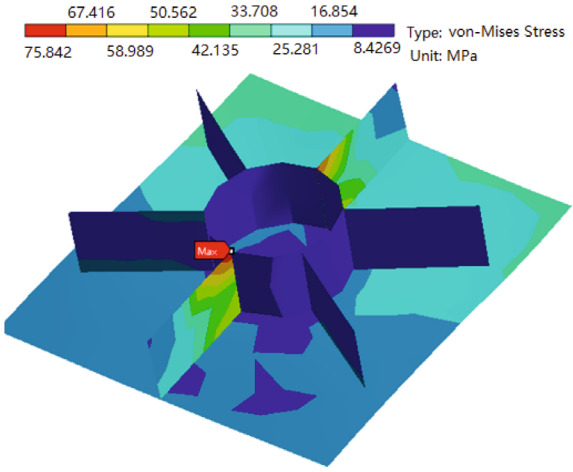


Fig. 9 Stress distribution of guide vanes

Fig. 10 Partial enlarged detail of stress



5 Conclusions

The fluid–structure coupling algorithm is used to simulate the airflow field and conduct static analysis of the structure of the first corner guide vane in a large wind tunnel. The following conclusions are drawn. When the airflow blows through the guide vane the flow velocity on the pressure side is relatively small compared to the suction side. The overall trend of the airflow velocity is to first increase rapidly and then decrease slowly. The airflow has the highest pressure on the windward side of the head skin of the guide vane, the positive pressure zone on the pressure side and the negative pressure zone on the suction side are formed. This is the main factor causing the largest deformation at the head skin of the guide vane. The maximum stress of the structure is less than the allowable stress of the material, which meet the strength requirements. On the basis of the above research, the alternating load on the structure when the wind blows through the surface of the guide vane can be studied and the fatigue life under this load can be evaluated.

Acknowledgements This work is supported by National Natural Science Foundation of China under Grant No.52275088 and the Fundamental Research Funds for the Central Universities under Grant (No.DUT22LAB506, No.DUT22LAB507).

References

1. The Training Teaching Material Editing Work Committee of the Chinese People's Liberation Army General Equipment Military (2003) Low speed wind tunnel experiment. National Defense Industry Press, Beijing, pp 42–43
2. Peiqing L, Yue C, Jin Z et al (2023) Numerical study on the influence of corner expansion ratio and guide vane installation angle on the flow field quality of BHAW wind tunnel. J Aerodyn 41(09):82–95
3. Li J, Chen B, Wu Z (2023) Aerodynamic and acoustic properties of a vertical suction wind tunnel guide vane. J Aerosp Eng 36(2). https://doi.org/10.1061/jaeeez.aseng-4687
4. Zhi-Gang Y, Ning M, Qi-Liang L et al (2007) Effect of solidity on wind tunnel corner vane losses. In: Proceedings of the 5th international conference on nonlinear mechanics
5. Lin Z, Can X, Yong H et al (2015) Research on flow control drag reduction in the corner section of low-speed wind tunnel. J Wuhan Univ Technol (Traffic Science and Engineering Edition) 39(1)

Open Access This chapter is licensed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License (http://creativecommons.org/licenses/by-nc-nd/4.0/), which permits any noncommercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if you modified the licensed material. You do not have permission under this license to share adapted material derived from this chapter or parts of it.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.



Research on Dynamic Envelope Calculation and Application of Metro Engineering Vehicles



Jinmei Xiao, Xuan Cai, Xiaobo Song, and Juntao Wang

Abstract Metro engineering vehicle is the key equipment to ensure the safe operation of metro system, and its overall dimension design is subject to strict restrictions imposed by metro gauge. This paper takes a specific metro engineering vehicle model as the object and 3 cross-sections susceptible to danger from the object for dynamic envelope calculations with two methods, the CJJ 96 standard-based method and the vehicle dynamic simulation-based method. The CJJ 96 standard method requires calculation of all possible offsets of the vehicle during operation and superimposition with vehicle contour, while the vehicle dynamic simulation method requires a multi-body dynamics model of the metro engineering vehicle and dynamics simulation thereof to obtain the vehicle's dynamic offset for dynamic envelope calculation. The dynamic envelope calculated with the CJJ 96 standard method exceeds the vehicle gauge and the metro engineering vehicle does interfere with the structures and equipment in actual operation, which means that the CJJ 96 standard method delivers results reflecting better engineering reality. However, the dynamic envelopes obtained by vehicle dynamics simulation method are all within the vehicle gauge, which is inconsistent with the actual situation. The metro engineering vehicle's overall dimensions are modified according to the results obtained with the CJJ 96 standard method. The modified envelope meets the requirements of vehicle gauge and no interference occurs in subsequent operation, which proves the effectiveness of the CJJ 96 standard method in dynamic envelope calculation.

Keywords Metro engineering vehicle · Gauge · Dynamic envelope · Dynamics

J. Xiao (✉) · X. Cai · X. Song
School of Automobile and Transportation, Chengdu Technological University, Chengdu, Sichuan, China
e-mail: xjmei2024@163.com

J. Wang
School of Mechanical Engineering, Southwest Jiaotong University, Chengdu, Sichuan, China

© The Author(s) 2026

H. Saman et al. (eds.), *The 9th International Conference on Advances in Construction Machinery and Vehicle Engineering*, Lecture Notes in Mechanical Engineering,
https://doi.org/10.1007/978-981-96-5527-4_38

1 Introduction

Metro engineering vehicles are widely used in urban metro overhaul and maintenance and are the key equipment to ensure the safe operation of metro systems. Metro engineering vehicles are subject to strict restrictions imposed by the metro gauge during operation. Therefore, comprehensive consideration shall be given to their geometric dimensions in the design process so as to avoid conflicts with the metro gauge. As the metro gauge is usually predetermined, the possible offset of the vehicle during operation shall be fully considered in the metro engineering vehicle design, that is, it is necessary to plot the vehicle's dynamic envelope. The International Union of Railways uses UIC 505-1 for railway gauge, and the railway gauge specified in this series of standards is dynamic [1]. In vehicle gauge calculation, based on the parameters of the vehicle subject to calculation, the calculated reduction is added to the vehicle's reference profile outwards to obtain the vehicle's maximum structural profile, and then the transverse offset and the maximum protruding contour value are added to obtain the vehicle's dynamic envelope. According to CJJ 96, the metro gauge standard currently adopted in China, the dynamic envelope is calculated with consideration on various influencing factors, and manufacturing-caused errors and car body offsets caused by various dynamic errors are included in gauge calculation [2]. To calculate vehicle dynamic envelope efficiently and accurately, Luo et al. proposed a metro vehicle gauge calculation method based on full dynamic envelope, which takes more displacement elements into consideration and will make better use of the space provided by metro gauge [3]. Hu et al. used MATLAB software to simulate and calculate the dynamic counters of metro vehicles running on actual railways to obtain the vehicles' dynamic envelopes more conveniently [4]. Dong Xiao et al. proposed a dynamic envelope correction algorithm based on accurate displacement and inclination detection techniques [5]. Zhang compared and analyzed the methods given in UIC 505 and CJJ 96 and the dynamics method, and concluded that the calculation results obtained with the UIC 505 and CJJ 96 standard methods were comparatively conservative and closer to the ultimate limit state [6]. Yi et al. carried out systematic research on the calculation principle and method of vehicle dynamic gauge in EN 15273, and provided illustrations with detailed calculation cases [7]. Based on the dynamics method, Men et al. analyzed the influence of crosswind speed on dynamic offset by taking AH metro vehicle as an example [8]. Ye et al. designed a comprehensive urban rail gauge design software based on the secondary development of AutoCAD2018-2022 drawing software to solidify the gauge design process and method [9]. Pandya studied computational geometry of fundamental interactions in static clearances [10].

In this paper, an actual metro engineering vehicle is taken as an example to compare and analyze different dynamic envelope calculation methods so as to provide a reference for the geometric dimension design of metro engineering vehicles.

2 Dynamic Envelope Calculation Methods

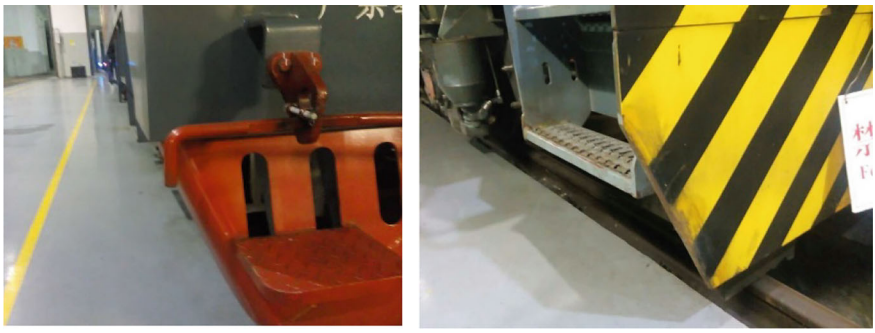
A company designed a metro engineering vehicle, but some components on the car body interfered with surrounding structures and equipment during vehicle operation, as shown in Fig. 1, indicating that the vehicle's dynamic envelope exceeded the vehicle gauge. To figure out the cause of such exceeding and formulate a vehicle dimensions modification scheme, this paper will use the CJJ 96 standard method and the dynamics simulation method to calculate the dynamic envelope of the metro engineering vehicle.

2.1 Calculation Method Based on CJJ96 Standard

According to CJJ96, the first step is to take a cross-section susceptible to danger on the vehicle and obtain the control point coordinates on the cross-section. After that, all possible offsets of the control points on each cross-section shall be calculated and then superimposed to the control point coordinate data to obtain new coordinate positions, which shall all be included in the dynamic envelope.

In CJJ96 standard, the following factors shall be considered for offset calculation:

(1) Offset caused by the vehicle's manufacturing error;
(2) Offset caused by the vehicle's maintenance limit;
(3) Vertical offset caused by the deflection difference between the vehicle's empty and loaded states;
(4) Lateral deflections caused by vehicle manufacturing and load asymmetry;
(5) Offset caused by the side roll of the vehicle's primary and secondary suspensions;
(6) Offset caused by geometric deviation, wear and elastic deformation of the track;



(a) cowcatcher (b) ladder

Fig. 1 Vehicle interfering with structures and equipment

(7) Displacement of bogie frame relative to wheelset;
(8) Relative displacement between car body and bogie frame;
(9) Car body yaw offset caused by the most unfavorable position on the track;
(10) Yaw deflection of the bogie wheelset caused by the most unfavorable running position on the track.

Taking the car body as an example, its transverse offset ΔXBP is calculated using Eq. (1):

$$\Delta X_{BP} = \left[\sqrt{\left(\frac{l-d}{2}\right)^2 + \Delta c^2} + (\Delta q_1 + \Delta q_2 + \Delta q_3) + (\Delta\omega_1 + \Delta\omega_2)\right]\frac{2n+a}{a} + \Delta e + \Delta M + \Delta X' + \Delta X_{B-xgpx} + \Delta X_{Bcp} \quad (1)$$

where:

l	the maximum track gauge including inner rail wear
a	the distance between bogie centers
d	the minimum outer distance of wheelset flange with the maximum wear
n	the distance from the cross-section under calculation to the adjacent central pin
Δc	the transverse position difference of the track center line
Δq1	transverse play clearance of bearings used in axle boxes
Δq2	offset caused by transverse elasticity of the wheel
Δq3	offset caused by transverse elasticity of primary suspension
Δw1	offset caused by radial clearance and wear of bogie center pin
Δw2	transverse elastic deformation of bogie secondary spring relative to nominal center position
Δe	transverse elastic deformation of track
ΔM	transverse manufacturing error value
ΔX′	offset compensation during curve widening check
ΔXB-xgpx	transverse offset of car body caused by suspension fault
ΔXBcp	transverse offset of car body lateral inclination.

2.2 Method Based on Dynamics Simulation

Based on multi-rigid body dynamics, a vehicle dynamics model may be established to simulate the actual motion attitude of the vehicle, so as to calculate the dynamic displacement of each position on the car body when the vehicle is running on the track. In the dynamics calculation method, the dynamic offset is the dynamic response of the moving vehicle to external excitation such as track irregularity and crosswind.

A metro engineering vehicle is a complex multi-body system composed of wheelsets, axle boxes, frames, car bodies and other components, which are connected by suspension systems and articulation systems. When building the dynamics model, since the elasticity of wheelset, frame, car body and other components is much smaller than that of suspension systems, they are all regarded as rigid bodies. In addition, the action of adjacent vehicles and the elastic deformation of rails are not considered. Because of non-linear factors in the bogie system of metro engineering vehicles, the vertical and transverse motions are closely coupled. Therefore, the transverse and vertical motions are integrated as a large system in the calculation model. The main parameters required for the dynamics model are shown in Table 1. The dynamics calculation is carried out by SIMPACK software. The multi-rigid body system dynamics model of the metro engineering vehicle built in the SIMPACK environment is shown in Fig. 2.

Table 1 Main parameters for dynamics calculation

No.	Parameter	Value
1	Distance between bogie centers (mm)	9800
2	Inflexible wheel base between bogies (mm)	2200
3	Transverse clearance of primary spring (mm)	2010
4	Transverse clearance of secondary spring (mm)	1910
5	Vertical rigidity of primary spring of each axle box (N/mm)	889
6	Vertical rigidity of secondary spring on each side of bogie (N/mm)	7830
7	Car body weight (kg)	31,834

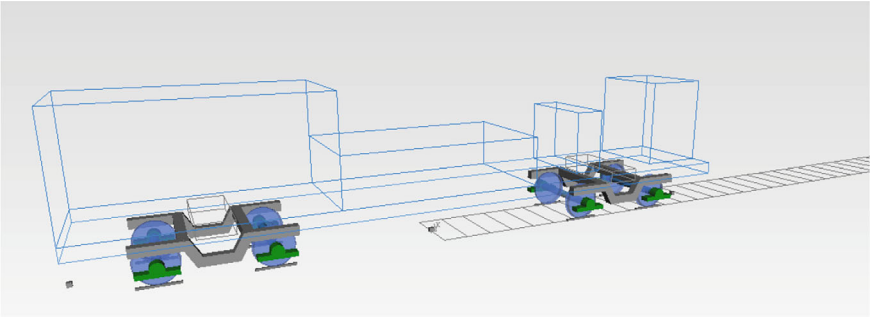


Fig. 2 Dynamics calculation model of metro engineering vehicle

3 Calculation of Dynamic Envelope

3.1 Selection of Cross-Section for Calculation and Vehicle Gauge

In dynamic envelope calculation, the geometric dimensions and dynamic offsets vary for each vehicle cross-section. Therefore, it is necessary to calculate the dynamic envelopes of multiple cross-sections. Usually, the critical cross-sections and those susceptible to danger are selected for calculation. In this paper, according to the structural characteristics of the metro engineering vehicle, 3 cross-sections in the front, middle and rear respectively are selected, as shown in Fig. 3. The views of the cross-sections taken are shown in Fig. 4.

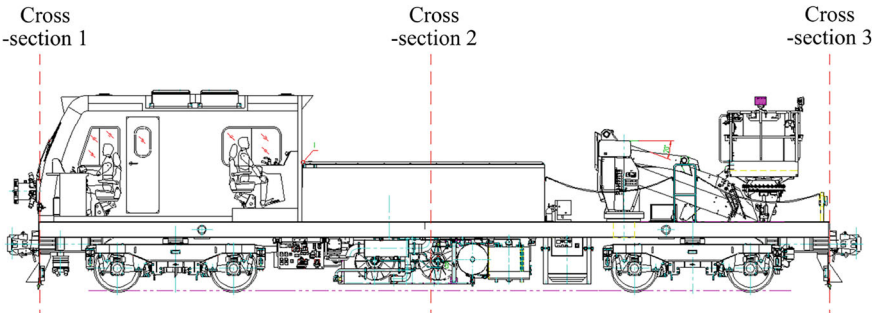
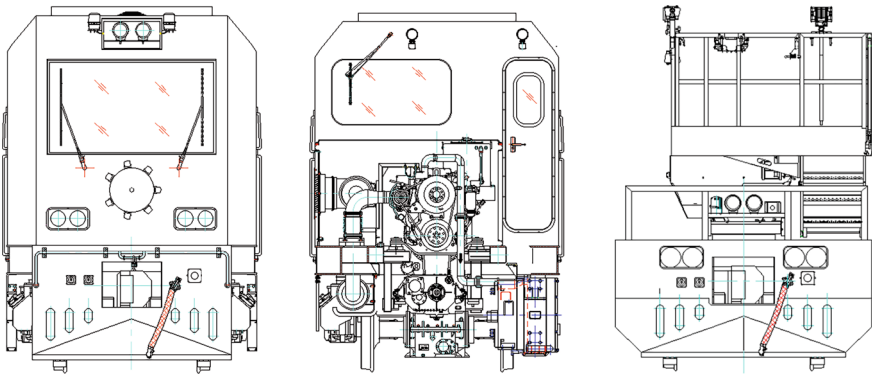


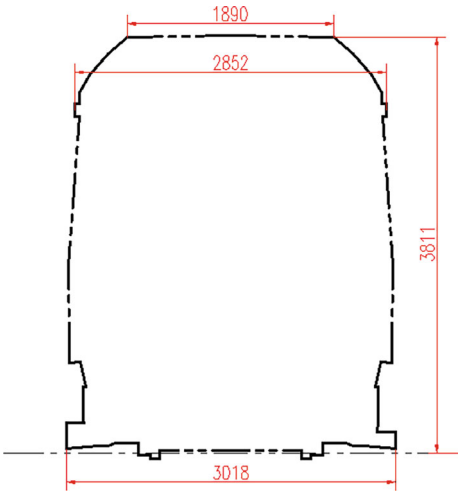
Fig. 3 Cross-sections taken for metro engineering vehicle calculation



(a) Cross-section 1 (b) Cross-section 2 (c) Cross-section 3

Fig. 4 Cross-section views of metro engineering vehicle

Fig. 5 Vehicle gauge



In order to ensure safe operation of the metro engineering vehicle, the dynamic envelope shall be within the vehicle gauge. The vehicle gauge of the vehicle operation section considered in the calculation of this paper is as shown in Fig. 5.

3.2 Results of CJJ96 Standard-Based Calculation

Using the dynamic envelope calculation method given in CJJ96, calculate the offset of each vehicle cross-section's control points and superimpose with the control points to plot out the dynamic envelope of each cross-section, as shown in Fig. 6. It can be seen from the figure that the dynamic envelope of Sect. 1 has exceeded the vehicle gauge at the lower right and other positions, which is completely consistent with the component interference on site. This indicates that the CJJ96 standard method accurately reflects the maximum envelope of vehicles running on the track. The dynamic envelopes of Sects. 2 and 3 are within the vehicle gauge, indicating that there will be no interference with structures and equipment at corresponding positions of vehicles.

3.3 Results of Dynamics Simulation Calculation

The dynamics simulation is carried out at a speed of 80 km/h by using the metro engineering vehicle's dynamic model, where the Level-5 track irregularity spectrum of the United States is adopted. The vehicle's dynamic envelope is obtained through calculation with the dynamic offsets of the vehicle's 3 cross-sections and the result

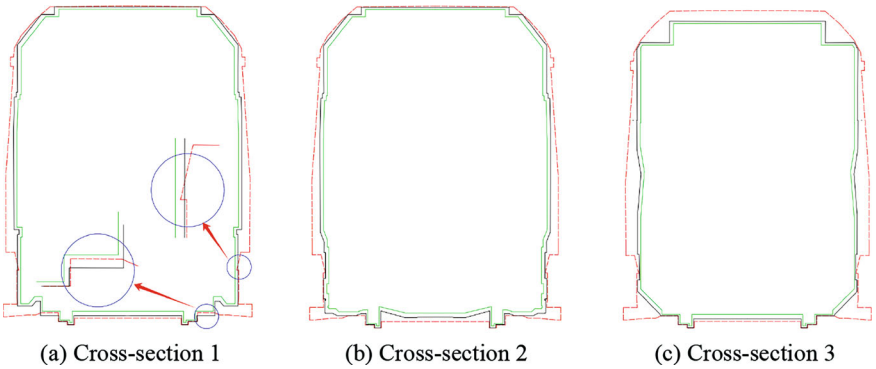


Fig. 6 Result of dynamic envelope calculation using CJJ96 standard-based method (cross-section outline in green, dynamic envelop in black, vehicle gauge in red)

is shown in Fig. 7. It can be seen from the figure that the dynamic envelope obtained by dynamics simulation is completely within the vehicle gauge. However, this is definitely inconsistent with the actual situation, indicating that the dynamic envelope obtained by dynamics simulation method is smaller than the actual envelope. This is mainly because that the most unfavorable working condition is not realized in the vehicle dynamics simulation. For this reason, the corresponding dynamic offset of each cross-section control point is comparatively small and the dynamic envelope obtained by simulation is smaller than that in actual vehicle operation.

Therefore, through the dynamic envelope calculation based on CJJ96 standard and dynamics simulation respectively, and after comparison with the actual operation state of vehicle, the author believes that the CJJ96 standard-based method or any other similar method should be preferred for calculation of vehicle dynamic envelope, and the results obtained thereby are more reliable and better fit with the actual situation, while the dynamics simulation method is only recommended as an auxiliary mean.

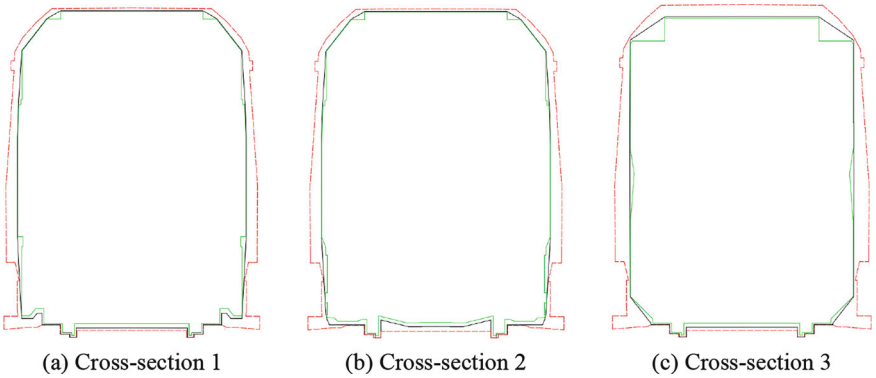


Fig. 7 Result based on dynamics simulation

4 Application of Dynamic Envelope Calculation Results

According to the calculation results before, it can be seen that the dynamic envelope of vehicle Sect. 1 has exceeded the vehicle gauge, and there is indeed collision interference with structures and equipment during the actual vehicle operation. Therefore, in order to make the metro engineering vehicle meet the use requirements, it is necessary to rectify the geometric dimensions of vehicle Sect. 1.

According to the results of CJJ96 standard-based calculation, the over-limit area of vehicle Sect. 1 is mainly located at the installation position of car ladder and cowcatcher. Therefore, the manufacturer has modified the car ladder and cowcatcher positions of the metro engineering vehicle. The specific modification is shown in Fig. 8. The modified vehicle is then subject to another calculation according to CJJ96 standard, and the calculation result is shown in Fig. 9. It can be seen that the dynamic envelope of the modified vehicle is completely within the vehicle gauge. When the modified vehicle runs on the track, there is no interference with structures or equipment. It shows that the modification is very effective and the dynamic envelope calculation method based on CJJ96 standard guides the vehicle design and improvement effectively.

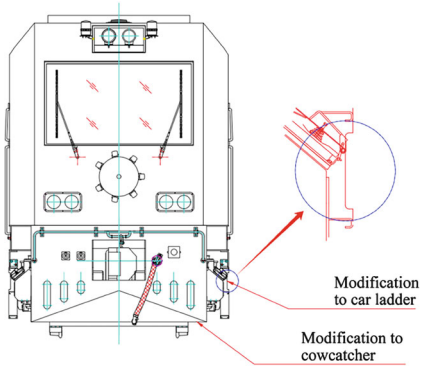


Fig. 8 Modification of vehicle geometry (cross-section outline in green, dynamic envelop in black, vehicle gauge in red)

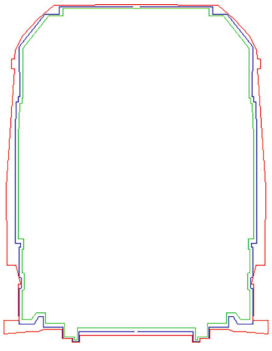


Fig. 9 Calculation results after modification (cross-section outline in green, dynamic envelop in black, vehicle gauge in red)

5 Conclusion

In this paper, 2 vehicle dynamic envelope calculation methods are compared and analyzed, and through comparison with the actual operation state of vehicle, the following conclusions are obtained:

(1) In the dynamic envelope calculation based on CJJ96 standard, the transverse offset of vehicle control points is superimposed under various unfavorable working conditions, thus reflecting the influence of various factors on the dynamic envelope to the greatest extent. The calculation results are basically consistent with the vehicle's actual dynamic envelope in operation. This method may identify the unreasonable design of vehicle geometric dimensions and guide the vehicle design improvement effectively.
(2) Because the dynamics simulation does not simulate the most unfavorable working conditions nor take the dimensional deviation caused by manufacturing errors and other factors into consideration, the dynamic envelope calculated thereby is smaller than that in actual vehicle operation. Therefore, it may not identify the unreasonable design of vehicle geometric dimensions effectively or accurately guide the vehicle design and improvement. Considering the conclusions above, it is suggested that the CJJ96 standard-based method or any other similar method should be preferred for vehicle dynamic envelope calculation.
(3) The dynamic simulation model of this paper did not consider the fault conditions of metro engineering vehicles, so the calculation results have certain limitations. In future research, the combination of vehicle fault conditions should be considered to make the calculated results more consistent with actual conditions.

Funding The research project of Chengdu Technological University (2021DF001).

References

1. UIC 505-1 (1977) Railway transport stock-rolling stock construction Ganges
2. CJJ/T 96-2018 (2018) Standard for Metro Gauge. China Architecture & Building Press, Beijing
3. Luo X (1997) Study on the fully dynamic envelope curve of subway car limiting Gauge. Rolling Stock 9:40–44
4. Hu P (2021) Research on method of obtaining vehicle dynamic envelope based on MATLAB and three-dimensional geometric transformation. Hoist Convey Machin 4:51–57
5. Dong X (2023) A correction algorithm for dynamic envelope of metro train based on detection technology. Rail Transp Equip Technol 1:20–24
6. Zhang Y (2020) Research on calculation method of rail vehicle Gauge. Dalian Jiaotong University
7. Yi X (2024) Research on calculation method of vehicle dynamic gauge based on EN 1527. Electric Locomotives Mass Transit Veh 47(3):66–71
8. Yonglin M (2024) Research on AH-model vehicle clearance based on dynamics simulation calculation. Urban Mass Transit 27(05):40–44
9. Ye G (2023) Research on software for clearance design of urban rail transit. Railway Technical Standards (Chinese and English) 5(12):8–16

10. Pandya A (2019) Exploring computational geometry of fundamental interactions in static clearance. Phys Essays 32(3):411–416

Open Access This chapter is licensed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License (http://creativecommons.org/licenses/by-nc-nd/4.0/), which permits any noncommercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if you modified the licensed material. You do not have permission under this license to share adapted material derived from this chapter or parts of it.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.



Ground Point Cloud Segmentation Method in Open-Pit Mine Scenarios Based on Quadtree Voxel Partition



Ruoyao Li, Yafei Wang, Shi Sun, Yichen Zhang, Mingyu Wu, and Zexing Li

Abstract In open-pit mines, segmentation of ground point clouds is crucial for the traversability analysis and path planning of autonomous mining trucks. However, mainstream point cloud segmentation methods often suffer from under-segmentation problems due to the heterogeneity of local features in mining environments and the indistinct boundaries between massif and ground. To address the above issues, we propose a novel ground point cloud segmentation method based on quadtree voxel partition. First, the method divides the point cloud of the single frame into voxels using a quadtree structure and adaptively adjusts the voxel composition through point cloud distribution. Subsequently, principal component analysis (PCA) is introduced for iterative parameters optimization and plane estimation within voxels. Finally, a multidimensional function is designed based on the verticality and flatness of the point cloud for likelihood estimation, thereby achieving ground point cloud segmentation. We evaluate the proposed method assessing precision and recall rates under field tests, and the results indicate that the proposed point cloud segmentation method achieves over a 5% improvement in precision compared with mainstream methods.

Keywords Lidar · Open-pit Mine · Quadtree voxel partition · Ground segmentation · Likelihood estimation

R. Li · Y. Wang (✉) · S. Sun · Y. Zhang · M. Wu · Z. Li
School of Mechanical Engineering, Shanghai Jiao Tong University, Shanghai, China
e-mail: wyfjlu@sjtu.edu.cn

Y. Wang
National Engineering Laboratory of Automotive Electronic Control Technology, Shanghai Jiao Tong University, Shanghai, China

© The Author(s) 2026

H. Saman et al. (eds.), *The 9th International Conference on Advances in Construction Machinery and Vehicle Engineering*, Lecture Notes in Mechanical Engineering,
https://doi.org/10.1007/978-981-96-5527-4_39

1 Introduction

Segmenting the ground of open-pit mines to delineate drivable areas is the key to achieving effective and safe route planning for autonomous mining trucks. 3D LiDAR has been widely used in ground segmentation tasks due to its high measurement accuracy, wide perception range, and strong robustness, which can provide environmental information characterized by dense point clouds [1]. Unlike structured environments represented by urban or high-way scenarios, there are significant differences in local point cloud features between different regions of open-pit mines [2]. In addition, mining environments will also dynamically change due to excavation operations [3]. Therefore, ground point cloud segmentation of autonomous mining trucks not only requires consideration of point cloud features in different local regions but also needs to meet the real-time requirements of mining operations.

At present, segmentation of ground point clouds mainly adopts model-based, feature-based, and learning-based approaches. Model-based methods are commonly used in structured scenarios with low undulation, where the entire region can be segmented based on models such as Random Sampling Consistency (RANSAC) [4]. Chen et al. introduce Gaussian processes to estimate the ground in point clouds [5]. Byun et al. adapt Markov random fields for point cloud segmentation of drivable areas [6]. Model-based methods are difficult to balance segmentation accuracy and runtime simultaneously and may overlook local features of point clouds. To improve the efficiency and segmentation accuracy of the algorithm, some methods improve the point cloud segmentation process based on geometric features. Himmelsbach et al. [7] perform line fitting based on elevation to segment the ground point cloud. Lim et al. [8] and Lee et al. [9] propose the patchwork series method, which effectively segments ground point clouds through concentric zone voxel segmentation, region-wise ground plane fitting, and adaptive ground likelihood estimation. Although the feature-based methods can adopt region partition to consider local point cloud features, the pattern of region partition is mainly manually predetermined, making it difficult to cope with open-pit mine scenarios with large changes in massif environments and significant differences between local features. In recent years, some methods have introduced deep learning to segment point clouds. Milioto et al. propose the RangeNet++ network for labeling the different components in the point cloud [10]. In GndNet [11], ground point clouds are identified in real-time by estimating ground elevation information based on grid representation. Deep learning based methods require a large amount of computing resources and often need precollection of training data [12], making them difficult to apply in open-pit mine scenarios.

In this paper, we propose a quadtree voxel partition based ground point cloud segmentation method to improve the efficiency and accuracy of segmentation process. The main contribution of the paper lies in utilizing the quadtree structure to adaptively carry out voxel partitioning based on point cloud distribution, achieving the consideration of the feature heterogeneity in different areas. Iterative parameter estimation is performed by principal component analysis (PCA) within each voxel,

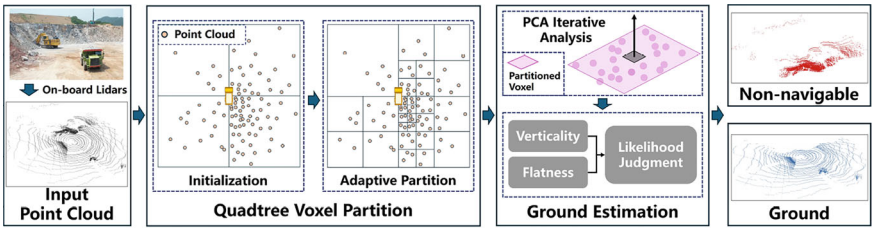


Fig. 1 Pipeline of the proposed ground segmentation method

and the multidimensional evaluation is designed based on the verticality and flatness to segment the ground point cloud. The results indicate that our method performs better computation time and segmentation accuracy for open-pit mine scenarios.

2 Methodology

2.1 Pipeline of Segmentation

The structure of the proposed ground point cloud segmentation method is illustrated in Fig. 1. This method comprises three main components: adaptive quadtree voxel segmentation, iterative PCA-based analysis of feature parameters, and point cloud likelihood judgment. The input, a three-dimensional LiDAR point cloud, undergoes these three stages to complete the segmentation of the ground point cloud, resulting in outputs of the ground point cloud and the non-navigable point cloud.

2.2 Quadtree-Based Adaptive Voxel Partition

For the point cloud input $\mathcal{P} = \{p_1, p_2, \ldots p_n\}$ with a perceived distance of r, we first initialize the root node V_0^0, which includes the square area of $\pm r$ within the x and y axes. Then, V_0^0 is divided into four leaf nodes $V_i^1, i = 0, 1, 2, 3$. Based on the partition of V_i^k, each node is recursively partitioned:

$$V_i^k = \bigcup V_j^{k+1} | \forall j \in 0, 1, 2, 3 : \text{num}(V_i^k) > \alpha_1, d(k) < \kappa, \text{num}(V_j^{k+1}) > \alpha_2 \quad (1)$$

where V_j^{k+1} denotes the leaf node partitioned from V_i^k, and $\text{num}(V_*)$ is the number of points in the voxel V_*, $d(k)$ is the depth of the depth of the quadtree structure after completing the kth partition. α_1 and α_1 are the quantity thresholds that specify the minimum number of points used for PCA, κ is the structure depth threshold

designed to control the complexity of quadtree structure. Further segmentation into four sub-voxels is only accessible if the voxels have enough points.

$\mathcal{P}$ can adaptively be partitioned into quadtree voxels based on point cloud distribution until one of the quantity or depth thresholds is not satisfied. Adopting the quadtree structure improves voxel retrieval speed and can adaptively determine whether to partition leaf nodes based on point cloud distribution. For point clouds with niform distribution, quadtree voxel partition can achieve a time complexity close to $O(n \log n)$. Even in the most complex case, the time complexity of quadtree partition does not exceed $O(n^2)$, which is better than that in the partition of Patchwork [8].

2.3 Iterative Estimation and Likelihood Judgement

For the input point cloud, the estimated plane normal vector $\boldsymbol{n}$ and the plane equation can be obtained through principal component analysis (PCA):

$$\boldsymbol{n}^T \cdot \overline{p} = -d \tag{2}$$

where $\overline{p}$ represents the centroid of the point cloud within the voxel.

Due to the sensitivity of the PCA method to outliers, we introduce the iterative approach like Patchwork++ [9] to estimate the ground point cloud. The initial ground estimation can be expressed as:

$$\widehat{\mathcal{G}}^0 = \{\, p_k \in V_{i,j,k,m} \Big| z(p_k) < z^{init}_{i,j,k,m} \,\} \tag{3}$$

where $z(p_k)$ represents the elevation of point p_k in the z-axis. $z^{init}_{i,j,k,m}$ is the initial height threshold.

When performing iterative estimation, the plane parameter $\boldsymbol{n}^0$, d^n and $\overline{p}^n$ is first obtained based on $\widehat{\mathcal{G}}^0$. Then, assuming that the plane estimation parameter obtained from the nth iteration is $\boldsymbol{n}^n$, d^n and $\overline{p}^0$, the estimation result of $n+1$th iteration can be obtained:

$$\widehat{\mathcal{G}}^{n+1} = \{p_k \in V_{i,j,k,m} \Big| \Big| d^n + (\boldsymbol{n}^n)^T p_k \Big| < \beta \,\} \tag{4}$$

β denotes the filter threshold. We adopt three iterations of PCA to eliminate the influence of discrete points on estimation and obtain robust algorithm performance. Therefore, our method designates $\boldsymbol{n}^3$, d^3 and $\overline{p}^3$ as the features for voxel $V_{i,j,k,m}$.

We introduce likelihood judgment to segment the type of point cloud in voxels based on the rough estimation $\widehat{\mathcal{G}}_{i,j,k,m}$, $\hat{\mathcal{N}}_{i,j,k,m}$ for ground and non-navigable areas. Assuming that the observation corresponding to a voxel is $\mathcal{X}\ominus_f$, and the corresponding feature parameter is $\theta\ominus_f$, the likelihood function can be constructed that

satisfies:

$$\hat{\mathcal{N}}_{i,j,k,m} = V_{i,j,k,m} - \widehat{\mathcal{G}}_{i,j,k,m} \tag{5}$$

$$L_{\mathcal{G}}(\theta_f \left| \mathcal{X}_{\ominus_f}\right.) = \Phi_{\mathcal{G}}(\theta_f) \cdot \Psi_{\mathcal{G}}(\theta_f)\, L_{\mathcal{N}}(\theta_f \left| \mathcal{X}_{\ominus_f}\right.) = \Phi_{\mathcal{N}}(\theta_f) \cdot \Psi_{\mathcal{N}}(\theta_f) \tag{6}$$

$\Phi_*(\theta_n)$ is the probability function for verticality, and $\Psi_*(\theta_n)$ is the probability function for flatness. Verticality expresses the degree of inclination of the estimation plane. The flatness function is calculated based on the average elevation and elevation distribution of point clouds within voxels. Based on $L_{\mathcal{G}}(\theta_f \left| \mathcal{X}_{\ominus_f}\right.)$ and $L_{\mathcal{N}}(\theta_f \left| \mathcal{X}_{\ominus_f}\right.)$, the final segmentation result can be determined:

$$\widehat{\mathcal{G}} = \{\widehat{\mathcal{G}}_l \left| L_{\mathcal{G}}(\theta_f \left| \mathcal{X}_{\ominus_f}\right.) > \mu_{\mathcal{G}}, L_{\mathcal{N}}(\theta_f \left| \mathcal{X}_{\ominus_f}\right.) < \mu_{\mathcal{N}}\right.\} \tag{7}$$

where $\mu_{\mathcal{G}}$ and $\mu_{\mathcal{N}}$ are the thresholds for classifying ground and non-navigable point clouds.

3 Experiments

To validate our method, field tests in open-pit mines were carried out using autonomous mining trucks. The truck is equipped with three 32-line LiDARs (RoboSense RS-32) positioned in the front left, front right, and rear. The data from three LiDARs are calibrated and fused as the input for the segmentation algorithm.

We denote the ground truth and estimation result of the ground point cloud and non-navigable point cloud as $\mathcal{G}$, $\mathcal{N}$, $\widehat{\mathcal{G}}$ and $\hat{\mathcal{N}}$. The results of segmentation can be categorized into $TP = \widehat{\mathcal{G}} \cap \mathcal{G}$, $FP = \widehat{\mathcal{G}} \cap \mathcal{N}$, $TN = \hat{\mathcal{N}} \cap \mathcal{N}$, and $FN = \hat{\mathcal{N}} \cap \mathcal{G}$. Subsequently, the precision and recall rate of the segmentation can be defined as:

$$P_{precision} = \frac{TP}{TP + FP}\, P_{recall} = \frac{TP}{TP + FN} \tag{1}$$

The visualized comparison results are shown in Fig. 2, including five different mining scenarios of loading and transportation. The proposed method, RANSAC [4], LineFit [7], and Patchwork++ [9] are compared with ground truth. The gray part represents the ground, the blue part represents the non-navigable areas, and the red part represents the difference between the segmentation result and the ground truth.

The average accuracy, recall, and running time obtained from field tests in multiple scenarios are shown in Table 1. By accounting for variations in point cloud distribution, the proposed method outperforms mainstream approaches in ground segmentation tasks, achieving an accuracy 5.6% higher than Patchwork++ [9]. In addition, the quadtree voxel structure enhances the real-time performance of the proposed method, with a runtime of under 30ms, meeting the requirements of open-pit mines.

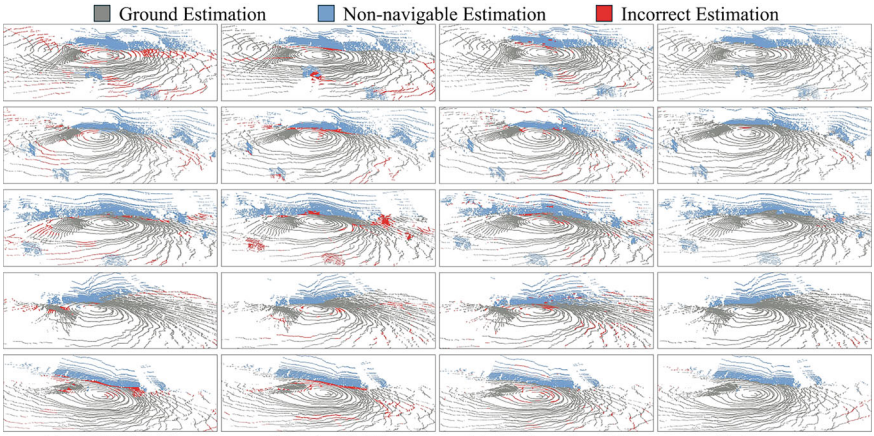


Fig. 2 Comparison of ground segmentation results in field tests

Table 1 Performance comparison with the mainstream methods

Method	Precision (%)	Recall (%)	Runtime (ms)
RANSAC [4]	85.92	73.38	21
LineFit [7]	83.19	75.21	29
Patchwork++ [9]	88.26	82.57	32
Proposed	93.88	92.93	20

4 Conclusion

This paper presents a novel quadtree voxel partition based ground segmentation method which iteratively employs estimation within each voxel to determine the ground areas through verticality and flatness. The experiment demonstrates that the proposed method outperforms other mainstream methods in computation time and segmentation accuracy. In the future, we will optimize voxel partition approaches to further improve the accuracy and real-time performance of ground segmentation.

Acknowledgements This work was Supported by the National Natural Science Foundation of China under Grant No. 52372417 and No. 52072243.

References

1. Liu X, Wang Y, Jiang K, Zhou Z, Nam K, Yin C (2022) Interactive trajectory prediction using a driving risk map-integrated deep learning method for surrounding vehicles on highways. IEEE Trans Intell Transp Syst 23(10):19076–19087

2. Meng D, Tian B, Pan Z, Cao D, Chen L (2021) Berm detection for autonomous truck in surface mine dump area. In: 2021 IEEE international intelligent transportation systems conference (ITSC), IEEE, pp 2829–2834
3. Li R, Wang Y, Li Z, Zhang Y, Zhang R (2024) Long-term map management in degraded scenarios based on pose reference and information density. In: 2024 IEEE intelligent vehicles symposium (IV), IEEE, pp 927–932
4. Chum O, Matas J (2008) Optimal randomized RANSAC. IEEE Trans Pattern Anal Mach Intell 30(8):1472–1482
5. Chen T, Dai B, Wang R, Liu D (2014) Gaussian-process-based real-time ground segmentation for autonomous land vehicles. J Intell Rob Syst 76:563–582
6. Byun J, Na KI, Seo BS, Roh M (2015) Drivable road detection with 3D point clouds based on the MRF for intelligent vehicle. In: Field and service robotics: results of the 9th international conference. Springer International Publishing, pp 49–60
7. Himmelsbach M, Hundelshausen FV, Wuensche HJ (2010) Fast segmentation of 3D point clouds for ground vehicles. In: 2010 IEEE intelligent vehicles symposium, IEEE, pp 560–565
8. Lim H, Oh M, Myung H (2021) Patchwork: concentric zone-based region-wise ground segmentation with ground likelihood estimation using a 3D LiDAR sensor. IEEE Robot Autom Lett 6(4):6458–6465
9. Lee S, Lim H, Myung H (2022) Patchwork++: fast and robust ground segmentation solving partial under-segmentation using 3d point cloud. In: 2022 IEEE/RSJ international conference on intelligent robots and systems (IROS), IEEE, pp 13276–13283
10. Milioto A, Vizzo I, Behley J, Stachniss C (2019) Rangenet++: Fast and accurate lidar semantic segmentation. In: 2019 IEEE/RSJ international conference on intelligent robots and systems (IROS), IEEE, pp 4213–4220
11. Paigwar A, Erkent Ö, Sierra-Gonzalez D, Laugier C (2020) GndNet: fast ground plane estimation and point cloud segmentation for autonomous vehicles. In: 2020 IEEE/RSJ international conference on intelligent robots and systems (IROS), IEEE, pp 2150–2156
12. Borges PV, Peynot T, Liang S, Arain B, Wildie M, Minareci M, Corke P et al (2022) A survey on terrain traversability analysis for autonomous ground vehicles: methods, sensors, and challenges. Field Robot 2(1):1567–1627

Open Access This chapter is licensed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License (http://creativecommons.org/licenses/by-nc-nd/4.0/), which permits any noncommercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if you modified the licensed material. You do not have permission under this license to share adapted material derived from this chapter or parts of it.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.



Research on Oil Film Pressure Distribution Simulation of Hydrodynamic Lubrication Sliding Bearing

Zhengxi Hu, Junbing Liao, Yingmiao Hu, Bin Peng, and Yongqiang Huang

Abstract The finite difference method is used to solve Reynolds equation and figure up the oil film pressure distribution of hydrodynamic lubrication sliding bearing. The effects of width-to-diameter ratio and eccentricity on the oil film pressure distribution were studied. Bearing capacity of the oil film and offset angle of plain bearing were solved, and the relationship between different width-to-diameter ratio, eccentricity and bearing capacity and offset angle were studied, and the experimental data were consistent with the simulation results, the correctness of the simulation data was verified. This work shows that the numerical calculation method can provide an important mean for the design of the hydrodynamic lubrication sliding bearing.

Keyword Hydrodynamic sliding bearing · Oil film pressure · Reynolds equation

1 Introduction

Slide bearings are widely used in mechanical fields as supporting elements of rotating shafts. With the rapid development of the mechanical industry, the bearings performance requirements are getting higher and higher. Hydrodynamic sliding bearing is to rely on the fluid lubrication film between the solid friction surface to produce pressure to withstand the external load and avoid solid contact with each other, thereby reducing friction resistance and protecting the solid friction surface [1–4]. With the development of computer technology, the method of computer analysis of physical problems has been adopted by more and more scientists. MATLAB is used by many researchers because of its high performance of numerical calculation and visualization capabilities. In the study of hydrodynamic plain bearings, it is one of the main

Z. Hu (✉) · J. Liao · B. Peng · Y. Huang
School of Artificial Intelligence, Hezhou University, Hezhou, China
e-mail: huzhengxi588193@126.com

Y. Hu
College of Mathematics and Computer Science, Northwestern Minzu University, Lanzhou, China

© The Author(s) 2026

H. Saman et al. (eds.), *The 9th International Conference on Advances in Construction Machinery and Vehicle Engineering*, Lecture Notes in Mechanical Engineering,
https://doi.org/10.1007/978-981-96-5527-4_40

contents to solve the oil film pressure distribution. To solve the oil film pressure, Reynolds equation, the basic equation of hydrodynamic lubrication, must be derived first. In recent years, many numerical calculation methods have been proposed, and the idea comes down to the solution of two-dimensional Reynolds equation [5–9]. However, experiments to verify the correctness of the results obtained by this method are rarely reported, and the numerical calculation method used by each person is different possible.

Based on the theory of finite width bearings and MATLAB software, this paper uses the finite difference method to solve the two-dimensional Reynolds equation, solve the oil film pressure distribution of plain bearings, and then solve the oil film bearing capacity, offset angle, etc. Finite Difference Method (FDM) is an approximate method to solve the numerical solution of differential equations. Its main principle is to carry out direct difference approximation to the differential terms in the differential equations, so as to transform the differential equations into algebraic equations. With the development of computer technology, it is possible to solve huge algebraic equations quickly and accurately, so it is gradually applied in a large number of applications. Up to now, finite difference method has become an important numerical solution method, which has a wide application background in the field of engineering.

At the same time, the influence of width-diameter ratio and eccentricity on the bearing capacity of the oil film and the variation rule of the offset angle is studied. Finally, the liquid pressure of the hydrodynamic lubrication sliding bearing is measured by the experimental method, and the correctness of the numerical calculation method is verified, which provides an important basis for the accurate solution of the static characteristic parameters of the oil film of the hydrodynamic lubrication sliding bearing, also for design and check of the plain bearing in the engineering practice.

2 Principles of the Solution of Renolds Equation

The Reynolds equation is reduced to a dimensionless form as follows:

$$\frac{\partial}{\partial \varphi} H^3 \cdot \frac{\partial p}{\partial \varphi} + \left(\frac{d}{L}\right)^2 \frac{\partial}{\partial \lambda} H^3 \cdot \frac{\partial p}{\partial \lambda} = \frac{\partial H}{\partial \varphi} \tag{1}$$

In the equation:

$H = 1 + \varepsilon \cos \varphi$; Epsilon $= e/c$; $c = R - r$; $\lambda = z/(L/2)$, p is the oil film pressure; ε is eccentricity; e is eccentric distance; c is the radius gap; d is the bearing diameter; L is the bearing width; and φ is the offset Angle.

According to (1), the calculated value of the dimensionless pressure P in the plain bearing mainly depends on the influence of the width-diameter ratio L/d and eccentricity ε. Based on the principle of finite difference method, and the method of calculating grid division. The oil film of a piece of bearing is divided into a grid along

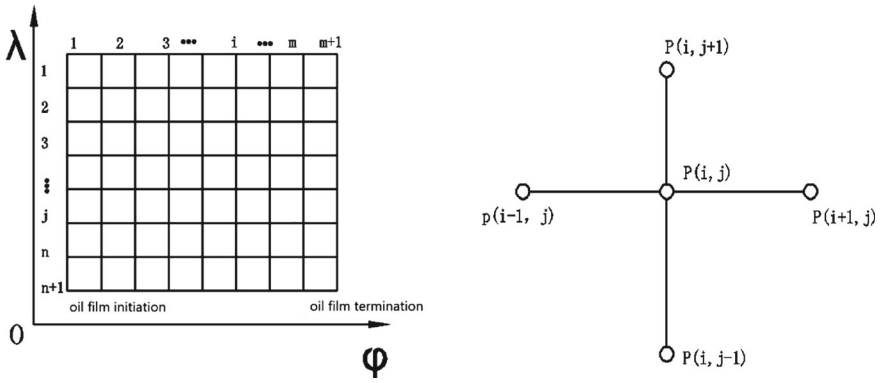


Fig. 1 Oil film grid division diagram

the coordinate axis (as shown in Fig. 1), which is to discrete the oil film block into a grid and divide it into m × n units. The number of columns along the φ direction is divided into i cells, and the bearing is also divided into m cells from number 1 to m + 1. The λ direction is divided by j numbers, the bearing is also divided into n cells from 1 to n + 1, so that each node position can obtain a two-dimensional coordinate, expressed by (i, j).

Therefore:

$$\frac{\partial \lambda}{\partial \phi} = \frac{\mathrm{P_{i,j+1}} - \mathrm{P_{i,j}}}{\Delta\phi \times \Delta\lambda} \tag{2}$$

$$\frac{\partial^2 \mathrm{P}}{\partial \phi^2} = \frac{\mathrm{P_{i,\,j+1}} - 2\mathrm{P_{i,\,j}} + \mathrm{P_{i,\,j-1}}}{\Delta\phi^2} \tag{3}$$

$$\frac{\partial^2 \mathrm{P}}{\partial \lambda^2} = \frac{\mathrm{P_{i+1,j}} - 2\mathrm{P_{i,\,j}} + \mathrm{P_{i-1,j}}}{\Delta\lambda^2} \tag{4}$$

Then the dimensionless form of the two-dimensional Reynolds equation is brought into the first derivative, and the equation is simplified and solved, and the pressure of the middle node can be calculated according to the nodes around, and the following is obtained:

$$P_{i,j} = A_1 P_{i+1,j} + A_2 P_{i-1,j} + A_3 P_{i,j+1} + A_4 P_{i,j-1} - A_5 \tag{5}$$

Thereinto, $A_1 = \frac{(n_1+n_2)}{n_6}$, $A_2 = \frac{(n_3+n_4)}{n_6}$, $A_3 = \frac{n_2}{n_6}$, $A_4 = \frac{n_4}{n_6}$, $A_5 = \frac{n_5}{n_6}$.

$$n_1 = 3 \times M_{i,j}^2 \frac{\left(H_{i,j+1} - H_{i,j}\right)}{\Delta\phi^2}, \quad n_2 = \frac{H_{i,j}^3}{\Delta\phi^2}$$

$$n_3 = \left(\frac{D}{L}\right)^2 \times 3 \times H_{i,j}^2 \times \frac{\left(H_{i,j+1} - H_{i,j}\right)}{\Delta\lambda}, \quad n_4 = \left(\frac{D}{L}\right)^2 \times \frac{H_{i,j}^3}{\Delta\lambda^2}$$

$$n_5 = 3 \times \frac{\left(H_{i,j+1} - H_{i,j}\right)}{\Delta\phi}, \quad n_6 = n_1 + 2 \times n_2 + 2 \times n_4 + n_3$$

The point-by-point relaxation iteration method is more convenient to solve and easier to program. According to the principle of point-by-point relaxation iteration method, Eq. (5) is rewritten as:

$$P_{i,j}^{(k)} = \vartheta\left[A_1 P_{i+1,j} + A_2 P_{i-1,j} + A_3 P_{i,j+1} + A_4 P_{i,j-1} - A_5 - P^{k-1}\right] + P_{i,j}^{k-1} \quad (6)$$

ϑ is the relaxation iteration factor, when it meets $\omega > 1$, it is the overrelaxation iteration method. The convergence speed inside is affected by the value of ω, and attention should also be paid to the value range of ϑ, which may not converge if it is too large. The general solution is to assign the value of the boundary condition to the pressure on each boundary node, and then take the initial value of the pressure of all points as the approximate value of the first iteration, and next step is to calculate the pressure value of the middle four points according to the pressure value on the node P(i, j),and the offset angle formula is:

$$\theta = -a\tan\left(\frac{W_x}{W_y}\right)$$

3 Analysis of MATLAB Solution

According to the principle of solving Reynolds equation described above and the boundary conditions mentioned in solving the equation, the discretized equation is solved by relaxation iterative method step by step.

Using software MATLAB for programming, the iterative method is used to calculate to under the following conditions: $\varepsilon = 0.9$, d/l = 1, m = 60, n = 40 and eccentricity $\varepsilon = 0.459$, d/l = 0.8, m = 60, n = 40, and take this data to obtain the dimensionless oil film pressure distribution of the two groups of data, as shown in Fig. 2. They are here as cases for us to study his laws. Other conditions are not listed in the paper because the figures are similar.

According to the two pictures of oil film pressure distribution images above, it can be seen that the distribution is the dimensionless oil film pressure of the 360° envelope Angle finite width radial bearing with a continuous parabolic distribution. The dimensionless oil film pressure shows an upward trend and gradually increases in a certain region where the circumference is $0 \leq \varphi < 30$. At a certain value of $20 < \varphi < 30$, the oil film pressure reaches its maximum value, and in the region of $20 < \varphi \leq 40$, the dimensionless oil film pressure gradually decreases until it drops to zero.

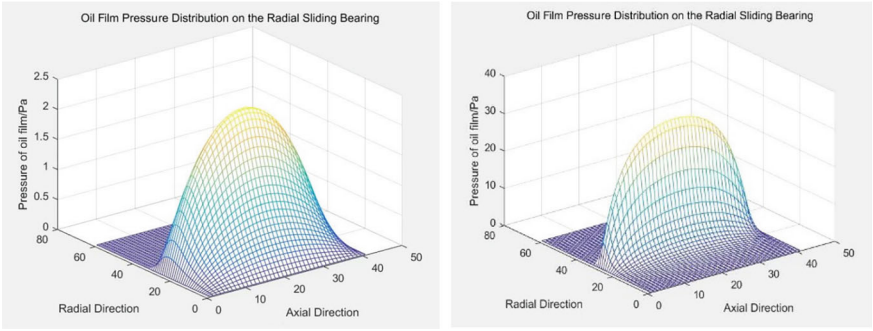


Fig. 2 Sliding bearing oil film pressure distribution

3.1 Influence of Width-to-Diameter Ratio and Eccentricity on Characteristics of Plain Bearings

Given the ratio of width to diameter, change the value of eccentricity, and then calculate the bearing capacity. The relationship between different width-diameter ratio and eccentricity as follows.

It can be clearly seen from Fig. 3 that the pressure of the oil membrane when the eccentricity tends to 1. The pressure of the oil film is equivalent to infinitely large, indicating that the sliding bearing oil bearing capacity, that is, the size of the countless load, is changed by the change of eccentricity. Of course, there is another situation that it is directly contacting, that is, direct contact between the neck and the surface of the bearing, and in this case, the loading capacity and the oil film reach a state of no relationship between them.

3.2 The Relationship Between the Ratio of the Biased Rate and the Offset Angle

To understand the relationship between the offset angle and the eccentricity rate, we can draw a relationship diagram, as shown in Fig. 4.

It can be seen from Fig. 4 that the eccentricity gets close to 1 but does not reach 1 yet, the eccentricity angle tends to 0°, which also indicates that the eccentricity angle is affected by eccentricity. The general trend is that the eccentricity angle decreases when the eccentricity increases; when the eccentricity is about 0.35, crossover points appear, and the influence of the width-diameter ratio on eccentricity increases, The larger the offset angle.

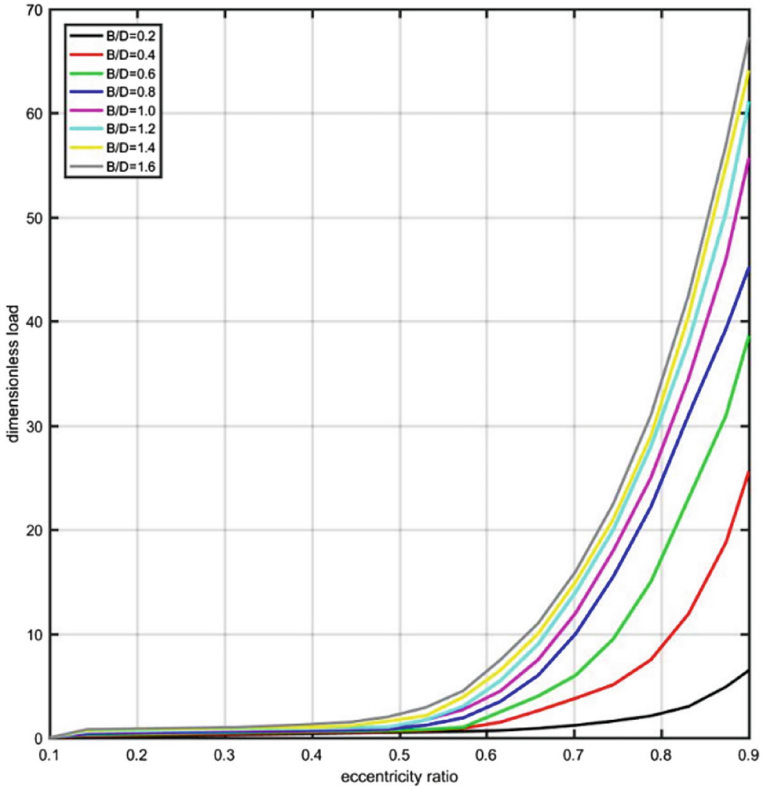


Fig. 3 Relationship curve between the bearing capacity and width ratio and eccentricity rate

4 Experimental Procedure

The test machine is CTHD-C series hydrodynamic lubrication sliding bearing experimental workbench, as shown in Fig. 5. The lubricating oil grade is L-AN32, the dynamic viscosity is 0.028 Pa s, the bearing inner diameter D = 80 mm, the effective width B = 125 mm. The pressure values of channel 1–8 were measured and recorded from left to right as shown in Table 1. The changing trend of the data are shown in Fig. 6 after they be processed. Obviously, the distribution pattern of oil film pressure was consistent with the classical plain bearing theory [10, 11]. The oil film pressure was directly related to the bearing width-to-diameter ratio and eccentricity, although different width-to-diameter ratio and eccentricity was tried, the oil film pressure distribution was similar, and the circumferential oil film pressure distribution starts from the initial angular position, and gradually increased, and then dropped sharply after reaching a certain peak, and it was like the parabolic shape.

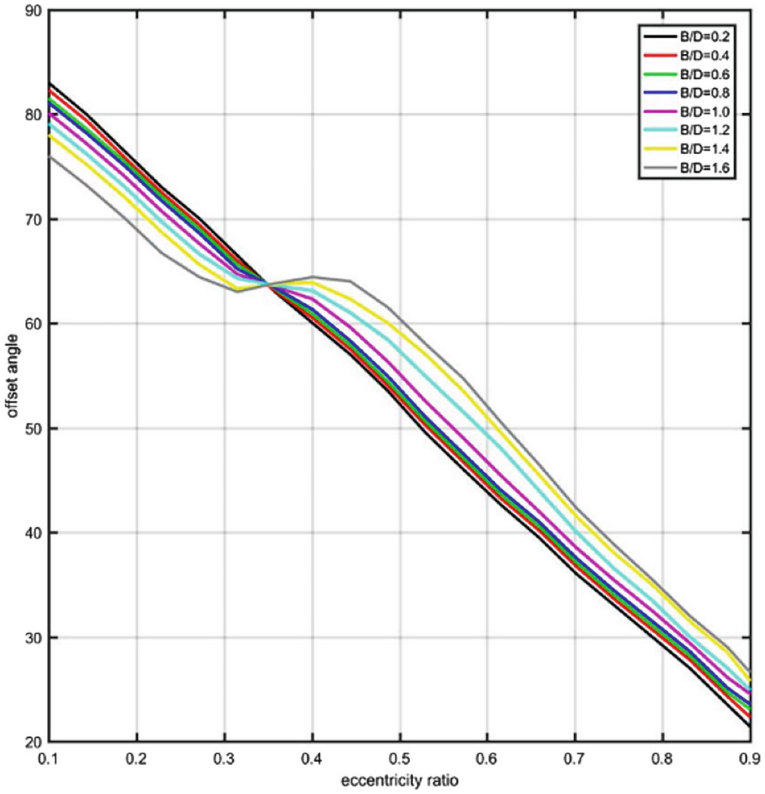


Fig. 4 The relationship between width diameter ratio and the offset angle



Fig. 5 Hydrodynamic lubrication sliding bearing experimental workbench

Table 1 The test value of the CTHD-C experimental workbench (axis speed is 100 rpm, and the pressure is 0.2 MPa)

Channel 1	Channel 2	Stress test value from left to right channel (Mpa)					
		Channel 3	Channel 4	Channel 5	Channel 6	Channel 7	Channel 8
0.187	0.3327	0.4195	0.486	0.4905	0.4112	0.2458	0.346

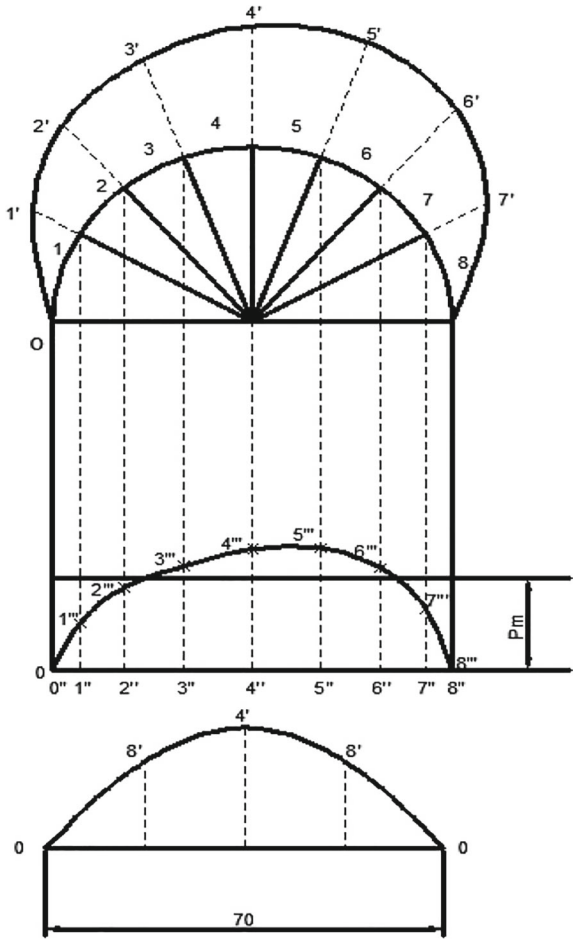


Fig. 6 Axial and radial oil film pressure experiment numerical distribution curve

5 Conclusions

Based on the simulation and experiment results presented above, the conclusions are obtained as below:

(1) Two-dimensional Reynolds equation was used to solve the oil film pressure distribution of hydrodynamic lubrication sliding bearing. Carrying capacity and offset angle of plain bearings according were according to solved by MATLAB programming based on the finite width bearing theory.
(2) The effects of different width-to-diameter ratio and eccentricity on oil film bearing capacity and offset angle were studied. When the width-to-diameter ratio is the same, the carrying capacity of the oil film also increases with the gradual increase of the eccentricity. On the other hand, when the eccentricity is the same, the carrying capacity of the oil film gradually decreases with the increase of the width-to-diameter ratio.
(3) The experimental data of CTHD-C workbench were used to verify that the simulation results were consistent with the experimental results, it indicated that the numerical calculation method was effective in design of hydrodynamic lubrication sliding bearings, and the simulation method could provide an important tool for design of hydrodynamic lubrication sliding bearings.

Acknowledgements All the work in this paper obtains great support from the Hezhou city central guide local science and technology development special fund project (contract number: HEKE ZY2023008) and Hezhou University project fund (No. HZUJS202206)

References

1. Wang Y, Chen H, Long R (2024) Mechanical modeling of friction phenomena in social systems based on friction force. 904
2. Summer F, Grtin F, Offenbecher M, Taylor S (2019) Challenges of friction reduction of engine plain bearings-Tackling the problem with novel bearing materials. J Tribol Int 131:238–250
3. Kadam KR, Banwait SS (2018) Performance characteristics of two-axial groove journal bearing for different groove angles. Ind Lubric Tribol 70(2)
4. Jiang W, Wang W (2019) Sliding bearing material testing machine under high temperature corrosion conditions. J Mater Sci Eng 692:012042
5. Kumar A, Sharma SC (2019) Textured conical hybrid journal bearing with ER lubricant behavior. Tribol Int 129:363–376
6. Saeidi F, Meylan B, Hoffmann P, Wasmer K (2016) Effect of surface texturing on cast iron reciprocating against steel under starved lubrication conditions: a parametric study. Wear 348–349:17–26
7. Yang J, Liu Z, Xie Y et al (2004) Research on dynamic boundary conditions for solving Reynolds equation. Heilongjiang Electric Power 5:346–349, 351
8. Hirani H, Suh NP (2005) Journal bearing design using multiobjec tive genetic algorithm and axiomatic design approaches. Tribol Int 38(1):481–485
9. Li Y-S, Ao L-B, Li L et al (2010) Dynamic analysis method for dynamic characteristic coefficient of plain bearing. J Mechan Eng 46(21):48–53
10. Yang J, Guo R, Tian Y (2008) Hybrid radial ba sis function/finite element modelling of journal bearing. Tribol Int 41:1169–1175
11. Zhang Y, Hei D, He J et al (2016) Approximate solution of nonlinear oil film force of finite-length plain bearing. Tribology 36(4):488–494

Open Access This chapter is licensed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License (http://creativecommons.org/licenses/by-nc-nd/4.0/), which permits any noncommercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if you modified the licensed material. You do not have permission under this license to share adapted material derived from this chapter or parts of it.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.



Bucket Drop Point Prediction for Excavators Integrating Point Cloud Information



Yong Wang, Xinhui Liu, Changsheng Liu, Yang Li, Qiushi Bi, and Zongwei Yao

Abstract Accurate prediction of the excavator bucket's landing point is a critical aspect of achieving intelligent operation. However, existing prediction methods generally suffer from low computational efficiency, vulnerability to environmental factors, and poor feasibility, making them inadequate to meet the dual demands of efficiency and prediction information completeness required for engineering construction. To address these issues, this paper proposes a bucket landing point prediction method that integrates point cloud information. First, IMU data is filtered, and a kinematic model of the working device is established to calculate the bucket's pose information. Then, LiDAR point cloud data is processed, and dynamic region point clouds are extracted based on the bucket's pose, followed by the reconstruction of the elevation contour, on which the key point distances are calculated. Finally, various experimental scenarios were constructed in Gazebo to comprehensively validate the proposed method through simulation. The results demonstrate that the proposed method meets the requirements for real-time and complete information in predicting the excavator bucket's landing point.

Keywords Excavator working device · Bucket drop point prediction · Patchwork++ · Kriging interpolation · Gazebo simulation

1 Introduction

Excavators, with their powerful digging capabilities and flexible operational performance, can efficiently complete various complex construction tasks, playing an indispensable role in areas such as mining, road and bridge construction, urban and

Y. Wang · X. Liu · C. Liu · Q. Bi · Z. Yao (✉)
Key Laboratory of CNC Equipment Reliability, Ministry of Education, School of Mechanical and Aerospace Engineering, Jilin University, Changchun, Jilin, China
e-mail: yzw@jlu.edu.cn

Y. Li
Liugong Changzhou Machinery Co., Ltd., Changzhou, Jiangsu, China

© The Author(s) 2026

H. Saman et al. (eds.), *The 9th International Conference on Advances in Construction Machinery and Vehicle Engineering*, Lecture Notes in Mechanical Engineering,
https://doi.org/10.1007/978-981-96-5527-4_41

rural infrastructure, and emergency rescue operations [1]. However, excavators face numerous challenges during operation: harsh construction environments with dust, strong vibrations, and high noise levels not only affect the health of the operators but can also lead to equipment failures; construction projects often have tight schedules, placing high demands on the operators; in high-risk working conditions, such as emergency rescue operations, the work site presents significant dangers, posing serious threats to the safety of the operators [2, 3]. To address these issues, achieving intelligent excavator operations has long been a key goal. Currently, excavator intelligence mainly includes two approaches: remote control [4] and autonomous driving [5]. Due to current technological and cost limitations, most users prefer remote-controlled excavators. This approach not only reduces the workload of the operators but also enhances safety and operational efficiency to a certain extent.

In remote-controlled excavation, accurately perceiving the conditions at the excavation site is essential. However, current excavation technologies can only transmit on-site images to a display screen. Presenting a three-dimensional environment on a two-dimensional screen significantly diminishes the sense of presence and spatial awareness, making it difficult for remote operators to experience a strong sense of immersion [6]. The perception of the work surface is far less clear and three-dimensional compared to direct on-site observation, making it challenging to accurately judge the distance between the bucket and the machine, the height relative to the ground, and the machine's orientation. As a result, operators cannot precisely locate the excavation point during operation, leading to a lower "bucket fill rate". Moreover, the inability to accurately judge the unloading point reduces work efficiency (e.g., during loading, the inability to properly distribute material in the truck bed leads to uneven loading and wasted truck resources). Additionally, the inability to clearly see ground undulations makes it difficult to assess leveling effectiveness during ground preparation. Therefore, accurately predicting the excavator bucket's drop point has become a critical technical challenge that must be addressed to overcome these issues.

Currently, the commonly used methods for measuring bucket drop points primarily include contact and non-contact measurements [7]. Contact measurement typically provides highly accurate and reliable results. Non-contact measurement, on the other hand, usually employs sensors like cameras and LiDAR, which do not require direct contact with the object being measured. These sensors can gather large amounts of data and, through processing and training, can obtain rich information. However, this method often requires significant computational resources. In practice, the working conditions of excavators are often harsh, and the site environment is unpredictable. Traditional contact measurement tools or single non-contact sensors may be damaged by factors such as rock impacts or machine vibrations. Therefore, employing multi-sensor fusion technology, which offers high reliability and is less affected by environmental changes, has become a necessary approach for achieving precise excavation in construction sites.

The integration of IMU (Inertial Measurement Unit) and LiDAR technologies is key to addressing the aforementioned challenges. By combining the efficiency and reliability of IMU in measuring distances with the ability of LiDAR to provide rich ground point information and remain less affected by environmental conditions, this

approach enables efficient and accurate bucket drop point prediction for excavators operating in complex environments. Therefore, to enhance the accuracy and reliability of drop point prediction, this paper proposes a bucket drop point prediction method that integrates point cloud information, aimed at improving the operational efficiency of remotely controlled excavators. The main contributions of this paper are as follows:

(1) A bucket drop point prediction method integrating point cloud information has been designed, enabling accurate prediction of drop points in complex operational environments.
(2) By efficiently identifying and extracting ground points, the method significantly reduces the impact of dynamic point clouds on subsequent data processing and analysis.
(3) A Gazebo simulation model has been established, allowing for comprehensive, efficient, and cost-effective validation of the proposed bucket drop point prediction method.

2 Related Research

2.1 Contact-Based Pose Estimation

Vahdatikhaki et al. [8] employed continuous analog potentiometers as joint angle sensors to achieve precise detection of the bucket pose for an intelligent excavator. Wind et al. [9] used angle sensors and draw-wire displacement sensors to perceive the bucket's pose, achieving precise excavation control for a hydraulic excavator. Kim et al. [10] developed a pose sensing system based on accelerometers and IMUs for excavators, significantly reducing pose sensing errors during bucket movement. Lee et al. [11] proposed a method that estimates the excavator's pose using forward kinematics and position using RTK, and validated the performance of this method by measuring errors during the horizontal, vertical, and rotational movements of the excavator's working device. Sumanas et al. [12] simulated excavator operations using a robotic arm and estimated joint angle information with four IMUs, demonstrating high applicability. Molaei et al. [13] determined the bucket's position using IMU measurements and forward kinematics, estimating excavation depth based on this information. Zheng et al. [14] combined data from RTK and inclinometers to obtain pose information for the excavator's undercarriage and working device. However, these contact-based measurement methods are heavily influenced by environmental factors, fail to capture rich environmental information, and can only acquire limited pose data, which imposes significant constraints on the overall accuracy and reliability of the system.

2.2 Pose Estimation Based on Deep Learning

Xu et al. [15] proposed a vision-based neural network system for estimating the pose of an excavator's working device. Liu et al. [16] combined machine vision and artificial neural networks to develop a real-time pose estimation method, providing a foundation for feedback control of the excavator's working device and significantly improving control accuracy and response speed. Shin et al. [17] developed a system that utilizes a single camera and AI-based image recognition algorithms to track the movement trajectory of an excavator bucket; compared to traditional methods that rely on multiple sensors, this system simplifies hardware configurations while significantly enhancing the accuracy and efficiency of position tracking. Yeom et al. [18] developed a vision-based pose perception system for guiding excavator operations, aiming to improve operational efficiency, with experiments demonstrating its ability to recognize the positions of the boom, arm, and bucket in real-time. Kashani et al. [19] employed a particle filter with distance transformation for spatial search and used the search results as the initial pose. They then applied the ICP (Iterative Closest Point) method to match the bucket point cloud with the physical model of the bucket to compute its pose. To address noise and occlusion in mining operations, Kashani et al. [20] proposed a point cloud-based pose perception method, using an improved ICP approach to calculate the bucket's pose, thus enhancing adaptability to the working environment. Considering the issues with the ICP method and its variants, Phillips et al. [21] accounted for the practical operation rules of mining excavators, making a series of pose assumptions for the bucket and selecting the most likely pose based on a Bayesian probability model as the final bucket pose. Kuris et al. [22] proposed a method using deep learning and inverse kinematics, extracting image coordinates of the excavator arm and bucket tip through deep learning for pose estimation from images captured by onboard cameras. Lundeen et al. [23] introduced an AR marker method to capture the geometric relationship between the camera and the marker, enabling precise pose calculation of the robotic arm. Xu et al. [15] utilized neural networks to estimate hydraulic cylinder displacement, which, combined with the geometric features of the excavator's working device, allowed for accurate pose calculation of the working device. Liang et al. [24] proposed a convolutional neural network-based method for estimating the pose of forward swinging actions. However, these methods often involve large computational demands and complex systems, which may result in slower response times, reduced real-time performance, and overall efficiency.

3 Background Work

Establishing the kinematic model and data preprocessing are fundamental tasks for predicting the drop point of an excavator bucket.

3.1 Kinematic Model

The D-H parameter method [25] is an effective approach for establishing the kinematic model of the excavator's working device. As shown in Fig. 1, the coordinate systems are defined as follows: the boom coordinate system is $O_1Y_1X_1Z_1$, the arm coordinate system is $O_2Y_2X_2Z_2$, the bucket coordinate system is $O_3Y_3X_3Z_3$, and the bucket tip coordinate system is $O_4Y_4X_4Z_4$. The boom coordinate system is set as the base coordinate. The parameters are defined as: the boom length is l_1, the arm length is l_2, the bucket length is l_3, the boom joint angle is θ_1, the arm joint angle is θ_2, the bucket joint angle is θ_3, and the bucket tip attitude angle is ξ.

The forward kinematics model of the working device, which is the transformation model from joint space to pose, can convert the joint angle information $[\theta_1, \theta_2, \theta_3]$ of the boom, arm, and bucket into the position and orientation angle information $[x, y, \xi]$ of the bucket's tooth tip.

The homogeneous transformation matrix $\mathbf{T}_{i+1}^{i}$ between adjacent coordinate frames is given by:

$$\mathbf{T}_{i+1}^{i} = \begin{bmatrix} c_i & -s_i & 0 & l_i c_i \\ s_i & c_i & 0 & l_i s_i \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \tag{1}$$

In the transformation matrix, if c_i represents $\cos\theta_i$ and s_i represents $\sin\theta_i$, then the transformation matrix between adjacent joint coordinate frames in the working device can be expressed as:

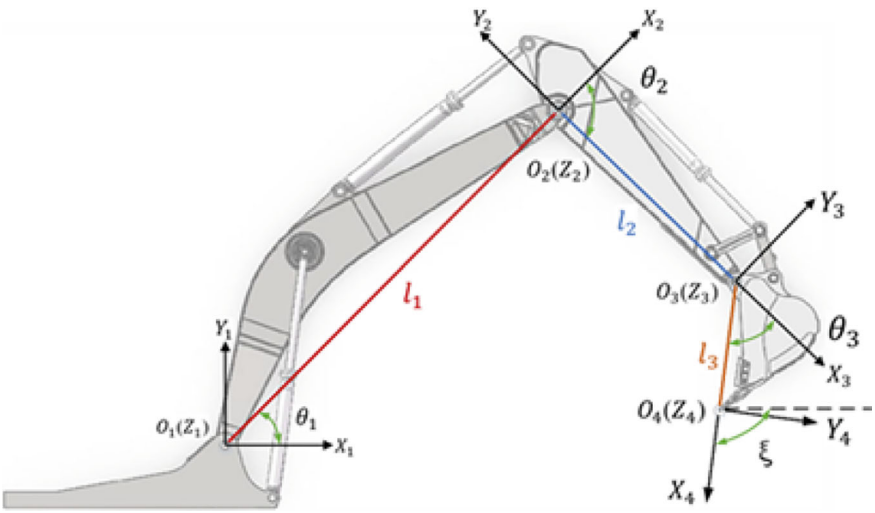


Fig. 1 The D-H coordinate system of the working device

$$\mathbf{T}_2^1 = \begin{bmatrix} c_1 & -s_1 & 0 & l_1c_1 \\ s_1 & c_1 & 0 & l_1s_1 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \tag{2}$$

$$\mathbf{T}_3^2 = \begin{bmatrix} c_2 & -s_2 & 0 & l_2c_2 \\ s_2 & c_2 & 0 & l_2s_2 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \tag{3}$$

$$\mathbf{T}_4^3 = \begin{bmatrix} c_3 & -s_3 & 0 & l_3c_3 \\ s_3 & c_3 & 0 & l_3s_3 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \tag{4}$$

Based on Eqs. (2) to (4), the transformation matrix $\mathbf{T}_4^1$ from the bucket tip coordinate system $O_4Y_4X_4Z_4$ to the base coordinate system $O_1Y_1X_1Z_1$ is given by:

$$\mathbf{T}_4^1 = \begin{bmatrix} c_{123} & -s_{123} & 0 & l_1c_1 + l_2c_{12} + l_3c_{123} \\ s_{123} & c_{123} & 0 & l_1s_1 + l_2s_{12} + l_3s_{123} \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \tag{5}$$

In the expression, c_{12} denotes $\cos(\theta_1 + \theta_2)$, s_{12} denotes $\sin(\theta_1 + \theta_2)$, c_{123} denotes $\cos(\theta_1 + \theta_2 + \theta_3)$, and s_{123} denotes $\sin(\theta_1 + \theta_2 + \theta_3)$.

The tip orientation angle ξ is:

$$\xi = \theta_1 + \theta_2 + \theta_3 \tag{6}$$

Based on Eqs. (5) and (6), the forward kinematics model of the working device is:

$$\begin{cases} x = l_1c_1 + l_2c_{12} + l_3c_{123} \\ y = l_1s_1 + l_2s_{12} + l_3s_{123} \\ z = 0 \\ \xi = \theta_1 + \theta_2 + \theta_3 \end{cases} \tag{7}$$

3.2 IMU Data Processing

The complex working environment of construction vehicles and strong vibrations can cause significant noise in IMU data. Data filtering is required to remove noise and restore the data as accurately as possible.

In this study, Kalman filtering [26] is used for state estimation of the target. The target state depends only on the previous state and the current observation. This approach significantly improves computational efficiency and meets the real-time requirements for vehicle pose calculation.

Through Eq. (8), the target can be transitioned from the previous moment to the current moment.

$$\mathbf{X}_n = \mathbf{A}_n \mathbf{X}_{n-1} + \mathbf{B}_n \mathbf{u}_n \tag{8}$$

In this equation, $\mathbf{X}_n$ represents the system's state at time n, $\mathbf{A}_n$ represents the transition matrix at time n, $\mathbf{B}_n$ represents the mapping relationship between the system input and the system state at time n, and $\mathbf{u}_n$ represents the system input at time n.

Equation (9) represents the system's observation equation. By knowing the initial state of the target and the inputs at each time step, this algorithm can be used to obtain the target's state information at any given time.

$$\mathbf{z}_n = \mathbf{H}_n \mathbf{X}_n \tag{9}$$

In this equation, $\mathbf{z}_n$ represents the observed value of the system at time n, and $\mathbf{H}_n$ represents the observation matrix at time n.

When estimating the state of a target, various types of noise inevitably influence the process. Therefore, it is necessary to introduce a covariance matrix to represent the uncertainty of the entire system. The covariance matrix of the system satisfies the following relationship:

$$\mathbf{P}_n = \mathbf{A}_n \mathbf{P}_{n-1} \mathbf{A}_n^T + \mathbf{B}_n \mathbf{Q}_n \mathbf{B}_n^T \tag{10}$$

In this equation, $\mathbf{P}_n$ represents the covariance matrix of the entire system at time n, and $\mathbf{Q}_n$ represents the covariance matrix corresponding to the uncertainty of the system input $\mathbf{u}_n$ at time n.

$$\mathbf{x}_n^+ = \mathbf{x}_n^- + \mathbf{K}_n (\mathbf{z}_n - \mathbf{H}_n \mathbf{x}_n^-) \tag{11}$$

$$\mathbf{P}_n^+ = (\mathbf{I} - \mathbf{K}_n) \mathbf{P}_n^- \tag{12}$$

In the formula, $\mathbf{I}$ represents the identity matrix, $\mathbf{x}_n^+$ and $\mathbf{x}_n^-$ represent the posterior and prior estimates of the system state, respectively. $\mathbf{P}_n^+$ and $\mathbf{P}_n^-$ represent the covariance matrices of the posterior and prior probabilities of the system state, respectively. $\mathbf{K}_n$ is the Kalman gain, which can be calculated using the following formula (13):

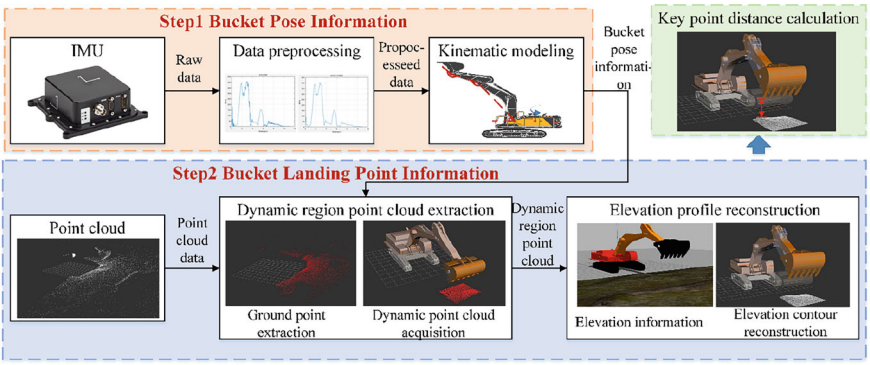


Fig. 2 The overall framework of the method

$$\mathbf{K}_n = \mathbf{P}_n^{-}\mathbf{H}^T(\mathbf{H}\mathbf{P}_n^T + \mathbf{R}_n)^{-1} \tag{13}$$

In the formula, $\mathbf{R}_n$ represents the covariance matrix corresponding to the uncertainty of the observed values.

4 Methodology Research

4.1 Methodology Framework

First, the IMU is used to obtain the angle information of the boom, arm, and bucket, and the IMU data is preprocessed using Kalman filtering. Then, based on the kinematic model of the working device, the pose information of the bucket is calculated. Next, the LiDAR point cloud data is processed to extract the dynamic region point cloud based on the bucket's pose. Following that, the elevation contour is reconstructed. Finally, the key point distances are calculated to determine the bucket drop point information. The overall framework of the method is shown in Fig. 2.

4.2 Dynamic Region Point Cloud Extraction

The point cloud obtained by LiDAR contains not only rich ground point information but also a large amount of dynamic object point cloud data, including the excavator's working device, bucket, and other components, as shown in Fig. 3. Therefore, to accurately analyze ground information, it is necessary to process this data to extract the ground points. Additionally, the point clouds of other dynamic objects must

be removed to obtain clearer and more accurate terrain data, thereby isolating the dynamic region point cloud.

The Patchwork++ algorithm [27] is an outstanding ground point cloud segmentation model known for its efficiency, accuracy, and robustness. It can quickly process large-scale point cloud data while maintaining high ground point recognition accuracy in complex terrains, as illustrated in the workflow in Fig. 4. This paper utilizes the Patchwork++ algorithm to extract ground point information, enhancing the overall effectiveness and efficiency of data processing.

(1) First, reflection noise removal is performed on the LiDAR point cloud data. Specifically, this step removes point cloud data that is too close to the LiDAR and filters out noise points that are lower than a predefined height threshold

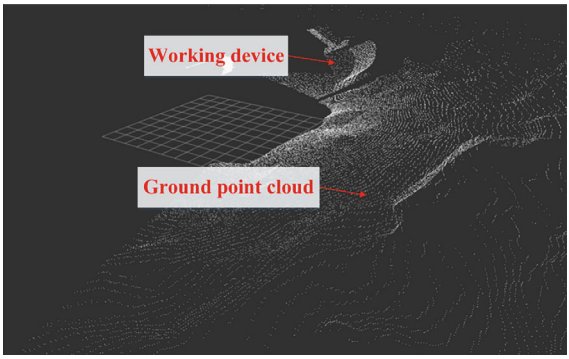


Fig. 3 Raw point cloud information

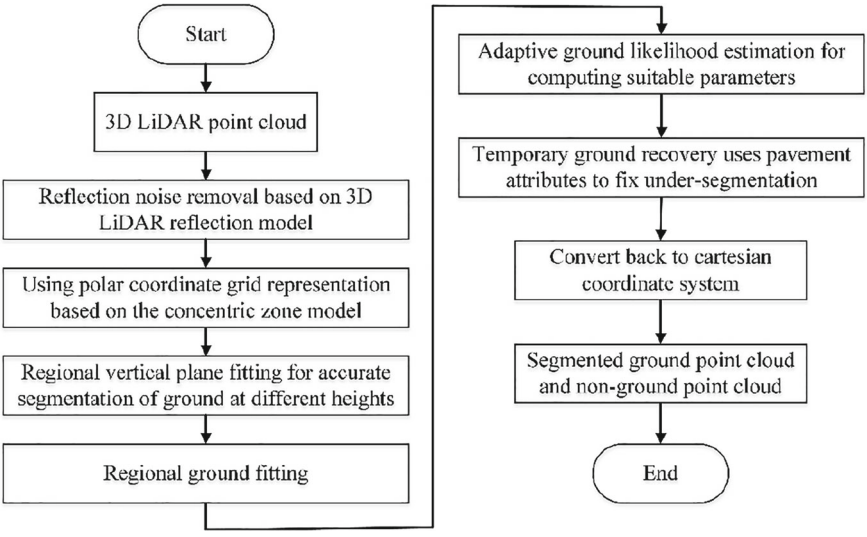


Fig. 4 Patchwork++ algorithm flowchart

and have an intensity below a certain threshold. This process helps improve the accuracy of subsequent data processing, ensuring the reliability and precision of the final results.

(2) Next, a concentric ring model is established to divide the point cloud into multiple regions of varying sizes. This model uses a central point as a reference and gradually expands outward to form a series of concentric rings, allowing the point cloud data to be processed region by region. This segmentation method aids in more effectively analyzing and processing the point cloud data, enhancing the overall efficiency and accuracy of the algorithm.

(3) Based on this, vertical plane fitting is performed for each region. This method improves the accuracy of plane fitting by identifying and removing vertical points that may interfere with the fitting process.

(4) Next, ground fitting is performed for each small region. This method allows for precise fitting of the ground shape and elevation within each small area, resulting in a more accurate representation of the terrain features.

(5) To more robustly determine the actual ground points, adaptive ground likelihood estimation is employed based on the previous ground segmentation results. This approach adjusts the model parameters to obtain appropriate ground parameters, providing a more accurate reflection of the actual ground conditions.

(6) Finally, to address the issue of under-segmentation in some ground areas, a temporary ground recovery method is used. This method corrects any ground points that were missed during the previous segmentation process, resulting in a more complete and accurate final ground model.

After filtering out redundant point cloud information other than the ground, it is necessary to extract the ground point cloud within the vertical dynamic region corresponding to the excavator bucket. The extracted ground point cloud and working device point cloud are shown in Fig. 5. By combining IMU data with the excavator's kinematic model, the spatial position coordinates (x, y, z) of the bucket tip can be obtained. Projecting this point vertically onto the ground yields the coordinates (x, y) in a two-dimensional plane. Using this point as the centroid and considering the geometric dimensions of the bucket, the point cloud within a square area is extracted, as shown in Fig. 6.

4.3 Elevation Contour Reconstruction

To aid operators in perceiving the bucket's pose and accurately locate digging points during remote control of an excavator, it is essential to convert the extracted point cloud data from the bucket's corresponding vertical dynamic area into an elevation map. Common methods for this process include nearest-neighbor interpolation, inverse distance weighting interpolation, kriging interpolation, bilinear interpolation, and cubic spline interpolation [28]. Among these, kriging interpolation is a statistical-based method that considers not only the distance between points but also

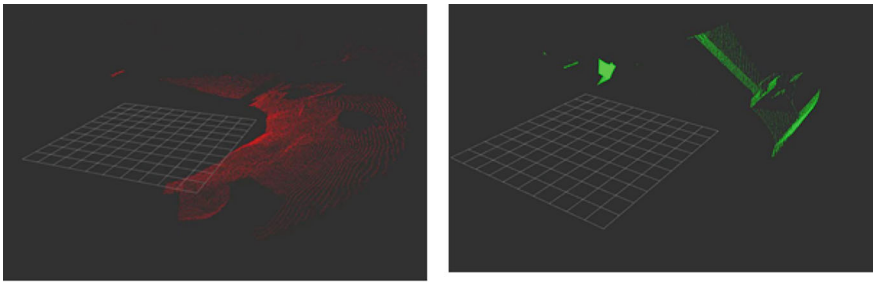


Fig. 5 Point cloud information: **a** Ground point cloud, **b** Working device point cloud

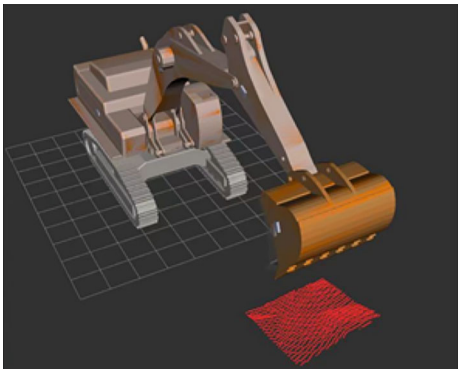


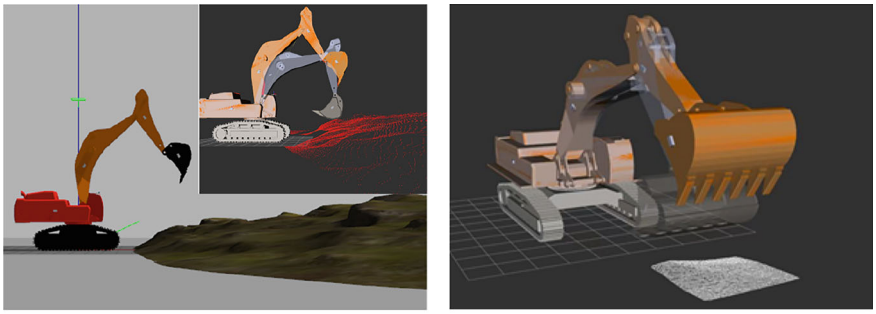
Fig. 6 Dynamic region point cloud

their spatial correlation, providing the optimal unbiased estimation. Compared to other interpolation methods, kriging offers the highest precision for LiDAR-scanned DEMs. Therefore, kriging is chosen as the preferred method in this study.

Assuming $Z(x, y)$ is a stationary random process with discrete values $Z(x_i, y_i)$, the Kriging estimate for the elevation value at the interpolation point (x_p, y_p) is given by:

$$Z\left(x_p, y_p\right) = \sum_{i=1}^{n} \beta_i Z(x_i, y_i) \tag{14}$$

In the equation, β_i are the unknown Kriging weights that must satisfy the condition $\sum_{i=1}^{n} \beta_i = 1$, and the weighted variance of the discrete points must satisfy $\mathrm{Var}\left[Z(x_i, y_i) - \sum_{i=1}^{n} \beta_i Z(x_i, y_i)\right] = \min$. Therefore, the elevation $Z(x_p, y_p)$ at the interpolation point is given by:



(a) (b)

Fig. 7 Elevation contour reconstruction: **a** Elevation information, **b** Dynamic region elevation contour reconstruction

$$Z(x_p, y_p) = M + [\sigma_{P1}, \sigma_{P2}, \cdots, \sigma_{Pn}] \begin{bmatrix} \sigma_{11} & \cdots & \sigma_{1n} \\ \vdots & \ddots & \vdots \\ \sigma_{n1} & \cdots & \sigma_{nn} \end{bmatrix} \begin{bmatrix} Z(x_1, y_1) - M \\ \vdots \\ Z(x_n, y_n) - M \end{bmatrix} \tag{15}$$

In the equation, M represents the median of the weights for the elevation value $Z(x_p, y_p)$.

The elevation information is shown in Fig. 7a, and the reconstructed elevation contour of the dynamic region is presented in Fig. 7b.

4.4 Key Point Distance Calculation

To enable the operator to perceive the spatial position of the bucket and its relation to the ground height, it is necessary to calculate the distance between the excavator's bucket and the ground within the corresponding vertical area in real-time. This helps the operator more accurately control the digging depth and operational precision during the excavation process, effectively avoiding mishaps and enhancing both operational efficiency and safety.

After obtaining the elevation map corresponding to the bucket's area, traverse the planar coordinate points (x, y) within this region to extract the maximum ground height $z_{max}(x, y)$ for each point. Assuming the bucket's spatial position coordinates are $z(x, y)$, the distance between the bucket's spatial coordinates and the highest ground point in this area is calculated using the distance formula, as shown in Fig. 8.

$$D = z(x, y) - z_{max}(x, y) \tag{16}$$

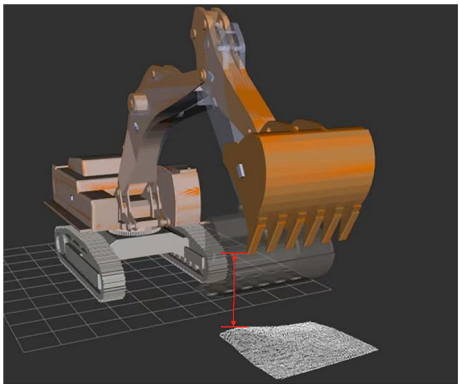


Fig. 8 Calculation of key point distance

5 Experimental Validation

5.1 Simulation Analysis

The simulation analysis was conducted on a computing device with the following main specifications: an R5-5600X processor, 16 GB of RAM, and the Ubuntu 20.04 operating system.

In this study, joint simulation experiments were conducted using the 923FN excavator model within Gazebo and MOVEIT. First, camera, IMU, LiDAR, and other sensor models were added to the 3D model of the excavator to facilitate subsequent configuration. The entire model was then exported as a URDF file compatible with the Gazebo platform, and the URDF file was modified according to the actual excavator model and the parameters of the sensors.

To facilitate manual adjustment of the bucket position, the excavator model was also configured in MOVEIT. The four rotating joints of the boom, arm, bucket, and rotation mechanism were added to the planning group, and ROS rotational controllers were configured for these joints. The integrated simulation model after configuration is shown in Fig. 9.

5.2 Results Analysis

To comprehensively evaluate the performance of the excavator bucket depth perception system, this study designed three key evaluation metrics. First, the vertical distance from the excavator bucket to the corresponding ground area was selected as the first evaluation metric. This metric serves as a direct measure of bucket positioning accuracy. By accurately calculating the error between the bucket's actual distance from the ground under various operating conditions, the accuracy of the bucket landing point prediction method can be effectively assessed. The second

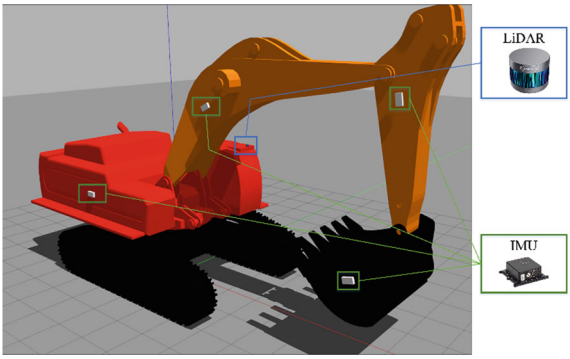


Fig. 9 Integrated simulation model

evaluation metric is the overlap of the bucket landing point area. This metric reflects the accuracy of the prediction method under different working conditions. A higher overlap indicates that the prediction method has strong adaptability and robustness in varying operational environments. The third evaluation metric is the time required for landing point prediction. This metric directly impacts the system's real-time performance and practicality. In intelligent excavation operations, the prediction of the bucket landing point not only needs to be highly accurate but also must be completed within a limited time to ensure the system's responsiveness during actual operations. This study designed an experimental scheme for three different simulation environments, using the aforementioned three evaluation metrics to assess the system's adaptability and stability under various working conditions and excavator postures. The results are shown in Fig. 10, with key metrics listed in Table 1.

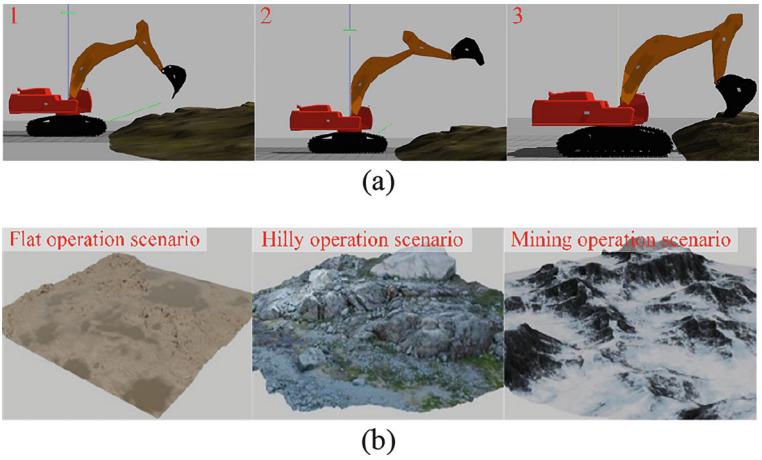


Fig. 10 Simulation scenes: **a** Excavator working device in three different postures, **b** Three different terrains

Table 1 Simulation results

Scenario	Postures	Distance calculation (mm)	Actual distance (mm)	Error (mm)	Area overlap ratio (%)	Time (ms)
1	1	3662	3642	20	92.1	12.5
	2	7390	7343	47	90.2	13.1
	3	2254	2269	15	93.3	11.7
2	1	3168	3138	30	93.2	15.2
	2	6789	6736	53	87.6	13.6
	3	1723	1744	21	96.7	14.2
3	1	2256	2222	34	86.9	14.3
	2	5593	5566	27	85.3	12.9
	3	698	717	19	88.6	13.4

The results show that under three different poses, the vertical error obtained does not exceed 60 mm (with a bucket size of $2650 \times 2050 \times 1960$ mm). This indicates that the proposed prediction method has high accuracy regardless of the pose, maintaining a small error range even in complex operational environments, ensuring the stability and reliability of excavator operation. Further analysis of the accuracy of the bucket's landing point reveals that the proposed method achieves an overlap rate of over 85% between the predicted landing point and the actual landing point area under different working conditions. This means that even in the presence of various variables and uncertainties, the method can still maintain high prediction accuracy, meeting the precision requirements of intelligent operations. The high-precision prediction of the bucket's landing point not only improves the efficiency of excavator operations but also reduces terrain damage, demonstrating significant practical application value.

6 Conclusion and Outlook

To address the problem of predicting the bucket's landing point in complex operational scenarios and to overcome the limitations of existing methods in terms of computational efficiency, environmental adaptability, and prediction accuracy, this paper proposes an efficient and accurate landing point prediction method that integrates point cloud information.

First, the angle information of the excavator's working device is obtained through the IMU, and the raw data is filtered. This data is then combined with the kinematic model of the working device to accurately calculate the spatial pose information of the bucket. Next, the point cloud data collected by the LiDAR is processed to separate the ground points from the working device points. Based on the bucket's pose, dynamic point clouds are extracted, elevation information is obtained, and elevation profiles are reconstructed to compute key point distances and predict the bucket's landing

point. Finally, simulation experiments were conducted on the Gazebo platform across multiple scenarios and working conditions to validate the application effectiveness of the proposed method. The experimental results demonstrate that the proposed method excels not only in prediction accuracy but also in computational efficiency, indicating that it can achieve rapid and accurate bucket landing point predictions in practical engineering vehicle operations, with promising engineering application prospects.

The proposed landing point prediction method integrating point cloud information still has areas for improvement. For example, it currently only displays elevation maps for the vertical ground area corresponding to the bucket, without considering the motion trajectory of the excavator's working device. Future work will explore integrating motion trajectories and neural network algorithms to further enhance the prediction of the bucket's landing point.

References

1. Yao Z, Zhao S, Tan X, Wei W, Wang Y (2023) Real-time task-oriented continuous digging trajectory planning for excavator arms. Autom Constr 152:104916
2. Feng C, Shen Y, Wang J, Wang Q, Suo Z, Su F (2024) Task-unit based trajectory generation for excavators utilizing expert operator skills. Autom Constr 158
3. Morita T, Ishikawa M, Naka R, Ogawa S (2021). Development and evaluation of automatic digging and pooling sand system for unlevel digging experiment. In: The proceedings of JSME annual conference on robotics and mechatronics (Robomec), 2P2–A02
4. Schönböck J et al (2022) From remote-controlled excavators to digitized construction sites—sciencedirect
5. Eraliev OMU, Lee KH, Shin DY, Lee CH (2022) Sensing, perception, decision, planning and action of autonomous excavators. Autom Constr
6. Richard B, Matthew G, Kianoosh S, Xie Y (2023) Machine learning full 3-D lower-body kinematics and kinetics on patients with osteoarthritis from electromyography. Biomed Eng Adv 5:100088
7. Tang J, Luo H, Chen W, Wong KY, Cheng JCP (2022) Imu-based full-body pose estimation for construction machines using kinematics modeling. Autom Constr 138
8. Vahdatikhaki F, Hammad A, Siddiqui H (2015) Optimization-based excavator pose estimation using real-time location systems. Autom Constr 56:76–92
9. Wind H, Renner A, Schaut S, Albrecht S, Sawodny O (2019) Comparison of joint angle, velocity and acceleration estimators for hydraulically actuated manipulators to a novel dynamical approach. Control Eng Pract 91:104118
10. Kim JH, Bae JH, Jung WY (2016) A study on position recognition of bucket tip for excavator. J Drive Control 13(1):49–53
11. Lee S, Park H, Chu B (2022) Excavator posture estimation and position tracking system based on kinematics and sensor network to control mist-spraying robot. IEEE Access 10:107949–107960
12. Sumanas M, Petronis A, Urbonis D, Januskevicius T, Bucinskas V (2021) Determination of excavator tool position using absolute sensors. In: 2021 IEEE open conference of electrical, electronic and information sciences (eStream). IEEE
13. Molaei A, Geimer M, Kolum A (2022) An approach for estimation of swing angle and digging depth during excavation operation. In: Proceedings of the international symposium on automation and robotics in construction (IAARC)

14. Zheng Y, Wang S, Liu Y, Xu Y, Li X, Chen G (2018) Research on refined 3D attitude model of smart construction machinery based on multi-sensor fusion. In: China satellite navigation conference
15. Xu J, Yoon HS (2019) Vision-based estimation of excavator manipulator pose for automated grading control. Autom Constr 98:122–131
16. Liu G, Wang Q, Wang T (2022) A new measurement method of real-time pose estimation for an automatic hydraulic excavator. In: 2022 IEEE/ASME international conference on advanced intelligent mechatronics (AIM)
17. Shin J, Park H, Jeong H, Jeong H, Chu B (2024) Image-recognition-based embedded system for excavator bucket tracking in construction sites. Int J Precis Eng Manuf 25(8):1689–1701
18. Yeom DJ, Yoo HS, Kim JS (2023) Development of a vision-based machine guidance system for hydraulic excavators. J Asian Architect Build Eng 22(3):1564–1581
19. Kashani AH, Owen WS, Himmelman N, Peter DL, Robert AH (2010) Laser scanner-based end-effector tracking and joint variable extraction for heavy machinery. Int J Robot Res 29(10):1338–1352
20. Kashani AH, Owen WS, Lawrence PD, Hall RA (2007) Real-time robot joint variable extraction from a laser scanner. In: IEEE international conference on automation & logistics. IEEE
21. Phillips TG, Mcaree PR (2018) An evidence-based approach to object pose estimation from lidar measurements in challenging environments. J Field Robot 35(6):921–936
22. Kurisu Y, Shigaki T, Araki N, Araki N (2022) Posture estimation system for excavator manipulator using deep learning and inverse kinematics. In: 2022 World Automation Congress (WAC)
23. Lundeen KM, Dong S, Fredricks N, Akula M, Seo J, Kamat VR (2016) Optical marker-based end effector pose estimation for articulated excavators. Autom Constr 65(May):51–64
24. Liang CJ, Lundeen KM, Mcgee W, Menassa CC, Lee SH, Kamat VR (2019) A vision-based marker-less pose estimation system for articulated construction robots. In: Automation in construction 104:80–94
25. Ding H, Sang Z, Li Z, Shi J, Wang Y, Meng D (2024) Trajectory planning and control of large robotic excavators based on inclination-displacement mapping. Autom Constr 158:105209
26. Yuan Z, Chen T, He J, Wu C, Wei J (2024) A dual-model adaptive Kalman filtering for remaining useful life prediction method based on feature fusion and online tsp recognition. Measurement 235
27. Lee S, Lim H, Myung H (2022) Patchwork++: fast and robust ground segmentation solving partial under-segmentation using 3D point cloud. In: 2022 IEEE/RSJ international conference on intelligent robots and systems (IROS). IEEE, pp 13276–13283
28. Kumar GA, Rao BS, Jagannadha Rao CVKVP (2022) Accuracy assessment of various interpolation techniques used for generation of precise geoidal undulations in lidar projects. J Ind Soc Rem Sens

Open Access This chapter is licensed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License (http://creativecommons.org/licenses/by-nc-nd/4.0/), which permits any noncommercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if you modified the licensed material. You do not have permission under this license to share adapted material derived from this chapter or parts of it.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.



Research on Longitudinal Speed Estimation Methods for Multi-wheel Skidding of Multi-axle Distributed Electric Vehicles Based on Wheel Speed Prediction



Mingyang Tong, Junqiu Li, Yongxi Yang, Xiaohan Li, and Wuze Wang

Abstract In order to solve the problem of estimating the longitudinal speed caused by wheel slippage of multi-axle distributed electric vehicles under off-road conditions, a longitudinal vehicle speed fusion estimation algorithm is proposed. Based on triple exponential smoothing and a long short-term memory network, the wheel speed is predicted, and an interactive multi-model method is used to fuse the output prediction results. In this paper, a noise covariance matrix adjusting method based on the trend of wheel speed change is adopted, which the actively adjusted Kalman filter (AAKF) for speed estimation is realized. The acceleration signal is not used ensuring the availability of the AAKF when the inertial measurement unit fails. The AAKF vehicle speed estimation results are fused to ensure the reliability of the algorithm. The simulation and real vehicle test results verify the accuracy of the proposed algorithm.

Keywords Longitudinal speed · Multi-axle distributed electric vehicles · Fusion estimation · Actively adjusted Kalman filter

1 Introduction

The longitudinal vehicle speed is a critical parameter of vehicle motion state, and the accuracy and efficiency of its estimation directly affect the performance of vehicle active safety control. As the sensors used for state estimation are more diverse, however, the safety risks posed by sensor failure are also extremely intractable. Thus, ensuring sensor redundancy is essential while striving for accurate estimation with the minimal number of sensors required.

M. Tong · J. Li (✉) · Y. Yang · X. Li · W. Wang
School of Mechanical Engineering, Beijing Institute of Technology, Beijing, China
e-mail: 1642990051@qq.com

© The Author(s) 2026

H. Saman et al. (eds.), *The 9th International Conference on Advances in Construction Machinery and Vehicle Engineering*, Lecture Notes in Mechanical Engineering,
https://doi.org/10.1007/978-981-96-5527-4_42

Currently, the vehicle speed estimation methods mainly include the dynamics-based, kinematic-based, and deep learning-based vehicle speed estimation methods [1]. Dynamics-based approaches require the construction of accurate vehicle models, which are suitable for specific conditions. Zeng et al. proposed a vehicle speed estimation method based on weak braking force for vehicle speed estimation under emergency braking conditions, which can obtain accurate vehicle speed estimation [2]. Peng et al. proposed a nonlinear observer based on the tire longitudinal force estimation and tire transverse friction model, which can provide accurate estimation results of longitudinal velocity, transverse velocity, and tire-pavement friction coefficient at the same time [3]. In [4, 5], improved Kalman filter methods are used to estimate the velocity of all-wheel drive vehicles based on high-precision dynamics models, and both methods work well. However, the estimation accuracy of dynamics-based is affected by the accuracy of the vehicle model and it is not very versatile [6]. The research based on kinematic methods mainly focuses on improving the accuracy of multi-sensor fusion estimation. Ding et al. combined wheel speed information, global positioning system (GPS), and inertial measurement unit (IMU) to compensate for the influence of road slope on speed estimation, ensuring the accuracy and reliability of vehicle speed estimation at a lower cost [7]. Based on the wheel speed signal and the motor speed signal, Pu et al. improved the accuracy of vehicle speed estimation under low speeds by switching different vehicle speed modes [8]. However, these methods depend on sensor performance and will be unrobust when sensors fail. The deep learning-based method has high estimation accuracy, but high computing resources of the controller is required. Thus, it is difficult to achieve on-board application in the short term [9, 10].

In this paper, a longitudinal vehicle speed fusion estimator is designed for multi-axle distributed electric drive vehicles to solve the speed estimation problem when multiple wheels of the vehicle are slipping. The core algorithm only uses pure wheel speed information, which ensures that the algorithm is still reliable when IMU fails. The simulation and real vehicle test results show that the proposed method can improve the accuracy and robustness of the estimation results under the premise of reducing the dependence on sensor information as much as possible. The kinematic model of the 8×8 multi-axle vehicle that is the object of study in this paper is shown in Fig. 1.

2 Longitudinal Vehicle Speed Fusion Estimation

The architecture of the longitudinal vehicle speed fusion estimation observer is shown in Fig. 2.

As shown in Fig. 2, the wheel speed u_i, acceleration a_x and a_y and yaw angular velocity $\dot{\psi}$, are fed into the observer by the filtering algorithm. Then, the interactive multi-model (IMM) predicts each wheel speed based on the historical information of the wheel speed. Through IMM prediction, the noise covariance matrix adjuster (NCMA) will dynamically adjust the noise covariance matrix Q and R. By adjusting Q

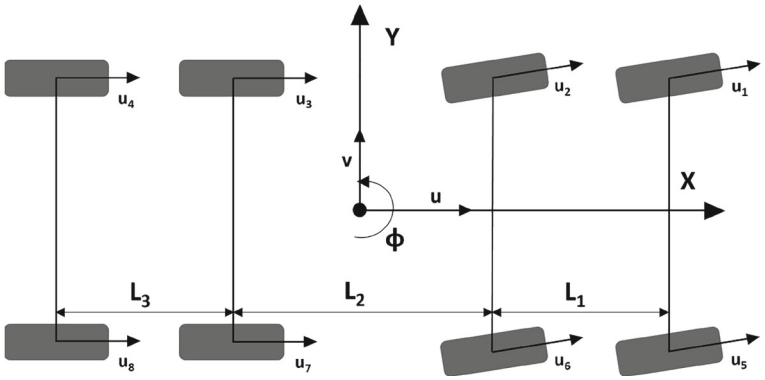


Fig. 1 8 × 8 multi-axis distributed electric drive vehicles

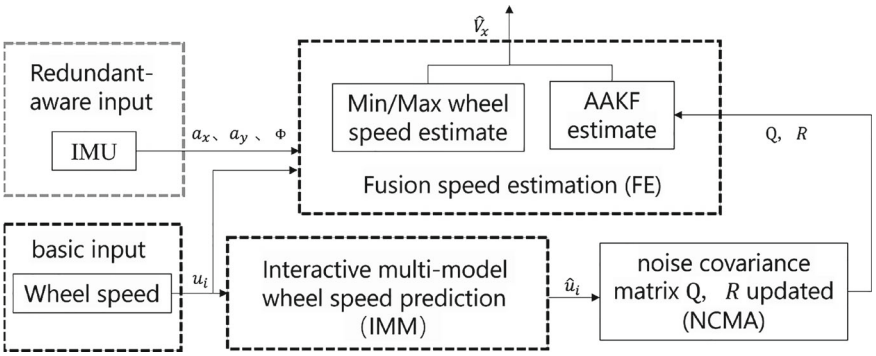


Fig. 2 Architecture of longitudinal speed observer

and *R*, the estimation results of AAKF are improved. Finally, a fusion estimation (FE) for output is used which fuses the result from AAKF with the minimum/maximum wheel speed estimation. It is worth mentioning that the IMU signal in this paper belongs to redundant perception information, and the core algorithm only relies on wheel speed signal processing to ensure the estimation accuracy and stability of the algorithm.

2.1 Actively Adjusted Kalman Filter (AAKF) Vehicle Speed Estimator

In order to describe the longitudinal velocity of the vehicle more accurately, considering the influence of the lateral and yaw movements of the vehicle on the longitudinal velocity, the differential equation of the longitudinal velocity can be defined as:

$$\dot{v}_x = \dot{\psi} v_y + a_x + w_x \tag{1}$$

where w_x denotes the process noise in the longitudinal velocity.

Regard the deviation between wheel speed and longitudinal velocity as measurement noise, so the measurement relationship between wheel speed and longitudinal speed can be obtained:

$$v_x = u_i + d_i \tag{2}$$

where d_i denotes the measured noise between the wheel-i and the longitudinal velocity.

In this paper, the measurement error caused by wheel slip is regarded as measurement noise, so the key to improving the accuracy of vehicle speed estimation is inhibiting d_i and w_x. The forward Eulerian method is used to discretize the Eqs. (1) and (2), and the discrete linear time-varying state space and measurement equation of longitudinal vehicle speed are established:

$$x(k+1) = A(k)x(k) + B(k)u(k) + W(k) \tag{3}$$

$$Y(k) = C(k)x(k) + D(k) \tag{4}$$

where the state variable is $x(k) = \begin{bmatrix} v_x(k) & v_y(k) \end{bmatrix}^T$, the state matrix is $A(k) = \begin{bmatrix} 1 & \dot{\psi}(k)dt \\ -\dot{\psi}(k)dt & 1 \end{bmatrix}$, the input matrix is $B(k) = \begin{bmatrix} dt & 0 \\ 0 & dt \end{bmatrix}$, the control variable is $u(k) = \begin{bmatrix} a_x(k) & a_y(k) \end{bmatrix}^T$, the measurement matrix is $C(k) = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}^T$, process noise matrix is $W(k) = \begin{bmatrix} q_x & q_y \end{bmatrix}^T$, the measured noise matrix is $D(k) = [d_1, \ldots, d_8]^T$, the measurements variables are $Y(k) = [u_1(k), \ldots, u_8(k)]^T$, $\dot{\psi}(k)$ is the yaw angular velocity of the vehicle at the current moment, dt represents the computational step.

When IMU data is lost, $A(k)$ degenerates to the identity matrix and $B(k)$ degenerates to the null matrix, the observer no longer depends on the IMU, and $x(k)$ will be updated by the observation equation.

According to $W(k)$ and $D(k)$, the process noise covariance matrix $Q(k)$ with size 2×2 and the measurement covariance matrix $R(k)$ with magnitude of 8×8 is designed, respectively.

$$Q(k) = diag(q_1, q_2) \tag{5}$$

$$R(k) = diag(r_1, r_2, r_3, r_4, r_5, r_6, r_7, r_8) \tag{6}$$

where q_i denotes the reliability of v_x and v_y, then r_i represents the reliability of wheel-i's speed measurement. The larger the q_i and r_i are, the lower the confidence of v_x, v_y and r_i will be.

According to Eqs. (3), (4) and (5), the actively adjust Kalman filter can be established as:

$$\begin{cases} \hat{x}(k) = A(k)x(k-1) + B(k)u(k-1) \\ \widehat{P}(k) = A(k)P(k-1)A^T(k) + Q(k) \\ K(k) = \widehat{P}(k)H^T(k)\Big(H(k)\widehat{P}(k)H^T(k) + R(k)\Big)^{-1} \\ \tilde{x}(k) = \hat{x}(k) + K(k)\big(Y(k) - C(k)\hat{x}(k)\big) \\ P(k) = (I - K(k)H(k))\widehat{P}(k) \end{cases} \tag{7}$$

The covariance matrix $Q(k)$ and $R(k)$ in classical Kalman filter are constant matrices. In this paper, by actively adjusting the $Q(k)$ and $R(k)$ based on the wheel speed information, AAKF can be achieved. Section 2.3 will explain how to adjust $Q(k)$ and $R(k)$ in detail.

Since the IMU is considered as a redundant backup in this paper, it is necessary to avoid using its data to judge wheel slippage. However, when more than half of the wheels are slipping, the slippage noise will be difficult to inhibit. Therefore, a fusion estimation (FE) is designed which fuses the minimum/maximum wheel speed estimation as well as the AAKF result as output to improve the applicability and reliability of the algorithm.

2.2 *Interactive Multi-Model (IMM) Wheel Speed Prediction*

Based on the wheel speed sequence information to predict the wheel speed in the future, the covariance matrix of the measured noise can be adjusted earlier before the wheel slipping. In this section, the methods of triple exponential smoothing (TES) and long short-term memory network (LSTM) fusion are used to predict wheel speed.

The TES for wheel speed prediction is established as follows:

$$\begin{cases} S(0) = u_i(0) \\ S(t) = \alpha \frac{u_i(t)}{c(t-L)} + (1-\alpha)(S(t-1) + b(t-1)) \\ b(t) = \beta(S(t) - S(t-1)) + (1-\beta)b(t-1) \\ c(t) = \gamma \frac{u_i(t)}{S(t)} + (1-\gamma)c(t-L) \end{cases} \tag{8}$$

where $u_i(t)$ is the measured value, which is the wheel speed measurement value at the current moment. $S(t)$ represents the wheel speed after exponential smoothing which is the state variable. $b(t)$ denotes the time trend variable, including the trend of wheel speed change at the current moment. Defined as the seasonal cycle description variable, $c(t)$ implies the periodic law of vehicle wheel speed. α, β and γ are the

exponential smoothing factors of the above three variables, respectively, indicating the confidence of the variables at the historical moment and the current moment.

Further the prediction results of single-wheel wheel speed based on TES can be obtained:

$$u_i{}^1(t+m) = (S(t) - mb(t))c((t - L - 1)\%(m - 1)) \tag{9}$$

where m is the predicted step size, $u_i{}^1(t+m)$ is the predicted result of the wheel speed after m step.

Using the method of cross-validation, the offline training is carried out based on the wheel speed data collected by the real vehicle. Then the optimal hyperparameters α, β, γ, m can be searched. The objective function of the regularized squared loss error is defined as follows:

$$J = min \sum_{t=m}^{end} \left(u_i{}^1(t) - u_i(t)\right)^2 + \lambda(\alpha, \beta, \gamma) \tag{10}$$

In practice, the periodic variation of wheel speed is complex and changeable, and the use of γ is certainly not enough to fully describe the law of wheel speed change. In order to make up for the shortcomings of the triple exponential smoothing method in periodic prediction, a long short-term memory network (LSTM) is used to improve the prediction accuracy. It captures and processes long-term dependencies in sequence data more effectively. LSTM for wheel speed forecast is shown below:

As shown in Fig. 3, $u_i(t)$ is the measured wheel speed at current moment, $u_i{}^2(t)$ is the predicted wheel speed of n steps in the future, $C(t)$ is the cell state which determines the output of $y(t)$, and $h(t)$ is the hidden layer.

The IMM is used to fuse the single-wheel speed prediction results of the TES model and the LSTM model to achieve efficient prediction of short-term and periodic changes of wheel speed, and the methods are as follows:

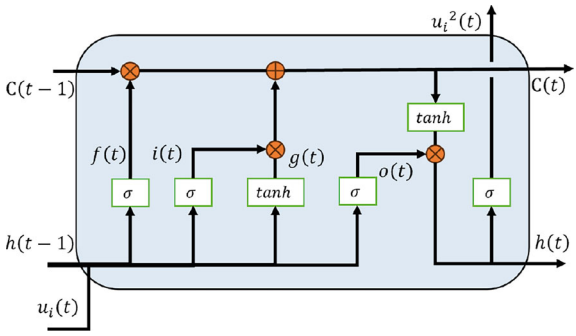


Fig. 3 LSTM architecture

$$\begin{cases} \Lambda^j(k) = \dfrac{\exp\left[\dfrac{{u_i}^j(k)^T \left(S^j(k)\right)^{-1} {u_i}^j(k)}{2}\right]}{\sqrt{2\pi\left|S^j(k)\right|}} \\ \mu^j(k) = \dfrac{\Lambda^j(k)c^j(k)}{\sum_{j=1}^{2}\Lambda^j(k)c^j(k)} \\ \hat{u}_i(k) = \sum_{j=1}^{2} u_i{}^j(k)\mu^j(k), j = 1{,}2 \end{cases} \tag{11}$$

where $j = 1$ indicates TES and $j = 2$ indicates LSTM, ${u_i}^j(k)^T$ are the wheel speed prediction results of each model, $S^j(k)$ is the corresponding prediction error covariance, $\Lambda^j(k)$ is the likelihood function of each model, $\mu^j(k)$ is the weight of the output, $c^j(k)$ is the probability of being in the model j after interaction.

Through the likelihood function of the two models, the weighted weights of the models are obtained. Finally, the wheel speed prediction results $\hat{u}_i(k)$ will be obtained based on weights fusion.

2.3 Design of Noise Covariance Matrix Adjuster (NCMA)

The general adaptive Kalman filter (AKF) can adjust the parameters according to the environmental changes under the condition of smooth operation of the vehicle, which can improve the accuracy of the vehicle speed estimation in the current environment. However, when the wheel speed changes sharply, especially when a slippage occurs, the estimated speed cannot track the real speed well because the dynamic response characteristics are relatively slow. Therefore, a new adjuster (NCMA) is designed to targeted adjust the noise covariance matrix $Q(k)$ and $R(k)$ in time when the vehicle state changes drastically, so as to reduce the adverse impact of measurement noise on the estimation algorithm.

The NCMA takes the prediction result of the wheel speed estimation estimated by IMM as input to judge the slippage of each wheel. According to the position of the slipping wheel-i, the diagonal element r_i in $R(k)$ will be marked. Then the noise covariance matrix gradient ∇r_i will be generated based on the degree of slippage. Through this gradient ∇r_i, the marked elements will get changed, and then the adjustment of the $R(k)$ matrix is realized.

When most of the wheels are slipped, it means that the results of the measurement equation become less reliable. At this time, the NCMA algorithm will begin to adjust the process covariance matrix $Q(k)$. The dynamic adjustment process of the NCMA is shown in Fig. 4.

The process of each step is as follows:

$$\begin{cases} \overline{u}(k) = \sum_{i=1}^{n} \frac{u_i(k)}{n} \\ \nabla u_i(k) = u_i(k) - \overline{u}(k), i = 1{,}2 \cdots 8 \\ u_{std}(k) = \sqrt{\frac{\sum_{i=1}^{n}(u_i(k) - \overline{u}(k))^2}{n-1}} \end{cases} \tag{12}$$

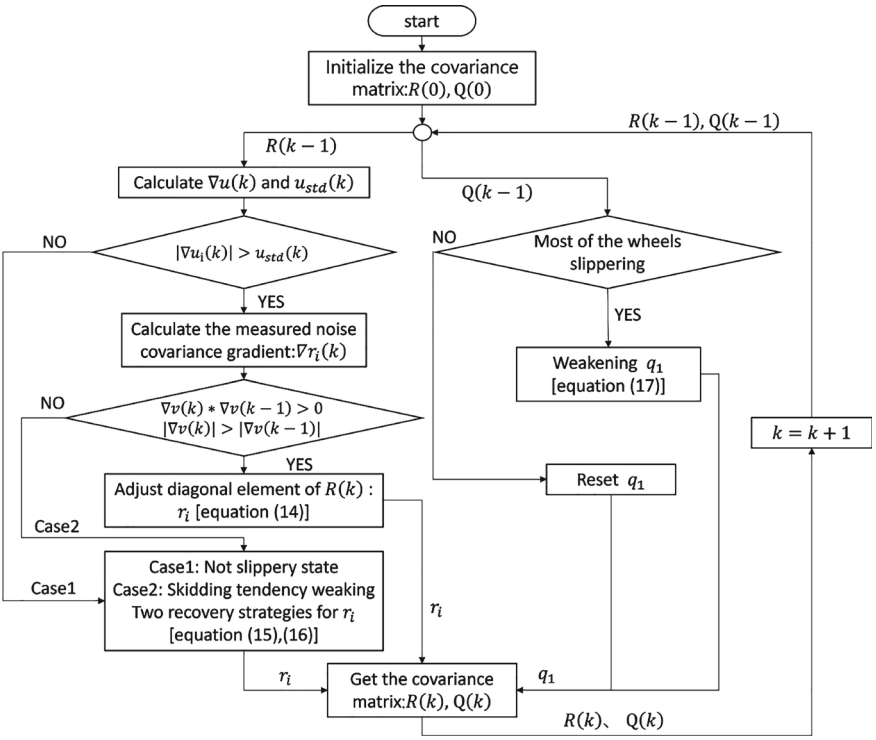


Fig. 4 Noise covariance matrix adjuster for $R(k)$ and $Q(k)$

$$\nabla r_i(k) = D * |(\nabla u_i(k) - \nabla u_i(k-1))| \tag{13}$$

$$r_i(k) = r_i(k-1) + \nabla r_i(k) \tag{14}$$

$$r_i(k) = \begin{cases} \max(0.5 * r_i(k-1) * B_1, 1), |\nabla u_i(k)| < 0.5 * u_{std}(k) \\ \max(r_i(k-1) * B_1, 1), 0.5 * u_{std}(k) < |\nabla u_i(k)| \end{cases} \tag{15}$$

$$r_i(k) = \max(r_i(k-1) * B_2, 1), |\nabla u_i(k)| < |\nabla u_i(k-1)| \tag{16}$$

$$q_1(k) = \max(q_1(k-1) * B_3, 1) \tag{17}$$

where $\overline{u}(k), u_{std}(k), \nabla u_i(k)$ represent the average wheel speed, the standard deviation of the wheel speed, and the difference between the wheel-i's speed and the average wheel speed, respectively. $\nabla r_i(k)$ denotes the covariance matrix gradient for $r_i(k)$ based on the trend of wheel slippage. D is the gradient factor, which quantifies the effect of the trend of $\nabla u_i(k)$ on $\nabla r_i(k)$. B_1, B_2 represent the attenuation scale factor of the measured noise covariance coefficient, designing for restoring the reliability

of the corresponding wheel speed. B_3 is the attenuation scale factor for the process noise matrix $Q(k)$.

The adjusting process of $R(k)$ is shown in Eqs. (13) and (14), which diagonal elements $r_i(k)$ are changed. The gradient accumulation strategy is used to gradually suppress the influence of skidding wheels on vehicle speed estimation by memorizing historical residual gradient information $r_i(k-1)$. Two proportional recovery strategies for measuring noise covariance coefficients are designed. As shown in Fig. 4, Eq. (15) is design for Case1 (Not slippery state) and the Eq. (16) is design for Case2 (Skidding tendency weaking) to recover $r_i(k)$, respectively.

The adjusting process of $Q(k)$ is shown in Eq. (17). When multiple (more than 4) wheels are detected to slip, the value of q_1 is reduced. And when the number of wheel slippage is less than 4, $q_1(k)$ returns to the initial value so that the measured value will be fully utilized.

3 Simulation Verification

In order to verify the effectiveness and real-time performance of the vehicle speed observer, a hardware-in-the-loop (HIL) simulation platform is built for algorithm testing. An 8×8 distributed electric drive vehicle dynamics model is established and ran on the HIL simulator. At the same time, a driver model is designed to simulate the driver's steering wheel angle and the output of pedal signals. The vehicle controller receives the driver's instruction and convert it into control command. The vehicle model receives commands via the Controller Area Network (CAN) communication system. The motion state information of the vehicle is fed back to the host computer by the HIL simulator, and the speed estimation algorithm of this paper is solved and verified by the host computer. The overall architecture of the simulation experiment is shown in Fig. 5.

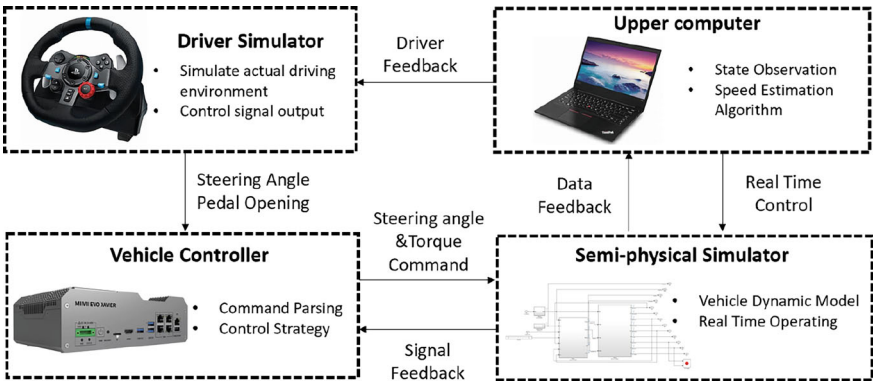


Fig. 5 Architecture of the simulation test bench

3.1 Simulation Under High-Adhesion Pavement

To verify the estimation accuracy of the FE under general road conditions, four types of simulation conditions are designed: direct acceleration, steering, braking and climbing. Set the road adhesion coefficient to 0.8, then verify the ability of the observer to track the actual vehicle speed. The simulation results are shown in Fig. 6.

In Fig. 6a–c, the vehicle starts accelerating, steering, and braking respectively at the first second. Figure 6d shows the continuous climbing condition. Within the specified time, the error of each simulation result is less than 1.5%. Moreover, the estimation accuracy of the observer for the straight acceleration condition is less than 0.2%. Therefore, this speed estimation method is suitable for the estimation of vehicle speed under general road conditions.

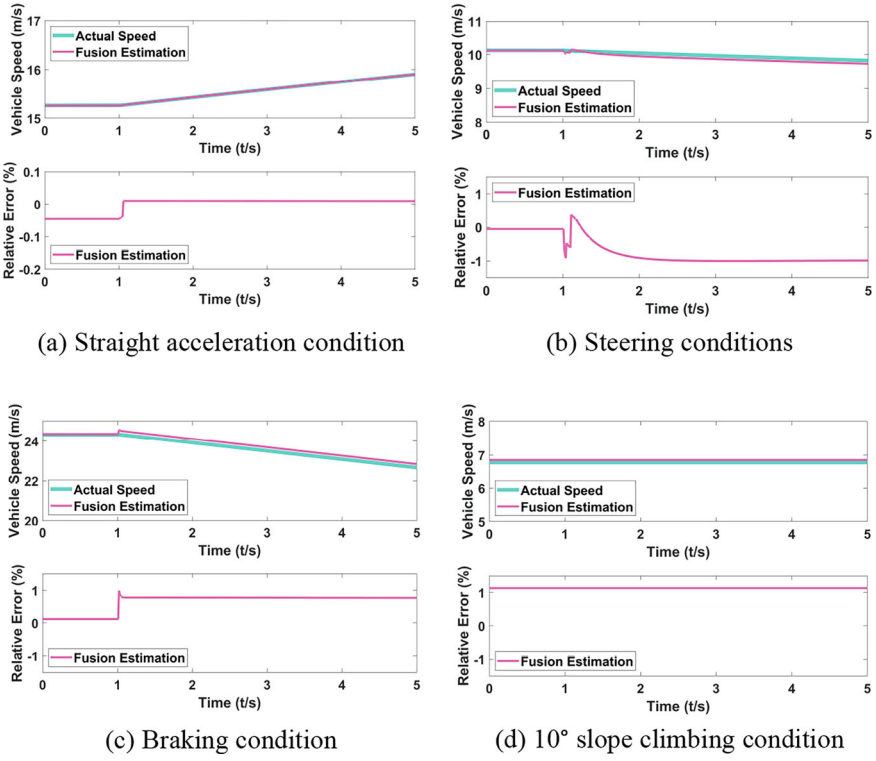


Fig. 6 Longitudinal speed estimation under high adhesion pavement conditions

3.2 Simulation Under Low Adhesion Pavement

In this paper, the influence of short-term wheel slip on vehicle speed estimation is mainly studied, and the vehicle speed estimation results under the condition of full wheels slippage are not considered. Therefore, only scenarios where the length of the low-attachment section is less than the length of the front axle to the rear axle of the vehicle will be discussed.

The simulated vehicle is a four-axle distributed electric drive vehicle, with the wheelbases between 1st and 2nd axles, 2nd and 3rd axles, and 3rd and 4th axles being 2, 3, and 2 m, respectively. First, the vehicle drive at a constant speed on the high-adhesion pavement with an adhesion coefficient of 0.8. Then when the 1st axle wheel gets into the low-adhesion pavement which the adhesion coefficient is 0.1, increasing the acceleration pedal opening, so the wheel of each axle will slip for a short time in turn. In order to study the slippage of only single-axle wheels and the simultaneous slippage of multi-axle wheels, the length of the low-adhesion road was set to 2m and 6m, respectively. The simulation results are shown in Fig. 7.

As shown in Fig. 7a, because the length of the low-adhesion road is 2 m which is less than the length between the axles, only one axle slips at the same time. The measurement error caused by the slipping wheel is completely filtered out by the algorithm, and the FE performs well.

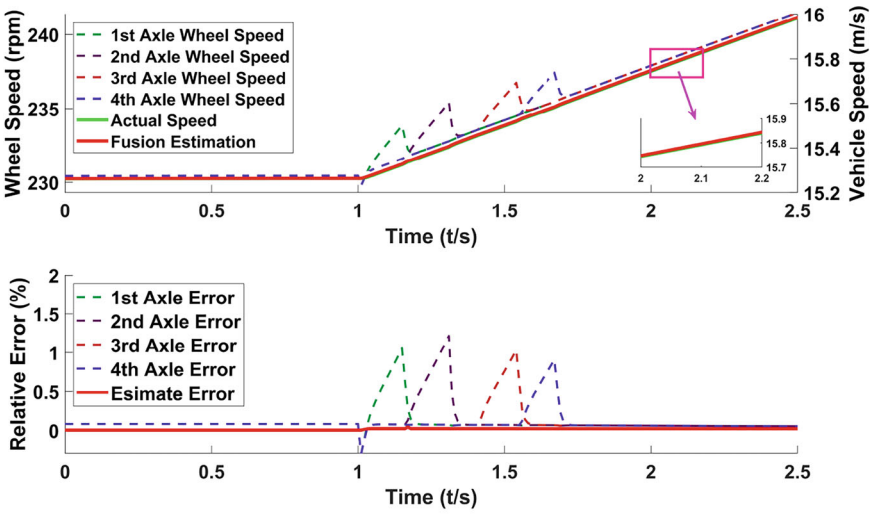
When the length of the low-adhesion road is 6 m, there are 6 wheels slipping at the same time in 1.25–1.45 s, However, the FE still tracks well to the true speed. As shown in Fig. 7b, there is a small deviation in the vehicle speed estimation between 1.15 and 1.25 s, but the FE quickly corrects it. In both condition the algorithm tracks the actual speed well, and the estimated effect is zoomed in the small pink box for show.

3.3 Simulation Under Acceleration Signal Lost

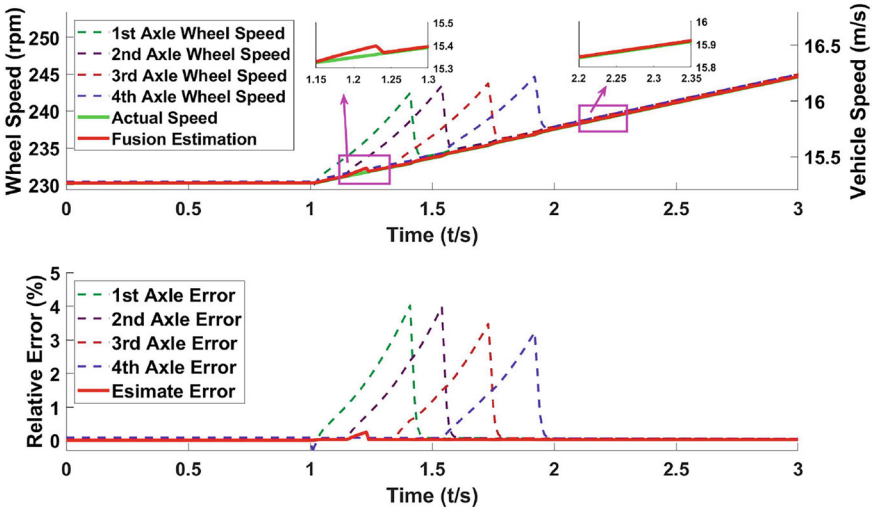
The FE algorithm realizes the dynamic adjustment of the noise covariance matrix without using IMU signals, maintaining a certain estimation accuracy. Therefore, the vehicle speed estimation will be carried out without acceleration (IMU) signal to verify the reliability of the algorithm.

Using the road condition same as Sect. 3.2, the length of the low-adhesion pavement is set to 6 m. Then compare the vehicle speed estimation results with and without acceleration signals.

Figure 8 shows the speed estimation results with and without acceleration signals. Except for the deviation between the two at 1.75 s–1.80 s, the estimation results of the two sides are basically the same at other times. The simulation results verify that the algorithm can still achieve accurate estimation results when the acceleration signal is lost.



(a) Result for 2m length of the low-adhesion pavement



(b) Result for 6m length of the low-adhesion pavement

Fig. 7 Vehicle speed estimation through a low-adhesion pavement of different lengths

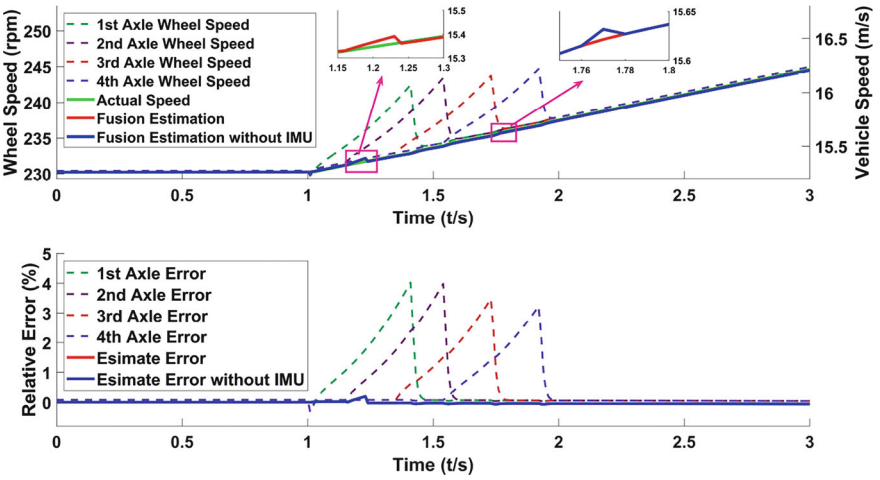


Fig. 8 Vehicle speed estimation under the condition of acceleration signal lost

4 Verified by Real Vehicle Test

According to the simulation verification results above, the verification of real vehicles is carried out. The experimental vehicle is an 8 × 8 multi-axle heavy-duty off-road vehicle, and the actual reference speed of the vehicle was obtained by GPS, which the measurement accuracy is 0.01 m/s.

The experiment take place on an unstructured road, where several instances of wheel slippage are observed. The method proposed in this paper (FE) and the conventional adaptive Kalman filter (AKF) method are used to estimate the vehicle speed. Each wheel speed and the estimation results of vehicle speed during the test process are shown in Fig. 9.

Figure 9a shows the variation of the wheel speed of each wheel. The vehicle starts to move at 3.1, and at 5.9–6.9 s, the vehicle leans right, and the wheels of the 2nd and 3rd axles on the left slip. At 7.2–8.4 s, the 2nd axle of the vehicle are suspended, so both wheels of the 2nd axle slip. In the 8.4–9.6 s, the road surface conditions are more complicated, and left wheels of the 2nd and 3rd axles have obvious slippage, while the right 2nd axle wheels have less slippage. At 9.6–10.4 s, the left 3rd axle wheel slips significantly. After that, the vehicle slows down gradually, and due to insufficient adhesion to the road surface, the left 2nd axle wheel gets locked in the 13.7–14.0 s.

Figure 9b shows the results of the vehicle speed estimation. In areas where the vehicle is running relatively smoothly, both methods track well to the actual speed. At 8.4–9.6 s, three wheels slippage occurs, and the FE method still tracks the real speed well, while the estimation method using AKF shows a large error. After that, the number of skidded wheels decreased, and the estimation accuracy of the FE method is improved, while the AKF method fails to track the target vehicle speed in time.

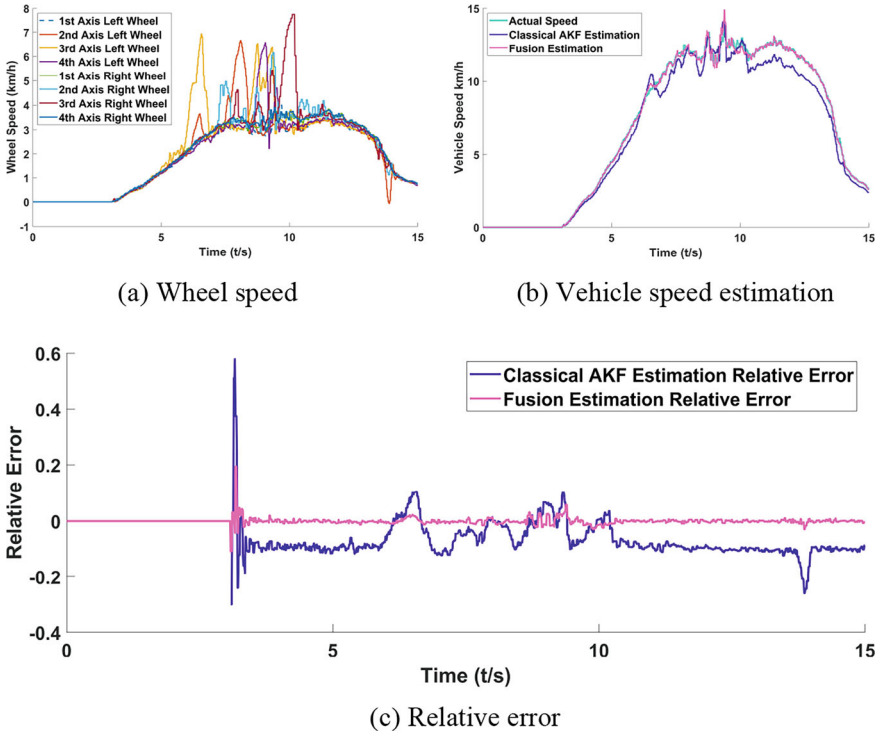


Fig. 9 Test results of real vehicles

Figure 9c shows the relative error of the vehicle speed estimate. In order to intuitively compare the accuracy of different vehicle speed estimation algorithms, the average estimation errors of the two methods are statistically calculated. The relative error of the FE method is 1.3%, while that of the AKF method is 7.5%. The real vehicle test proves the estimation algorithm in this paper can still get high accuracy when multiple wheels have large slippage.

5 Conclusion

In this paper, an actively adjusted Kalman filter (AAKF) algorithm based on the wheel speed prediction is proposed, which dynamically adjusts the weight of the noise covariance matrix by predicting the trend of wheel slip. The impact of wheel slip on vehicle speed estimation can be suppress through this method. And in order to improve the scope of application of the AAKF, a fusion estimation (FE) is used which fuses the result from AAKF with the minimum/maximum wheel speed estimation. The innovative non-use of acceleration (IMU) signal successfully avoids the problem

of algorithm failure due to IMU failure. Experiment results show that the proposed algorithm can ensure high estimation accuracy with/without acceleration signal.

Acknowledgements The research of this article was supported by the National Key Research and Development Program of China (2021YFB2501801).

References

1. Editorial Department of China Journal of Highway and Transport (2023) Review on China's automotive engineering research progress: 2023. China J Highw Transp 36(11):1–192
2. Zeng X, Qian Q, Song D, Gao H, Wu J (2022) Estimation method of longitudinal vehicle speed under emergency braking condition. Automot Eng 44(09):1425–1436
3. Peng Y, Chen J, Ma Y (2019) Observer-based estimation of velocity and tire-road friction coefficient for vehicle control systems. Nonlinear Dyn 96(01):363–387
4. Wu D, Du K, Liu H and Liu X (2023) Longitudinal speed estimation of multi-axle distributed drive vehicle based on federal Kalman filter. In: CVCI, Changsha, China, pp 1–7
5. Zhang Y, Li M, Zhang Y, Hu Z, Sun Q, Lu B (2022) An enhanced adaptive unscented Kalman filter for vehicle state estimation. IEEE Trans Instrum Meas 77:1–12
6. Chu W, Li S, Jiang Q, Liu L, Luo Y (2011) Speed estimation for all-wheel drive vehicles based on multi-information fusion. Automot Eng 33(11):962–966
7. Ding X, Wang Z, Zhang L, Wang C (2020) Longitudinal vehicle speed estimation for four-wheel-independently-actuated electric vehicles based on multi-sensor fusion. IEEE Trans Veh Technol 69(11):12797–12806
8. Pu Z, Tang L, Shangguan W, Wang W and Jiang K (2023) Research on the estimation of vehicle speed under low-speed conditions based on multi-sensor information. Autom Eng 45(07):1235–1243+1275
9. Choi S, Shon H, Huh H (2024) Interpretable vehicle speed estimation based on dual attention network for 4WD off-road vehicles. IEEE Trans Intell Veh 9(01):151–164
10. Gao Z, Wen W, Tang M, Zhang J, Chen G (2022) Estimation of vehicle motion state based on hybrid neural network. Automot Eng 44(10):1527–1536

Open Access This chapter is licensed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License (http://creativecommons.org/licenses/by-nc-nd/4.0/), which permits any noncommercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if you modified the licensed material. You do not have permission under this license to share adapted material derived from this chapter or parts of it.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.



Research on Visual SLAM Integrating Feature Method and Direct Method in Sparse Texture Environment



Yuxin Cui, Chunpeng Zhang, Tao Ni, Dingxuan Zhao, and Hongyan Zhang

Abstract Aiming at the difficulty in feature extraction and matching in visual SLAM in a sparse texture environment, this paper proposes a sparse semi-direct SLAM that integrates feature method and direct method. Based on the reprojection error model and photometric error model, a theoretical model for sparse semi-direct SLAM method is established. Based on the ORB-SLAM2, a software of sparse semi-direct SLAM method is programmed. Experiments were conducted on multiple public datasets containing sparse texture environments, and the results showed that the sparse semi-direct SLAM proposed in this paper can achieve simultaneous localization and mapping in sparse texture environments, and its tracking robustness has been greatly improved. Compared with ORB-SLAM2, the maximum error of trajectories on fr3/structure_notexture_far dataset has been reduced by 65.06%, the root mean square error has been reduced by 36.08%, and other errors have also been significantly reduced. Overall, the trajectory accuracy is better than ORB-SLAM2.

Keywords Simultaneous localization and mapping · Sparse texture environment · Feature method · Direct method

1 Introduction

Simultaneous Localization and Mapping (SLAM) technology refers to the movement of a mobile robot in an unknown environment at an unknown location [1]. During the movement, the robot observes the surrounding environment through its own sensors, locates its position based on the environment, and performs incremental map construction. Using a camera as sensor is called visual SLAM. Unlike satellite

Y. Cui · C. Zhang · H. Zhang (✉)
School of Mechanical and Aerospace Engineering, Jilin University, Changchun, China
e-mail: zhanghy@jlu.edu.cn

T. Ni · D. Zhao
College of Vehicles and Energy, Yanshan University, Qinhuangdao, Hebei, China

© The Author(s) 2026

H. Saman et al. (eds.), *The 9th International Conference on Advances in Construction Machinery and Vehicle Engineering*, Lecture Notes in Mechanical Engineering,
https://doi.org/10.1007/978-981-96-5527-4_43

positioning commonly used in daily life, SLAM has a very high positioning accuracy, reaching centimeter level. Secondly, satellite positioning is usually available in outdoor open environments. In scenarios such as inside buildings, caves, tunnels, oceans, etc., due to signal obstruction, satellite positioning doesn't work. Because SLAM positioning is calculated from its own measurements, it can be used indoors. Furthermore, the positioning obtained through the SLAM system includes position and orientation, with six dimensional information, which can be extended to more application scenarios such as augmented reality.

There have been many landmark achievements in the field of visual SLAM research. Visual SLAM can be roughly divided into visual front-end, optimization back-end, map construction, and loop detection. In the visual front-end, those relying on feature point extraction and descriptor matching is called the feature method front-end. Among various SLAM schemes that use feature method in the front-end, The ORB-SLAM series proposed by the University of Zaragoza is a peak in the development of visual SLAM [2]. It innovatively uses three threads: tracking, local mapping, and loop closing to complete SLAM. The data exchange within the system is clear, making it very suitable for secondary development; Those not relying on features or descriptors is called the direct method front-end. Researchers from the Technical University of Munich proposed DSO, which expands the parameter range of joint optimization and includes camera exposure related parameters in the optimization framework [3]. The SVO proposed by the University of Zurich estimates inter frame motion by minimizing photometric error in the direct method, which runs very fast [4]. However, SVO and DSO only contain visual odometry parts and are not a complete SLAM system.

The environment in the real world is complex and diverse, including both rich texture and sparse texture scenes. The feature method SLAM highly relies on the extraction and matching of environmental feature point, so it is easy to lose in sparse texture environments. Although the direct method SLAM can overcome the localization difficulties caused by texture sparsity, it is highly constrainted by motion speed, lighting conditions, etc., and the disadvantage of tracking failure in high-speed motion is a common feature of all direct method SLAM.

In this paper, we propose an optimized algorithm that combines the feature method with direct method based on ORB-SLAM2 to improve the robustness of visual SLAM systems in sparse texture environments.

2 Theoretical Model of Sparse Semi-Direct Method

The SLAM problem can be represented by a general motion equation and an observation equation:

$$\begin{cases} x_k = f(x_{k-1}, u_k, w_k) \\ z_{k,j} = h(x_k, y_j, v_{k,j}) \end{cases} \tag{1}$$

The observation equation usually refers to the camera pinhole imaging model. Through the above equation, SLAM is modeled as a state estimation problem, which intuitively means the probability of the current motion data and observation data appearing is the highest. This is a maximum likelihood estimation. Assuming that the noise in the equation follows a Gaussian distribution $N(0, Q_{k,j})$ $N(0, Q_{k,j})$, The maximum likelihood estimate can be expressed as:

$$x = \arg\min(e(x_k)^T Q_{k,j}^{-1} e(x_k)) \tag{2}$$

where e_k e_k represents the error in the x_k x_k state, including motion error e_1 e_1 and observation error e_2 e_2.

$$e(x_k) = [e_1, e_2]^T = \begin{bmatrix} x_k - f(x_{k-1}, u_k) \\ z_{k,j} - h(x_k, y_j) \end{bmatrix} \tag{3}$$

The solution of pose in visual SLAM mainly involves observation errors and ignores the motion error.

The feature method SLAM solves the pose by optimizing the reprojection error, while the direct method SLAM solves the pose by optimizing the photometric error [5]. According to how much the utilization of image information, the direct method is divided into dense direct method, semi dense direct method, and sparse direct method [6]. In this paper, we propose a sparse semi-direct method [7], which implement the feature method in normal environment, while implement the direct method in sparse texture environment, so that guarantee the SLAM in any case.

3 Algorithm Framework of Sparse Semi-Direct SLAM

On the basis of ORB-SLAM2, improvements have been made by using sparse semi-direct method as the second method for recovering pose besides relocalization, and also as an option for normal tracking. The main improvement is in the tracking thread. The details of the optimized algorithm framework are shown in Fig. 1.

4 Experiment and Analysis

The TUM RGB-D dataset is a set of RGB-D datasets and real trajectory data recorded by the Technical University of Munich in Germany, used to evaluate visual odometry and visual SLAM systems. For different scenarios and application purposes, the dataset is divided into several series, including Structure vs. Texture sequences.

The fr3/nostructure_notexture_far dataset contains a series of surface images of wooden boards, with little visible structure or texture. The fr3/structure_notexture_

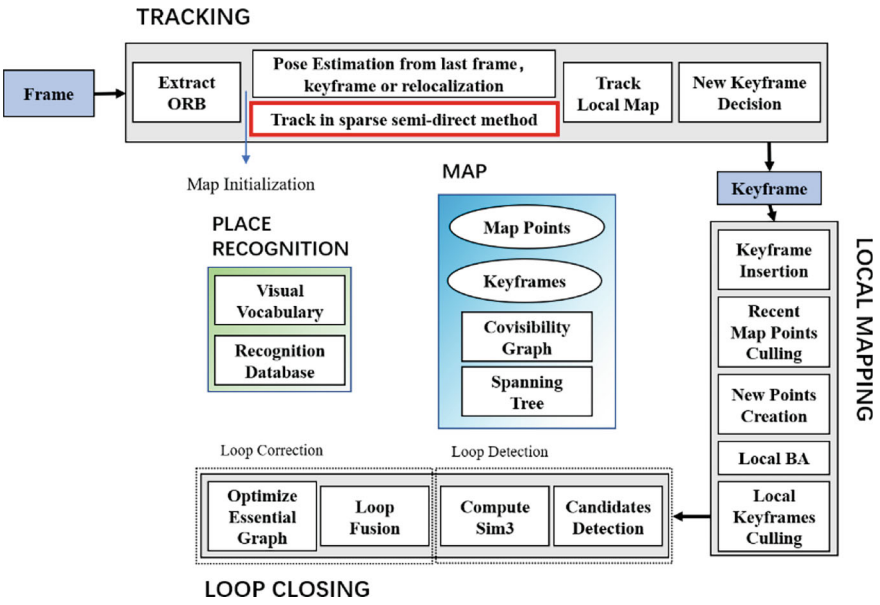


Fig. 1 Algorithm framework (red part is the optimization made different from the original system)

far was collected at approximately 1 m position, to capture images of wooden boards covered with plastic film with almost no texture. The RGB images are shown in Fig. 2.



Fig. 2 fr3/nostructure_notexture_far dataset (left), fr3/structure_notexture_far dataset (right)

4.1 Comparison Test of Robustness

Taking fr3/nostructure_notexture_far as an example, when running in ORB-SLAM2, tracking lost occurs due to the inability to match a sufficient number of feature pairs. The running effect is shown in Fig. 3.

In the right image of Fig. 3, the blue box represents the keyframes during the motion process, the green box represents the current position of the camera, and the red point cloud is a sparse point cloud map composed of map points. It can be seen that the system can only maintain tracking in the initial period of motion. The results of our method on fr3/nostructure_notexture_far is shown in Fig. 4. It can be seen that the optimized algorithm runs stably and the tracking can be maintained throughout the entire motion.

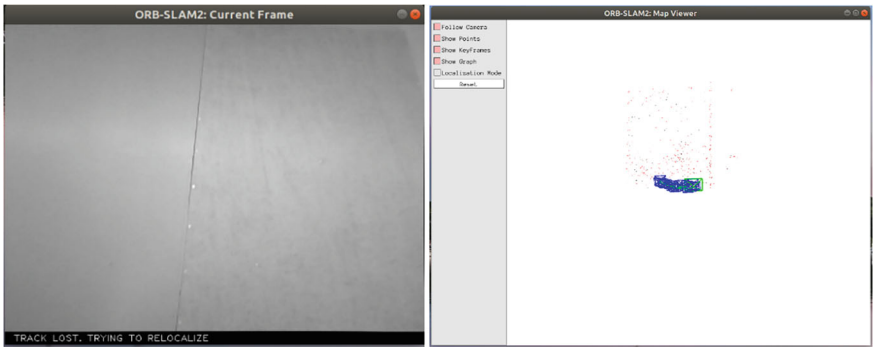


Fig. 3 ORB-SLAM2 runs on fr3/nostructure_notexture_far

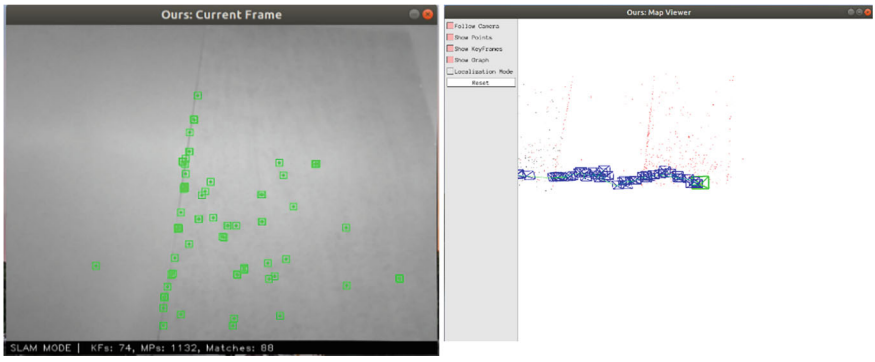


Fig. 4 Optimized algorithm runs on fr3/nostructure_notexture_far

4.2 *Trajectory Accuracy Test*

In order to verify the ability to ensure trajectory accuracy while improving tracking robustness, a comparison of trajectory accuracy was conducted with the ORB-SLAM2 on fr3/structure_notexture_far dataset. Figure 5 shows the absolute trajectory error comparison before and after algorithm optimization.

The comparison of absolute trajectory errors data before and after optimization on fr3/structure_notexture_far dataset is shown in Table 1.

By analyzing the chart, it can be seen that on fr3/structure_notexture_far dataset, the optimized algorithm significantly reduces multiple errors and improves trajectory accuracy.

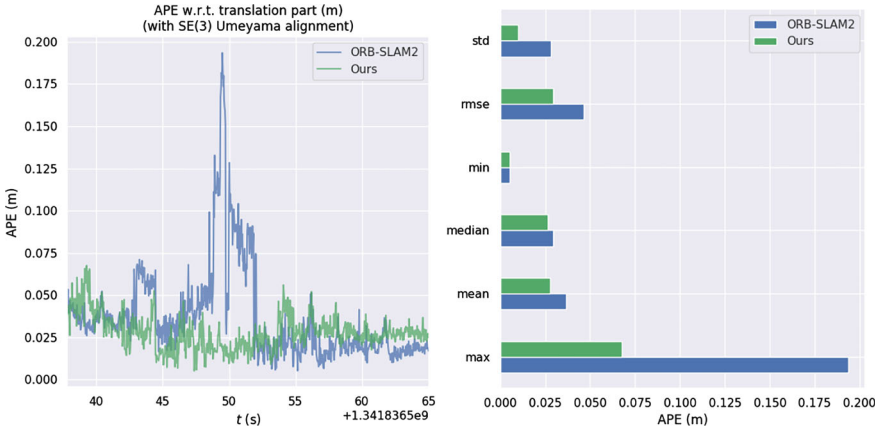


Fig. 5 Comparison of absolute trajectory accuracy of algorithms before and after optimization on fr3/structure_notexture_far dataset

Table 1 Absolute trajectory errors before and after optimization on fr3/structure_notexture_far

Error	Ours	ORB-SLAM2
Standard deviation (m)	**0.0101173**	0.0281625
Root mean square error (m)	**0.0296196**	0.0463373
Minimum error (m)	**0.0051874**	0.0052983
Median error (m)	**0.02666**	0.0295876
Mean error (m)	**0.0278381**	0.0367969
Maximum error (m)	**0.0675859**	0.193434

Bold means that the data is superior to the comparative method

5 Conclusion

In this paper, we optimize the tracking thread module based on ORB-SLAM2, integrate direct method with feature method, develop the theoretical model of sparse semi-direct SLAM, construct the algorithm framework, and obtain a sparse semi-direct SLAM algorithm.

Through experimental verification, it was found that the optimized algorithm significantly improved the robustness of tracking in sparse texture environments. The trajectory accuracy comparison experiments were conducted on sparse texture public dataset. It was found that on fr3/structure_notexture_far dataset, the optimized algorithm reduced the maximum error by 65.06% and the root mean square error by 36.08% compared to ORB-SLAM2, and all other errors were also relatively lower as well. The overall trajectory accuracy of the optimized algorithm is better than ORB-SLAM2.

References

1. Durrant-Whyte H, Bailey T (2006) Simultaneous localization and mapping: part I. IEEE Robot Autom Mag 13(2):99–110
2. Mur-Artal R, Tardós JD (2017) ORB-SLAM2: an open-source SLAM system for monocular, stereo, and RGB-D cameras. IEEE Trans Robot 33(5):1255–1262
3. Engel J, Koltun V, Cremers D (2018) Direct sparse odometry. IEEE Trans Pattern Anal Mach Intell 40(3):611–625
4. Forster C, Pizzoli M, Scaramuzza D (2014) SVO: fast semi-direct monocular visual odometry. In: 2014 IEEE international conference on robotics and automation (ICRA). Hong Kong, China, pp 15–22
5. Chen Z, Sheng W, Yang G et al (2018) Comparison and analysis of feature method and direct method in visual SLAM technology for social robots. In: 13th world congress on intelligent control and automation (WCICA), pp 413–417. Changsha, China
6. Zhang D, Liu Y, Zhang H (2022) Slam algorithm based on semi-direct visual odometry. In: 2nd international conference on algorithms, high performance computing and artificial intelligence (AHPCAI). Guangzhou, China, pp 14–18
7. Bu SH, Zhao Y, Wan G et al (2017) Semi-direct tracking and mapping with RGB-D camera for MAV. Multimed Tools Appl 76(3):4445–4469

Open Access This chapter is licensed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License (http://creativecommons.org/licenses/by-nc-nd/4.0/), which permits any noncommercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if you modified the licensed material. You do not have permission under this license to share adapted material derived from this chapter or parts of it.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.



Effect of Typhoon Load on the Attachment Height of a Rigid Attached Tower Crane



Xiangxiang Wang, Hongsheng Zhang, and Yue Wang

Abstract In view of the collapse accidents of most high-attached tower cranes under typhoons, this paper makes an in-depth analysis on the strength of tower cranes under typhoons and discusses the maximum attached height of attached tower cranes under typhoons. First, based on the average wind speed profile of typhoon exponential law, the exact calculation formula of the maximum bending moment of the tower crane under ideal conditions and the practical calculation formula under different loads are derived. Secondly, the overhang section of the attached tower crane is equivalent to the elastic support model, and the exact calculation formula of the maximum bending moment subjected to this model and the practical calculation formula under different loads are derived. Finally, the principle of strength equality is proposed, stating that the maximum bending moment of the ideal tower crane equals that of the attached overhang tower crane. This allows us to establish the relationship between the maximum height of the overhang section and the maximum independent working height of the tower crane. The maximum height coefficient of the overhang section is further compared to evaluate the strength and stability of the overhang section in the typhoon state more comprehensively. In general, assuming the axial force is 0, the overhang height ratio under a typhoon can be checked according to the overall stability formula. In the case of different coaxial force coefficients, the ratio of overhead height will be different.

Keywords Tower crane · Typhoon · Attachment distance · Maximum bending moment · Overhang height

X. Wang · H. Zhang · Y. Wang (✉)
School of Mechatronics Engineering, Harbin Institute of Technology, Harbin, Heilongjiang, China
e-mail: wang_yue@hit.edu.cn

© The Author(s) 2026

H. Saman et al. (eds.), *The 9th International Conference on Advances in Construction Machinery and Vehicle Engineering*, Lecture Notes in Mechanical Engineering,
https://doi.org/10.1007/978-981-96-5527-4_44

1 Introduction

With the rise of high-rise buildings, the use of attached tower cranes is becoming more widespread. The issue of attachment has garnered attention from engineering and technical personnel. When to start attaching, the influence of typhoons on attaching distance, and the maximum tower overhang height after attaching are all common problems in engineering, which are also the main contents to be discussed in this paper. Therefore, the maximum attachment height of a tower crane under a typhoon is studied to improve the safety performance of construction cranes in typhoons and improve the prevention and control level of major construction hazards.

Existing tower cranes' strength and stability calculations are based on specification [1] and related research is carried out. Luo et al. [2] obtained the cantilever height coefficients of tower cranes under different nodes by numerically calculating Euler's critical internal equality equations. Lin et al. [3, 4] verified the correctness of the cantilever height coefficient of the tower crane by using Matlab. Xu [5] analyzed the combination of the most unfavorable loads under various working conditions on the force of the attachment point of the assembled crane and proposed multiple methods to improve the attachment safety of the assembled structure tower crane. Based on the second-order theory, Liu [6]conducted a nonlinear modeling of the tower structure of the flexibly attached tower crane and improved the analytical expression of the internal force and equivalent stiffness of the rod of the flexibly attached tower crane by using the exact differential equation method. The methods derived in these literatures are all based on Euler critical force equalization, without considering the toppling strength of the tower crane. However, Huang et al. [7, 8] calculated that the failure ratio of the tower crane collapse attachment device was 15.20% through investigation and relatively complete data, emphasizing that the height of the tower crane cantilever end was too large, which was one of the main reasons for the collapse accident. Wang et al. [9] studied the force coefficient related to the height and distance of the tower crane through CFD simulation and analysis and calculated the corrected impact force by using ANSYS restart technology. Wang et al. [10, 11] obtained effective measures to reduce tower crane height and lightweight structure through simulation analysis.

Therefore, the existing studies primarily focus on calculating the height coefficient of the attached overhang section by equating the Euler critical force between the cantilever beam and the attached overhang section in the overall stability of tower cranes. However, few studies have addressed the strength of tower cranes under the action of typhoon exponential wind velocity profiles. This paper aims to address the strength problem of tower cranes by establishing the relationship between the maximum bending moment of the ideal tower crane and the attached overhang tower crane. Additionally, it seeks to determine the relationship between the maximum height of the overhang section and the maximum independent working height of the tower crane. Furthermore, the typhoon attachment height coefficients under different nodes are also revealed.

2 Calculation of the Maximum Bending Moment of the Ideal Cantilever Beam Model of the Tower Crane

The attached tower crane often needs to be attached because the height of the independent working state makes it difficult to meet the working requirements. Therefore, the maximum distance between the first attachment point and the ground (rail surface) should be the maximum allowable tower height of the tower crane under the independent working state after deducting the height of the jacking sleeve. The tower crane's maximum independent working height is determined by the stiffness, strength, and stability of the tower. The structure of the tower crane's tower body is made of lattice members that are subject to vertical and horizontal bending. The overall stability control formula for this type of bending structure in the tower crane design code is more strict than the strength control formula. The tower is typically a lattice truss with a symmetrical section, so there are no relatively strong or weak axes. It functions as a compression and bending member as a whole. The load condition of the tower body is shown in Fig. 1.

The condition under which the edge fibers yield to the limit state is:

$$\frac{N}{\mathrm{A}} + \frac{M_{\max}}{W_1} \leq \sigma_s \tag{1}$$

where N is the Axial force on the structure, A is the cross-sectional area of the member, $M_{\max}$ is the maximum bending moment in the member, W_1 is the flexural modulus of the maximum compressive fiber in the plane of bending moment.

For the cantilever beam structure, the maximum bending moment is at its bottom, including the first-order bending moment and the second-order bending moment.

$$M_{\max} = M_q + N\delta \tag{2}$$

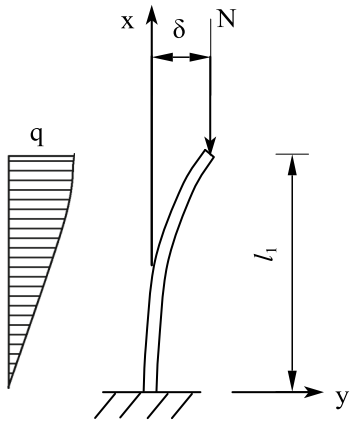


Fig. 1 Model of tower crane under typhoon condition

where M_q is the bending moment generated by the wind load on the component, which is a first-order bending moment, $N\delta$ is the bending moment produced by the axial force on the flexure deformation, which is called the second-order bending moment, δ is the end position shift.

When the external load is known, the independent height of the tower crane l_1 can be determined by the overall stability checking Eq. (1) of the bending member, so that the ideal position of the first attachment point can be theoretically determined. The ideal attachment distance of the second attachment point l_2 thereafter shall be determined by the maximum overhang height above the first tower body attachment point. When determining the attachment support point, we follow a principle: the maximum bending moment at the end of the extension section of the tower body after the establishment of the attachment support is consistent with the maximum bending moment at the end of the maximum height section in the independent working state. The dimensionless coefficient $\Omega = N/N_{cr}$ is introduced to represent the magnitude of axial force. $\Omega = 0.01 \sim 0.2$

The exponential law average wind velocity profile q is used to describe the wind load at the height of the wind pressure unit:

$$q = q_0\left(\frac{x}{10}\right)^{0.6}, q_0 = 0.625v_s^2 b\varphi(1+\xi) \tag{3}$$

where v_s is the calculated wind speed, b is the tower shift width, φ is the filling rate of the tower crane, ζ is the wind reduction factor of a tower crane.

The differential equation of deflection of the cantilever column is as follows:

$$EIy'' = M + N(\delta - y) + Q(l_1 - x) + \int_x^{l_1} q_0(t/10)^{0.6} \times (t-x)dt \tag{4}$$

Let $k^2 = N/EI$, and $M = 0$, $Q = 0$, so Eq. (5) is reduced to:

$$y'' = k^2(\delta - y) + \frac{k^2}{N} \times 0.251q_0(0.240x^{2.6} - 0.625xl^{1.6} + 0.385l^{2.6}) \tag{5}$$

where $0.240x^{2.6} - 0.625xl^{1.6} + 0.385l^{2.6}$ is simplified by an infinite series:

$$-0.0578l^{7.8} - 0.208l^{6.8}(x-l) - 0.270l^{5.8}(x-l)^2 - 0.144l^{4.8}(x-l)^3 - 0.0216l^{3.8}(x-l)^4 \tag{6}$$

Equation (6) is reduced to:

$$\begin{aligned} y'' = k^2(\delta - y) + \frac{0.0145q_0k^2}{N}(&-1.558f^{7.8} - 0.231f^{6.8}x \\ &+ 0.562f^{5.8}x^2 - 0.999f^{4.8}x^3 - 0.374f^{3.8}x^4) \end{aligned} \tag{7}$$

The differential Eq. (8) is solved and converted from the boundary condition: $x = 0, y = 0, x = l_1, y = \delta_1$ and $x = 0, y' = 0$ to obtain:

$$\delta_1 = \frac{q_0}{-N\cos(kl)}\left(\begin{array}{l}\dfrac{0.130327l^{3.8}}{k^4} - \dfrac{0.1358l^{5.8}}{k^2} + 0.0145l^{7.8} \\ +\dfrac{0.08691l^{5.8}\sin(kl)}{k^3} - \dfrac{0.00334l^{7.8}\sin(kl)}{k} \\ -\dfrac{0.1303l^{3.8}\cos(kl)}{k^4} - \dfrac{0.01629l^{5.8}\cos(kl)}{k^2} + 0.0006l^{7.8}\cos(kl)\end{array}\right) \tag{8}$$

At this time, we can obtain Eq. (3) to calculate the maximum bending moment of the tower body:

$$M_{\max} = \int_0^l q_0(x/10)^{0.6} \times x dx + N\delta_1 = \frac{q_0 l_1^{2.6}}{10.35} + N\delta_1 \tag{9}$$

3 Calculation of Elastic Support Stiffness and Maximum Bending Moment of Equivalent Bending Column

The attached structure of the upper structure of the attached tower crane can be equivalent to the flexure cantilever column with elastic support at the root, as shown in Fig. 2c. The root stiffness depends on the number of legs, span, axial pressure, and section size of the lower multi-span continuous pressure rod.

The tower body attachment structure in Fig. 2a is divided into two parts equivalent to the upper and lower parts, bounded by the attached support points: the lower tower body provides equivalent rotational stiffness C for the upper tower body, and the upper cantilever section is equivalent to a cantilever beam with elastic rotational support. What we need to do now is to determine the equivalent rotational stiffness C provided by the lower tower and the maximum bending moment of the cantilever segment on this elastic support.

The model used in this paper is similar to that in reference2. For node n = 1, the root rotational stiffness is:

$$C = \frac{EI}{l_1}\frac{\varepsilon\sin\varepsilon - \varepsilon^2\cos\varepsilon}{2(1-\cos\varepsilon) - \varepsilon\sin\varepsilon} \tag{10}$$

where $\varepsilon = l_1\sqrt{N/(EI)}$

From the theory of vertical and horizontal bending, we can determine the relationship between the rod end Angle φ_1 and the moment M of the structure in Fig. 2b as follows:

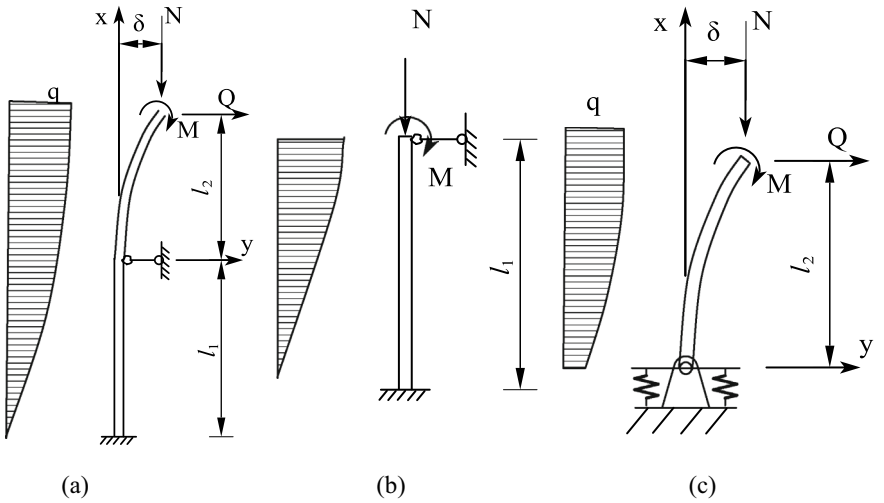


Fig. 2 The attachment structure diagram of the tower body

$$M = C\varphi_1 \tag{11}$$

The bottom rotational stiffness C of the equivalent cantilever structure is determined, and the differential equation of cantilever column deflection is obtained as follows:

$$EIy'' = M + N(\delta - y) + Q(l - x) - \int_x^{l_2} q_0[(t + l_1)/10]^{0.6} \times (t - x)dt \tag{12}$$

The deflection differential equation of the simplified equivalent cantilever structure is as follows:

$$EIy'' = N(\delta_2 - y) - \int_x^{l_2} q_0[(t + l_1)/10]^{0.6} \times (t - x)dt \tag{13}$$

Let $k^2 = N/EI$, the shift of the end of the overhang section be shown in the appendix by converting the boundary conditions $x = 0, y = 0, x = l_2, y = \delta_2$ and $x = 0, y' = [N\delta_1 + \frac{q_0 l_2}{6.37} \times (l_1 + l_2)^{1.6} - \frac{q_0}{16.56} \times (l_1 + l_2)^{2.6} + \frac{q_0}{16.56} \times l_1^{2.6}]/C$.

At this time, we can obtain Eq. (3) to calculate the maximum bending moment at the root of the overhanging section tower:

$$M_{\max 2} = \int_0^{l_2} q_0 \left(\frac{x + l_1}{10}\right)^{0.6} \times x dx + N\delta_2 = \frac{l_1^{2.6} q_0}{16.56} - \frac{q_0 (l_1 + l_2)^{2.6}}{16.56} + \frac{q_0 l_2 (l_1 + l_2)^{1.6}}{6.37} + N\delta_2 \tag{14}$$

Let Eqs. (10) and (15) equal, divide both sides by $q_0 l_1^{2.6}$, denoting the adhesion height ratio as the coefficient $\beta = l_2/l_1$, denoting the difference of the second order bending moment as W_2, the solution can be obtained:

$$0.03623 + 0.060386(1 + \beta)^{2.6} - 0.157\beta(1 + \beta)^{1.6} = \frac{N}{q_0 l_1^{2.6}}(\delta_2 - \delta_1) = W_2 \tag{15}$$

It is easy to know that the Euler critical force of a cantilever column with length l_1 is $N_{cr} = \pi^2 EI/(2l_1)^2$. $\varepsilon_1 = \pi\sqrt{\Omega}/2$, $\varepsilon_2 = \pi\beta\sqrt{\Omega}/2$ can be easily obtained, and the rotational stiffness C can be solved by substituting Eq. (11), as shown in Table 1.

The key to solving Eq. (16) is the calculation of W_2, and the expression of W_2 when $Cl_1/EI = 3.997$ is:

$$W_2 = \frac{0.3\beta^4(\beta + 3)(\beta + 1)^{3/5} \sin\left(1.571\beta\sqrt{\Omega}\right)}{(\beta + 1)^3\left(3\sin\left(1.571\beta\sqrt{\Omega}\right) + 8\cos\left(1.571\beta\sqrt{\Omega}\right)\right)} \tag{16}$$

It shows that the subtraction of the second-order bending moment corresponds to $\Omega = 0.01$ and $Cl_1/EI = 3.997$. And since the approximate range of the ratio coefficient of adhesion height is $\beta = 0.7 \sim 0.8$, when $\cos(0.15708\beta) \approx 0.99$, $\sin(0.15708\beta) \approx 0.12$ and $\Omega = 0.01$ can be obtained in Eq. (17)

$$W_2 = \frac{0.00467941\beta^4(\beta + 3)(\beta + 1)^{3/5}}{(\beta + 1)^3} \tag{17}$$

Equation (16) is reduced to:

$$0.03623 + 0.0604(1 + \beta)^{2.6} - 0.157\beta(1 + \beta)^{1.6} = \frac{0.004679\beta^4(\beta + 3)(\beta + 1)^{3/5}}{(\beta + 1)^3} \tag{18}$$

The solution to the positive number has understanding: $\beta \approx 0.769042$.

Table 1 Optimal attachment spacing ratio of attached supports under different loads

Ω	0(N = 0)	0.01	0.05	0.1	0.15	0.2
Cl_1/EI		3.997	3.984	3.967	3.950	3.933
β	0.77489	0.76904	0.7710	0.773226	0.802352	0.794104

Then, by using the above similar method, we can derive the rotational stiffness of the equivalent structure under the load condition $\Omega = 0.05, 0.1, 0.15, 0.2$ in turn, and then obtain the height ratio coefficient β of the overhang section under the same maximum bending moment.

This shows that the attachment height of the tower crane is related to the axial force, when the axial force coefficient $\Omega < 0.1$, taking into account the calculation error, the ratio of the maximum overhang height to the independent height $\beta \approx 0.77$, when the axial force coefficient $\Omega > 0.1$, the ratio of the maximum overhang height to the independent height $\beta \approx 0.80$. From this, we can get the optimal ratio of the maximum independent height of the attached tower crane to the attachment distance and overhang height under different loads.

4 Conclusion

Based on the deflection differential equation and the equivalent elastic support model of the attached structure, the maximum independent height is equivalent to the maximum bending moment of the tower body during the attached operation. An equation can be established for the maximum overhang height ratio with respect to the axial force coefficient ratio. By applying this principle, the maximum overhang height can be determined for various axial force coefficient ratios corresponding to actual engineering scenarios. The specific conclusions are as follows:

(1) When the ratio of overhang height to independent height is 0.77489, the bottom of the attached structure model and the bottom of the independent height tower structure model bear the same maximum bending moment.
(2) Under different axial force coefficients, the ratio of the maximum independent height of the attached tower crane to the optimal attachment distance and overhang height is different. When the axial force coefficient ratio is less than 0.1, the maximum overhang height is smaller, and when the axial force coefficient ratio is greater than 0.1, the maximum overhang height is larger.
(3) When the axial force coefficient ratio is less than 0.1, the maximum overhang height ratio calculated considering the maximum bending moment is greater than the maximum overhang height ratio of 0.742 calculated considering the need for overall stability in literature 2, which is similar to the independent height ratio 0.77 of the multi-point attached support.

Acknowledgements This research is supported by the Guangxi Science and Technology project of an open competition mechanism to select the best candidates (Grant no: 2022JBGS039).

References

1. GB/T 13752-2017. Code for the design of tower cranes [S]
2. Luo B, Lu NL, Xia YJ (2006) Determination of reasonable attachment distance and maximum overhang height of tower crane. Constr Mach (10):12–15+82
3. Lin B, Liu SG, Cong J et al (2019) Research on the calculation method of attachment distance and cantilever height of tower crane. Constr Mechanization 40(02):18–20
4. Lou KM (2019) Analysis of crane stability against overturning. China Equipment Eng (03):102–103
5. Xu XB (2021) Analysis of tower crane attachment problem in assembly structure. Constr Mechanization 45–46+74
6. Liu JH (2019) Research on analysis method of tower body structure of tower crane with attached flexibility. Harbin Institute of Technology
7. Huang MN, Xie GD, Feng T et al (2017) Investigation and analysis of the collapse of tower crane in Xiamen during Typhoon "Meranti". Constr Mechanization 38(01):23–28+44
8. Guo YM (2014) The harm of typhoon to tower crane and its defensive measures. Build Safety 29(10):51–55
9. Wang H, Liu Y, Xu ZD et al (2023) Wind-induced buffeting comfort assessment of tower cranes considering the wake effect of super-high bridge towers. J Wind Eng Ind Aerodyn 240
10. Wang Y (2023) Analysis and plan of emergency measures of tower crane under typhoon based on Ansys. Lifting Transp Mach 23:62–67
11. Huang JT, Gao CR, Wang GT et al (2023) Optimization analysis of boom structure of flat head tower crane based on Ansys workbench. Lifting Transp Mach 16:37–42

Open Access This chapter is licensed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License (http://creativecommons.org/licenses/by-nc-nd/4.0/), which permits any noncommercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if you modified the licensed material. You do not have permission under this license to share adapted material derived from this chapter or parts of it.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.



Dynamics Coupling Model and Stability Analysis of Special Vehicles Under Operating Conditions



Xiaoyu Sun, Jiandong Li, and Chengqian Su

Abstract In order to analyze the dynamics laws of special vehicle driving conditions, this paper established a six-bridge special vehicle model, which covers vertical acceleration, lateral acceleration, and lateral Angle speed, pitch Angle speed and yaw speed in the vehicle rotating motion. The translational and rotational motion of the vehicle under driving conditions were studied. A vertical model is built to study vertical acceleration in translational motion, pitch Angle speed and lateral Angle speed in rotational motion, and a lateral model is built to study lateral acceleration in translational motion and yaw speed in rotational motion. The longitudinal speed of the longitudinal model is introduced, and the longitudinal speed is dynamically coupled with the vertical model and the lateral model through the tire model. A special vehicle dynamic coupling model was established, and the road excitation and vehicle steering shaft Angle input generated by random white noise method were used to analyze the dynamic coupling model, and the vehicle stability under the driving conditions of the special vehicle was studied.

Keywords Multi-axle special vehicles · Driving conditions · Dynamic coupling · Vehicle simulation design · Tire models · Vehicle stability

1 Introduction

The multi-axle special vehicle in the driving condition of the environment is more complex, in the face of bad road conditions and road multi-bend steering two situations, the accuracy of the special vehicle driving dynamics model put forward higher requirements, in view of the above two situations, the special vehicle driving dynamics model is required to cover the road input and steering input. And the coupling relationship is used in the driving dynamics modeling of the multi-axle

X. Sun (✉) · J. Li · C. Su
Beijing Institute of Machinery and Equipment, Beijing, China
e-mail: sunxiaoyu0819@126.com

© The Author(s) 2026

H. Saman et al. (eds.), *The 9th International Conference on Advances in Construction Machinery and Vehicle Engineering*, Lecture Notes in Mechanical Engineering,
https://doi.org/10.1007/978-981-96-5527-4_45

special vehicle to explore the vehicle stability under the road input and steering input under the driving conditions of the special vehicle.

At the same time, the structure of Special Vehicles is increasingly complex. Along with the coupling of electromechanical and hydraulic systems, the cockpit [1], electromechanical equipment, hydraulic devices [2] and intelligent driving sensing and identification equipment [3] all put forward higher requirements for the accuracy of the model under driving conditions, and put forward higher requirements for the stability of the vehicle under driving conditions in the stability of the stability of the vehicle. To achieve the task of safe driving under complex and changeable road conditions is also the basic demand of multi-axle Special Vehicles such as intelligent off-road vehicles. Therefore, the requirements for special vehicle running stability have been improved, and the vehicle running stability includes vehicle ride comfort and handling stability. The research on ride comfort and handling stability needs to regard the vertical, horizontal and longitudinal of the vehicle as a coupled whole to have better research significance. Therefore, the research on special vehicle dynamic coupling under driving conditions has far-reaching significance.

Vehicle dynamic coupling is mainly classified according to different research objectives. The research on driving and braking direction tends to study the longitudinal model of the vehicle, the research on the steering system tends to pay more attention to the lateral model of the vehicle, and for the design requirements of complex driving roads or highways, it is often concerned about the vertical model. According to different research objectives, the research on the system and its interaction in various directions mainly includes longitudinal coupling, vertical coupling and three-way coupling models, among which the three-way coupling model is the most helpful to the simulation response of the vehicle, and the processing methods of coupling are constantly updated.

Li Chunming from the North Vehicle Research Institute studied the longitudinal and vertical coupling dynamics model of tracked vehicles [4], and quantitatively analyzed the influence of different types of road surface on vehicle driving power, but the steering application scenario was lacking. Li Shaohua from Shijiazhuang Railway University studied the three-way coupling modeling of heavy vehicles [5], and compared it with the FVP model, but the load analysis of vertical and lateral coupling was lacking. HRGETIC took tire lateral force as a random variable for model calculation [6, 7], but this calculation is accidental and has no coupling significance. Li Zihan established LuGre tire to reflect the influence of coupling characteristics of tire slip rate and side deflection Angle on vehicle lateral force [8]. A kind of vehicle lateral and longitudinal coupling cooperative optimization controller is designed, but only the vehicle handling stability is concerned. Li Shaosong from Changchun University of Technology studied the stability of vehicles in extreme working conditions [9], and proposed the application of MPC method in lateral front wheels to improve the stability of vehicles. Lai Fei of Chongqing University of Technology established a three-way coupling model of 18-DOF vehicle, but did not consider the system response of Angle input [10].

In this paper, starting from the vehicle driving dynamics, a three-way coupling vehicle dynamics model is established, which can cover the road input and steering

input under driving conditions, explore the vehicle's driving stability and operational stability, and conduct stability region analysis to obtain the stability region speed range under vehicle driving conditions.

2 Modeling of Vertical Vehicle Model

In this paper, the simulation and modeling method of the whole vehicle system composed of road–tire–suspension–body is used to study the vertical dynamic characteristics of the multi-axle vehicle. From the contact point between the wheel and the road as the excitation of the road to the vehicle simulation model, the tire vertical force is transmitted to the vehicle suspension system, and the suspension system is transmitted to the body, and the whole simulation data of the body is obtained. The structure diagram of the vehicle vertical model is shown in Fig. 1.

The road excitation can be used to describe the relief of the road surface. In the simulation, the road excitation is mainly used as the input, and the vertical displacement caused by the road surface is input into the vehicle model, and the output of the suspension and chassis models in the simulation modeling method in this paper

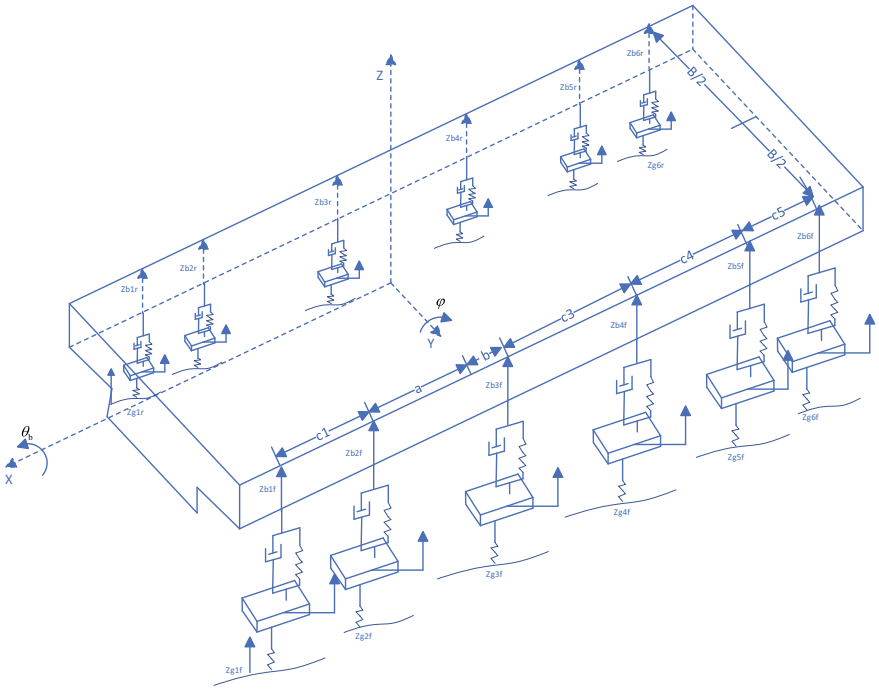


Fig. 1 Modeling of the vertical model of the vehicle

is observed. Therefore, the displacement power spectral density method is selected as the input of the vertical model in this paper.

For the road excitation file generated in this paper, the road excitation input of the left and right wheel of axis i is represented Z_{gif}. Since the time frequency in the power displacement spectral density method established in this paper is influenced by the joint action of the speed and spatial frequency, different longitudinal speed Settings will lead to different vertical excitation input of the road.

In the vertical modeling of Special Vehicles in this paper, the vertical displacement Z_{bif} at the connection point between suspension and body can be calculated by the vehicle side pitch Angle θ_b, roll Angle φ and centroid acceleration Z_b. Since the centroid of the special vehicle model selected in this paper is between the second bridge and the third bridge and on the center line of the left and right wheels, it can be obtained by the formula (1). In the formula, B is the distance between wheels; *li* is the horizontal distance from the i bridge to the body's center of mass.

$$\begin{cases} Z_{bif} = Z_b + l_i \cdot \sin(\theta_b) + \frac{1}{2}B\varphi (i = 1, 2) \\ Z_{bif} = Z_b + l_i \cdot \sin(\theta_b) - \frac{1}{2}B\varphi (i = 3, 4, 5, 6) \\ Z_{bir} = Z_b - l_i \cdot \sin(\theta_b) + \frac{1}{2}B\varphi (i = 1, 2) \\ Z_{bir} = Z_b - l_i \cdot \sin(\theta_b) - \frac{1}{2}B\varphi (i = 3, 4, 5, 6) \end{cases} \tag{1}$$

According to the mechanical characteristics and topological relationship of the vehicle, the vertical model of the vehicle is established, and the mathematical expression is shown in the formula (2).

$$\begin{cases} F_{sij} = C_{sij}(\dot{Z}_{wij} - \dot{Z}_{bij}) + K_{sij}(Z_{wij} - Z_{bij}) \\ m_{wij}\ddot{Z}_{wij} = K_{tij}(Z_{gij} - Z_{wij}) + F_{sij} \\ m_b\ddot{Z}_b = \sum_{i=1}^{6}\sum_{j=l}^{r}(F_{sij}) \\ I_p\ddot{\theta}_b = \sum_{i=1}^{2}\sum_{j=l}^{r}(l_i \cdot F_{sij}) - \sum_{i=3}^{6}\sum_{j=l}^{r}(l_i \cdot F_{sij}) \\ I_f\ddot{\varphi} = \sum_{i=1}^{6}((F_{sif} - F_{sir}) \cdot \frac{B}{2}) \end{cases} \tag{2}$$

In the formula (2), m_b is the body mass; I_p the pitch moment of inertia of the vehicle around the central axis; If is the rolling moment of inertia of the vehicle around the central axis; K_{sij} is the spring stiffness coefficient of the suspension; C_{sij} is the damping factor of suspension; K_{tif} is the stiffness coefficient of the tire; F_{sij} is the suspension output force; m_{wij} is the suspension unsprung mass. The specific vehicle vertical modeling structure diagram is shown in Fig. 1.

In the simulation solution process, the simulation subsystem is constructed according to the input vehicle side pitch Angle θ_b, roll Angle φ and centroid acceleration Z_b, and the vertical displacement Z_{bif} at the connection point between the suspension and the body of the variable is obtained; The road excitation model is used to input Z_{gif} and Z_{bif} to construct the state space expression: determine the state

variable X as shown in the formula (3).

$$X = \left(Z_{wij}\ \dot{Z}_{wij} \right)^T \tag{3}$$

The input Z_{gif} of the road excitation model and the vertical displacement Z_{bif} at the connection point between suspension and body are taken as the control variable U, and the formula can be written as the state space expression of: to solve the problem, the displacement (2), $\dot{X} = AX + BU$. Z_{wif} of the vehicle suspension unsprung mass block is obtained, and the obtained Z_{wif} is combined with the input Z_{gif} of the road excitation model and the vertical displacement Z_{bij} at the connection point between suspension and body is obtained. By solving the linear second-order ordinary differential equation, the vehicle pitch Angle θ_b, roll Angle φ and centroid acceleration Z_b are solved. Finally, the solved vehicle pitch Angle θ_b, roll Angle φ and centroid acceleration Z_b are input into the simulation subsystem to complete the closed-loop model solution.

3 Lateral Model

In this paper, a lateral model is established to analyze the stability of the vehicle under driving conditions including steering conditions, and the dynamic response of the lateral model indexes such as lateral acceleration and yaw velocity under driving conditions is obtained from the lateral model, which supports the subsequent establishment of the dynamic coupling model in this paper. In the lateral model, the rotation Angle of the vehicle drive shaft is taken as the input, and the rotation value of each vehicle is determined by the Ackermann steering theorem. The lateral dynamics of the vehicle are analyzed, and the model simulation results under the driving condition of the lateral model are obtained.

In this paper, the Angle of the first axis is taken as the input, and the four steering axes (the first, second, fourth and fifth axes) in the six axes of the special vehicle are input through the Ackermann steering model to complete the simulation of the lateral model of the special vehicle.

The lateral model structure diagram is shown in Fig. 2. The distance from each axis to the center of mass can be expressed by the distance between the wheels, l_i represents the distance from the four active steering axes to the center of mass as shown in the formula (4).

$$\begin{cases} l_1 = c_1 + a \\ l_2 = a \\ l_4 = c_3 + b \\ l_5 = c_3 + c_4 + b \end{cases} \tag{4}$$

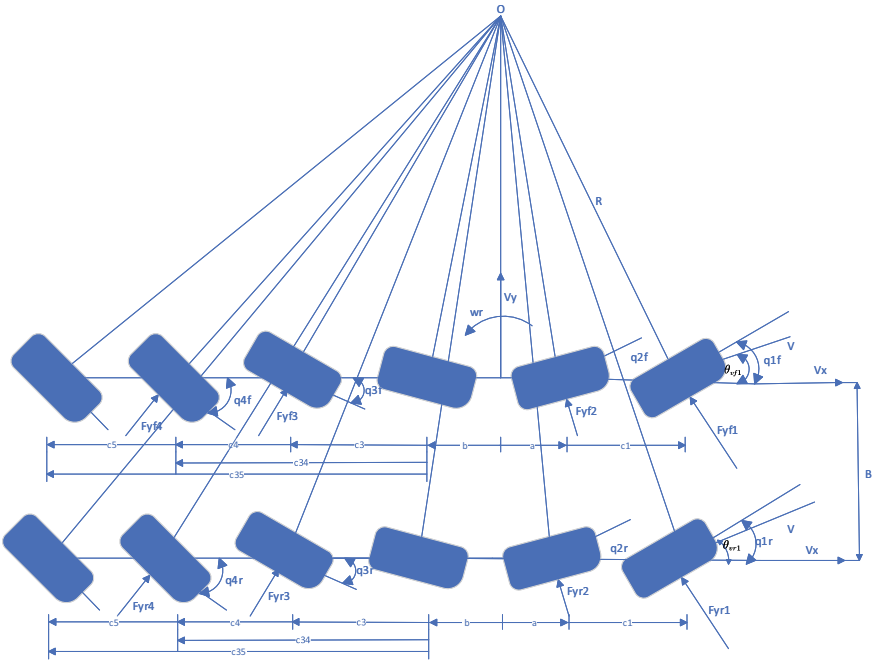


Fig. 2 Modeling of the lateral model of the vehicle in the linear tire model the lateral

According to Ackermann's steering theorem, the Angle of the inner wheel q_{if} is shown in the formula (5).

$$\frac{l_1}{\tan(q_{1f})} = \frac{l_2}{\tan(q_{2f})} = \frac{l_4}{\tan(q_{4f})} = \frac{l_5}{\tan(q_{5f})} \tag{5}$$

Therefore, there are four quantitative relations between the side angles of the front wheels as shown in the formula (6).

$$q_{if} = \arctan\left(\frac{\tan(q_{1f}) \cdot l_i}{l_1}\right) \tag{6}$$

For the outer inner wheel Angle q_{ir} is shown in the formula (7).

$$\frac{l_i}{\tan(q_{ir})} = B + \frac{l_i}{\tan(q_{if})} \tag{7}$$

Therefore, the relationship between the number of wheel side angles is shown in the formula (8).

$$q_{ir} = \arctan\left(\frac{\tan(q_{1f}) \cdot l_i}{B \cdot \tan(q_{1f}) + l_1}\right) \tag{8}$$

Therefore, q_{if} can be used to represent q_{ir}, and then q_{1f} can be used to represent q_{if}, q_{ir}, The vehicle lateral model modeling diagram is shown in Fig. 2.

3.1 Vehicle Lateral Dynamics Analysis

In the lateral force analysis of the vehicle, ay is the lateral acceleration at the mass center of the multi-axis vehicle, F_{yfi} is the lateral force exerted by the ground on the four driving front wheels of the special vehicle, and F_{yri} is the lateral force exerted by the ground on the four driving front wheels of the special vehicle, as shown in the formula (9).

$$ma_y = \sum_{i=1}^{2}(F_{yfi}+F_{yri}) + \sum_{i=4}^{5}(F_{yfi}+F_{yri}) \tag{9}$$

For the decomposition of the lateral acceleration of the vehicle, the formula (10) can be obtained, where wr is the yaw angular speed of the vehicle's center of mass.

$$a_y = \dot{V}_y + V_x \cdot wr \tag{10}$$

Force of the vehicle is described as the product of the lateral deflection Angle and the offset stiffness of the tire. The formulas (11) and (12) are obtained by combining the formulas (9) and (10).

$$\begin{cases} m_b(\dot{V}_y + V_x \cdot wr) = C_{afi}\theta_{if} + C_{ari}\theta_{ir} \\ \theta_{if} = q_{if} - \arctan\frac{V_y + l_i \cdot wr}{V_x} \\ \theta_{ir} = q_{ir} - \arctan\frac{V_y + l_i \cdot wr}{V_x} \end{cases} \tag{11}$$

$$\dot{V}_y = \frac{1}{m_b}(C_{afi}\theta_{if} + C_{ari}\theta_{ir}) - m_b \cdot V_x \cdot wr \tag{12}$$

For the torque balance equation around the vertical Z-axis as shown in the formula (13). I_z represents the moment of inertia of the special vehicle around the Z-axis.

$$\begin{cases} I_z\dot{w}r = \sum_{i=1}^{2}(M_{yfi} + M_{rfi}) - \sum_{i=4}^{5}(M_{yfi} + M_{rfi}) \\ M_{yfi} = F_{yfi} \cdot \cos q_{if} \cdot l_i \\ M_{rfi} = F_{yri} \cdot \cos q_{ir} \cdot l_i \end{cases} \tag{13}$$

Under the condition of corner step input, the simulation data of lateral velocity and yaw velocity of the center of mass of the vehicle are obtained. The simulation unit of lateral velocity is m/s, and the simulation unit of yaw velocity of center of mass is rad/s.

4 Coupling Structure

Perform transient corner step input and vertical road excitation input for Special Vehicles, use Ackermann model to describe the corner input of four active steering axes, and give the "magic formula" vehicle lateral speed and yaw Angle speed curve after the tire coupling the steering system with the vertical model. Data acquisition and analysis of each part of the vehicle model are carried out to determine the stable speed interval of the special vehicle. The coupling structure diagram is shown in Fig. 3.

In this paper, the tire in the side model is free rolling condition, and the tire has no side movement. The Pacejka tire model is an important model for vehicle tire modeling. Based on the Pacejka tire model, the tire model part of the vehicle side model is established in this paper. Bakke and Pacejka proposed the Magic Formula model for the first time, integrating the typical physical quantities of tire dynamics with trigonometric functions. In the side model, the relationship between the side

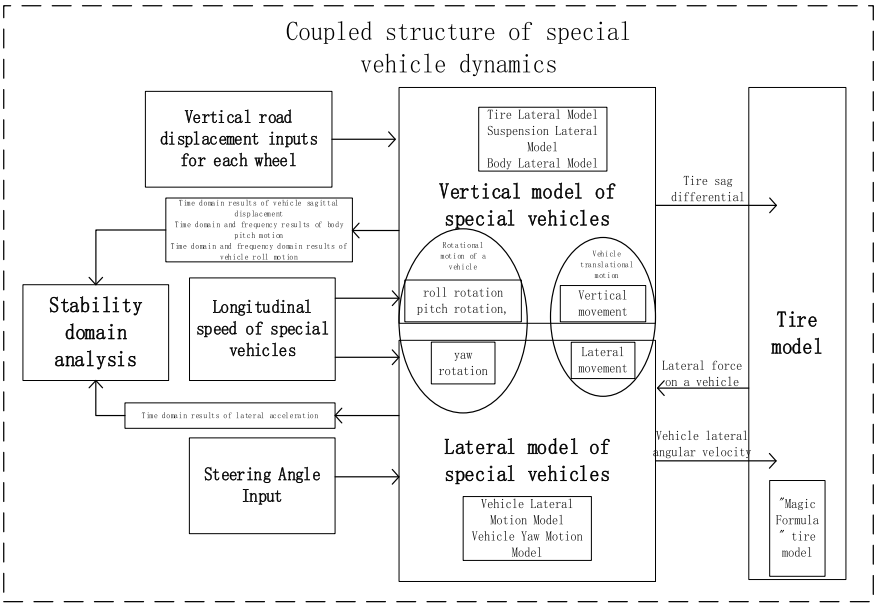


Fig. 3 Structural diagram of the coupling model of special vehicle dynamics

force and the side yaw Angle was established by using the Magic Formula model, and the vertical load relationship in the formula parameters was used. To complete the coupling of lateral model and vertical model of vehicle. "Magic Formula" mathematical expressions such as formula (14).

$$\begin{cases} y(x) = D\sin\{C\arctan(Q(x))\} \\ Q(x) = Bx - E(Bx - \arctan Bx) \\ Y(x) = y(x) + S_v \\ x = X + S_h \end{cases} \tag{14}$$

where Y(x) represents the lateral force; x is the side deflection Angle. D is the curve peak factor, C is the shape factor, B is the stiffness factor, E is the curvature factor, Sh is the horizontal offset, Sv is the vertical offset, is the wheel camber, Q(x) is the intermediate variable. The fitting parameters of the magic formula are shown in Table 1.

Input side deflection Angle of tire model and tire vertical load input in vertical model to get wheel side deflection force, mass distribution and center of mass movement of vehicle are not studied in this paper.

The dynamic load of the vehicle in the process of driving is processed. The dynamic load is determined by the road condition and driving speed. The shape variable of the vehicle tire can be read in the vertical model of the vehicle for different roads and different driving speeds (Fig. 4).

Table 1 "Magic Formula" tire fitting parameters

a1	a2	a3	a4	a5	a6	a7	a8	a9	a10	a11	a12
−22.1	1011	1078	1.82	0.208	0	−0.354	0.707	0.028	0	14.8	0.022

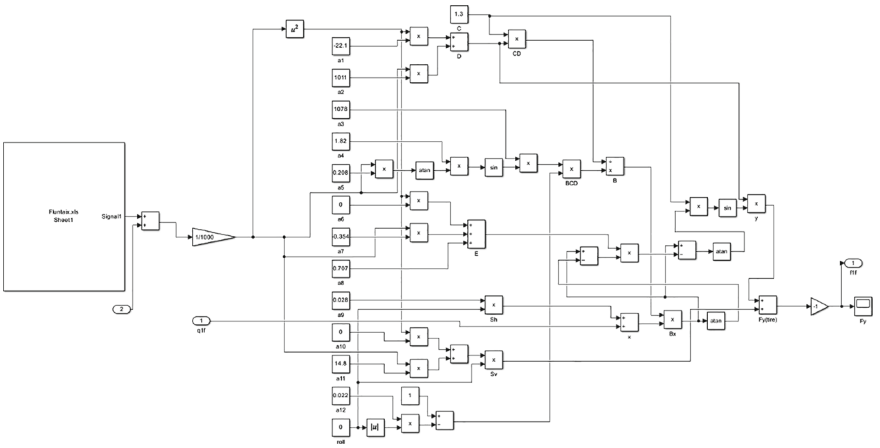


Fig. 4 MATLAB/Simulink simulation structure diagram of "Magic Tire" formula

In MATLAB/Simulink software, the tire load data of the entire vertical model is exported from Simulink in the form of excel, with the suffix.xlsx. Call From Spreadsheet to import the data list module, as input, import into the magic tire formula in the side model as the tire dynamic load input.

5 Simulation Calculation

In the simulation calculation, the tire dynamic load output without vertical model and tire dynamic load output with vertical model are divided into uncoupled model and coupled model. Both of them use transient Angle as input to test vehicle stability under driving conditions.

When the simulation time reaches 1s, the transient Angle input is adopted with the step input of 5 degrees and 30 degrees respectively. The lateral acceleration and yaw Angle velocity curves of the vehicle are obtained under the longitudinal speed of the vehicle at 10*km/h*, 20*km/h*, 40*km/h*, 60*km/h*, 80*km/h* and 120*km/h*. The simulation unit of the vehicle lateral acceleration is m/s^2, and the simulation unit of the yaw Angle speed of the center of mass is rad/s. In order to directly reflect the coupling effect of the coupled system, the vehicle dynamic response of the coupled model and the uncoupled model is compared in this paper.

According to the national standard GBT13047-91 for the steering transient response test of the national standard 5.3 requirements, specifies the stability of the steering transient response integral rules, but because the total mass of this paper is greater than 6*t* Special Vehicles, this paper adopts the lateral acceleration is not greater than 0.4*g* method to evaluate the stability of the lateral model.

5.1 Simulation Results of Uncoupled Model

The uncoupled model does not input the tire dynamic load from the vertical source, and the vertical tire load is only composed of the vehicle tire static load, and the steering input is consistent with the longitudinal speed conditions. The lateral dynamic response of the vehicle is shown in Figs. 5, 6, 7 and 8.

It can be seen from the uncoupled model diagram that when the transient Angle is 5°, it can be considered that except for the speed of 60, 80 and 120 km/h, the other three Special Vehicles in the simulation meet the stability requirements, and the speed of 60 km/h is slightly greater than 0.4 g. When the transient Angle is 30°, only the Special Vehicles travel at 10, 20 and 40 km/h meet the stability requirements in the simulation. The maximum lateral acceleration of the vehicle at 60 km/h exceeds 0.5 g, and the maximum lateral acceleration of the vehicle at 80 km/h exceeds 0.6 g. At 120 km/h, the lateral acceleration of the vehicle is greater than 0.9 g and close to 1.0 g.

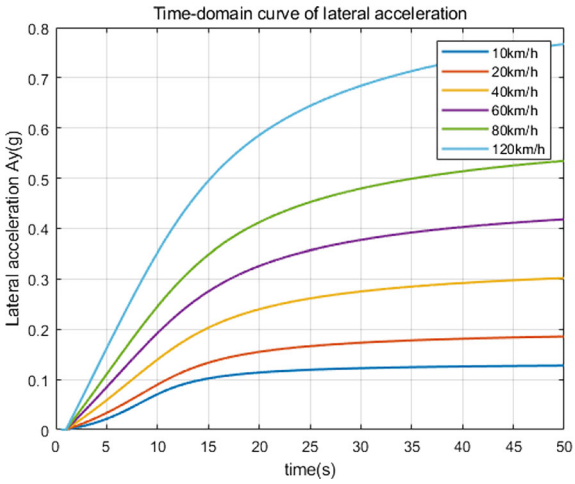


Fig. 5 Time-domain curve of lateral acceleration at an uncoupled rotation angle of 5 degrees

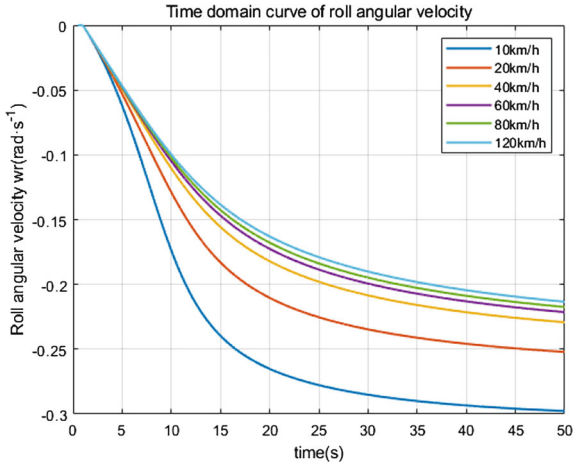


Fig. 6 Time-domain curve of centroid yaw velocity at an uncoupled rotation angle of 5 degrees

5.2 Simulation Results of Coupled Model

After receiving the type D road excitation input, the coupling model extracted the vertical load data generated under the road excitation at different longitudinal speeds and input the time domain data of tire load under the driving conditions of the vehicle longitudinal speed of 10, 20, 40, 60, 80 and 120 km/h into the coupling system. The simulation calculation is completed. The lateral dynamic response of the vehicle is shown in Figs. 9, 10, 11 and 12.

It can be seen from the coupling diagram that when the transient Angle is 5°, the vehicle speed of 120 km/h can be in an unstable state, and 80 km/h is slightly greater than 0.4 g. The other four Special Vehicles running at different speeds in the

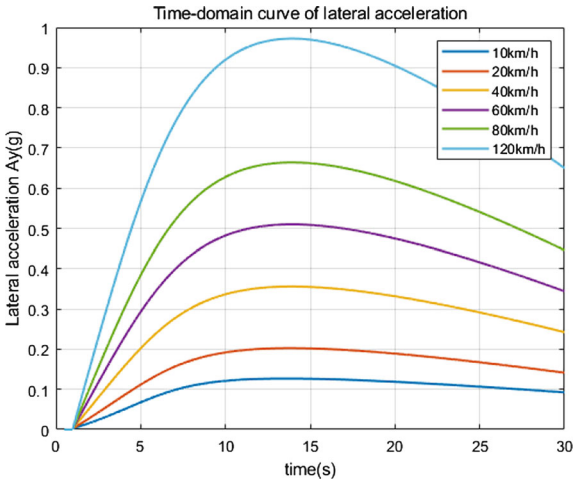


Fig. 7 Time-domain curve of lateral acceleration at an uncoupled rotation angle of 30 degrees

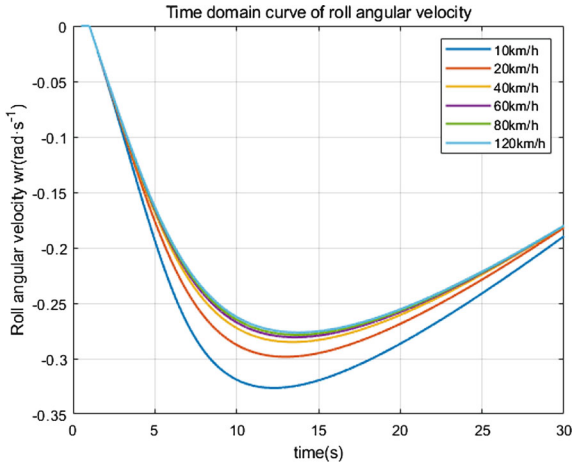


Fig. 8 Time-domain curve of centroid yaw velocity at an uncoupled rotation angle of 30 degrees

simulation all meet the stability requirements. When the transient Angle is 30°, only the Special Vehicles travel at 10, 20 and 40 km/h meet the stability requirements in the simulation. The maximum lateral acceleration of the vehicle at 60 km/h exceeds 0.4 g, and the maximum lateral acceleration of the vehicle at 80 km/h is close to 0.6 g. At 120 km/h, the maximum lateral acceleration of the vehicle is close to 0.9 g.

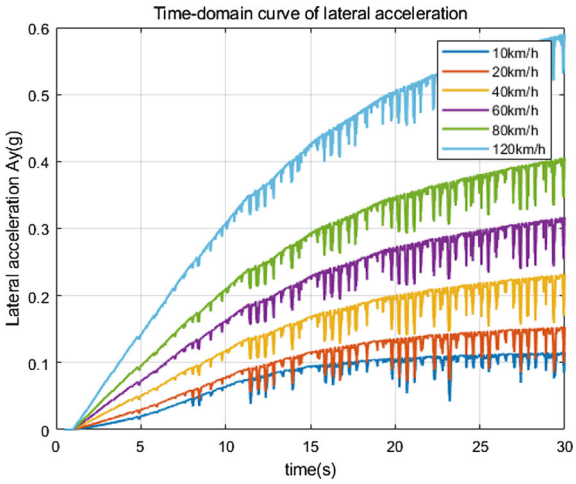


Fig. 9 Time-domain curve of lateral acceleration at a coupling angle of 5°

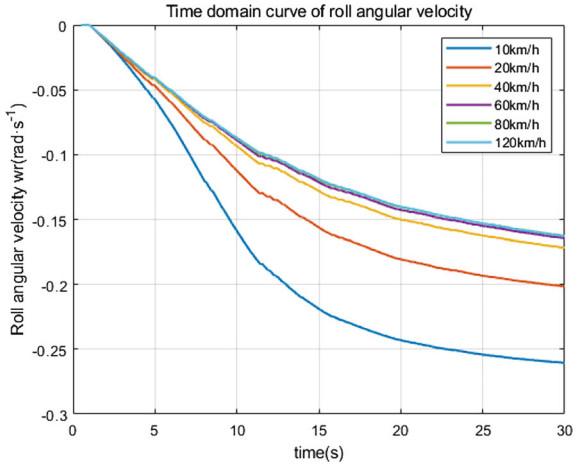


Fig. 10 Time-domain curve of centroid yaw velocity at a coupling angle of 5°

5.3 Comparison of Simulation Results Between Uncoupled Model and Coupled Model

In the uncoupled simulation in this paper, it can be considered that the stability region of the special vehicle driving condition is below 60 km/h when the transient Angle is 5°, and below 40 km/h when the transient Angle is 30°. In the coupling simulation in this paper, the stability region of the special vehicle driving condition of the coupling model is below 80 km/h when the transient Angle is 5°, and below 40 km/h when the transient Angle is 30°. Considering the vertical vehicle model, the stability region boundary of the special vehicle under the driving condition is different, so it can be

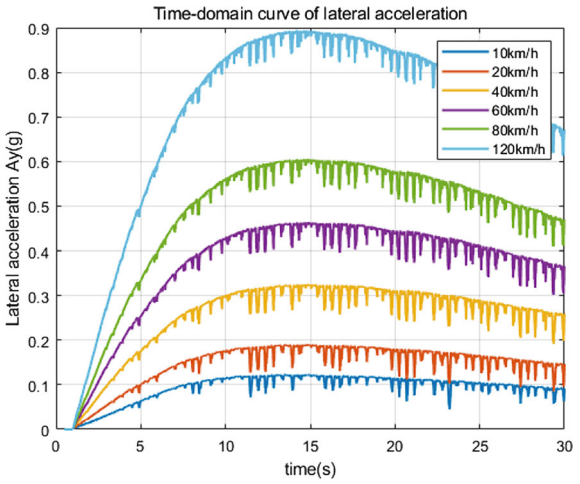


Fig. 11 Time-domain curve of lateral acceleration at a coupling angle of 30°

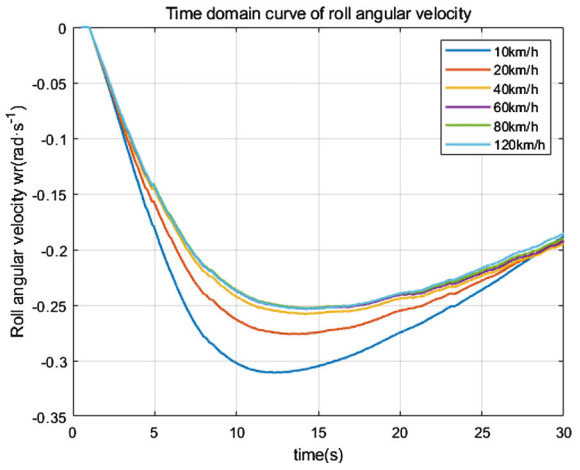


Fig. 12 Time-domain curve of centroid yaw velocity at a coupling angle of 30°

proved that the external input expanded by the coupling model has a better analysis and research effect on the vehicle stability.

6 Concluding Remarks

The special vehicle coupling dynamics model established in this paper can reflect the influence of road input and steering input on vehicle stability in the driving condition. The model covers longitudinal speed, steering system, suspension and body displacement changes, and the three are coupled to obtain the lateral acceleration and yaw velocity of the vehicle under driving conditions, and then the stability conditions of

the special vehicle under driving conditions are found and quantitative analysis is obtained. The application scenario of vehicle stability judgment is expanded through the coupling model. It can be combined with longitudinal speed, vertical input of road surface and vehicle steering input to judge the stability of Special Vehicles under driving conditions and has a guiding role in the design of indicators under vehicle driving conditions such as center of mass distribution, suspension parameters, body weight, tire vertical and lateral stiffness selection in the design process of Special Vehicles.

Multi-axle Special Vehicles in the driving conditions of the environment is more complex, in the face of adverse road conditions and road multi-turn steering two situations, the accuracy of the special vehicle driving dynamics model put forward higher requirements, in view of the above two situations, the special vehicle driving dynamics model is required to cover the road input and steering input. And the coupling relationship is used in the driving dynamics modeling of the multi-axle special vehicle to explore the vehicle stability under the road input and steering input under the driving conditions of the special vehicle.

References

1. Liu Q, Liu X, Gu J, Liu Y (2024) Simulation analysis of ride comfort of multi-axle vehicle based on a rigid-flexible coupling model. Missiles Space Veh 47(2):69–73
2. Duan P, Yu L, Li Y (2024) Optimization design of steering system for special semi-trailer. Hydraul Pneumatics Seals 44(7):104–106
3. Li B, Han Z, Chen J, Wang H (2024) Robust control of lateral stability for intelligent driving electric vehicles. Transp Energy Conservation Environ Protection 20(01):47–52
4. Li C, Wu W, Guo Z, Yuan S, Chen S (2021) Longitudinal and vertical coupled dynamic model and power characteristics of tracked vehicle. Acta Armamentarii 42(03):449–458
5. Li S, Yang S, Chen L (2014) Modeling and dynamic analysis of a non-linear heavy vehicle with three-directional coupled motions. J Vib Shock 33(22):131–138
6. Ray LR (1997) Nonlinear tire force estimation and road friction identification: simulation and experiments. Automatica 33(10):1819–1833
7. Hrgetic M, Deur J, Ivanovic V et al (2014) Vehicle sideslip angle EKF estimator based on nonlinear vehicle dynamics model and stochastic tire forces modeling. SAE Int J Passenger Cars Mech Syst 7(1):86–95
8. Li Z, Chen H, Wang P (2019) Coordinated longitudinal and lateral stability vehicle control based on combined-slip tire model in the model predictive control framework. J Tongji Univ (Natural Science) 47(S01):1–8
9. Li S, Guo K, Qiu T, Chen H, Wang G, Cui G (2020) Stability control of vehicle with active front steering under extreme conditions. Autom Eng 42(02):191–198
10. Lai F (2020) Modelling and simulation analysis of the unified model of vehicle longitudinal-lateral and vertical dynamics. J Mach Des 37(05):37–43

Open Access This chapter is licensed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License (http://creativecommons.org/licenses/by-nc-nd/4.0/), which permits any noncommercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if you modified the licensed material. You do not have permission under this license to share adapted material derived from this chapter or parts of it.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.



Dynamic Analysis of Crawler Traveling Mechanism Steering Performance Analysis



Yongran Sui, Xin Wang, Zijun Li, and Junli Pi

Abstract A certain crawler crane was taken as the research object, and the multi-body dynamics model of the traveling mechanism is established based on RecurDyn software, and the simulation analysis is carried out for three typical working cases, namely, straight-line, hill-climbing, and steering-in-Place, so as to obtain the driving torque of the driving wheels. The simulation values are compared with the theoretical calculations. It is verified that the simulation model is correct. On the basis of this, the simulation of steering conditions is carried out to analyze the influence of different track gauges, steering radii and rotational speeds on the steering slip force. The results show that the steering slip force increases with the increasing of the track gauge within a certain range, and the steering slip force decreases with the increasing of the steering radius, and the turning speed.

Keywords Traveling mechanism · Steering performance · Slippage · Multi-body dynamics

1 Introduction

Crawler cranes are widely used in electric power construction, bridge construction, water conservancy and other industrial fields due to their advantages of small grounding specific pressure, strong climbing ability, and the ability to travel with loads [1]. As an important part of crawler crane, traveling mechanism play the role

Y. Sui · X. Wang (✉)
Dalian University of Technology, Dalian, China
e-mail: wangxbd21@163.com

Z. Li
Zoomlion Heavy Industry Science and Technology Co, Ltd, Changsha, China

J. Pi
Beijing Materials Handling Research Institute Co, Ltd, Beijing, China

© The Author(s) 2026

H. Saman et al. (eds.), *The 9th International Conference on Advances in Construction Machinery and Vehicle Engineering*, Lecture Notes in Mechanical Engineering,
https://doi.org/10.1007/978-981-96-5527-4_46

of bearing and providing traction. However, its structure is complex, and the resistance to contact with the ground when traveling is difficult to determine accurately. In particular, the resistance when steering is large and exists slip force, which is difficult to calculate.

Ma [2] analyzed the effects of track pre-tensioning force, steering radius, and road type on track steering stability of tracked vehicles during steering. Mu [3] investigated the change rule of steering center offset variation in the steering transient of tracked vehicles and analyzed its effect on the steering stability; Meng [4] studies the calculation method of steering angular velocity and radius of the machine under the consideration of slip conditions; Zhang [5] analyzed the influence rule of physical properties of wetland soil on the steering performance of crawler bulldozer.

The traveling mechanism of a certain type of crawler crane is taken as the research object to analyze its structural principle, establish its dynamic model and verify its correctness, and analyze the influence of different factors on its steering performance for the steering working conditions.

2 Modeling of Traveling Mechanism

The traveling mechanism mainly consists of driving wheel, driven wheel, supporting wheels, carrier sprockets, crawler plates, grinding irons and crawler frame, and its structural form is shown in Fig. 1.

The traveling mechanism is connected to the frame of the vehicle through the crawler frame. The driving wheel and driven wheel are connected to both ends of the crawler frame through pins. The driving wheel is driven by hydraulic motor to rotate and engage with the crawler plate to turn the rotation into linear motion to realize the movement of the whole machine. Several supporting wheels are arranged between the driving and driven wheels, the crawler frame and the crawler plates, which roll with the driving wheels during operation to support the crane and the load weight [6].

The model parameters of the traveling mechanism of 180-ton crawler crane established in this paper are shown in Table 1.

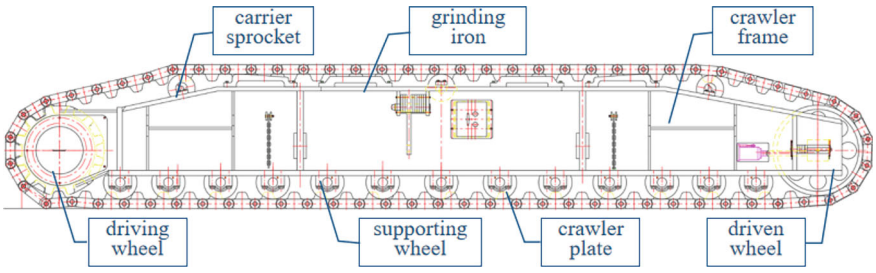


Fig. 1 Sketch of travel mechanism

Table 1 Traveling mechanism parameters

Parameter name	Value
Weight (t)	111
Track gauge × grounding length (mm)	6000 × 7250
Velocity (km/h)	0.5/1.0
Driving wheel diameter (mm)	850
Crawler plate pitch (mm)	253.5
Crawler plate width (mm)	1000

Table 2 Constraints setting

Component name	Constraints setting	Note
Crawler frame	Fixed Joint	With chassis
Driving wheel	Revolute Joint	With crawler frame
Driven wheel	Revolute Joint	With crawler frame
Supporting wheel	Revolute Joint	With crawler frame
Carrier sprocket	Revolute Joint	With crawler frame
Crawler plate	Bushing	With crawler plate

The constraint relationships between the components are determined based on the actual operation of the traveling mechanism, as shown in Table 2.

The setting of the contact relationship between the components of the traveling mechanism is shown in Table 3.

Crawler cranes are often used to operate on firm, smooth roads. So non-sunken, rigid surfaces are created to simulate the surfaces on which crawler cranes often work [7].

The dynamic model of the crawler crane carriage is shown in Fig. 2.

Table 3 Contact setting

Contact components	Contact relationship
Crawler plate and driving wheel	Solid Contact
Crawler plate and driven wheel	Geo Contact
Crawler plate and supporting wheel	Geo Contact
Crawler plate and carrier sprocket	Geo Contact
Crawler plate and ground	Solid Contact
Crawler plate and grinding irons	Geo Contact

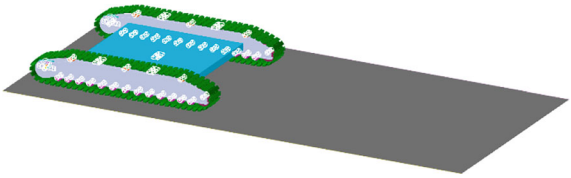


Fig. 2 Crawler crane carriage dynamic model

3 Model Verification

In order to verify the correctness of the dynamic model, three typical working conditions are selected to derive the driving torque and compared with the theoretical calculation results to verify the correctness of the model.

3.1 Verification of Straight-Line Modeling

The total resistance of the track in the flat road straight traveling condition is calculated as follows:

$$F_1 = F_f + F_g + F_i \tag{1}$$

Knowing the radius of driving wheel pitch circle R_q, the driving wheel torque T_1 is:

$$T_1 = \frac{F_1 R_q}{2} = \frac{114.7 \times 0.425}{2} = 24.37\,\text{kN m} \tag{2}$$

The simulation was carried out under straight traveling conditions and the driving torque curve was obtained as shown in Fig. 3.

The simulated value is 22.93 kN m, and the theoretical value is 24.37 kN m. The error between the above two is only 5.9%, which indicates the dynamics model is correct.

3.2 Verification of Hill-Climbing Condition Model

The total resistance of the track in the hill-climbing condition is calculated as follows:

$$F_2 = F_f + F_g \cos\alpha + F_s \tag{3}$$

The unilateral driving wheel torque T_2 is:

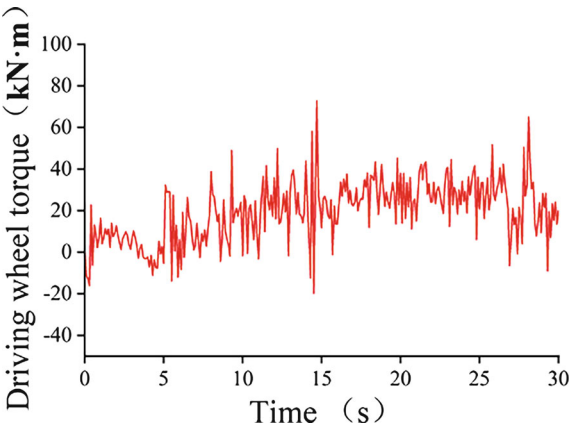


Fig. 3 Straight traveling condition driving torque

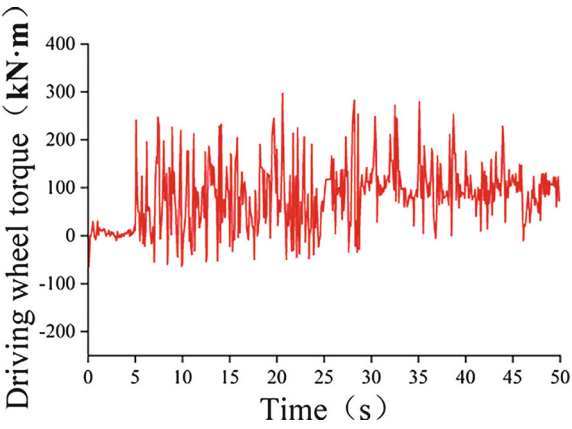


Fig. 4 Hill-climbing condition driving torque

$$T_2 = \frac{F_2 R_q}{2} = \frac{476.24 \times 0.425}{2} = 101.2\,\text{kN m} \tag{4}$$

The simulation is carried out under the hill-climbing condition, and the driving torque curve obtained is shown in Fig. 4. The simulated value is 105.56 kN m, the theoretical value is 101.2 kN m. The error between the above two is only 4.3%, which indicates the dynamics model is correct.

3.3 Verification of Steering-in-Place Condition Model

The maximum torque of the drive wheels in the steering-in-Place condition is:

$$T_{\max} = \max\{F_{q1}, F_{q2}\} R_q = 437.38 \times 0.425 = 185.89\,\text{kN m} \tag{5}$$

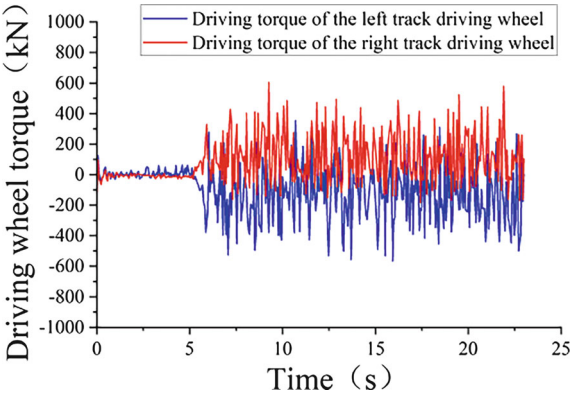


Fig. 5 Steering-in-place condition driving torque

The simulation is carried out under the steering-in-Place condition, and the simulation results of the driving torque variation with time are obtained as shown in Fig. 5.

The simulated value is 171.23 kN m, the theoretical value is 185.89 kN m. The error between the above two is only 7.9%, which indicates the dynamics model is correct.

4 Steering Performance Dynamics Analysis

The horizontal interaction between the track and the ground is a very complex physical phenomenon when the traveling mechanism is steered, and many factors can have an effect on this horizontal interaction, such as the overall structural form of the track, the contact state between the track and the ground, and so on [8, 9]. This paper focuses on the influence of different track gauges, different steering radii and different rotational speeds on the steering slip force during steering traveling.

Dynamic simulation analysis is carried out for the traveling mechanism with different track gauges, different steering radii and different rotational speed during steering process. And the extreme values of the amplitude of the steering slip force between the crawler plate and the ground are extracted as shown in Fig. 6.

In Fig. 6a, it can be seen that the larger the track gauge, the larger the steering slip force, and also the larger the difference of traveling speeds between the two sides. A 10% increase in track gauge increases the slip force by about 5%, which is a linear change. In Fig. 6b, it can be seen that the steering slip force decreases with the increase of the steering radius. This is due to the fact that the smaller the steering radius, the sharper the torsional movement, and the more serious the damage of the crawler to the ground, and thus the greater the lateral resistance of the crawler, and the greater the steering slip force. A 10% increase in steering radius increases the slip force by about 1%. The change of the steering radius does not have as big an

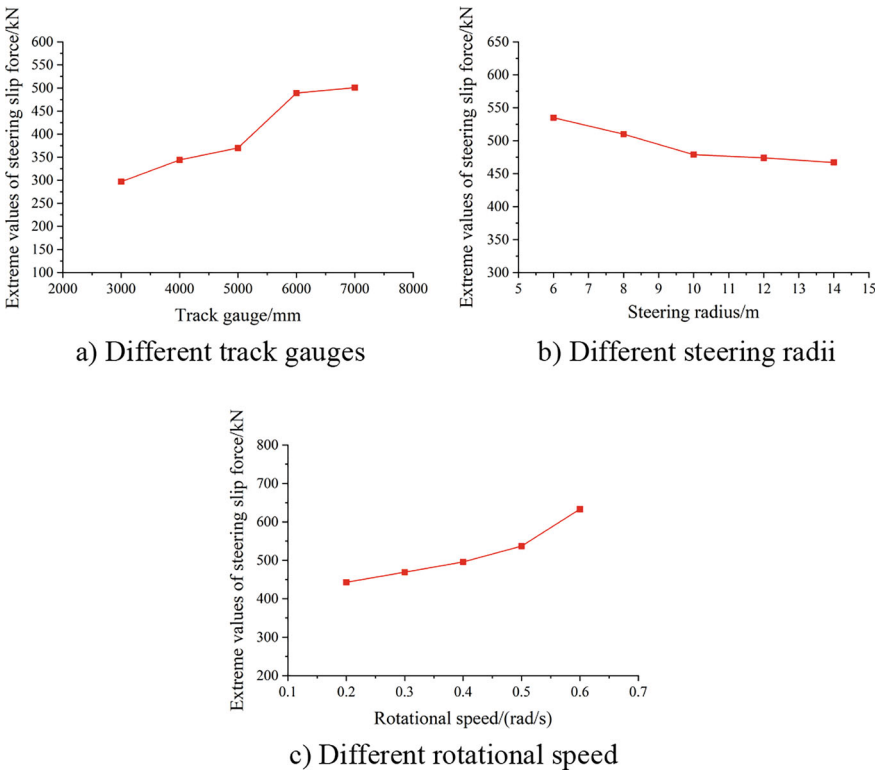


Fig. 6 Plot of extreme values of steering slip force

effect on the steering slip force as the track gauge. In Fig. 6c, it can be seen that the steering slip force increases as the rotational speed increases. The higher the rotational speed, the sharper the torsional movement of the traveling mechanism, the higher the resistance from the ground, and the higher the steering slip force. A 10% increase in rotational speed increases the slip force by about 2%, which is a linear change.

5 Conclusion

From the calculation results, in a certain range, the larger the track gauge of the traveling mechanism, the larger the steering slip force between the crawler plate and the ground. Accordingly, the size of the track gauge can be reasonably selected in the design and development.

In a certain range, the steering slip force of the traveling mechanism decreases with the increase of steering radius and increases with the increase of rotational speed. The conclusion can provide some guidance for the actual steering operation.

Comparing the three factors on the steering slip force, the track gauge has a little more influence on the slip force, followed by the rotational speed, and the steering radius is the smallest.

In the future, the effect of different ground surfaces, different track plates and other parameters on the steering slip force can also be investigated.

Acknowledgements This work is supported by National Natural Science Foundation of China under Grant No.52275088 and the Fundamental Research Funds for the Central Universities under Grant (No. DUT22LAB506, No.DUT22LAB507).

References

1. Zhou J (2009) Research on steering resistance moment and steering radius of crawler crane. Dalian University of Technology
2. Ma X, Chen Y, Liu X et al (2012) Steering performance analysis of tracked vehicle based on multi-body dynamics simulation. Mech Des 29(06):52–57
3. Mu L, Wu K, Pan G (2024) Transient analysis and stability study of slope steering of tracked vehicle. Contemp Agric Mach (04):43–44+47
4. Meng X (2013) Research on dynamic characteristics of telescopic boom crawler crane traveling with load. Dalian University of Technology
5. Zhang Y, Zhao L, Dong K et al (2023) Steering performance analysis of flat-frame wetland-type tracked bulldozer based on RecurDyn. J Beihua Univ (Natural Science Edition) 24(02):269–274
6. Bing L, Xidong H, Jianguo L et al (2021) Finite element analysis of the chassis in telescopic crawler crane, 2021-01-01
7. Wang Y, Jin C, Yang T et al (2024) Analysis and experimental investigation of steering kinematics of driven steering crawler harvester chassis. Agriculture-Basel 14(1)
8. Kormanek M, Dvorak J (2021) Ground pressure changes caused by MHT 8002HV crawler harvester chassis. Croatian J Forest Eng 42(2):201–211
9. Pei T, Yan B, Dong H et al (2024) Research on the influence of heavy-duty tracked vehicles parameters on steering driving process. Int J Heavy Veh Syst 31(2)

Open Access This chapter is licensed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License (http://creativecommons.org/licenses/by-nc-nd/4.0/), which permits any noncommercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if you modified the licensed material. You do not have permission under this license to share adapted material derived from this chapter or parts of it.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.



Optimization of Formation Topology Based on the Leader–Follower Method



Linhan Lin, Ning Chen, Mingxing Fan, Shiyao Cheng, Hanteng Chen, and Yuhan Hou

Abstract This study focuses on optimizing the formation topology in multi-robot systems using the leader–follower method. The primary objective is to enhance system observability and navigation accuracy. First, the chapter introduces the importance of observability in leader–follower systems, particularly analyzing how particle filter algorithms can control the positioning error of the leader robot. The Fisher Information Matrix (FIM) and Cramér-Rao Lower Bound (CRLB) are employed to evaluate the formation topology for single and multiple follower robots. By constructing an evaluation function based on these metrics, the optimal formation topology is identified to minimize localization errors and improve overall system performance. Simulation experiments are conducted to verify the effectiveness of the proposed optimization strategies, demonstrating significant improvements in obstacle avoidance and navigation precision. This chapter provides a solid foundation for improving the efficiency and reliability of multi-robot formation control in complex environments.

Keywords Leader–follower · Multi-agent · Formation

1 Introduction

Inspired by biology and human society, multi-robot systems have successfully mimicked the cooperative characteristics of collective behavior in nature and have been applied in various fields. In manufacturing, multi-robot collaborative operations enable flexible configurations and efficient production on production lines. Intelligent transportation systems utilize collaboration between robots to facilitate smooth traffic flow and reduce congestion. Multi-agent systems can collect data more extensively and accurately in environmental monitoring and resource exploration. On the

L. Lin · N. Chen (✉) · M. Fan · S. Cheng · H. Chen · Y. Hou
School of Marine Equipment and Mechanical Engineering, Jimei University, Xiamen, Fujian, China
e-mail: cn1972@jmu.edu.cn

© The Author(s) 2026

H. Saman et al. (eds.), *The 9th International Conference on Advances in Construction Machinery and Vehicle Engineering*, Lecture Notes in Mechanical Engineering,
https://doi.org/10.1007/978-981-96-5527-4_47

Informative battlefield, multi-agent systems can work together to complete complex combat tasks, significantly improving operational effectiveness.

2 Related Work

Based on different formation structures and working characteristics, researchers have developed and designed various typical multi-robot formation control methods. These include common methods such as leader–follower control [1], behavior-based control [2], virtual structure control [3], and consensus-based control methods.

The leader–follower strategy was initially introduced and applied to multi-robot formation control. Under this framework, Desai et al. used a relative distance and angle control method and achieved exponential stability in the system through feedback linearization [4]. For applications in aircraft formations, Koo et al. employed centralized planning and analyzed scenarios of followers taking off one by one versus simultaneously [5]. In terms of formation stability, Tanner et al. conducted theoretical research, analyzing error propagation through LFS (leader–follower systems) gains [6]. Liu et al. modeled multi-robot formations and implemented motion control based on the relative positions of leaders and followers [7]. Mastellone et al. applied the Lyapunov feedback control method to multi-robot formations and verified it practically [8]. From a dynamic perspective, Tran V. P. et al. explored the leader–follower mode in multi-robot formations and used backstepping to generate dynamic formation error equations [9]. Pan Z. et al. proposed a multi-mobile platform formation control method combining leader–follower strategy with the artificial potential field method to resolve issues in traditional strategies where formation control and obstacle avoidance were not unified [10].

3 Optimization of Formation Topology Based on the Leader–Follower Method

In multi-robot formation control, collaborative localization fundamentally relies on relative observations between the leader and follower robots. The topology of the formation, established between these robots, plays a crucial role in determining the system's overall navigation accuracy. This raises a key question: how can the formation topology be optimized to enhance navigation performance?

From the perspective of positioning accuracy, in a leader–follower cooperative localization system with a single leader robot, maintaining the same formation topology over time could result in the system approaching an unobservable state. To address this, an effective strategy is to adjust the trajectories of the leader and follower robots, ensuring that they maintain relative displacement and, thus, enhance system observability.

This chapter aims to optimize the formation topology of multi-robot systems with a smaller number of robots, focusing on improving the system's observability and overall localization precision.

3.1 Study on Positioning Algorithms for Wheeled Leader Robots

Leader Robot Positioning Based on Particle Filters. In leader–follower formation control systems, the leader robot plays a pivotal role. First, it is crucial to address the positioning accuracy of the leader robot, as it directly affects the overall system's performance. When the leader robot system is observable, the particle filter localization algorithm can be employed to control the positioning error within a bounded range. This ensures that the trajectory of the leader robot closely follows the preplanned path. The algorithm operates within a Bayesian framework, approximating the robot's current state using a set of weighted particles. This approach improves both the accuracy and stability of the leader robot's control and navigation system.

Observability Analysis of the Leader Robot System. To enhance the performance of multi-robot formation control, it is crucial to ensure the local weak observability of the system. This allows the system to transmit more observation information, thereby improving the observer's accuracy in estimating the system's state. Conversely, poor observability will directly impact the effectiveness of formation control.

According to the rank criterion for nonlinear systems, the observability of the system can be determined by calculating the rank of the observability matrix. Specifically, we need to calculate the linear independence of the rows of the observability matrix by computing the Lyapunov derivatives.

$$L_0 h_k s, z = \gamma_{ij1} = \arctan\left(\frac{y_{j1} - y_i}{x_{j1} - x_i}\right) - \theta_{vi} \tag{1}$$

In the formula, s represents the state variable, and z represents the observation variable.

By calculating the Lyapunov derivative, we can conclude that the system has local weak observability. The row vector *matrix* M_{ij1}, composed of the gradient of the zero-order Lyapunov derivative and the gradient of the first-order Lyapunov derivative of γ_{ij1}, is expressed as:

$$M_{ij1} = \begin{bmatrix} -\Delta y_{ij1} & \Delta x_{ij1} & -R_{ij1}^2 \\ \sin\theta_i & -\cos\theta_i & \Delta x_{ij1}\cos\theta_i + \Delta y_{ij1}\sin\theta_i \\ -2\Delta x_{ij1} & -2\Delta y_{ij1} & 1 \end{bmatrix} \tag{2}$$

The rank of the matrix M *i*n Eq. (2) is 2 when the following two conditions are met:

(1) The velocity of all leader robots satisfies $v_i > 0$, where $i = 1, \ldots, n$,
(2) $\gamma_{ij1} \neq 0$, ensuring that the leader robot does not move in a straight line along the line connecting the leader robot and the observed landmark.

When conditions 1 and 2 are satisfied, the matrix M_{ij1} in Eq. (3) can be simplified through elementary row transformations to:

$$M_{ij1} \Rightarrow \begin{bmatrix} 1 & 0 & \Delta y_{ij1} \\ 0 & 1 & -\Delta x_{ij1} \\ 0 & 0 & 0 \end{bmatrix} \tag{3}$$

From Eq. (3), it can be seen that the rank of the matrix M_{ij1} is given by rank $M_{ij1} = 1$. The observability analysis of the navigation robot regarding two landmarks is similar to the observability analysis when the navigation robot observes a single landmark; it is necessary to calculate the rank of the observability matrix.

From the zero-order Lyapunov derivative gradient of $\begin{bmatrix} \gamma_{i_1} \\ \gamma_{i_2} \end{bmatrix}$ to the first-order Lyapunov derivative gradient, we can form the row vector matrix M_{ij12} expressed as:

$$\begin{bmatrix} -\Delta y_{ij1} & x_{ij1} & -R^2_{i2j1} \\ -\Delta y_{ij2} & x_{ij2} & -R^2_{i2j2} \\ \sin\theta_i & -\cos\theta_i & \Delta x_{ij1}\cos\theta_i - \Delta y_{ij1}\sin\theta_i \\ \sin\theta_i & -\cos\theta_i & \Delta x_{ij2}\cos\theta_i - \Delta y_{ij2}\sin\theta_i \\ -2\Delta x_{ij1} & -2\Delta y_{ij1} & 0 \\ -2\Delta x_{ij2} & -2\Delta y_{ij2} & 0 \end{bmatrix} \tag{4}$$

The rank of the matrix M_{ij12} in Eq. (4) is 3 when the following conditions are met:

(1) $v_i > 0$ for i = 1,2,…, n.
(2) $\gamma_{ij1} \neq 0$, $\gamma_{ij2} \neq 0$, meaning the navigation robot does not move in a straight line along the line connecting itself and the observed landmark.
(3) $\gamma_{ij1} \neq \gamma_{ij2}$, indicating that the navigation robot and the two observed landmarks are not collinear.

When n robots can directly observe two different landmarks, the rank of the system becomes $3n$, indicating that this multi-robot system is full rank and possesses local weak observability. Additionally, when the robots can directly observe more than two different landmarks, the multi-robot system can acquire richer observational information, resulting in improved observability.

3.2 Optimization of Formation Topology for a Single Follower Robot

Fisher Information Matrix and Cramér-Rao Lower Bound. To optimize the formation topology of a single follower robot, we employ the Fisher Information Matrix (*FIM*) and the Cramér-Rao Lower Bound (CRLB). These metrics quantify the amount of information the system has about the relative positions of the robots. By constructing an evaluation function based on these metrics, we can assess how changes in the topology influence the system's localization accuracy.

The Cramér-Rao lower bound does not focus on specific parameter estimation methods; rather, it indicates the optimal parameter estimation that can be achieved based on existing information. Let the observation vector be X, and its probability density function with respect to the parameter θ be $p(x|\theta)$. Under certain regularity conditions, any unbiased estimator $\hat{\theta}$ of θ must satisfy the Cramér-Rao inequality:

$$Cov\left[\widehat{\theta}\right] = E\left[\left(\theta - \hat{\theta}\right)\left(\theta - \hat{\theta}\right)'\right] \geq CRLB = FIM^{-1} \tag{5}$$

where *FIM* is the Fisher Information Matrix, which is defined as:

$$FIM = E\left\{\left[\frac{\partial lnp(x|\theta)}{\partial \theta}\right]\left[\frac{\partial lnp(x|\theta)}{\partial \theta}\right]^T\right\} \tag{6}$$

This inequality indicates that the covariance of the estimation error must be greater than or equal to the inverse of the Fisher Information Matrix. Therefore, this study uses the Fisher Information Matrix as a key indicator to describe the system's accuracy. Specifically, the magnitude of the Fisher Information Matrix reflects the amount of information contained in the state variables during the measurement process, thereby quantifying the system's uncertainty. A larger Fisher Information amount signifies that the model itself provides more information, which helps reduce uncertainty and improve prediction accuracy.

Construction of the Evaluation Function. This study considers the formation topology between a single follower robot and multiple leader robots. Each follower robot can only communicate with the leader robots through its own communication device, and there is no information exchange among the follower robots. Based on this premise, the navigation coordinate system of the system is established, as shown in Fig. 1.

For the robots within the formation, we disregard their shapes and sizes, treating them as points in the navigation coordinate system. Let the projected coordinates of the leader robot in the two-dimensional plane be $p_i = [x_i, y_i]$ and the coordinates of the follower robot be $p_i = [x_i, y_i]^T$. The actual distance between the leader robot and the follower robot is given by:

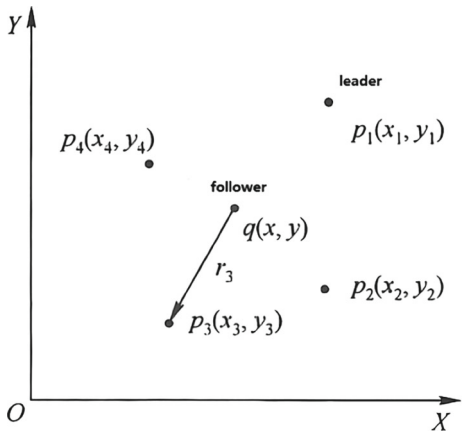


Fig. 1 Position of leader and follower

$$r = p - q\sqrt{(x_i - x)^2 + (y_i - y)^2} \tag{7}$$

where, we introduce measurement noise w, where w is Gaussian white noise:

$$L = p - q + w = \sqrt{(x_i - x)^2 + (y_i - y)^2} + w \tag{8}$$

To evaluate the formation using the Fisher Information Matrix later, it is necessary to obtain the likelihood function of the follower robot's position. When the measurement equation conforms to a Gaussian distribution, its likelihood function can be expressed as:

$$P(q) = \frac{1}{\sqrt{(2\pi)^n |\Sigma|}} \exp\left(-\frac{1}{2}(L - R)^T \Sigma^{-1} (L - R)\right) \tag{9}$$

where n is the number of leader robots, and R and L represent the actual and measured distances, respectively. Taking the natural logarithm of the above likelihood function yields:

The Fisher Information Matrix is defined as:

$$FIM(q) = E\left[\nabla \ln(p(q)) \nabla \ln \ln(p(q))^T\right] \tag{10}$$

Based on the above derivation, the formation topology can be determined by maximizing the determinant of the Fisher Information Matrix.

Leader Formation Topology Optimization. In the previous section, we obtained the Fisher Information Matrix $FIM(q)$ for a single follower mobile robot, which contains information about the measurement data of an unknown parameter matrix. By taking the logarithm of its determinant, we can derive the evaluation function for

the formation topology, which can be optimized to determine the topology of the leader mobile robots.

Assuming there are n leader mobile robots and one follower mobile robot in a two-dimensional space, the position coordinates of the leader mobile robots are given by $P_i = \left[\mathrm{x}_i, \mathrm{y}_i\right]^T$ for $i = 1, 2, 3,\ldots,n$, while the coordinates of the follower mobile robot are $q = \left[x, y\right]^T$. Additionally, we consider a measurement noise vector $W = [w_1, w_2, \ldots, w_n]^T$, where w_i represents the measurement noise of the i-th leader mobile robot with respect to the follower mobile robot.

First, considering the independent and identically distributed nature of the measurement noise, we define the distance r_i between the follower robot and each leader robot, calculated by the following equation:

$$r_i = \sqrt{(x - x_i)^2 + (y - y_i)^2} \tag{11}$$

Based on this definition, the Fisher Information Matrix $FIM\,(q)$ is further constructed as follows:

$$FIM\,(q) = \sum_{i=1}^{n} \left(\frac{\partial r}{\partial x} \right)^2 \tag{12}$$

Without loss of generality, we choose the follower robot as the origin of the coordinate system, i.e., $\mathrm{q} = [0, 0]^T$. Under this setup, the Fisher Information Matrix $FIM\,(q)$ can be further transformed into:

$$FIM\,(q) = \sum_{i=1}^{n} \begin{bmatrix} \cos^2(a_i) & \sin(a_i)\cos(a_i) \\ \sin(a_i)\cos(a_i) & \sin^2(a_i) \end{bmatrix} \tag{13}$$

where α_i is the angle between the vector from the follower mobile robot to the i-th leader mobile robot and the positive x-axis in the counterclockwise direction.

Through further mathematical derivation, an important conclusion can be reached: to maximize the determinant of the Fisher Information Matrix, the vectors $A = [\cos(\alpha_1), \cos(\alpha_2), \ldots, \cos(\alpha_n)]$ and $B = [\sin(\alpha_1), \sin(\alpha_2), \ldots, \sin(\alpha_n)]$ should be orthogonal in space. This implies that the formation topology is related to the relative angles between the leader and follower robots, rather than the distances between them.

Based on this analysis, a geometric optimization solution is proposed: in the formation structure, the n leader robots should be arranged evenly around a single follower robot, forming an equiangular distribution with an angle of $\frac{2\pi}{n}$ between adjacent vectors.

Since there are no distance constraints, the n leader mobile robots can be uniformly arranged along the circumference of a circle with radius r, centered at the follower mobile robot, thus satisfying the requirements for the topology of the leader mobile robots.

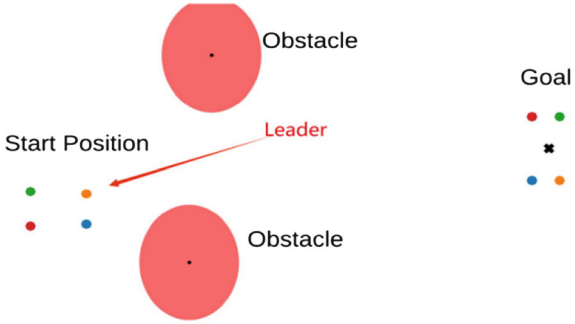


Fig. 2 The distribution of the starting point, endpoint, and obstacles of the formation

4 Static Obstacle Experiment

4.1 Simulation Experiment

Based on the above analysis, a simulation environment has been established in MATLAB. As shown in Fig. 2, the initial positions of the four mobile robots are displayed, with the robot in the upper right corner acting as the leader of the formation, while the others are followers. The entire formation, led by the leader, moves towards a predetermined target position. The process of navigating around obstacles is illustrated in Fig. 3. The entire process took 75 s, and the robots did not collide with any obstacles.

4.2 Real Vehicle Experiment

As shown in Fig. 4a shows the initial state of the mobile robot formation in a "line" topology. The robots maintain the "line" topology as they continue to move forward. Figures 4c–f depict the process of formation change. From the figures, it can be seen that the formation transforms into a triangular shape.

As shown in Fig. 5a, the initial formation of the team is a triangular shape. In Fig. 5b–d, as the formation passes through two cylindrical obstacles, it tightens and continues to move forward in a triangular topology. In Fig. 5e, f, the formation fully clears the obstacles, and the shape returns to its original state.

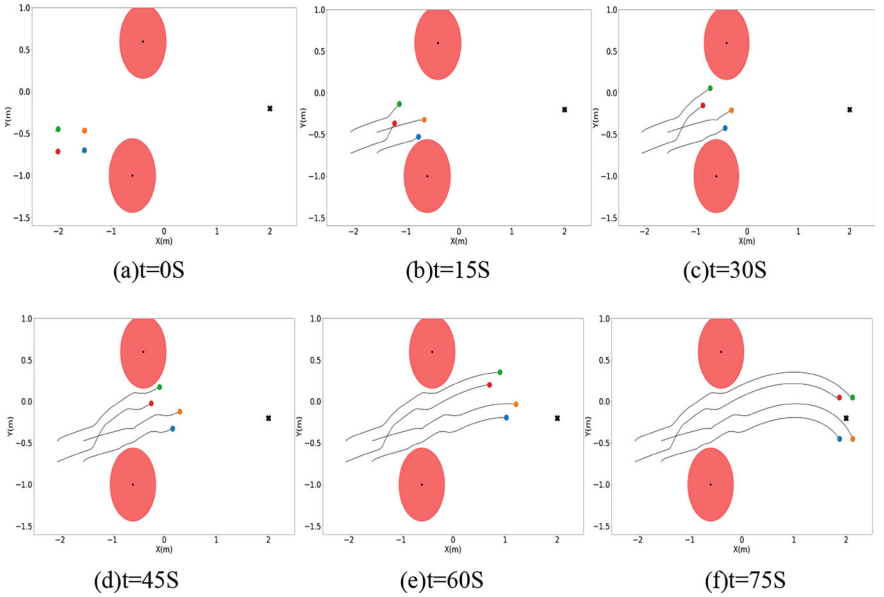


Fig. 3 The process of the robot formation passing through two static obstacles and reaching the target position

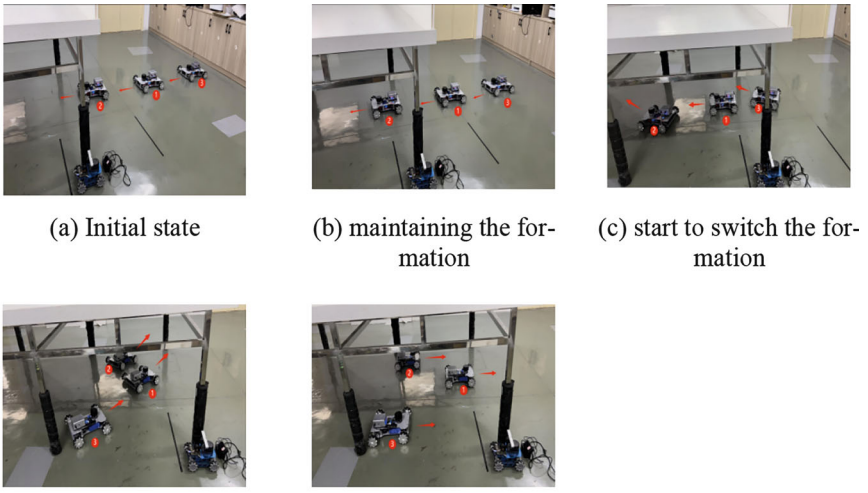


Fig. 4 Formation switching experiment

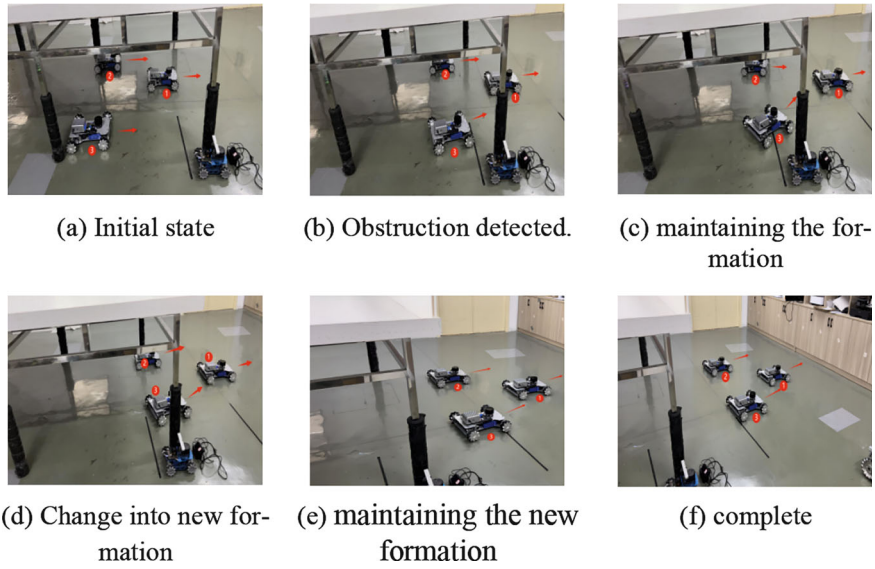


Fig. 5 Formation maintenance experiment

5 Conclusion

This study is based on the optimization of the topology structure of leader–follower formations. Firstly, the observability of the leader robot system is analyzed, and a particle filtering positioning algorithm is employed to control the positioning error within a bounded range, allowing the trajectory of the leader robot to closely match the pre-planned trajectory. This is achieved by using a set of random particles with non-negative weights to approximate the current state of the robot. Through this approach, we aim to achieve a more accurate and stable leader robot formation control system. The Fisher Information Matrix and the Cramér-Rao lower bound are introduced, and based on this, an evaluation function is constructed to optimize the formation topology structure for both single follower robots and multiple follower robots. The effectiveness of these optimizations is validated through simulation experiments.

However, this study still has some shortcomings, if the leader goes offline abruptly in maintaining formation followers will lose track of the target. Base on this, the next step of the study will focus on addressing this issue.

Acknowledgements The authors gratefully acknowledge the support of the Xiamen Production and Research Post-subsidy Project (No. 2023CXY0316), Fujian Provincial Department of Science and Technology- University-industry cooperation project (No. 2024H6012).

References

1. Tahoun AH, Arafa M (2022) Adaptive leader–follower control for nonlinear uncertain multi-agent systems with an uncertain leader and unknown tracking paths. ISA Trans 131:61–72
2. Pastrana MA, Santos KR, de Farias ABC et al (2022) Data-driven control and behavior-based control applied to a SISO mobile robot//2022 Latin American Ro-botics Symposium (LARS), Brazilian Symposium on Robotics (SBR), and 2022 Workshop on Robotics in Education (WRE). IEEE 2022:395–400
3. Chen X, Huang F, Zhang Y et al (2020) A novel virtual-structure formation control design for mobile robots with obstacle avoidance. Appl Sci 10(17):5807
4. Desai JP, Ostrowski J, Kumar V (1998) Controlling formations of multiple mobile robots: Proceedings. In: IEEE International conference on robotics and automation (Cat. No. 98CH36146)
5. Koo TJ, Shahruz SM (2001) Formation of a group of unmanned aerial vehicles (UAVs). In: Proceedings of the 2001 American control conference. (Cat. No. 01CH37148)
6. Tanner HG, Pappas GJ, Kumar V (2004) Leader-to-formation stability. IEEE Trans Robot Autom 20(3):443–455
7. Shi-Cai L, Da-Long T, Guang-Jun L (2007) Formation control of mobile robots with active obstacle avoidance. Acta Automatica Sinica 33(5):529–535
8. Kress-Gazit H, Conner DC, Choset H et al (2008) Courteous cars. IEEE Robot Autom Mag 15(1):30–38
9. Tran VP, Garratt MA, Petersen IR (2020) Switching formation strategy with the directed dynamic topology for collision avoidance of a multi-robot system in uncertain environments. IET Control Theory Appl 14(18):2948–2959
10. Pan Z, Zhang C, Xia Y et al (2021) An improved artificial potential field method for path planning and formation control of the multi-UAV systems. IEEE Trans Circuits Syst II Express Briefs 69(3):1129–1133

Open Access This chapter is licensed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License (http://creativecommons.org/licenses/by-nc-nd/4.0/), which permits any noncommercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if you modified the licensed material. You do not have permission under this license to share adapted material derived from this chapter or parts of it.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.



Multi-body Rigid-Flexible Coupling System Dynamic Analysis of Spent Fuel Crane



Xuyang Cao, Wanlong Shu, Chengjie Zhang, and Kai Lin

Abstract The vibration of the main beam structure of the spent fuel crane and the swing of the payload will affect the safety of the main beam structure and the positioning accuracy of the payload. A multi-body rigid-flexible coupling system dynamics model composed of flexible main girder, trolley and payload is established. The second type of Lagrange equation is used to derive the dynamics differential equations, and the vibration modes of the main beam are solved by the Newmark-β method. The effects of the system's multi-parameters on the vibration characteristics of the main beam and the swing characteristics of the payload are analyzed in simulation. The results show that the increase of the mass of the lifting trolley and the payload significantly affects the vibration amplitude of the main beam and the swing angle of the payload, and the first-order mode of the main beam dominates the vibration process; the running speed of the lifting trolley has no significant effect on the vibration amplitude of the main beam.

Keywords Spent fuel crane · Rigid-flexible coupling · Lagrange equation · Vibration characterization

1 Introduction

Spent fuel pool crane (referred to as spent crane) is a special crane used to handle fuel loading and unloading in the spent fuel pool, which has a structure similar to a bridge crane [1], with a large tonnage and large span. Modern special lifting equipment is gradually developing towards large-scale and intelligentization, and the requirements for high lifting capacity, transportation efficiency and safety performance [2] of bridge cranes are increasing, so it is of great significance to study the coupled vibration characteristics of the crane main beam-trolley-payload system and the swing characteristics of the payload.

X. Cao · W. Shu (✉) · C. Zhang · K. Lin
Dalian University of Technology, Dalian, China
e-mail: shuwl9914@163.com

© The Author(s) 2026

H. Saman et al. (eds.), *The 9th International Conference on Advances in Construction Machinery and Vehicle Engineering*, Lecture Notes in Mechanical Engineering,
https://doi.org/10.1007/978-981-96-5527-4_48

The coupling system of moving mass and beam structure [3, 4] is a class of widely used research models. In coupled system modeling [5] many scholars use HHT-α method [6], generalized Dosin method to describe the dynamic behavior of the axle system, in beam structure modeling related scholars based on the Euler-Bernoul beam theory [7] to establish a single-beam structural numerical model of the banked bridge [8], and there are other scholars using the Preissmann method [9] to study the speed of the beam parameters and the moving mass and other working conditions on beam vibration [10]. Based on this, this paper takes the spent fuel crane as the research object, establishes the main beam-trolley-crane weight multi-body rigid-flexible coupling system dynamics model, adopts the second type of Lagrangian equation to establish the multi-body rigid-flexible coupling system mechanics model, and utilizes the Newmark-β method to solve the system differential equation system, and researches the influence of different working conditions on the structural vibration characteristics of the main beam as well as the swing characteristics of the payload.

2 Rigid-Flexible Coupling System Model

2.1 Modeling of System Dynamics Equations

The multi-body rigid-flexible coupling system model of the trolley-crane-main beam structure of the crane system is shown in Fig. 1. The main beam structure adopts Euler beam model, which is simplified as a simply supported beam.

Let the deflection at x on the main beam structure at moment t be $w(x, t)$, then the vector $\vec{\boldsymbol{v}}_b$ in the vertical direction at each point on the main beam structure is,

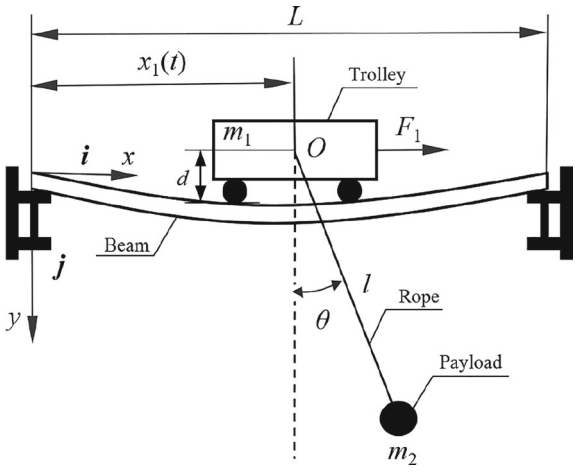


Fig. 1 Dynamics model of multi-body rigid-flexible coupling system

$$\vec{\boldsymbol{v}}_b = \dot{x}_b\boldsymbol{i} + \dot{y}_b\boldsymbol{j} = \frac{\partial w(x,t)}{\partial t}\boldsymbol{j} \tag{1}$$

The trolley moves on the main beam structure with coordinates in the vertical direction linked to the deflection of the main beam structure at that point, and the velocity vector of the trolley is,

$$\vec{\boldsymbol{v}}_1 = \dot{x}_1\boldsymbol{i} + \dot{y}_1\boldsymbol{j} = \dot{x}_1(t)\boldsymbol{i} + \frac{\partial w(x_1,t)}{\partial t}\boldsymbol{j} \tag{2}$$

The position coordinates of the payload (x_2, y_2) are,

$$\begin{cases} x_2 = x_1(t) + l\sin\theta \\ y_2 = w(x_1,t) + l\cos\theta - d \end{cases} \tag{3}$$

The velocity vector $\vec{\boldsymbol{v}}_2$ of the payload is obtained from Eq. (3) as,
Let the deflection at x

$$\vec{\boldsymbol{v}}_2 = \dot{x}_2\boldsymbol{i} + \dot{y}_2\boldsymbol{j} = (\dot{x}_1(t) + l\dot{\theta}\cos\theta)\boldsymbol{i} + (\frac{\partial w(x_1,t)}{\partial t} - l\dot{\theta}\sin\theta)\boldsymbol{j} \tag{4}$$

The vibration of the main beam can be solved by the modal superposition method, and the deflection of the main beam is calculated by the following equation,

$$w(x,t) = \sum_{i=1}^{N}\phi_i(x)q_i(t) \tag{5}$$

where, $\emptyset_i(x) = sin\left(\frac{i\pi x}{L}\right)$ is the modal function of the main beam, and the order of the main beam is taken as Nth order, i is the ith order of the mode, and $q_i(t)$ is the generalized coordinate corresponding to the ith order of the mode.

The total kinetic energy T and potential energy V of the multi-body coupling system is the sum of the kinetic energy and potential energy of the main beam, the trolley and the payload, which are calculated by the following equations,

$$\begin{aligned} T = {} & \frac{1}{2}m_1\left(\dot{x}_1^2 + \left(\dot{x}_1\sum_{i=1}^{N}\phi_i\prime(x_1)q_i + \sum_{i=1}^{N}\phi_i(x_1)\dot{q}_i\right)^2\right) + \frac{1}{4}m_aL\sum_{i=1}^{N}\dot{q}_i^2 \\ & + \frac{1}{2}m_2\left(\begin{array}{l} \dot{x}_1^2 + 2\dot{x}_1 l\dot{\theta}\cos\theta + (l\dot{\theta})^2 + \left(\dot{x}_1\sum_{i=1}^{N}\phi_i\prime(x_1)q_i + \sum_{i=1}^{N}\phi_i(x_1)\dot{q}_i\right)^2 \\ -2l\dot{\theta}\sin\theta\left(\dot{x}_1\sum_{i=1}^{N}\phi_i\prime(x_1)q_i + \sum_{i=1}^{N}\phi_i(x_1)\dot{q}_i\right) \end{array}\right) \end{aligned} \tag{6}$$

$$V = \frac{EI\pi^4}{4L^3}\sum_{i=1}^{N} i^4 q_i^2 - m_1 g\sum_{i=1}^{N}\phi_i(x_1)q_i - m_2 g\left(\sum_{i=1}^{N}\phi_i(x_1)q_i + l\cos\theta\right) \tag{7}$$

The trolley-crane-main beam multi-body rigid-flexible coupling system is a typical nonlinear underdriven dynamical system, which can be described by Lagrangian equation,

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}}\right) - \frac{\partial T}{\partial q} + \frac{\partial V}{\partial q} = Q \tag{8}$$

where T is the total kinetic energy of the coupling system, V is the total potential energy of the coupling system, q is the generalized coordinate of the coupling system, q is the generalized velocity of the coupling system, and Q is the generalized force of the coupling system.

x_1, θ and q_i are selected as the generalized coordinates of the coupling system, and (6) and (7) are applied to the Lagrange equation to obtain the set of differential equations for the dynamics of the multi-body coupling system. It can be assumed that: $\sin\theta = \theta$, $\cos\theta = 1$, and the simplified system of equations is as follows,

$$\begin{aligned} &m_1\ddot{x}_1 + m_2\left(\ddot{x}_1 + l\ddot{\theta} - l\theta\ddot{\theta}\sum_{i=1}^{N}\phi_i{}'(x_1)q_i - l\dot{\theta}^2\left(\theta + \sum_{i=1}^{N}\phi_i{}'(x_1)q_i\right)\right) \\ &+(m_1+m_2)\left(\begin{array}{l} \ddot{x}_1\left(\sum_{i=1}^{N}\phi_i{}'(x_1)q_i\right)^2 + 2\dot{x}_1\sum_{i=1}^{N}\phi_i{}'(x_1)q_i\sum_{i=1}^{N}\phi_i{}'(x_1)\dot{q}_i \\ -g\sum_{i=1}^{N}\phi_i{}'(x_1)q_i + \dot{x}_1^2\sum_{i=1}^{N}\phi_i{}''(x_1)q_i\sum_{i=1}^{N}\phi_i{}'(x_1)q_i \\ +\sum_{i=1}^{N}\phi_i{}'(x_1)q_i\sum_{i=1}^{N}\phi_i(x_1)\ddot{q}_i \end{array}\right) = F_1 \end{aligned} \tag{9}$$

$$m_2\left(\begin{array}{l} -l\theta\left(\ddot{x}_1\theta\sum_{i=1}^{N}\phi_i{}'(x_1)q_i + \dot{x}_1^2\sum_{i=1}^{N}\phi_i{}''(x_1)q_i + 2\dot{x}_1\sum_{i=1}^{N}\phi_i{}'(x_1)\dot{q}_i\right) \\ +\ddot{x}_1 + l^2\ddot{\theta} + l\theta\left(g - \sum_{i=1}^{N}\phi_i(x_1)\ddot{q}_i\right) \end{array}\right) = F_\theta \tag{10}$$

$$(m_1+m_2)\begin{pmatrix}\ddot{x}_1\phi_j(x_1)\sum_{i=1}^{N}\phi_i\prime(x_1)q_i+2\dot{x}_1\phi_j(x_1)\sum_{i=1}^{N}\phi_i\prime(x_1)\dot{q}_i\\ +\dot{x}_1^2\phi_j(x_1)\sum_{i=1}^{N}\phi_i^{''}(x_1)q_i+\phi_j(x_1)\sum_{i=1}^{N}\phi_j(x_1)\ddot{q}_i\end{pmatrix}+\frac{EI\pi^4j^4}{2L^3}q_j$$
$$+\frac{1}{2}m_aL\ddot{q}_j+m_2\left(-l\theta\ddot{\theta}\phi_j(x_1)-l\dot{\theta}^2\phi_j(x_1)\right)-(m_1+m_2)g\phi_j(x_1)=0 \tag{11}$$

2.2 Numerical Solution of Systems of Equations

In the vibration response analysis of the main beam, generally the first few orders of modes play a major role in the vibration pattern of the main beam, so only the first few orders of modes of the main beam can satisfy the accuracy in solving. The system of Eqs. (9), (10) and (11) can be written in the following matrix form,

$$M\ddot{Z}+C\dot{Z}+KZ=F \tag{12}$$

In the formula, $Z=\left[x_1,\theta,q_1,q_2,\ldots,q_N\right]$.

The above system of equations is a second order system of ordinary differential equations with variable coefficients, which is solved numerically using ode45 solver in MATLAB. The vibration modes of the main beam are solved numerically using the Newmark-β method. The values of the parameters are $\gamma=0.25$, $\beta=0.5$, and $\Delta t=0.01$.

3 Dynamics Simulation Solution and Analysis

The dynamic characteristics of multi-body coupling system are studied and analyzed as an example of a certain type of crane, and the main parameters of this crane are shown in the following Table 1.

3.1 Constant Drive Force Model Simulation

Taking the driving force $F_1=8\times10^3$N, and Fig. 2 shows the vibration modes of the main beam structure of the first 6 orders under the mass of the two different trolleys, and it can be seen that the vibration amplitude of the main beam in the first order is much larger than that of the other orders, and the vibration value of the main beam is close to 0 after the 2nd order. The result proves that it is reasonable to take only the

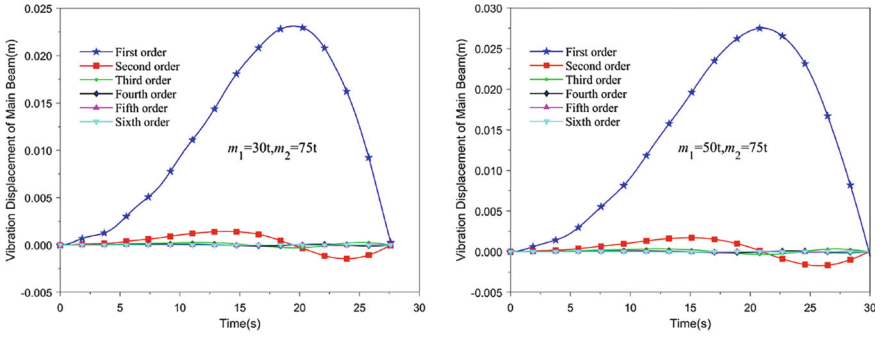
Table 1 Crane model parameters

Parameter name	Notation	Parameter value
Trolley mass	m_1	3–5 × 10^4 kg
Trolley speed	v_1	0–25 m/min
Lifting mass	m_2	5–15 × 10^4 kg
Modulus of elasticity	E	2.06 × 10^{11} N/m^2
Moment of inertia	I	0.12m^4
Length of main beam	L	30 m
Main beam unit mass	m_a	1.5 × 10^4 kg
Air damping coefficient	μ	0.015

first few orders of the main beam modes. Comparing the two working conditions, it can be seen that the vibration amplitude of the main beam modes increases with the increase of the mass of the trolley.

Take the mass of trolley $m_1 = 3 \times 10^4$ kg, 5×10^4 kg, and the mass of payload $m_2 = 7.5 \times 10^4$ kg, 1×10^5 kg, combine the mass of trolley and the mass of payload two by two to carry out the simulation solution, and get the results of the main beam mid-point deflection in the four conditions as shown in Fig. 3, and it can be concluded from the comparison that, when the mass of the trolley and the mass of the payload increase, the mid-point deflection of main beam of the main beam will both increase.

Figure 4 shows the pendulum angle of different lifting mass when the mass of the trolley is certain. It can be seen from these two figures that when the mass of the trolley and the mass of the payload increase, the swing angle of the payload decreases.



a) Trolley mass condition I b) Trolley mass condition II

Fig. 2 The first 6 orders of vibration modes of main beam

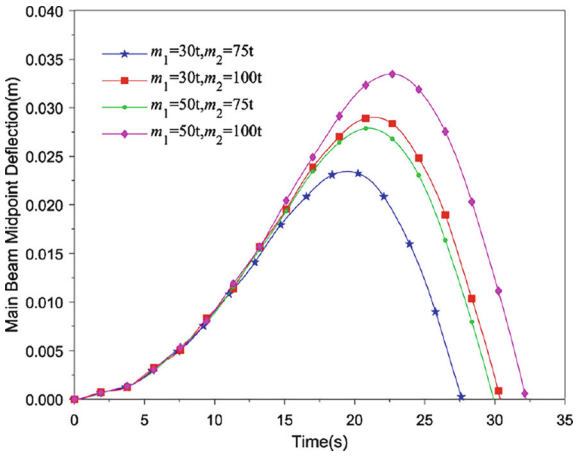
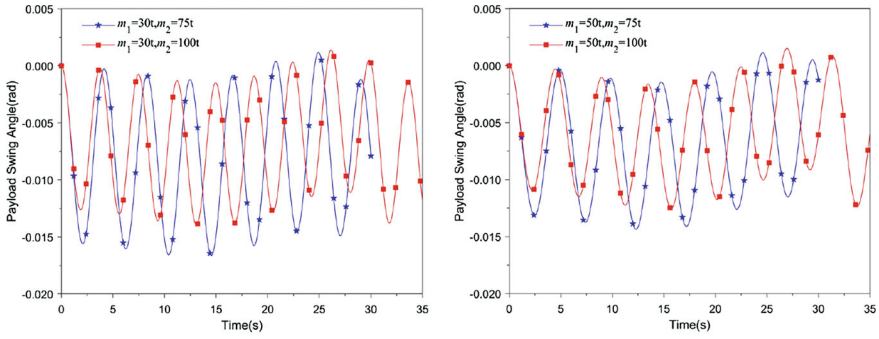


Fig. 3 Deflection at midpoint of main beam

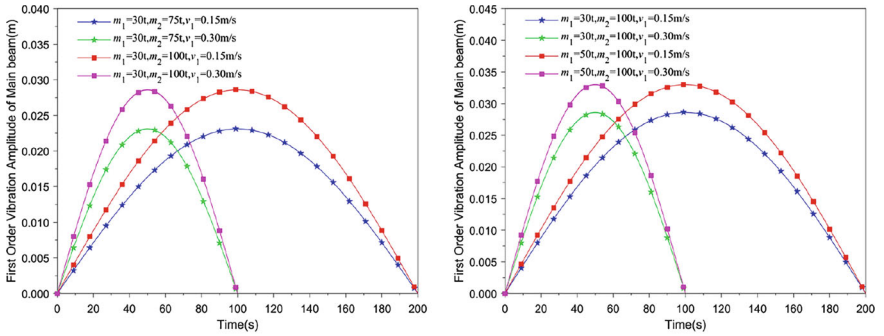


a) Working condition I **b)** Working condition II

Fig. 4 Swing angle of payload

3.2 Constant Trolley Speed Model Simulation

Figure 5 shows the first-order vibration modes of the main beam obtained by letting the trolley pass through the main beam at a certain speed and changing the mass of the trolley and the mass of the payload. From the figure, it can be seen that the vibration amplitude of the first order of the main beam increases with the increase of the mass of the trolley and the mass of the payload, while the speed of the trolley has almost no effect on the vibration amplitude of the main beam.



a) Working condition I **b)** Working condition II

Fig. 5 First order vibration modes of the main beam

3.3 Variable Rope Length Model Simulation

The length of the crane wire rope affects the swing of the payload, take the length of the wire rope l = 10, 16, 22 m. The simulation results are shown in Fig. 6. From the results, the following conclusions can be obtained: with the increase of the length of the wire rope, the pendulum angle of the payload increases accordingly, and the pendulum angle of the payload under the three kinds of wire rope lengths has a big difference in the pendulum amplitude.

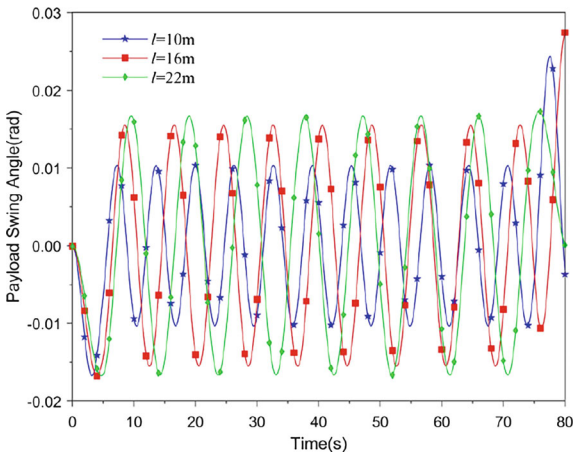


Fig. 6 Swing angle of payload with different rope lengths

4 Conclusion

In this paper, a multi-body rigid-flexible coupling system dynamics model of the spent fuel crane is established by using the Lagrange equation, and the effects of the system parameters on the vibration characteristics of the main beam and the swinging characteristics of the payload are investigated through dynamics simulation. The results show that the modal vibration amplitude of the main beam structure increases with the increase of the mass of the trolley and the mass of the payload, and the swinging angle of the payload decreases; the running speed of the lifting trolley has no significant effect on the modal vibration of the main beam structure; and the increase of the length of the wire rope also causes the increase of the swinging angle of the payload.

Acknowledgements This research is supported by National Natural Science Foundation of China No.52275088 and the Fundamental Research Funds for the Central Universities (No. DUT22LAB506, No. DUT22LAB507).

References

1. Pang Y, Pang Y (2020) Analysis of electrical control system of spent fuel pool operator and special gripper in nuclear power plant. Mach Manag Dev 35(10):276–278
2. Calimanescu I, Stan LC (2020) Seismic qualification by analysis of an overhead crane for nuclear industry. In: Proceedings of SPIE—the international society for optical engineering. Constanta, Romania: SPIE, 2020: Maritime University of Constanta; Ministry of Research and Innovation; University Politehnica of Bucharest
3. Tong M, Shen Y (2020) Research on coupling vibration of wheel-rail beam system in the high-speed quayside container crane. Mech Adv Mater Struct 27(17):1473–1482
4. Homaei H, Stoura CD, Dimitrakopoulos EG (2024) Extended modified bridge system (EMBS) method for decoupling seismic vehicle-bridge interaction. In: Earthquake Engineering and structural dynamics
5. Wang SQ, Wan X, Wang X, Tong et al (2023) Calculation and monitoring of vibration response of large-span cable-stayed bridge under vehicle dynamic loading. Sci Technol Eng 23(31):13556–13565
6. Sun PF (2024) Dynamic response analysis of axle coupling system under seismic excitation. Jilin University
7. Bettinelli L, Stollwitzer A, Fink J (2023) Numerical study on the influence of coupling beam modeling on structural accelerations during high-speed train crossings. Appl Sci-Basel 13(15):8746
8. Wang L (2022) Identification of structural dynamic parameters based on time-varying analysis of axle coupled system. Southeast University
9. Hu J, Hu W, Xiao C et al (2022) Coupling vibration of simply-supported damping beam carrying a moving mass. Acta Mech Solida Sin 35(6):972–978
10. Chen Q, Cheng W, Gao L et al (2020) Dynamic response of a gantry Crane's beam subjected to a two-axle moving trolley. Math Probl Eng 2020:1–10

Open Access This chapter is licensed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License (http://creativecommons.org/licenses/by-nc-nd/4.0/), which permits any noncommercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if you modified the licensed material. You do not have permission under this license to share adapted material derived from this chapter or parts of it.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.



Dynamics Modeling and Simulation of Dual-Crane Cooperative Lifting System



Xuyang Cao, Chengjie Zhang, Wanlong Shu, and Baokang Yan

Abstract The dual-crane cooperative lifting system demonstrates significant advantages in lifting large and complex components. However, due to the inertial effects of the payload and the dynamic motion characteristics of the lifting equipment, it is prone to causing the payload to swing, which affects the system's stability and safety. To address this issue, this paper establishes a dynamic model for the dual-crane cooperative lifting system based on the Lagrangian equation. Through simplified assumptions, the complete nonlinear dynamic equations are derived, and the model's validity is verified using Adams simulation software. The simulation results show that the established model can accurately describe the system's dynamic behavior and reveal the impact of wire rope length and initial angle on the payload's swinging characteristics. Further analysis indicates that a longer steel wire rope increases the swing period and amplitude of the payload, while a larger initial angle significantly raises its swinging frequency and amplitude.

Keywords Dual-crane cooperative lifting · Dynamic model · Lagrangian equation · Adams

1 Introduction

As engineering projects grow larger, traditional single-crane lifting methods can no longer handle oversized, heavy components. To address this, dual-crane cooperative lifting systems have been developed (see Fig. 1), allowing two cranes to work together to lift heavy loads. This improves load capacity, efficiency, and safety. However, the payload's inertia and the dynamic characteristics of the equipment often cause swinging, reducing positioning accuracy and increasing safety risks. Therefore, suppressing payload swing during lifting has become a key challenge in ensuring system stability and safety.

X. Cao · C. Zhang (✉) · W. Shu · B. Yan
Dalian University of Technology, Dalian, China
e-mail: cjzhang@mail.dlut.edu.cn

© The Author(s) 2026

H. Saman et al. (eds.), *The 9th International Conference on Advances in Construction Machinery and Vehicle Engineering*, Lecture Notes in Mechanical Engineering,
https://doi.org/10.1007/978-981-96-5527-4_49



Fig. 1 Dual-crane cooperative lifting

In recent years, scholars have conducted in-depth research on the dynamic modeling and anti-swing control of dual-crane cooperative lifting systems, with significant studies focused on dual drones [1, 2], dual automobile cranes [3, 4], dual overhead cranes [5–8], and dual ship-mounted crane systems [9, 10]. To address common issues in the dynamic modeling of dual-crane cooperative lifting systems, this paper abstracts the lifting equipment as a trolley at the end position and establishes a mathematical dynamic model for the system based on the Lagrangian equation. The correctness of the mathematical model is verified using the dynamic analysis software Adams, and the swinging behavior of the payload under different working conditions is explored.

2 Dynamic Modeling

As shown in Fig. 2, two cranes with masses m_1 and m_2 (referred to as trolleys hereafter) work together to transport a cylindrical payload with mass m, length l, and radius r. The driving forces of the trolleys are F_1 and F_2, and their positions in the global coordinate system oxy are x_1 and x_2, respectively. The two suspension points of the payload, a and b, are connected to the trolleys via two steel wire ropes of lengths l_1 and l_2. The angles between the steel wire rope and the vertical direction are θ_1 and θ_2, and the angle between the payload's central axis and the horizontal direction is θ_3.

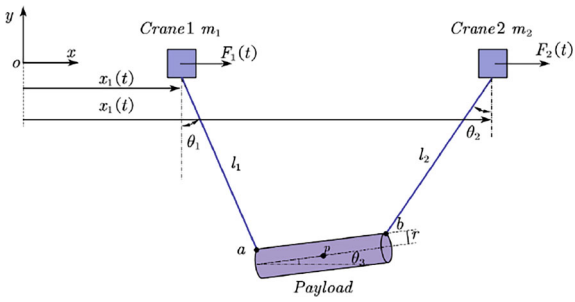


Fig. 2 Physical diagram of the dual-crane cooperative lifting system

2.1 Dynamic Analysis

The dynamic model of the system is established using the Lagrangian equation, as follows:

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}_k}\right) - \frac{\partial L}{\partial q_k} = Q_k \tag{1}$$

where L is the Lagrangian, $L = T - V$. q_k and Q_k are the generalized coordinates and generalized forces of the system, respectively. T represents the kinetic energy of the system, and U is the potential energy. Based on geometric relationships, the coordinates of point a in the global coordinate system are:

$$\begin{aligned} x_a &= x_1 + l_1 \sin\theta_1 \\ y_a &= -l_1 \cos\theta_1 \end{aligned} \tag{2}$$

The coordinates of point b in the global coordinate system are:

$$\begin{aligned} x_b &= x_2 + l_2 \sin\theta_2 \\ y_b &= -l_2 \cos\theta_2 \end{aligned} \tag{3}$$

The coordinates of the center of mass of the payload are:

$$\begin{aligned} x_p &= \frac{1}{2}(x_1 + x_2 + l_1 \sin\theta_1 + l_2 \sin\theta_2) + r\sin\theta_3 \\ y_p &= -\frac{1}{2}(l_1 \cos\theta_1 + l_2 \cos\theta_2) - r\cos\theta_3 \end{aligned} \tag{4}$$

The system's kinetic energy T and potential energy U are given by:

$$\begin{aligned} T &= \frac{1}{2}m_1 x_1^2 + \frac{1}{2}m_2 x_2^2 + \frac{1}{2}m(x_p^2 + y_p^2) \\ U &= mgy_p \end{aligned} \tag{5}$$

By choosing the generalized coordinates $q_k = \begin{bmatrix} x_1 & x_2 & \theta_1 & \theta_2 & \theta_3 \end{bmatrix}^T$, the corresponding generalized forces are $Q_k = \begin{bmatrix} F_1 & F_2 & 0 & 0 & 0 \end{bmatrix}^T$. Combining Eqs. (1)–(5), the specific dynamic equations for the dual-crane cooperative lifting system are derived. For the generalized coordinate x_1:

$$\begin{aligned} &\left(\frac{m}{4} + m_1\right)\ddot{x}_1 + \frac{m\ddot{x}_2}{4} + \frac{l_1 m\cos(\theta_1)\ddot{\theta}_1}{4} + \frac{l_2 m\cos(\theta_2)\ddot{\theta}_2}{4} \\ &- \frac{rm\cos(\theta_3)\ddot{\theta}_3}{2} - \frac{l_1 m\sin(\theta_1)\dot{\theta}_1^2}{4} - \frac{l_2 m\sin(\theta_2)\dot{\theta}_2^2}{4} + \frac{rm\sin(\theta_3)\dot{\theta}_3^2}{2} = F_1 \end{aligned} \tag{6}$$

For the generalized coordinate x_2:

$$\begin{aligned}&\frac{m\ddot{x}_1}{4}+\Big(\frac{m}{4}+m_2\Big)\ddot{x}_2+\frac{l_1 m\cos(\theta_1)\ddot{\theta}_1}{4}+\frac{l_2 m\cos(\theta_2)\ddot{\theta}_2}{4}\\&-\frac{rm\cos(\theta_3)\ddot{\theta}_3}{2}-\frac{l_1 m\sin(\theta_1)\dot{\theta}_1^2}{4}-\frac{l_2 m\sin(\theta_2)\dot{\theta}_2^2}{4}+\frac{rm\sin(\theta_3)\dot{\theta}_3^2}{2}=F_2\end{aligned}\tag{7}$$

For the generalized coordinate θ_1:

$$\begin{aligned}&\frac{l_1 m\cos(\theta_1)\ddot{x}_1}{4}+\frac{l_1 m\cos(\theta_1)\ddot{x}_2}{4}+\frac{l_1^2 m\ddot{\theta}_1}{4}+\frac{l_1 l_2 m\cos(\theta_1-\theta_2)\ddot{\theta}_2}{4}\\&+\frac{l_1 rm\cos(\theta_1-\theta_3)\ddot{\theta}_3}{2}+\frac{l_1 l_2 m\sin(\theta_1-\theta_2)\dot{\theta}_2^2}{4}\\&+\frac{l_1 rm\sin(\theta_1-\theta_3)\dot{\theta}_3^2}{2}+\frac{l_1 gm\sin(\theta_1)}{2}=0\end{aligned}\tag{8}$$

For the generalized coordinate θ_2:

$$\begin{aligned}&\frac{l_2 m\cos(\theta_2)\ddot{x}_1}{4}+\frac{l_2 m\cos(\theta_2)\ddot{x}_2}{4}+\frac{l_1 l_2 m\cos(\theta_1-\theta_2)\ddot{\theta}_1}{4}+\frac{l_2^2 m\ddot{\theta}_2}{4}\\&+\frac{l_2 bm\cos(\theta_2-\theta_3)\ddot{\theta}_3}{2}-\frac{l_1 l_2 m\sin(\theta_1-\theta_2)\dot{\theta}_1^2}{4}\\&+\frac{l_2 rm\sin(\theta_2-\theta_3)\dot{\theta}_3^2}{2}+\frac{l_2 gm\sin(\theta_2)}{2}=0\end{aligned}\tag{9}$$

For the generalized coordinate θ_3:

$$\begin{aligned}&-\frac{rm\cos(\theta_3)\ddot{x}_1}{2}-\frac{rm\cos(\theta_3)\ddot{x}_2}{2}+\frac{l_1 rm\cos(\theta_1-\theta_3)\ddot{\theta}_1}{2}+\frac{l_2 rm\cos(\theta_2-\theta_3)\ddot{\theta}_2}{2}\\&+r^2 m\ddot{\theta}_3-\frac{l_1 rm\sin(\theta_1-\theta_3)\dot{\theta}_1^2}{2}-\frac{l_2 rm\sin(\theta_2-\theta_3)\dot{\theta}_2^2}{2}+rgm\sin(\theta_3)=0\end{aligned}\tag{10}$$

In summary, a two-input, five-output underactuated nonlinear dynamic equation for the dual-crane cooperative lifting system has been established. Since this system exhibits highly complex nonlinearity, and internal constraints affect the system's solution, it is necessary to simplify the mathematical model by incorporating these constraints to find the solution to the dynamic equations.

2.2 Mathematical Model Simplification

As shown in Fig. 2, the system has several geometric constraints:

$$\begin{cases} l_1 \sin\theta_1 - l_2 \sin\theta_2 + l\cos\theta_3 - (x_2 - x_1) = 0 \\ l_1 \cos\theta_1 - l_2 \cos\theta_2 - l\sin\theta_3 = 0 \end{cases} \tag{11}$$

From Eq. (11), it is clear that out of the five state variables, two are dependent variables that vary according to the other three variables. Therefore, by choosing $\overline{q} = \begin{bmatrix} x_1 & x_2 & \theta_1 \end{bmatrix}^T$ as the independent variables, θ_2 and θ_3 can be expressed as functions of $\overline{q}$.

$$\begin{cases} \theta_2 = g(\overline{q}) = g(x_1, x_2, \theta_1) \\ \theta_3 = h(\overline{q}) = h(x_1, x_2, \theta_1) \end{cases} \tag{12}$$

Taking the time derivative of Eq. (12), we get:

$$\begin{cases} \dot{\theta}_2 = g_1\dot{x}_1 + g_2\dot{x}_2 + g_3\dot{\theta}_1 \\ \dot{\theta}_3 = h_1\dot{x}_1 + h_2\dot{x}_2 + h_3\dot{\theta}_1 \end{cases} \tag{13}$$

In Eq. (13):

$$\begin{aligned} g_1 = -g_2 = \frac{\cos\theta_3}{l_2\cos(\theta_2 - \theta_3)}, h_1 = -h_2 = \frac{\sin\theta_2}{l\cos(\theta_2 - \theta_3)} \\ g_3 = \frac{l_1\cos(\theta_1 - \theta_3)}{l_2\cos(\theta_2 - \theta_3)}, h_3 = \frac{-l_1\sin(\theta_1 - \theta_2)}{l\cos(\theta_2 - \theta_3)} \end{aligned} \tag{14}$$

The system's dynamic equations have many degrees of freedom and nonlinear characteristics, making it challenging to obtain accurate numerical solutions. Therefore, it is necessary to simplify the equations based on practical conditions. Assuming the two trolleys operate in the same state, $\dot{x}_1 = \dot{x}_2 = \dot{x}$, and the two wire ropes have equal lengths, $l_1 = l_2 = l$, the initial angles of the steel wire ropes are $\theta_{10} = -\theta_{20}, \theta_{30} = 0$, where θ_{10}, θ_{20} and θ_{30} are the initial values of θ_1, θ_2 and θ_3, respectively. Since the trolleys move slowly in practical scenarios, the payload oscillates slightly around the equilibrium point, meaning $\theta_1 \approx -\theta_2, \theta_3 \approx 0$. Based on Fig. 2, we can derive the following:

$$\begin{cases} \sin\theta_1 = -\sin\theta_2 \approx \sin\theta_{10}, \cos\theta_1 = \cos\theta_2 \approx \cos\theta_{10} \\ g_3 = \frac{\cos(\theta_1 - \theta_3)}{\cos(\theta_2 - \theta_3)} \approx 1, h_3 \approx -\frac{2l_1\sin\theta_{10}}{l} \end{cases} \tag{15}$$

Substituting Eq. (15) into Eq. (12), we get $\dot{\theta}_1 = \dot{\theta}_2$. Further, we can derive:

$$\begin{aligned}\theta_2 &= \theta_1 - 2\theta_{10} \\ \theta_3 &= -\frac{2l\sin\theta_{10}}{l}(\theta_1 - \theta_{10})\end{aligned} \tag{16}$$

Substituting Eqs. (15) and (16) into Eq. (8), we obtain the simplified dynamic equation for the system as follows:

$$\begin{aligned}&\frac{2l_1\sin\theta_{10}}{l}(l\cos\theta_{10} - \frac{2rl_1\sin\theta_{10}}{l})\ddot{x} + (l_1\cos\theta_{10} - \frac{2rl_1\sin\theta_{10}}{l})^2\ddot{\theta}_3 \\ &+ g(l_1 + r(\frac{2l_1\sin\theta_{10}}{l})^2)\theta_3 = 0\end{aligned} \tag{17}$$

3 Adams Dynamic Analysis Comparison

To verify the accuracy of the simplified dynamic equations, it is necessary to build an Adams model. Based on the physical diagram of the dual-crane cooperative lifting system shown in Fig. 2, a model of trolleys, drums, steel wire ropes, and the payload is created in the Adams dynamic analysis software, as shown in Fig. 3.

In MATLAB and Adams, the length of the steel wire rope is set to $l_1 = l_2 = 10m$, and the initial angles are $\theta_{10} = -\theta_{20} = 10^\circ$ and $\theta_{30} = 0$. As shown in Fig. 4, the simulation analysis reveals that the numerical model and the Adams simulation results are highly similar, indicating that the established dynamics equations are accurate. This lays a foundation for further exploration of the swing characteristics of the payload under different working conditions.

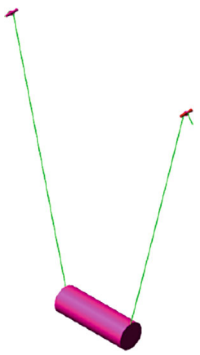
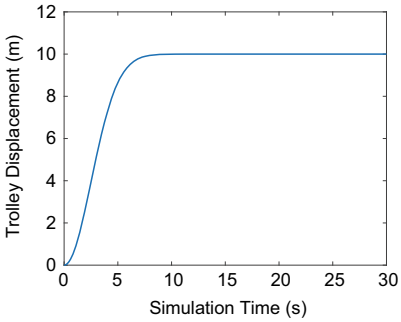
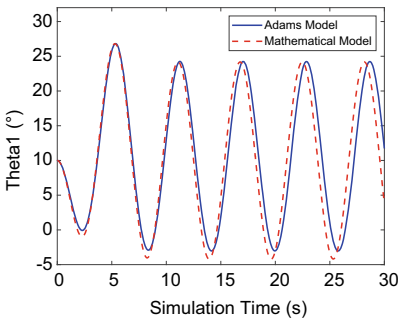


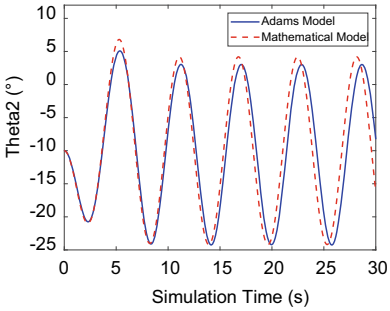
Fig. 3 Adams model of the dual-crane cooperative lifting system



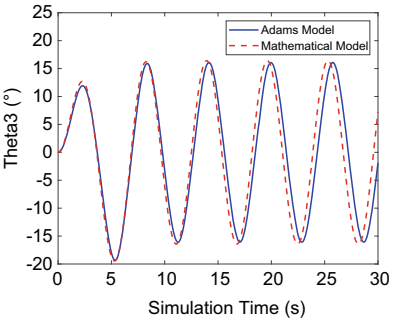
(a) Trolley displacement $x(t)$



(b) Left steel wire rope angle θ_1



(c) Right steel wire rope angle θ_2



(d) Payload deflection angle relative to the horizontal direction θ_3

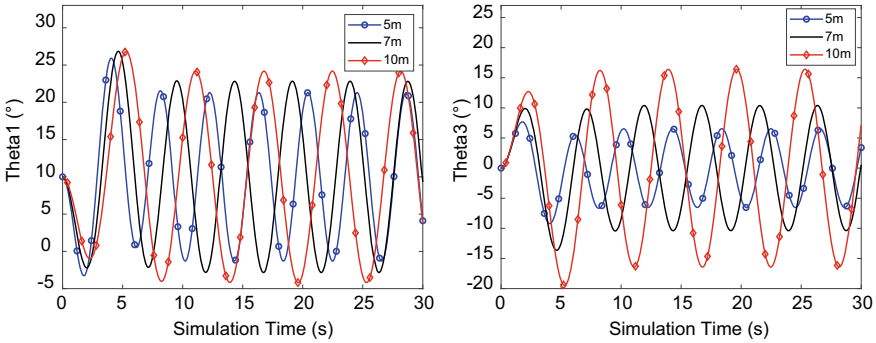
Fig. 4 Comparison of simulation results between the numerical and Adams models

4 Numerical Simulation Analysis

In the dual-crane lifting system, the payload's swing is influenced by factors such as the length of the steel wire rope and the initial angle. This section conducts a simulation analysis to study their impact on the swing behavior.

4.1 Rope Length Variation

As shown in Fig. 5, with a 10° initial angle and steel wire rope lengths of 5, 7, and 10 m, the time-varying curves of θ_1 and θ_3 are plotted. The results show that as the rope length increases, the swing period lengthens and the amplitude grows. Choosing the right rope length enhances stability and safety.



(a) Left steel wire rope angle θ_1

(b) Payload deflection angle relative to the horizontal direction θ_3

Fig. 5 Angle variations with steel wire rope lengths of 5, 7, and 10 m

4.2 Initial Angle Variation

As shown in Fig. 6, with a steel wire rope length of 10 m and initial angles of 5°, 10°, and 15°, the response curves of θ_3 are plotted. The results show that as the initial angle increases, the swing amplitude grows and the period shortens. Reducing the initial angle helps control swing and improves stability and safety.

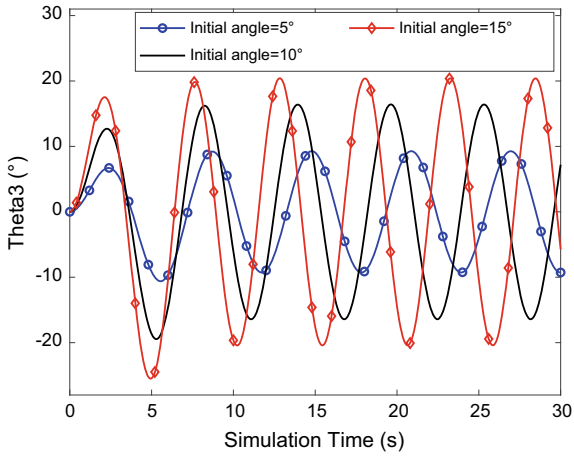


Fig. 6 θ_3 variation curves at initial angles of 5°, 10°, and 15°

5 Conclusion

This paper establishes a dynamic model of the dual-crane cooperative lifting system based on the Lagrangian equation, and verifies and simulates the model using MATLAB and Adams software. The results show that the simplified dynamic model can effectively describe the system's dynamic characteristics, providing a theoretical foundation for control strategy design. By analyzing the impact of steel wire rope length and initial angle on the payload's swing, it is found that a longer steel wire rope increases the swing period and amplitude, while a larger initial angle increases the swing frequency and amplitude.

Acknowledgements This work is supported by National Natural Science Foundation of China under Grant (No.52275088) and the Fundamental Research Funds for the Central Universities under Grant (No.DUT22LAB506,No.DUT22LAB507).

References

1. Huang JW, Wang YK, Ni LL (2024) Design of cooperative lifting and anti-swing control for dual drones. Comput Integr Manuf Syst 1–18
2. Zhang P, Fang YC, Liang X et al (2019) Modeling and nonlinear coordination control of a dual drones lifting system. In: Technical committee on control theory, Chinese association of automation, pp 809–814
3. Zi B, Zhou B, Qian S (2017) Dynamic modeling and analysis of dual automobile cranes flexible cable parallel equipment under luffing motion. J Mech Eng 53(07):55–61
4. Liu ZQ, Fu Y, Sun N et al (2022) Collaborative anti-swing hoisting control for dual rotary cranes with motion constraints. IEEE Trans Industr Inf 18(9):1
5. Vaughan J, Yoo J, Knight N et al (2013) Multi-input shaping control for multi-hoist cranes. IEEE, pp 3449–3454
6. Huang J, Zhu K (2021) Dynamics and control of three-dimensional dual cranes transporting a bulky payload. In: Proceedings of the institution of mechanical engineers. Part C, J Mech Eng Sci 235(11):1956–1965
7. Zhao XS, Huang J (2019) Distributed-mass payload dynamics and control of dual cranes undergoing planar motions. Mech Syst Signal Process 126:636–648
8. Lu B, Fang Y, Sun N (2020) Adaptive output-feedback control for dual overhead crane system with enhanced anti-swing performance. IEEE Trans Control Syst Technol 28(6):2235–2248
9. Qian YZ, Hu D, Chen YZ et al (2023) Programming-based optimal learning sliding mode control for cooperative dual ship-mounted cranes against unmatched external disturbances. IEEE Trans Autom Sci Eng 20(2):969–980
10. Hu D, Qian YZ, Fang YC et al (2021) Modeling and nonlinear energy-based anti-swing control of underactuated dual ship-mounted crane systems. Nonlinear Dyn 106(1):323–338

Open Access This chapter is licensed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License (http://creativecommons.org/licenses/by-nc-nd/4.0/), which permits any noncommercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if you modified the licensed material. You do not have permission under this license to share adapted material derived from this chapter or parts of it.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.



Study on the Lean Analysis of Low Stretch for Automobile Sheet Metal



Jinwu Wang, Binhan Xu, Cancan Wang, and Kangkang Yan

Abstract In order to emerge the surface low stretch of auto body panel, we study the mechanism, dent test methods and influencing factors of low stretch on the auto body panel. As the traditional sheet metal forming CAE analysis software has not the function of the specific low stretch surface defect analysis at present, we propose low stretch surface defects diagnostic methods and solutions, and develop a special defect analysis software module. Through analyzing the drawing processes on the auto outer door panel, we diagnose and resolve low stretch defect on the auto body panel, and prove that the method of analysis and software module are effective. And then we introduced some methods to optimize the adjustment amount of metal flow and the stamping dies for low stretch.

Keywords Automobile sheet metal · Low stretch · Lean analysis · Formability index

1 Introduction

With the increasing degree of automotive lightweight, the thickness of the automotive sheet metal is gradually decreased, especially for some large sweep surface parts in large size, simple and flat shape, small drawing depth such as hood outer panels,

J. Wang (✉) · B. Xu · C. Wang
School of Mechatronic Engineering, Changchun University of Technology, Changchun, Jilin Province, China
e-mail: wangjinwu@ccut.edu.cn

B. Xu
e-mail: xubinhan@ccut.edu.cn

C. Wang
e-mail: wangcancan@ccut.edu.cn

K. Yan
China Shanxi Automobile Group Co., Ltd R&D Center, Xi'an, Shanxi Province, China

© The Author(s) 2026

H. Saman et al. (eds.), *The 9th International Conference on Advances in Construction Machinery and Vehicle Engineering*, Lecture Notes in Mechanical Engineering,
https://doi.org/10.1007/978-981-96-5527-4_50

roof, auto outer door panel outer panel, car door panels and so on. Due to various process conditions, forming can not produce enough plastic deformation, hardening of stamping is not enough, so the product can not meet the requirements of dent resistance, i.e. the existence of stamping defects in low stretch surface. Although the existing sheet metal forming CAE analysis software can simulate the automobile panel drawing process and calculate the strain distribution after forming, but for the users who has not enough CAE analysis and actual production experience, they still can not complete the quantitative analysis of the low stretch surface defects. This leads to the defects to retain until the production stage, results the delay of the production schedule, and increases cost of production. Therefore, the study on the lean diagnosis and resolution technology for the low stretch on automobile sheet metal has important and practical significance.

2 The Mechanism of Low Stretch and Test of Dent Resistance Ability

2.1 The Mechanism of Low Stretch

Strain magnitude affects usage performance on large, flat/sweep areas. The usage performance of large, flat/sweep areas of stamped products includes stiffness, dent resistance, and oil canning load. From the viewpoint of mechanical structure, the stiffness and dent resistance are related to the yield strength; while the oil canning load is related to both the yield strength and first type of residual stress. When bake-hardenable sheet metals are used, the critical strain value is also required to further increase the yield strength during painting processes. If deformation in the large, flat/sweep areas is less than the critical strain value, they can not be hardened enough to satisfy the dent resistance requirement [1–3].

2.2 Test of Dent Resistance Ability

Dent resistance ability of automotive panel of is a sheet metal forming, after the load against external shocks (like flying stone, gravel and other heavy impact, etc.) dent in the ability of the plastic surface. As the automobile parts have some dent resistance requirements, different parts of the dent resistance tests to evaluate through the dent, the test principle is shown in Fig. 1. After the forming of automotive panels fixed on the test bench, with spherical punch in the vertical direction to exert pressure on the panels, measuring records generated by the plastic panels on the dent depth and applied load size, Fig. 2 is a hood outside the plate dent test. Dent performance is often part of the evaluation, the other is a certain depth of the dent the size of

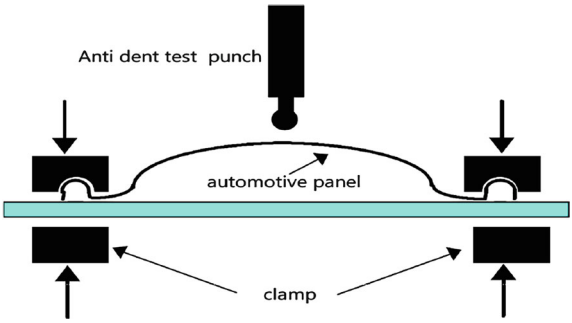


Fig. 1 Dent test

the required external load, after an evaluation method is usually given dent depth 0.1 mm, two test methods were dent test curve shown in Figs. 3 and 4 [4].



Fig. 2 Dent test of hood outer panel

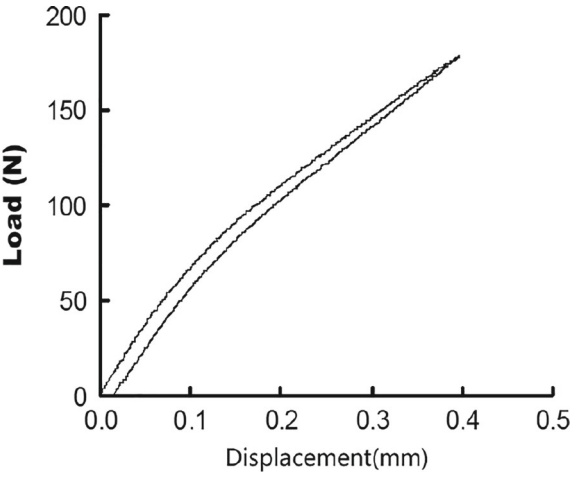


Fig. 3 Dent test curve of fixed load

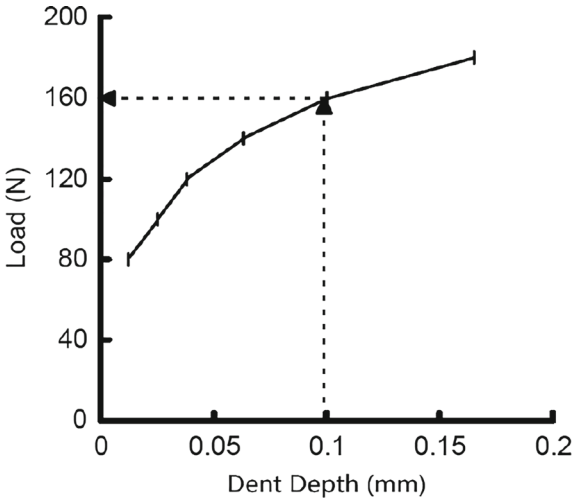


Fig. 4 Dent test curve of fixed indentation depth

3 Lean Analysis of Low Stretch Surface Defects

3.1 The Process of Lean Analysis

Lean analysis of low stretch surface defects is in the deformation zone, combines with field analysis technology to analyze the deformation process qualitatively and quantitatively. Complete the accurate diagnosis of the low stretch surface defects according to the given failure criterion of forming analysis. And then in accordance with principles of metal flow, control the amount of size and mobility trends by regulating the flow of metal stamping variables and reach purpose of eliminating defects. The process of the lean analysis for the low stretch surface defects is completed by six steps, and achieves the function such as the introduction of finite element data, the establishment of formability map and confirming the deformation zone, deformation analysis, shape analysis, defect solutions and output analysis.

3.2 The Diagnosis Method of Low Stretch Surface Defects

Calculate formability index: Search and determine the minimum area to find out the deformation field unit to represent the entire state of deformation. In this paper, we define the equivalent strain of the most small deformation element as the formability index of the lean analysis for low stretch surface defects. Judge the formability state: In the process of the lean analysis for low stretch surface defects, the system automatically judge the formability state according to the criterion of judging defect

Table 1 Mechanical performance parameters of the auto outer door panel material

Yield stress σ_s/MPa	Elastic modulus E/GPa	Poisson's ratio μ	Hardening exponent n	Hardening exponent K/MPa	Anisotropy coefficient r
129.5	198.5	0.33	0.25	521.5	1.3

and the value of the safety factor in Table 1 to achieve the diagnosis of low stretch surface defects on the automobile sheet metal.

3.3 Solution of the Low Stretch Surface Defects

If the formability state is unsafe, the stamping produces low stretch surface defects, and the system automatically gives the solution of the defect. The low stretch surface defects is caused by insufficient deformation, therefore, it needs to increase the deformation in the defect part through the adjusting the amount of metal flow from the perspective of metal flow to achieve the purpose of eliminating defects [5].

The adjusting amount of metal flow takes place in the deformation field and the deformation zone. There are three kinds of adjustment direction: along the direction of principal strain, along the direction of second strain, along the direction of linear strain paths. The adjusting amount of strain takes place in the range between the minimum amount and maximum amount of adjustment. The minimum amount of adjustment is the up limit of the warning area, the maximum amount of adjustment is the up limit of the 5% of the deformation zone [6].

4 Lean Analysis Software Module and Its Application of Low Stretch Surface Defects

4.1 Lean Analysis Software Module of Low Stretch Surface Defects

Develop the CAE analysis software of low stretch surface defects for automobile sheet metal according to the lean analysis process. Software module can achieve the lean diagnosis for low stretch surface defects and give defect solutions automatically according to diagnosis. Simulate the drawing process of the auto outer door panel by finite element simulation, and solve the low stretch surface defects of the simulation results. At the same time, we simulate the forming process of the same parts with a finite element simulation in the same process conditions applying DYNAFORM software, and compare the results of the DYNAFORM and the module we developed. The result shows that DYNAFORM can not accurately indicate low stretch surface

defects since there is no special analysis function of the low stretch surface defects in DYNAFORM, and the methods and solutions we proposed can effectively indicate and solve low stretch surface defects on the automobile sheet metal [7, 8].

4.2 Engineering Applications

Simulate the drawing process of the auto outer door panel by finite element simulation, and solve the low stretch surface defects of the simulation results. The methods and solutions we proposed can effectively indicate and solve low stretch surface defects on the automobile sheet metal. Simulate and calculate the drawing process of the auto outer door panel with DYNAFORM, use the module we developed to further diagnose the simulation results. The thickness of the sheet metal is 0.9 mm, and the material of the sheet metal is cold-rolled deep drawing low carbon steel BUFD_CR4, the mechanical performance parameters of the sheet metal are in Table 1. The finite element model of the auto outer door panel is shown in Figs. 5 and 6.

First, analyze the formed auto outer door panel using traditional post-processing mode, both the module we developed and DYNAFORM can not judge the low stretch surface defects of the auto outer door panel from directly from thickness, strain distribution and FLD, as shown in Figs. 7 and 8.

Analyze the simulation results of the low stretch surface defects using the module we developed, we find that the center of the auto outer door panel is the area with low stretch surface defects. The minimum deformation is the 24,532-th element, and the equivalent strain is 0.65% which is far below the minimum deformation 2.25% of steel required [9, 10]. Therefore, diagnosis results determine this region occur the

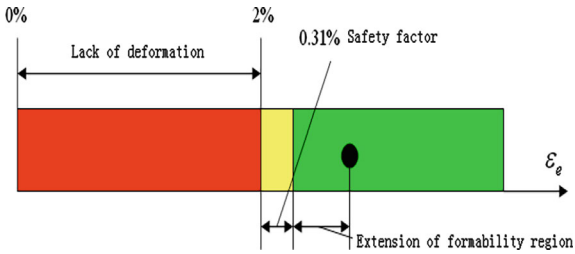


Fig. 5 Diagnosis of low stretch surface defects

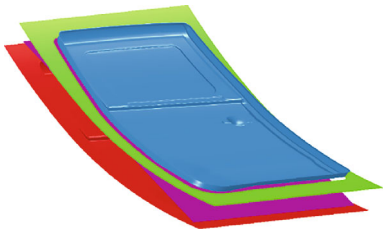


Fig. 6 Finite element model of the drawing process for auto outer door panel

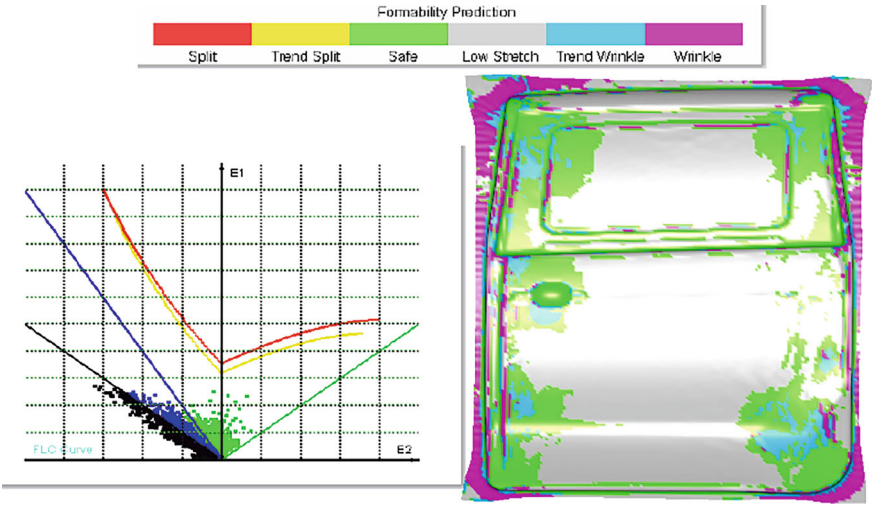


Fig. 7 FLD of the auto outer door panel in the CAE module

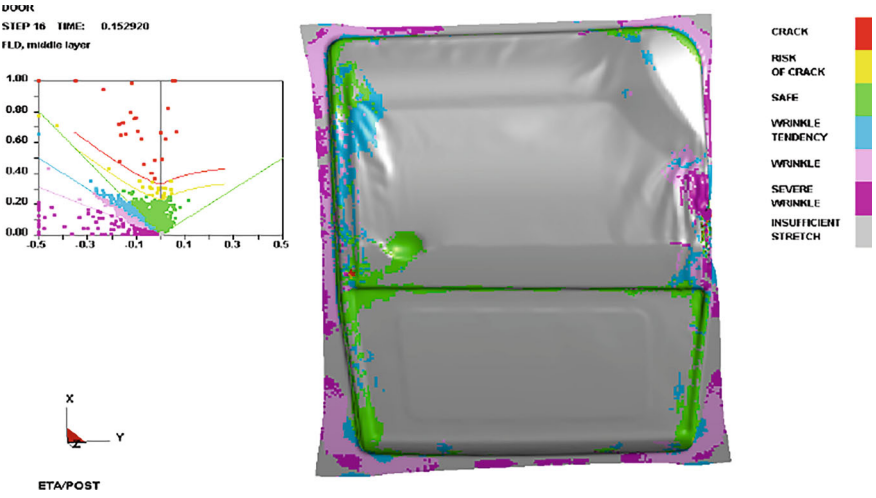


Fig. 8 FLD of the auto outer door panel in DYNAFORM

low stretch surface defects which is consistent with the actual dent. In order to solve low stretch surface defects, the module calculate the minimum strain adjustment amount is 1.65%, and the maximum strain adjustment amount is 4.25% as shown in Fig. 9. The length of this area along linear strain path direction is 472 mm. So we can calculate the range of metal flow adjustment is 7.79–20.1 mm. This means that we need to adjust the blank holder force or drawbead to make metal deformation add

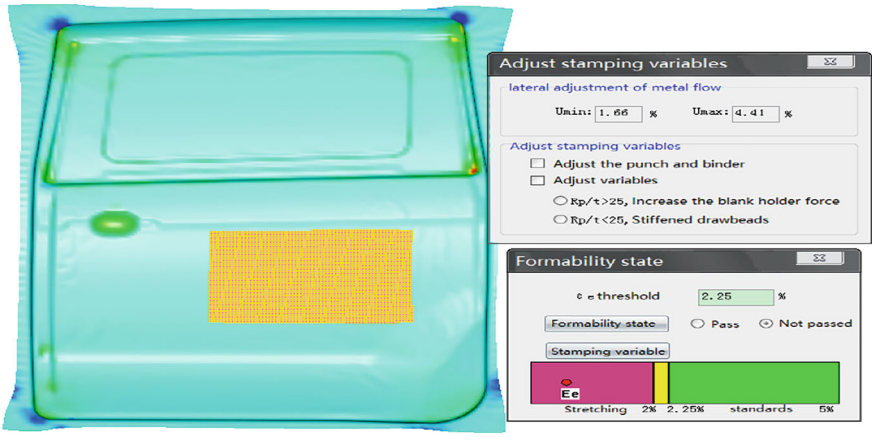


Fig. 9 Diagnosis and metal flow adjustment amount of the low stretch surface defects for the auto outer door panel

7.79–20.1 mm along linear strain path direction, or makes the metal into the mold cavity to reduce the 8–21.3 mm.

5 Conclusions

During researching low stretch surface defects of automobile sheet metal, we proposed the method based on formability and calculation methods of both measure index and severity of low stretch surface defects for the shortcoming of traditional program. According to the above mentioned programs within the low stretch surface defects diagnosis, we developed a CAE module base on COM component technology, and solved the problems which the traditional CAE software can not accurately predict and solve low stretch surface defects. Simulate a drawing process of auto outer door panel by applying the developed CAE module which shows the method we proposed is accurate, and then we introduced some method to optimize the adjustment amount of metal flow and the stamping dies for low stretch surface defects. Meanwhile, it gives suggestion of metal flowing adjustment for the corresponding defects, which pave the way for improving equipments for experimental die and achieving once stamping without defects.

Acknowledgements This research work was supported by the Science and Technology Development Project of Jilin Province (No.20240101089JC).

References

1. Xu Y (2004) Modern formability measurement and applications. Cincinnati, Henser Garden Publications Inc. LNCS Homepage. http://www.springer.com/lncs. Last accessed 2016/11/21
2. Xu Y (2004) Universal formability technology and applications. J Mater Process Technol 151(1–3):119–125
3. Xu Y (2006) Modern formability: measurement, analysis and applications. Hanser Gardner Publications Cincinnati, Cincinnati
4. Zhou B, Liu B, Zhang S et al (2021) Microstructure evolution of recycled 7075 aluminum alloy and its mechanical and corrosion properties. J Alloy Compd 879:160407
5. Mandatsy Moungomo JB, Kibouka GR, Nganga-Kouya D (2021) Influence of heat treatments, parameters and machining conditions on machinability in the turning of recycled 6061 aluminum alloy. J Manuf Mater Process 5(2):37
6. Wu Y, Wang H, Ban C (2021) Effect of Fe content on the microstructure and properties of hot-extruded 6061 Aluminum Alloy. J Phys Conf Ser 1986(1):012011 (IOP Publishing)
7. Al-Helal K, Lazaro-Nebreda J, Patel JB et al (2021) High-shear de-gassing and de-ironing of an aluminum casting alloy made directly from aluminum end-of-life vehicle scrap. Recycling 6(4):66
8. Maat N, Nor MKM, Ho CS (2021) Characterisation of mechanical properties, damage progression and fracture modes of recycled aluminium alloys AA6061 reinforced alumina oxide. Int J Integr Eng 13(7):215–225
9. Raabe D, Ponge D, Uggowitzer PJ, Roscher M, Paolantonio M, Liu C, Antrekowitsch H, Kozeschnik E, Seidmann D, Gault B, De Geuser F (2022) Making sustainable aluminum by recycling scrap: the science of "dirty" alloys. Progr Mater Sci 128:10947
10. Ho CS, Mohd Nor MK (2021) Tensile behaviour and damage characteristic of recycled aluminium alloys AA6061 undergoing finite strain deformation. Proc Inst Mech Eng C J Mech Eng Sci 235(12):2276–2284

Open Access This chapter is licensed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License (http://creativecommons.org/licenses/by-nc-nd/4.0/), which permits any noncommercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if you modified the licensed material. You do not have permission under this license to share adapted material derived from this chapter or parts of it.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.



A Ground Segmentation Algorithm for Slopes



Zeyu Liu, Dawei Li, and Dawei Jiang

Abstract When autonomous vehicles are driving in complex environments, the presence of uneven ground and obstacles can impact the safety of the vehicle. Traditional obstacle detection algorithms overlook the impact of sloped terrain on vehicle path planning. In environmental perception, determining whether a slope is passable has become a hot and challenging issue. This paper proposes a 3D LiDAR-based ground segmentation algorithm. This method relies on data obtained from experimental equipment for terrain information collection systems ahead of the vehicle, enabling it to distinguish between ground points and non-ground points in steep terrain. Experiments show that this method achieves good results in both static and dynamic slope detection.

Keywords Obstacle detection · Slope terrain · Environmental perception style

1 Introduction

1.1 A Subsection Sample

At present, the sensors often used in environmental sensing technology are camera and lidar [1]. Due to the relatively low price of cameras, there are many people who do cameras and algorithms, and the research is relatively in-depth, but its shortcomings are very obvious, it is unable to obtain accurate three-dimensional information, and it is also vulnerable to the interference of light environment [2]. Lidar can well obtain the three-dimensional characteristics and accurate distance information of surrounding objects, and overcome the disadvantage that sensors such as cameras are easily affected by the environment such as light. At the same time, lidar has the advantages of small volume, fast speed measurement, high reliability and accuracy

Z. Liu (✉) · D. Li · D. Jiang
Beijing Institute of Mechanical Equipment, Beijing, China
e-mail: intelligentxh@gmail.com

© The Author(s) 2026

H. Saman et al. (eds.), *The 9th International Conference on Advances in Construction Machinery and Vehicle Engineering*, Lecture Notes in Mechanical Engineering,
https://doi.org/10.1007/978-981-96-5527-4_51

of information acquisition, so lidar is widely used in intelligent vehicle environment sensing system. At present, lidar obstacle detection is mainly based on deep learning and clustering. Although the method based on deep learning has high detection accuracy, the training cycle is long and the training data is difficult to obtain, so the algorithm design is difficult; Clustering method is widely used in detection tasks because of its strong stability and real-time. Scholars at home and abroad have done a lot of research on lidar clustering detection algorithms, which can be divided into feature-based and ground projection based.

M. HIMMELSBACH proposed a method to distinguish the road surface and detect obstacles by grid division and then linear fitting in each ray direction [3]. This method has the advantages of short time-consuming and good detection effect in indoor environment, but its disadvantage is that the detection effect is poor in outdoor complex road environment. Ye Gang et al. Projected the three-dimensional LIDAR point cloud data onto the two-dimensional grid map, calculated the height difference of each grid, obtained the information occupied by the grid, and then extracted the obstacles by using the clustering method of connected area analysis [4]. This method is simple and easy to reproduce, but its accuracy is insufficient. When a dynamic object appears, each grid passed by the dynamic object will be regarded as an obstacle. Kamme proposed a ground detection algorithm based on the maximum height difference in grid point cloud [5]. If the maximum height difference in the grid is greater than a threshold, it is marked as a non ground point. Because this method only calculates the height difference between point clouds in each grid, it is efficient, but it is prone to under segmentation, and can not well separate ground points and non ground points. 3D convolution neural network is directly used for obstacle detection. Although 3D point cloud can be directly used for obstacle information extraction, because the point cloud data obtained by lidar is sparse, 3D convolution neural network generally carries out convolution neural network in three dimensions in order to obtain better effect. Although this method can achieve good results, the calculation cost is too high and the real-time performance is very poor.

Generally speaking, the existing ground segmentation algorithms have good segmentation effect in indoor environment and poor segmentation effect in outdoor complex terrain, ignoring the influence of slope terrain on ground segmentation; In addition, the existing ground segmentation algorithms have poor real-time performance and can not complete the ground segmentation task well in low-density point clouds; The obstacle detection algorithm based on deep learning is not suitable for complex field environment.

In light of the aforementioned issues, this paper proposes a ground segmentation method for slopes, which improves segmentation accuracy and also exhibits good performance on sloped road surfaces.

2 Ground Segmentation Algorithm for Slopes

2.1 Algorithm Theory

In the three-dimensional point cloud data of lidar, the point with the lowest height is generally the ground point by default. However, in the real road, especially in some mountainous cities, the slope is everywhere, and a single plane model is not applicable. The algorithm in this paper is to fit the plane in each region to deal with the road with steep slopes. There are two methods to fit the plane: the first is random uniform sampling, and the second is singular value decomposition (SVD). Compared with random uniform sampling, singular value decomposition is more suitable. Because the LIDAR point cloud is sparse, the singular value decomposition can also fit the plane well when the number of point clouds is small, as shown in Fig. 1.

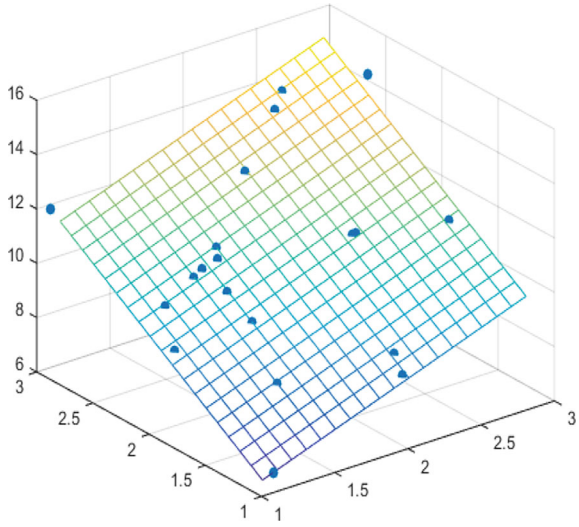
It is known that the lidar data is a three-dimensional point cloud, and a large number of ground points fit the plane equation:

$$ax + by + cz = d \tag{1}$$

In order to conveniently obtain the distance from the plane to the origin, it is necessary to add constraints:

$$a^2 + b^2 + c^2 = 1 \tag{2}$$

Fig. 1 Fitting effect drawing



The center of gravity coordinates of the three-dimensional point cloud are obtained, and it is obvious that the plane must pass through the center of gravity coordinates, that is

$$a\bar{x} + b\bar{y} + c\bar{z} = d \tag{3}$$

Subtract (3) from (1) to get

$$a(x_i - \bar{x}) + b(y_i - \bar{y}) + c(z_i - \bar{z}) = 0 \tag{4}$$

Written in matrix form:

$$\begin{bmatrix} x_1 - \bar{x}\ y_1 - \bar{y}\ z_1 - \bar{z} \\ x_2 - \bar{x}\ y_2 - \bar{y}\ z_2 - \bar{z} \\ x_3 - \bar{x}\ y_3 - \bar{y}\ z_3 - \bar{z} \end{bmatrix} \begin{bmatrix} a \\ b \\ c \end{bmatrix} = 0 \tag{5}$$

Let matrix A be.

$$\mathbf{A} = \begin{bmatrix} x_1 - \bar{x}\ y_1 - \bar{y}\ z_1 - \bar{z} \\ x_2 - \bar{x}\ y_2 - \bar{y}\ z_2 - \bar{z} \\ x_3 - \bar{x}\ y_3 - \bar{y}\ z_3 - \bar{z} \end{bmatrix} \tag{6}$$

Let column vector X be

$$\mathbf{X} = \begin{bmatrix} a \\ b \\ c \end{bmatrix} \tag{7}$$

Then,

$$\mathbf{AX} = 0 \tag{8}$$

In an ideal case, (8) is equivalent to (4), but there is always a radar point cloud that is not on the plane during fitting, and the sum of the minimum distances from the point cloud to the plane is required to be as small as possible, so the fitting objective function

$$\min \|\mathbf{AX}\| \tag{9}$$

According to (2)

$$\|\mathbf{X}\| = 1 \tag{10}$$

Singular value decomposition is performed on A to obtain:

$$\mathbf{A} = \mathbf{UDV}^{\mathrm{T}} \tag{11}$$

where D is a diagonal matrix, U and V are unitary matrices. According to the properties of unitary matrix,

$$\|\mathbf{AX}\| = \left\|\mathbf{UDV}^{\mathrm{T}}\mathbf{X}\right\| = \left\|\mathbf{DV}^{\mathrm{T}}\mathbf{X}\right\| \tag{12}$$

where $V^{\mathrm{T}}X$ is a column matrix, and because the diagonal element of D is singular value, the last diagonal element is the minimum singular value if and only if

$$\mathbf{V}^{\mathrm{T}}\mathbf{X} = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \tag{13}$$

then $\|\mathbf{AX}\|$ be a minimum, and

$$\mathbf{X} = \begin{bmatrix} a \\ b \\ c \end{bmatrix} = \mathbf{V}\begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \tag{14}$$

where (a, b, c) is the normal vector of the plane $\mathbf{n}$.

2.2 Algorithm Procedures

First, in order to explain the process in detail, several important parameters will be used next.

S is a seed point set, which is a matrix about three-dimensional point cloud.

K is the number of times of optimal iterative fitting. The higher the number, the higher the accuracy.

NLPR is used to select the number of lowest point representative (LPR).

Thseed is the threshold used to judge whether the point cloud belongs to the seed point. When the height Z of the point cloud is less than the height of, the point is added to the seed point set s.

Thdist plane distance threshold: calculate the distance from each point in the point cloud to the orthogonal projection of the fitted plane. If the distance from this point to the plane is less than the threshold, this point belongs to the ground, otherwise it belongs to non ground points.

The distance from point h of the model to the plane.

H is the LPR adaptive distance threshold. Due to the complex road conditions in the field environment, if the distance threshold is too large, some obstacles may be mistakenly classified as ground points, which will affect the effect of ground segmentation. On the contrary, if the distance threshold is too small, some roads with

small slope will be regarded as obstacles, which will also affect the effect of ground segmentation. Therefore, it is necessary to set the LPR adaptive distance threshold to avoid over segmentation or incomplete segmentation. For each sub region, in order to fully combine the characteristics of terrain, the LPR adaptive distance threshold can be expressed as:

$$H = \frac{\sum_{i=1}^{N} \sigma_i}{N} \tag{15}$$

Algorithm flow chart shown in Fig. 2. The specific steps are as follows:

Step1: cut and filter the obtained LIDAR point cloud data first. The higher area will have a certain impact on the ground segmentation, so cut the higher radar point cloud data; In order to avoid the influence of the vehicle itself on the radar reflection, the point cloud around the vehicle is cut and down sampled.

Step 2: divide the point cloud data according to the forward direction of the vehicle and a certain distance, divide the point cloud data into multiple groups, and sort each group of point cloud data according to the Z direction, i.e. height.

Step3 in the input point cloud group (which has been sorted according to the Z direction, i.e. height), select the points with small height as the seed point set to describe the initial plane model of ground segmentation.

Step 4 determines a plane model and calculates the distance of the orthogonal projection from the point cloud to the plane. If it is less than the height threshold, it is a ground point, otherwise it is a non ground point. Put the newly added ground points into the seed point set and continue iterative optimization. Where the plane model is

$$ax + by + cz = d \tag{16}$$

and its normal vector is

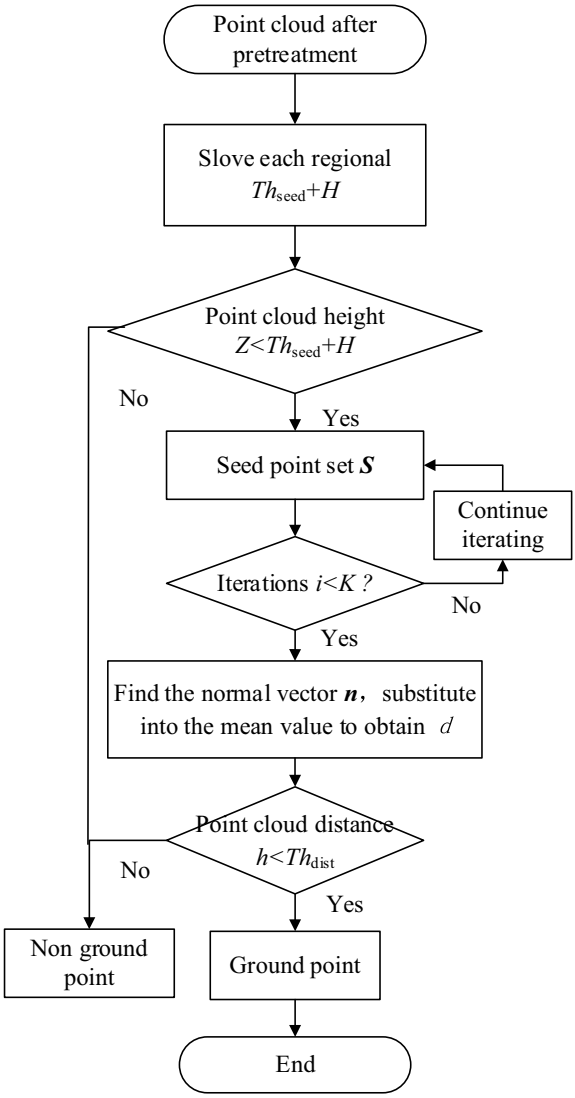
$$\mathbf{n} = [a, b, c]^T \tag{17}$$

3 Experiment

This experiment first verifies the accuracy of the algorithm on the structured road. As shown in Fig. 3, the actual scene is a flat experimental site of the school. The surrounding environment is with houses, trees, walls and iron fences, and the obstacles are bilateral bridges.

As shown in Fig. 4, it is the point cloud segmentation diagram of experimental scene 1, in which the red represents ground points and the black represents non ground points. The segmentation effect is obvious, which clearly distinguishes between walls

Fig. 2 Algorithm flow chart



and bilateral bridges. Although the double-sided bridge is lower than the wall and other obstacles, its characteristics are very obvious. This experiment proves that this method and ground fitting segmentation method are also suitable for structured pavement. This experiment proves that this method and ground fitting segmentation method are also suitable for structured pavement.

Next, this experiment verifies the accuracy of the algorithm on the unstructured pavement of the slope. In the actual scene II, as shown in Fig. 5, the slope ground is surrounded by grass, trees and iron fences.



Fig. 3 Experiment scenario 1



Fig. 4 Point cloud segmentation diagram

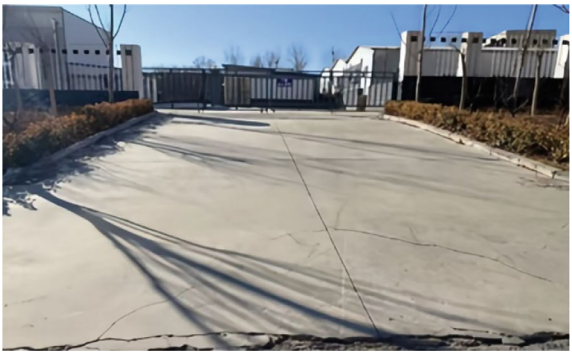


Fig. 5 Experiment scenario 2



Fig. 6 Comparative diagram of experimental results

The segmentation results of the two point clouds for the scene are shown in Fig. 6, where red represents ground points and black represents non-ground points. This experiment demonstrates that the method is applicable to unstructured pavements with slopes. The algorithm presented in this paper is more robust than traditional algorithms, capable of segmenting ground with slopes, and maintains excellent segmentation performance on both flat and sloped terrains. Additionally, the presence of dynamic obstacles does not affect the segmentation results.

4 Conclusion

This paper presents a slope ground segmentation algorithm based on 3D LiDAR. Initially, to mitigate the impact of vehicles on the experiment and focus on the region of interest, the point cloud data is trimmed and filtered. Subsequently, the voxel grid method is employed to downsample the point cloud, thereby reducing the volume of data that needs to be processed. Building upon this foundation, multi-region ground segmentation is carried out. Characterized by multiple iterations and robust performance, this method effectively minimizes the influence of outliers on the experimental results. Finally, through testing and comparative experiments conducted on both structured and unstructured road surfaces with slopes, it is verified that the proposed algorithm can accurately segment the ground under sloped conditions and remains unaffected by dynamic obstacles.

References

1. Ni T, Li W, Zhang H et al (2019) Pose prediction of autonomous full tracked vehicle based on 3D sensor. Sensors (Basel, Switzerland) 19(23)

2. Zermas D, Izzat I, Papanikolopoulos N (2017) Fast segmentation of 3D point clouds: a paradigm on LiDAR data for autonomous vehicle applications. In: 2017 IEEE International Conference on Robotics and Automation (ICRA), IEEE. https://doi.org/10.1109/ICRA.2017.7989591
3. Chen J, Wang Y, Dai K et al (2019) Fast segmentation of 3-D point clouds based on ground plane state tracking. IEEE. https://doi.org/10.1109/CVCI47823.2019.8951690
4. Chen T-H, Chang TS (2022) RangeSeg: range-aware real time segmentation of 3D LiDAR point clouds. IEEE Trans Intell Veh 7(1):93–101
5. Steinke N, Goehring D, Rojas R (2024) GroundGrid: LiDAR point cloud ground segmentation and terrain estimation. IEEE Robot Autom Lett 9(1):420–426. https://doi.org/10.1109/LRA.2023.3333233

Open Access This chapter is licensed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License (http://creativecommons.org/licenses/by-nc-nd/4.0/), which permits any noncommercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if you modified the licensed material. You do not have permission under this license to share adapted material derived from this chapter or parts of it.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.



Sparse Feature-Oriented 3D SLAM Mapping with Semantic Fusion



Zhangjie Ding, Tao Li, Linyuan Xie, and Yongming Bian

Abstract In order to achieve the measurement of the 3D environment, a LiDAR SLAM 3D mapping algorithm for weak geometric road environments oriented to sparse features and fusing semantic features is built. The extraction of ground point cloud is optimized for the distortion-corrected point cloud data. The characteristics of columnar structures and roadside structures in sparse feature environments are analyzed, and the corresponding semantic feature extraction scheme is designed to implement the LiDAR odometry based on the above features. A certain detection strategy is proposed to construct loop closure constraints to avoid invalid and frequent loop closure detection. A global descriptor-based loop closure detection method is introduced for loop closure drift in large scale range. In the back-end of the SLAM algorithm, keyframes are matched with local maps, and the LiDAR, IMU, and loop closure constraints constructed above are added into the factor map to optimize the algorithm, and a SLAM 3D mapping algorithm is finally constructed for the fusion of semantic features in sparse feature environments.

Keywords Sparse feature · SLAM · Semantic fusion

1 Introduction

SLAM 3D mapping technique is an effective means to solve the difficulty of measuring spatial features in 3D scenes. Among them, odometry based on LiDAR geometric feature points is the key part of SLAM mapping algorithm, and the method can effectively construct odometry and achieve high accuracy results [1, 2]. However,

Z. Ding · L. Xie · Y. Bian (✉)
School of Mechanical Engineering, Tongji University, Shanghai, China
e-mail: ymbianmail@163.com

T. Li
CCCC Second Harbor Engineering Company Ltd, Wuhan, Hubei, China

CCCC Highway Bridge National Engineering Research Centre Ltd, Beijing, China

© The Author(s) 2026

H. Saman et al. (eds.), *The 9th International Conference on Advances in Construction Machinery and Vehicle Engineering*, Lecture Notes in Mechanical Engineering,
https://doi.org/10.1007/978-981-96-5527-4_52

in the weak geometric scenes with sparse features orientated in this paper, such as the environment of internal roads in industrial parks, due to the lack of large elevations in the space and the existence of objects only at lower heights and in the close range, the feature points extracted by this method will have a large number of point clouds located on objects such as weeds, bushes, uneven land, leaves, etc. or invalid point clouds, and this type of point cloud is a relatively unstable feature point cloud. It is not easy to be observed repeatedly during continuous motion, thus it is easy to cause wrong matching associations and reduce the accuracy of SLAM construction.

To address this problem, this paper constructs a 3D mapping algorithm by extracting the most common semantic features in roads. In a road scene with sparse features, there is often a lack of large planes in the space with medium or high heights, and thus feature points can only be concentrated in the lower height range. In such environments, columns such as trees and street lights, as well as road edges and ground features are very common and stable geometric features [3, 4]. The distribution of these three features is hierarchical; they are orderly and easily observed objects in the road, and the point clouds located on them are concentrated and have obvious structural information, and are less affected by the time of the day, season, weather, etc., and more stable [5]. The above features are conducive to improving the accuracy of point cloud alignment. Therefore, this paper constructs a fused semantic LiDAR odometry by extracting the above structural features [6, 7].

2 Fusion of Semantic Structures for LiDAR Odometry

2.1 Optimization of Ground Point Clouds Segmentation

A large number of candidate ground points are first obtained by the angle thresholding method. Since some roads are not ideal horizontal planes but slightly lower on both sides and slightly higher in the middle, the ground point cloud obtained in the first step is divided into two planes, the left side and the right side, according to the vehicle forward direction. Afterwards, the ground plane is fitted by introducing the RANSAC algorithm, and finally the ground point cloud is extracted. RANSAC fitting is performed separately for the left and right planes and distance thresholds are set to exclude points that do not satisfy the plane fitting as outlier points.

2.2 Semantic Structure Feature Extraction Scheme Design

In the environment of suburban roads, rural roads, or roads inside industrial parks, columnar objects are common semantic features with obvious structures in sparse environments, which mainly include columnar objects such as orderly arranged street lights, a large number of trees, and road signs. The LiDAR point cloud distributed on

these columnar objects is densely distributed in the longitudinal direction, which is more stable than the corner points on some surfaces, less susceptible to the influence of changes in the sensor's viewing angle, more advantageous in dynamic environments, and less prone to be lost or matched incorrectly, and also able to constrain the error in the horizontal direction of the laser odometry well, which improves the stability of the accurate odometry and the construction of the maps in sparse feature environments.

In road environments with sparse features, point clouds distributed in and around columns tend to have the following structural characteristics:

(1) Point clouds located at the edges of the bollards have a more significant distance difference from the point clouds on their outer edges.
(2) Point clouds distributed on the columns in the same harness are denser.
(3) Point clouds of the column tends to cluster upwards along the vertical ground direction in space.

The block diagram of columnar point clouds extraction process is shown in Fig. 1.

Road edges are another stable and common semantic feature structure in internal road environments in industrial parks. The extraction of road edge features means extracting the point clouds of point cloud line bundles as they cross the road edge. This type of point cloud located on the surface of the road boundary serves as a reference for the pavement area, which has the characteristics of continuity and stability, and is also able to improve the accuracy of map construction.

The point cloud located at the roadside will have some spatial characteristics different from the point cloud on the road surface:

(1) For the point clouds on the road boundary surface on the same beam, the height difference between their neighbors in the vertical direction is significantly higher than that of their neighbors on the road surface, and the angle between the neighbors at the upper and lower boundaries of the road edge will change.
(2) The point cloud on the road surface is vertically between the ground plane and the upper surface of the road.

The block diagram of roadside point clouds is shown in Fig. 2.

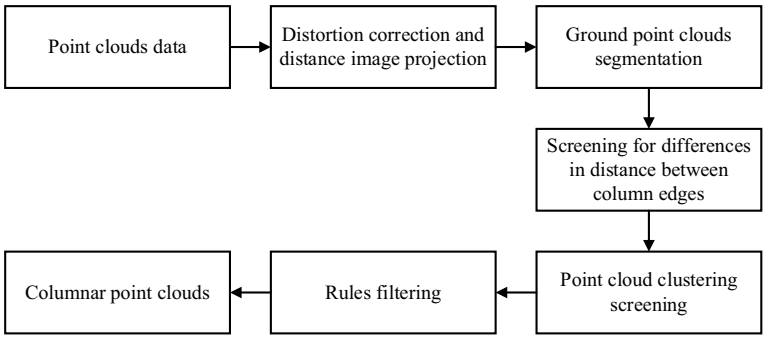


Fig. 1 Block diagram of columnar point clouds extraction process

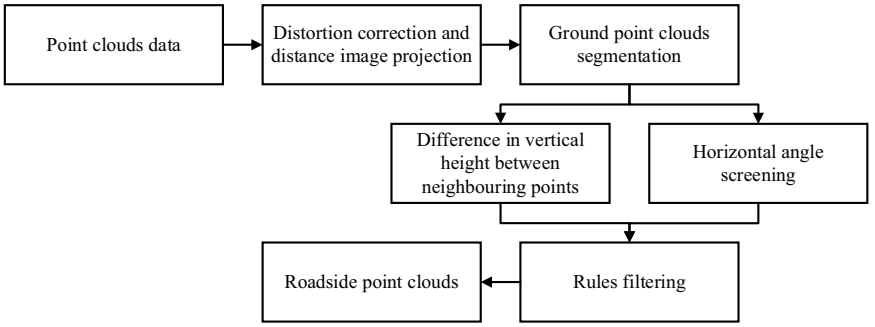


Fig. 2 Block diagram of roadside point clouds extraction process

2.3 Feature Matching Associations for Fusion Semantics

The above implements the extraction of three types of point clouds: ground, columnar and roadside semantic features in a road environment with sparse features. These features can be repeatedly observed in two consecutive frames of the point cloud during the continuous motion of the LiDAR. Therefore, the position transformation obtained by correlation matching of the same category of feature point clouds in two consecutive frames is the position change of the carrier, and then the laser odometer is constructed through the consecutive position change.

For the matching and correlation of the same type of feature data, the point to line and point to surface correlation methods are usually used. The above three types of features extracted in this paper can be expressed as line and surface structures. Linear structure presents the characteristic of extending along a particular direction, while faceted structure unfolds in two directions. In this paper, from the characteristics of the feature point cloud itself, the three types of features are directly associated and matched in different ways; the columnar and roadside point clouds are matched and associated with the linear structure, while the ground point cloud is more suitable for the faceted structure.

For the columnar features, the point cloud presents a short bundle of multiple dense point clouds with parallel upward distribution in space. A point p_a in the point cloud of the current frame is selected to find the nearest neighbor p_b corresponding to the previous frame, and the nearest point of this nearest neighbor in the adjacent line bundles to construct the point-to-line data association. The coordinates of the point under the LiDAR coordinate system are denoted as follows, respectively. The point-to-line association matching equation is denoted as:

$$d_e = \frac{\left|\left(X^L_{(k,p_a)} - X^L_{(k-1,p_b)}\right) \times \left(X^L_{(k,p_a)} - X^L_{(k-1,p_c)}\right)\right|}{\left|\left(X^L_{(k-1,p_b)} - X^L_{(k-1,p_c)}\right)\right|} \tag{1}$$

For roadside features, the distribution of the point cloud in space is a line structure stretching along the road, and the selection of the nearest points of the neighboring

line bundles can reflect the line structure of the roadside, and the same point to line data association is taken through the above process.

For the ground feature point cloud, the point p_i of the current frame is selected with the nearest neighbour p_j corresponding to the previous frame, and the two nearest points p_l and p_m on the adjacent line bundle to construct the point-to-face data association. The coordinates of points p_i, p_j, p_l, p_m under the LiDAR coordinate system are noted as: $X^L_{(k,p_i)}$, $X^L_{(k-1,p_j)}$, $X^L_{(k-1,p_l)}$, $X^L_{(k-1,p_m)}$. The point-to-face association matching equation is expressed as:

$$d_h = \frac{\left|\left(X^L_{(k,p_i)} - X^L_{(k-1,p_j)}\right) \times \left(X^L_{(k-1,p_j)} - X^L_{(k-1,p_l)}\right) \times \left(X^L_{(k-1,p_j)} - X^L_{(k-1,p_m)}\right)\right|}{\left|\left(X^L_{(k-1,p_j)} - X^L_{(k-1,p_l)}\right) \times \left(X^L_{(k-1,p_j)} - X^L_{(k-1,p_m)}\right)\right|} \tag{2}$$

During the point-to-face data association process, the selection of similar feature points that are not co-linear can help to avoid the degradation of the plane composed of three points into a straight line. The above point-to-line and point-to-face data association is illustrated in Fig. 3, with the green line representing the previous frame point cloud and the dark brown line representing the current frame point cloud.

The matching equations for the inter-frame features are obtained by the above process, and inter-frame six-degree-of-freedom bit-pose $[p_x, p_y, p_z, \theta_{roll}, \theta_{pitch}, \theta_{yaw}]$ transform needs to be optimally solved. Remember that the bit-pose transform from frame $k-1$ to frame k is T^k_{k-1}, then:

$$X_k = T^k_{k-1} X_{k-1} \tag{3}$$

The line-plane association matching constraint equations with respect to T^k_{k-1} are respectively:

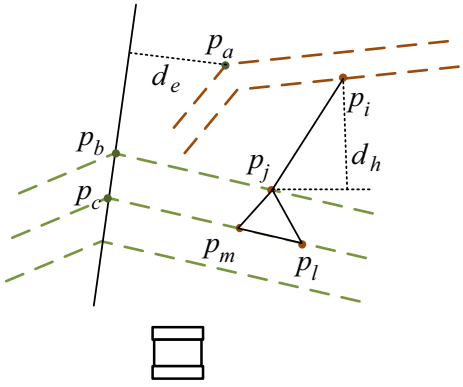


Fig. 3 Fusion of semantic features for matching associations

$$f_e\left(T_{k-1}^k\right) = d_e \tag{4}$$

$$f_h\left(T_{k-1}^k\right) = d_h \tag{5}$$

Uniformly recorded as:

$$f\left(T_{k-1}^k\right) = d \tag{6}$$

Each row in f represents the feature point cloud and d contains the corresponding distance relationship. The solution of the above equation is a nonlinear least squares optimisation problem. Firstly the Jacobi matrix J of the matrix f with respect to the variable T_{k-1}^k is computed:

$$J = \frac{\partial f}{\partial T_{k-1}^k} \tag{7}$$

After obtaining the minimum value of d, the position can be optimized in two iterative steps with the help of the following equation:

$$T_{k-1}^k \leftarrow T_{k-1}^k - \left(J^T J + \lambda \mathrm{diag}\left(J^T J\right)\right)^{-1} J^T d \tag{8}$$

where λ is the coefficient factor of the Levenberg–Marquardt method.

The two-step optimization to accelerate the solution of the T_{k-1}^k transform is performed by the Levenberg–Marquardt method, where the IMU can provide the initial value of the bit-pose for the above solution, and the Levenberg–Marquardt method interpolates prior to the Gaussian Newton and gradient descent methods, which is a confidence domain method. Compared to the Gaussian Newton method, the Levenberg–Marquardt method is more robust and in most cases converges quickly and iteratively even if the unfolding point is very far from the final minimum. The flow of the LM two-step optimization algorithm is schematically shown in Fig. 4. T_{k-1}^k contains six degrees of freedom poses, and for the ground feature points, the above three degrees of freedom poses are solved first by the first step of the LM algorithm because of better constraints on $\left[t_z, \theta_{roll}, \theta_{pitch}\right]$. While reducing the number of optimisation iterations, the LM algorithm provides initial values for the constraint solving of $\left[t_x, t_y, \theta_{yaw}\right]$ in the second step, thus speeding up the solving speed and finally obtaining a more accurate result of the position transformation [8, 9].

2.4 Loop Closure Detection Based on Scan Context

The Scan Context (SC) [10] based method transforms the 3D point cloud into a 2D image with rotational invariance, i.e., column vector height invariance, according to the idea of dimensionality reduction. As this paper is oriented to the outdoor scene with sparse features, the road environment is mostly a turning road and there are

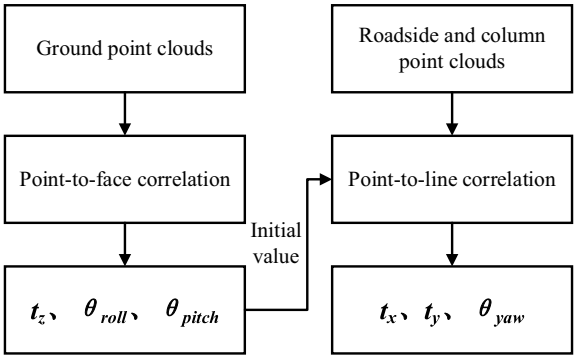


Fig. 4 Schematic flow of the LM two-step optimization algorithm

more trees on both sides of the road, which have more stable ground plane features and rich columnar features, and the land on both sides of the road is uneven, with large feature variability in the height direction. Therefore, this paper introduces the Scan Context method in loop detection to improve the computational efficiency and ensure sufficient accuracy.

3 Sparse Feature Environment Mapping Experiment

The Lidar used is RS-LiDAR-16, set the rotational speed of 600 rpm, then the point cloud scanning frequency of 10 Hz. 9-axis IMU is used, model CH110, the angular velocity and acceleration output frequency of IMU is set to 100 Hz, and the internal parameter calibration of IMU is required before use, which is achieved by Allan variance is set to 100 Hz, the internal parameter calibration of the IMU needs to be carried out before use, which is achieved by Allan variance, and the IMU in this paper has been calibrated in advance. The software platform for the experiments of graph building algorithms is Ubuntu 18.04-Melodic operating system.

The scenario chosen for the test is a small field multi-loop scenario, which is a satellite image of a driving school's subject 2 training field in Shanghai, which has a length of about 200 m in the horizontal direction and a width of about 60 m. The length of the experimental trajectory in Scenario 1 is about 900 m. It is found from the satellite image and the fieldwork that the test environment is characterized by both smooth planar features of the road as well as the road edge features on the two sides of the road, which This is conducive to the extraction of planar and semantic features.

The absolute trajectory error data for the ablation experiments for the small-field multi-loop scenario are shown in Table 1. Comparing the constructed trajectories without any semantic feature extraction, the absolute trajectory error is significantly larger than the result after extracting the semantic features in all indicators except the minimum value. When only column features or only roadside features are extracted, the average value of absolute trajectory error is about 0.21 m. When two types of

Table 1 Critical values of absolute trajectory errors

Selected features	Average value (m)	Minimum value (m)	Median (m)	Maximum value (m)	Standard deviation (m)	Root mean square deviation (m)
No extraction	0.266156	0.014323	0.240273	1.118884	0.131168	0.296722
Columnar features	0.209563	0.021944	0.205791	0.455042	0.086263	0.226623
Roadside features	0.208071	0.007440	0.199692	0.461292	0.087534	0.225734
Columnar and roadside features	0.183198	0.010224	0.176629	0.426834	0.077690	0.198990

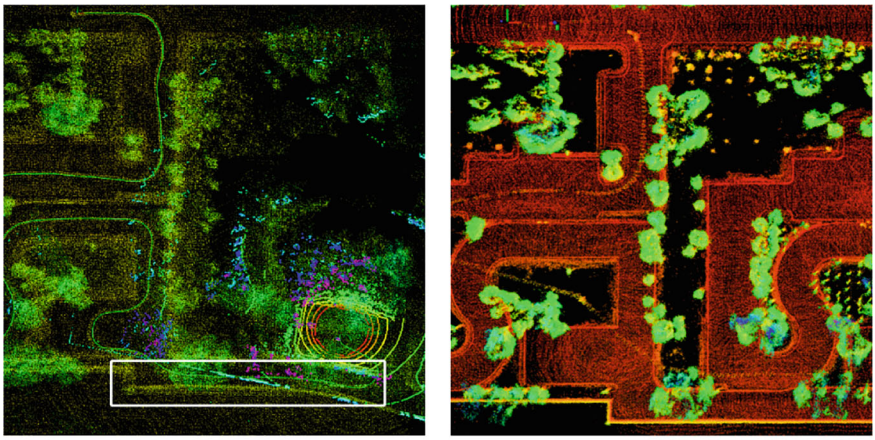
semantic features are extracted at the same time, the average value of the absolute trajectory error is about 0.18 m, and the rest of the indicators except the minimum value are the smallest values in the same category.

Comparison of the map building effect shown in Fig. 5a is the map building effect of the A-LOAM algorithm, due to the poor processing of A-LOAM on the rotation of the mobile carrier, the point cloud map building in the feature part of the fence wall produces a large re-shadowing, as shown by the white box in Fig. 5a, and the mobile carrier at this time is in the rotational movement of the larger angular velocity, so that the construction of the point cloud map as a whole also produces a larger angle of the rotation error Therefore, the map constructed by A-LOAM also has a large rotational error, which ultimately leads to the failure of the map construction. In addition, A-LOAM does not differentiate between ground and non-ground feature points, so the quality of the map is not high.

Figure 5b shows the local map building effect of this paper's algorithm, and it can be clearly seen that the point cloud building effect of the fence part is better, with small ghosting and good consistency. In addition, thanks to the fusion of IMU and LiDAR, the map does not drift when the mobile carrier rotates at a large angular velocity, and the distinction between different features in the environment, such as roadside, tree trunks, foliage, fences, etc., is more obvious, and the quality of the map is higher.

As shown in Table 2 is the comparison of absolute trajectory error data for multiple loop scenarios in small sites. The absolute trajectory errors built by the algorithm of this paper have been improved in terms of mean and median. Comparing with LeGO algorithm, it can be clearly seen that this paper can effectively reduce the trajectory error and improve the accuracy of map building by extracting semantic features.

Combined with Table 3, the relative trajectory error of this paper's algorithm in the small field multi-loop scenario performs better in terms of the maximum value than with the LeGO algorithm, and the rest is basically the same. This is mainly



(a)A-LOAM (b)Algorithm raised in this paper

Fig. 5 Comparison of the effects of mapping

Table 2 Comparison of absolute trajectory errors

Algorithm	Average value (m)	Minimum value (m)	Median (m)	Maximum value (m)	Standard deviation (m)	Root mean square deviation (m)
LeGO	0.256887	0.025299	0.224565	0.823636	0.136874	0.291076
Ours	0.183198	0.010224	0.176629	0.426834	0.077690	0.198990

due to the fact that the environment of the small field multi-loop scenario has only sparse features such as trees, road edges, and land to construct the odometer, so the variation of the error does not vary greatly.

As shown in Table 4 is a comparison of the different number of loop closure detections obtained and the absolute trajectory error of each data before and after the loop closure detection strategy is adopted. Before adopting the loop closure detection strategy, 106 loop closure detections were successfully detected and loop closure detection constraints were constructed. After adopting the loop closure detection strategy, 80 loop closure detections were successfully detected, and the number of

Table 3 Comparison of relative trajectory errors between LeGO and our algorithm

Algorithm	Average value (m)	Minimum value (m)	Median (m)	Maximum value (m)	Standard deviation (m)	Root mean square deviation (m)
LeGO	1.946215	0.031516	1.971228	3.732737	0.528276	2.016638
Ours	1.944769	0.006048	1.971706	1.971706	0.523781	2.014069

Table 4 Comparison of relative trajectory errors before and after adopting the strategy

Number of loop closure detections	Average value (m)	Minimum value (m)	Median (m)	Maximum value (m)	Standard deviation (m)	Root mean square deviation (m)
80	0.209581	0.009636	0.206118	0.490318	0.080638	0.224559
106	0.184940	0.018815	0.175789	0.780831	0.109177	0.214761

loop closure detections were significantly reduced, which saves the resource overhead. In the end, the reduction in the number of loopback detections is about 28% with a slight reduction in the accuracy of the build trajectory and a small impact, reducing the number of ineffective and frequent loop closure detections by a certain amount.

4 Conclusion

This paper presents research and experiments on mapping algorithms for road scenes with sparse features. The experimental scenario, i.e., a road scene with multiple loops in a small field, is introduced. Evaluation metrics are employed to analyze the mapping trajectory. Ablation experiments are conducted to validate the effectiveness of semantic features. Mapping experiments are performed in this scenario to assess the performance of the proposed algorithm, and comparisons with other algorithms demonstrate its superiority in improving mapping trajectory accuracy. Additional experiments verify the effectiveness of the loop closure strategy, which reduces the number of invalid detections without compromising accuracy. However, the robustness of LiDAR mapping algorithms is low in weak geometry environments, and fusion of more sensors, such as LiDAR-camera fusion for mapping tasks, can be considered in the future.

Acknowledgements Supported by the National Key Research and Development Program of China (Grant No. 2023YFF0613200).

References

1. Badue CF (2021) Self-driving cars: a survey. Expert Syst Appl 165:113816
2. Lu YF (2021) Real-time performance-focused on localization techniques for autonomous vehicle: a review. IEEE Trans Intell Transp Syst 23(7):6082–6100
3. Sandström EF (2023) Point-slam: dense neural point cloud-based slam. In: Proceedings of the IEEE/CVF international conference on computer vision. IEEE, Paris, pp 18387–18389
4. Liso LF (2024) Loopy-slam: dense neural slam with loop closures. In: Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. IEEE, Seattle, pp 20363–20373

5. Wang HF (2023) Co-slam: joint coordinate and sparse parametric encodings for neural real-time slam. In: Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. IEEE, Vancouver, pp 13293–13302
6. Wang ZF (2021) Pole-like objects mapping and long-term robot localization in dynamic urban scenarios. In: International conference on robotics and biomimetic. IEEE, Sanya, pp 998–1003
7. Chiang KWF (2020) Performance enhancement of INS/GNSS/refreshed-SLAM integration for acceptable lane-level navigation accuracy. IEEE Trans Veh Technol 69(3):2463–2476
8. Ran HF (2022) Surface representation for point clouds. In: Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. IEEE, New Orleans, pp 18942–18952
9. Boulch AF (2022) Poco: Point convolution for surface reconstruction. In: Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. IEEE, New Orleans, pp 6302–6314
10. Kim GF (2018) Scan context: egocentric spatial descriptor for place recognition within 3d point cloud map. In: 2018 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS). IEEE, Madrid, pp 4802–4809

Open Access This chapter is licensed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License (http://creativecommons.org/licenses/by-nc-nd/4.0/), which permits any noncommercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if you modified the licensed material. You do not have permission under this license to share adapted material derived from this chapter or parts of it.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.



Research on the Dynamics Model of Specialized Heavy-Duty Multi-axle Electric Drive Vehicles Based on SIMPACK



Weize Wang, Xuelei Wu, and Xia Zuo

Abstract Specialized heavy-duty multi-axle electric drive vehicles are commonly used in aerospace engineering to transport weapons and strategic materials, whose mechanical structure and dynamic performance emerging are always a hot spot in the field of aerospace and mechanical fields. At the same time, SIMPACK, a robust multi-body dynamics simulation software, is seldom utilized in the modeling of passenger vehicles and lacks established methodologies for modeling specialized heavy-duty multi-axle electric drive vehicles. Addressing this gap, this study investigates modeling approaches for a 8 × 8 heavy-duty vehicle, focusing on its unique structure and operational principles. Using the modular methodology, the vehicle was divided into distinct components—cab, frame, suspension, wheels, steering, and powertrain—to construct a comprehensive vehicle model in SIMPACK. Simulations under various operating conditions were performed, and the simulated data were validated against empirical test data to ensure model accuracy. Results indicate that the dynamic indicators of the simulated vehicle fall within an acceptable range. This study provides a reference framework for future dynamic research on specialized heavy-duty multi-axle electric drive vehicles.

Keywords Aerospace and mechanics · Specialized heavy-duty multi-axle electric drive vehicles · Multi-body dynamics modeling

1 Introduction

Specialized heavy-duty multi-axle electric drive vehicles are commonly used in aerospace engineering to transport weapons and strategic materials, offering strong transport capacity and high efficiency. This makes them a key focus in aerospace research, with their mechanical structure and dynamic performance emerging

W. Wang · X. Wu (✉) · X. Zuo
Beijing Institute of Space Launch Technology, China Academy of Launch Vehicle Technology, Beijing, China
e-mail: 15041405988@163.com

© The Author(s) 2026

H. Saman et al. (eds.), *The 9th International Conference on Advances in Construction Machinery and Vehicle Engineering*, Lecture Notes in Mechanical Engineering,
https://doi.org/10.1007/978-981-96-5527-4_53

as primary areas of study. The transport environments in aerospace engineering often involve large payloads and variable conditions, which makes real vehicle testing costly and time-consuming [1]. However, with advancements in simulation technology, virtual prototyping now provides a novel method for performance development and validation [2].

By using multi-body dynamics software to construct a vehicle simulation model, researchers can not only investigate the vehicle's dynamic characteristics but also evaluate and validate its performance under various operating conditions, thereby saving significant time and resources. Therefore, selecting an appropriate modeling approach and creating a model that accurately reflects real vehicle characteristics are critical for researching specialized heavy-duty multi-axle electric drive vehicles suited to aerospace applications.

In previous studies, Kim et al. [3] developed a 26-degree-of-freedom dynamics simulation platform, proposing a torque-coordination control algorithm for an eight-wheel hub motor electric drive vehicle. Fauroux et al. [4] modeled a 6X6 all-terrain electric vehicle, enhancing steering efficiency by adjusting the normal ground contact forces on the wheels. Domestically, by establishing a linear two-degree-of-freedom model of a multi-axle vehicle, Li et al. [5] performed the steady-state and transient analyses of the steering characteristics of a six-axle vehicle. Simulation results qualitatively show that the use of multi-axle steering technology can effectively improve the steering maneuverability of the vehicle. Cheng et al. [6] focused on a five-axle specialized vehicle, utilizing ADAMS to create a refined full-vehicle dynamics model, which included braking, suspension, and steering systems, to enhance longitudinal braking safety and vertical stability.

Previous studies either specialized in certain vehicle attributes or developed full-vehicle models with parameters tailored to specific vehicle types, limiting broader applicability. Therefore, research on dynamics models for specialized heavy-duty electric drive vehicles is crucial, along with the development of rapid, accurate, and efficient modeling methodologies.

This study addresses these issues by building a full-vehicle dynamics model for a 8X8 heavy-duty multi-axle electric drive vehicle using SIMPACK. The model integrates subsystems for the frame, suspension, wheels, steering, and powertrain, with a generalized parameter template to facilitate rapid modeling in future applications. Different driving conditions, such as stationary acceleration, hill climbing, and braking, are used to verify the model's applicability.

2 Modeling Methodology for Specialized Heavy-Duty Multi-Axle Electric Drive Vehicles

2.1 Modular Methodology

The modular methodology divides complex systems into multiple independent, interchangeable components based on their functions or structures, simplifying management, design, assembly, and operation. Each module can be simulated independently and reconfigured to meet diverse product requirements, reducing development costs [7]. In research on specialized heavy-duty multi-axle electric drive vehicles, modular modeling remains underutilized.

For multi-axle electric drive vehicles, modular methodology simplifies the complex process, eliminating redundant tasks. A "module" here represents a "functional sub-structure" aligned with the vehicle's structure and functions. The full vehicle is divided into five modules: frame, suspension, wheels, steering, and powertrain. After individually constructing each module's model, these are assembled into a complete vehicle model. By replacing modules or adjusting parameters, researchers can reconfigure the model for new vehicle types or validate specific structures using a parameterized approach.

2.2 Parameterized Modeling

In conventional vehicle design, component dimensions and positions are frequently adjusted, while topological relationships between components remain largely unchanged [8].

By consolidating module parameters into a unified parameter set, it becomes easier to modify individual parameters and propagate updates across the design, achieving rapid updates and assessments. This parameter set can also be imported into other models, streamlining the modification process and significantly reducing development time [9].

Following this approach, the parameters of each module in the multi-axle electric drive vehicle model were consolidated into a parameter set, facilitating the application of complete data sets in subsequent modeling.

3 Modeling Software for Specialized Heavy-Duty Multi-axle Electric Drive Vehicles

SIMPACK, a kinematic and dynamic simulation software for electromechanical systems, is predominantly applied to railway vehicle simulations, while its use in heavy-duty multi-axle electric drive vehicle modeling remains limited [10]. Utilizing SIMPACK in this domain provides foundational work for the software's broader applications.

SIMPACK's unique subsystem functions and parameter set configurations support both modular and parameterized modeling needs. As one of the few real-time capable dynamic simulation tools, it enables immediate simulation without exporting models or code, facilitating real-time validation and adjustment of vehicle models. SIMPACK can also connect with driving simulators and test rigs, enabling human-in-the-loop testing and enhancing driver experience for specialized heavy-duty electric drive vehicles.

4 Model Construction for an 8 × 8 Heavy-Duty Multi-axle Electric Drive Vehicle

Upon defining the modeling methodology and software, the dynamics model of a specific 8 × 8 heavy-duty multi-axle electric drive vehicle was constructed. The vehicle was divided into cab, frame, suspension, wheels, steering, and powertrain modules. These were then assembled into a full-vehicle dynamics model using the UG model as a foundation. Advanced SIMPACK modules, including inputs, excitations, and expressions, along with force elements like springs and torques, were used to build the vehicle's powertrain model, forming a complete dynamics model for the specialized heavy-duty multi-axle electric drive vehicle.

In multi-body dynamics systems, the connection topology between structures is a critical element influencing model complexity and accuracy, as well as computation time and efficiency [11]. The primary topology for the 8 × 8 heavy-duty vehicle model is illustrated in Fig. 1, and the full-vehicle model developed in SIMPACK is shown in Fig. 2.

As shown in the figure above, the main model is divided into eight substructures: cab, frame, axles one through four, steering system, and powertrain system. During the modeling process, a bottom-up approach was used: each sub-model was first created individually and then integrated into the main model. The joint between the vehicle body and the ground is represented by Hinge 19, a standard vehicle motion joint that provides six degrees of freedom (three translational and three rotational along each axis). The road surface is modeled as a flat terrain, with a signal generator for "road access points" linked to the corresponding signal receiver in the tire module. This setup enables the tires to detect the road surface and trajectory, responding accordingly with contact and bumps during operation.

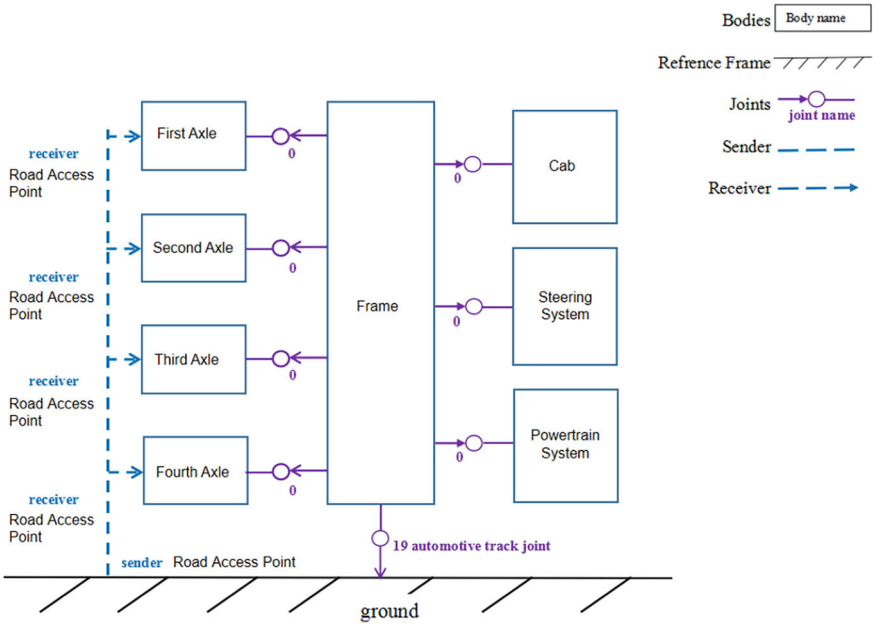
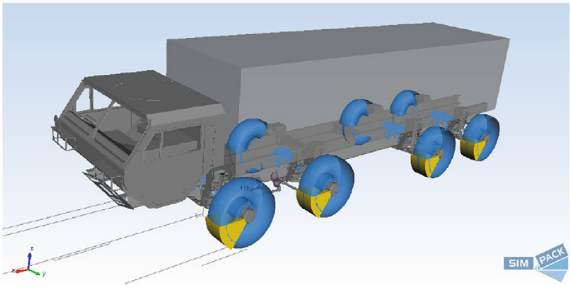


Fig. 1 Primary topology of the 8 × 8 heavy-duty vehicle model

Fig. 2 Full-vehicle model of the heavy-duty vehicle



Following the parameterized modeling method, basic vehicle parameters—such as mass, total length, width, height, stationary acceleration time, and full-load braking distance—are grouped into a "Vehicle Overall Parameters" set. Similarly, positional and dimensional parameters for each component are grouped into respective parameter sets, which are automatically imported with each substructure when added to the main model.

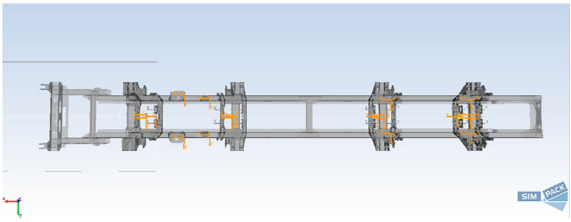


Fig. 3 Frame structure and connection points

4.1 Development of the Frame and Cab Dynamics Model

The frame serves as the foundational structure of the vehicle, providing support and mounting points for other components. In the model configuration, the frame represents the integrated structure that supports and secures other vehicle parts, simultaneously serving as the attachment point for various joints.

The cab was imported directly as a UG (Unigraphics) model and positioned to align with the designated frame joint points. The connection points between the frame, axles, and upper-body-mounted components are illustrated in Fig. 3.

4.2 Development of the Dynamics Model for Axles One to Four

The components of Axles One through Four include the suspension, wheel hubs, and tires. The second axle is configured as a mirror image of the first axle and requires reconfiguration to achieve symmetry, while the third and fourth axle share an identical structure with the first axle and can be positioned in the main model by referencing the first axle and connecting it via joints at the corresponding locations.

The tires are imported as UG (Unigraphics) models and introduced as substructures within each axle model. Road contact points are defined on the tires. Given the structural complexity of the suspension components, they are also imported as UG models, with a density set to 7850 kg/m^3.

The hydropneumatic springs are simplified to standard spring elements, with stiffness and damping values inputted as parameters. Taking the first axle as an example, the completed model with articulated joints and designated joint markers is shown in Fig. 4.

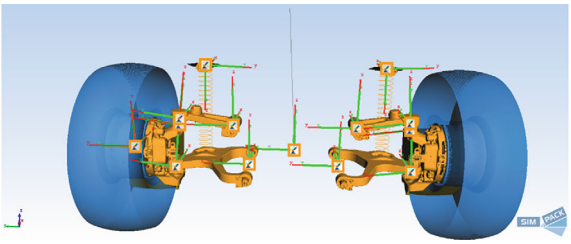


Fig. 4 The first axle model diagram

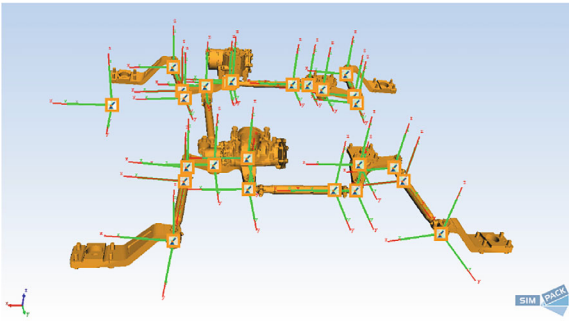


Fig. 5 Steering system model with joint connections

4.3 Development of the Steering System Dynamics Model

The steering system is modeled by importing UG (Unigraphics) models, with the left front axle's power steering unit as the primary reference point for articulating the connections between components.

The vehicle operates in an all-wheel steering mode. The steering mechanism functions as follows: the steering wheel directly transmits a 1:1 angle ratio to the left power steering unit on the front axle. The power steering unit has a fixed steering ratio of 23:1, which is transmitted to the wheels via the steering linkage rods. A control unit, C-Driver, is implemented with a steering wheel transmission ratio of 23 and a control mode set to PID control. The complete model with articulated joints and constraints is shown in Fig. 5.

4.4 Development of the Powertrain System Dynamics Model

The powertrain system is central to an electric vehicle, primarily responsible for converting electrical power into mechanical power, thereby driving the mechanical transmission and transferring power to the wheels to enable vehicle motion. Modeling and simulation of this system are critical in the full-vehicle modeling process. SIMPACK's unique force element library conveniently simulates motor

torque and enables dynamic power adjustment based on throttle input through the use of expressions, facilitating the control of vehicle drive power.

This 8 × 8 multi-axle electric drive vehicle utilizes a distributed drive system, consisting of eight wheel-edge electric drive units. During the modeling process, simplifications were made: Force Element 13 simulates the torque transmission of the drive shaft, allowing the omission of the physical drive shaft structure. The torque values from the motor's external characteristic curve are multiplied by the wheel-edge gearbox ratio, simulating the torque increase effect of the gearbox, which also allows the omission of the physical gearbox structure. Consequently, the differences between drive shaft types across axles can be disregarded, simplifying the system to eight identical motor structures.

Construction and Configuration of Inputs, Excitations, and Expressions

In SIMPACK, force settings often follow a sequence of "Input—Excitation—UE—Expression—Force Element."

Accordingly, motor external characteristic curves and throttle opening curves are defined as inputs. An excitation referencing the throttle curve is then created, followed by a u-vector to facilitate subsequent expression calls. The motor torque expression is:

UELE ($UE_throttleposition_Value)* IFCTNX(JOINTST (%J_rotor,1,1)*30/pi, &I_external characteristic, 0, −1, −1)*(−1).

Model Construction, Joint, and Force Element Configuration

The motor structure retains only the stator and rotor components, with the stator simplified as a hollow cylinder and the rotor as a solid cylinder. To generate motor torque, Force Element 50 is applied between the stator and rotor, with the torque expression set to "Expression for Ly," referencing the torque expression defined earlier.

The force element responsible for transmitting motor torque is configured at the axle level. In each axle's sub-model, Force Element 13 is created, starting at the motor rotor and ending at the drive shaft connection point on the wheel hub. An overview of the entire system is shown in Fig. 6.

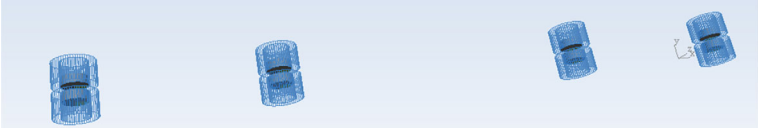


Fig. 6 Powertrain system model diagram

5 Simulation Validation of the Dynamics Model

After constructing the multi-body dynamics model for the multi-axle heavy-duty electric drive vehicle, simulations were conducted under various conditions to validate its dynamic performance and braking characteristics. These simulations were then compared against real-world test data to confirm the accuracy of the model.

For dynamic performance, acceleration and hill-climbing tests were simulated. For braking performance, an emergency braking test was conducted. The road surface friction coefficient was set to 0.7 across all tests, and throttle opening curves and force element settings were adjusted according to each specific condition.

5.1 Dynamic Performance Testing

Start-Up Acceleration Performance Simulation

A simulation was conducted to accelerate the vehicle model from a stationary position to a speed of 60 km/h, recording the time and distance required. The throttle was pressed 2.5 s after the start of the simulation, reaching maximum throttle opening at 3 s. The throttle opening curve was set accordingly, with the road surface in SIMPACK configured as a flat terrain. According to the vehicle design specifications, the 0–60 km/h acceleration time should not exceed 25 s. Real-world test data showed an acceleration time of 15.00 s. Post-simulation, vehicle velocity and position data were obtained, as shown in Figs. 7 and 8.

As shown in the figure, the vehicle's acceleration time from a standstill to 60 km/h is calculated as $(19.585 - 2.500) = 17.085$ s, with an acceleration distance of 140.945 m. This meets the design specifications and aligns closely with the real-world test data, indicating that the model's acceleration performance is satisfactory.

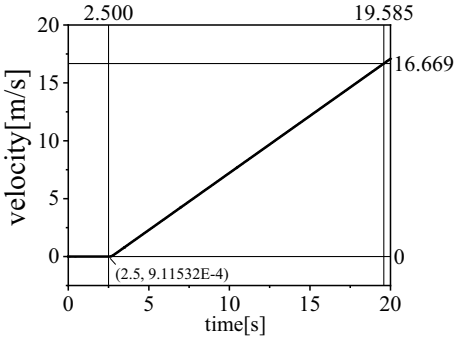


Fig. 7 Vehicle velocity–time curve

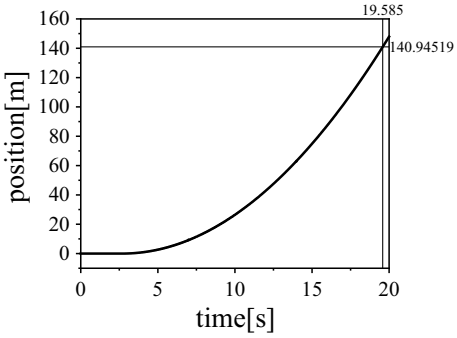


Fig. 8 Vehicle position-time curve

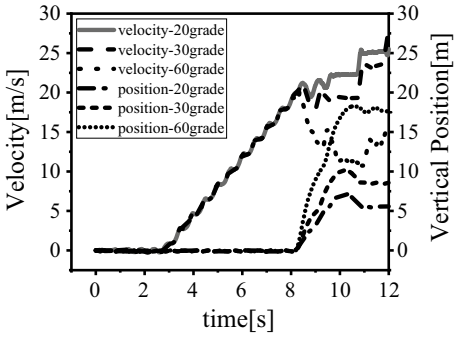


Fig. 9 Vehicle speed-time and vertical position-time curves

Hill-Climbing Performance Simulation.

The vehicle's ability to maintain a stable speed while climbing is crucial for ensuring safe transport operations. In real-world testing, the vehicle successfully ascended a 60% grade. In SIMPACK, road surface angles were set to simulate slopes with gradients of 20, 30, and 60%. The model vehicle started from a standstill on a flat section leading to the slope. The resulting velocity and vertical position over time are shown in Fig. 9.

The figure shows that, although vehicle speed decreases as it climbs the slope, it stabilizes overall and successfully completes the ascent on each gradient. The maximum grade climbed is at least 60%, consistent with real-world test results, confirming that the model meets the vehicle's hill-climbing performance requirements.

5.2 Braking Performance Test

To validate the vehicle's braking performance, an emergency braking simulation was conducted. The vehicle traveled along the centerline of the test track at an initial

speed of 30 km/h, then performed an emergency brake to bring it to a stop. During this process, the vehicle's position and speed were recorded. The designed braking distance for an emergency stop from 30 km/h is specified to be $\leq$ 12.0 m. In real-world testing, the actual initial braking speed was 31.4 km/h, with a measured braking distance of 8.71 m and a braking time of 2.30 s.

The braking mechanism is implemented by applying a torque opposite to the direction of travel between the steering knuckle and the wheel hub, combined with friction to achieve braking and bring the vehicle to a complete stop.

The brake pedal is pressed at 6 s, reaching maximum pressure by 7 s. The resulting speed, displacement, and acceleration versus time curves are shown in Figs. 10 and 11.

The figure shows that during the emergency braking from an initial speed of 30 km/h, the braking time was $(7.865 - 6.0) = 1.865$ s, and the braking distance was $(57.1264 - 48.9102) = 8.2162$ m. These results meet the design requirements for braking and align closely with real-world test values.

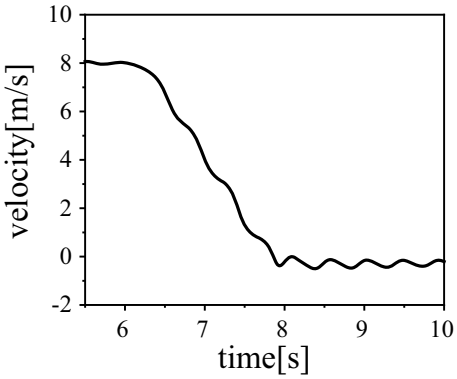


Fig. 10 Vehicle velocity–time curve

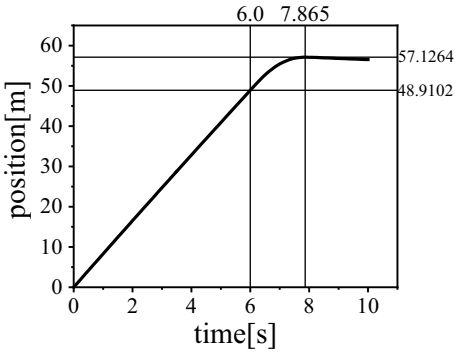


Fig. 11 Vehicle position-time curve

6 Conclusion

With SIMPACK, a multi-body dynamics model was developed for a specific 8 × 8 specialized heavy-duty multi-axle electric drive vehicle. By applying modular and parameterized modeling techniques, a streamlined, efficient, and adaptable modeling process was established for specialized heavy-duty multi-axle electric vehicles. The model underwent simulations for stationary acceleration, hill-climbing, and braking scenarios, allowing for an assessment of its dynamic performance and braking capabilities. These simulation results were compared to design values and real-world test data, validating the model's applicability. This study offers a methodological reference for building simulation models and analyzing dynamic characteristics of specialized heavy-duty multi-axle electric drive vehicles.

Acknowledgements This article is one of the interim achievements of the major project of the National Natural Science Foundation of China, "Overall Design and System Verification Technology for Ultra-large/Heavy-load Transportation System" (Theory and System Control Technology for Modular Distributed Electrically-driven Heavy-duty Vehicles Design).

References

1. Jing X (2007) Research and development of vehicle virtual driving system based on virtual reality technology. Shandong University of Technology, Master's thesis
2. Wei YS, Yang XL, Wang F (2004) Virtual reality and system simulation. National Defense Industry Press, Beijing
3. Kim W, Yi K, Lee J (2011) Drive control algorithm for an independent 8 in-wheel motor drive vehicle. J Mech Sci Technol 25:1573–1581 [Crossref]
4. Fauroux JC, Vaslin P (2010) Modeling, experimenting, and improving skid steering on a 6 × 6 all-terrain mobile platform. J Field Robot 27:107–126 [Crossref]
5. Li Z, Liu XX, Li L (2022) Simulation of the steering characteristics of multi-axle vehicles based on a linear two-degree-of-freedom model. Shandong Sci 35:99–106
6. Cheng HJ, Gao L, Liu ZH, Liu Q (2021) Dynamic modeling and braking performance optimization of multi-axle special vehicle. J Vibration Shock 40:241–248+289
7. Dhar S, Hohnstadt W, Green J (2002) Integrated modular methodology—philosophy and strategy to build full vehicle finite element model. SAE Technical Paper, 2002-01-3072 [Crossref]
8. Li N, Gao WM, Dai Y (2008) Optimization design of car body structure based on implicit parametric model. Automot Eng 30:857–860
9. Xie JY, Chen YS, Shen GM (2023) Full-parametric platform and modular modeling for electric vehicle body design with performance bandwidth study. Chinese J Autom Eng 13(1):94–103
10. Fu XT (2004) Expert kinetic analysis software—SIMPACK. Intell Manuf 3:69–70
11. Miu BR, Luo R, Wang Z (2010) SIMPACK kinetic analysis advanced tutorial. Southwest Jiaotong University Press, Chengdu

Open Access This chapter is licensed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License (http://creativecommons.org/licenses/by-nc-nd/4.0/), which permits any noncommercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if you modified the licensed material. You do not have permission under this license to share adapted material derived from this chapter or parts of it.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.



Research on Monocular Vision Indoor Construction Positioning Method Based on YOLOv5



Hang Yang, Bo Sun, Weize Lin, Taozheng Wu, Gangju Wang, and Guangjun Liu

Abstract Currently, indoor renovation construction technology is developing towards intelligence and automation, creating an urgent need for mobile robots to participate in high-pollution and high-risk construction operations. To address the issues of insufficient robustness and accuracy in indoor construction robot positioning due to relatively small indoor spaces and weak signals, this paper proposes a monocular vision indoor construction positioning and distance measurement method. It designs an improved YOLOv5 target detection algorithm with Canny edge detection operator for distance measurement and positioning. Experiments were conducted in a specific indoor construction environment by measuring the distance between the cross point projected by the laser emitter on the construction robot and the cross point generated by the laser level. The experimental results indicate that the calculated values from the algorithm differ from the actual measurements by no more than 8 mm, achieving an accuracy of 96%. This demonstrates that the proposed method has good performance in indoor construction positioning.

Keywords Indoor construction · Monocular vision · YOLOv5 · Distance measurement and positioning

1 Introduction

From the perspective of modern industrial development trends, there is a general inclination within the industry to assign tasks in high-pollution and high-risk environments to industrial robots. In large buildings, construction robots can also be employed for indoor renovation, particularly in elevated work areas. To ensure that

H. Yang · B. Sun (✉) · W. Lin · T. Wu · G. Liu
School of Mechanical Engineering, Tongji University, Shanghai, China
e-mail: bo.sun@tongji.edu.cn

G. Wang
Shanghai Engineering Research Center for Safety Intelligent Control of Building Machinery, Shanghai Research Institute of Building Sciences, Shanghai, China

© The Author(s) 2026

H. Saman et al. (eds.), *The 9th International Conference on Advances in Construction Machinery and Vehicle Engineering*, Lecture Notes in Mechanical Engineering,
https://doi.org/10.1007/978-981-96-5527-4_54

construction robots achieve good positioning and installation accuracy as well as high construction efficiency, there is a need to redevelop and optimize their positioning systems, enhancing the robustness and precision of the overall positioning system.

Currently, the mainstream and mature indoor positioning technologies include Assisted Global Positioning System (AGPS), Ultra-Wideband (UWB) technology, Radio Frequency Identification (RFID) technology, LiDAR positioning technology, and computer vision positioning technology. Common visual servo control schemes mainly include five types: Position-Based Visual Servoing (PBVS), Image-Based Visual Servoing (IBVS), 2.5D Visual Servoing (2.5DVS), Local Visual Servoing (PVS), and Shortest Path Visual Servoing (SPVS) [1]. Mohammad Hazrul and colleagues developed a positioning and autonomous navigation technology that integrates cameras and ultrasonic sensors on Autonomous Mobile Robots (AMRs) to detect and avoid obstacles [2]. This technology can comprehensively assess the distance and size of obstacles based on their shape and proximity to the robot. Joseph Redmon and his team at the University of Washington proposed a new object detection framework called YOLO (You Only Look Once) to address the issue of traditional computer vision object detection methods needing to reclassify objects in advance [3], which hinders real-time processing of sensor-captured image information. YOLO requires only a single neural network to directly delineate predicted bounding boxes and identify the probabilities of target categories from a complete image in one go. Yoshitaka Matsuda and colleagues developed a vision servo system for a mobile robotic arm based on USB-transmitted images from industrial cameras [4], achieving a high success rate in remote control operations and strong control robustness. Wu and others proposed a distributed networked visual servo control system based on Real-Time Transport Protocol (RTP) to address the problem of the small proportion of effective information sampled in commonly used visual servo systems [5], effectively improving the accuracy and robustness of the control system.

Therefore, based on existing research achievements and shortcomings, this paper proposes an improved YOLOv5 object detection algorithm that incorporates the Canny edge detection operator [6]. This algorithm can achieve relatively accurate localization between construction robots and target points, and it is relatively lightweight with good generalization capabilities, making it applicable to other fields.

2 Theory of Distance Measurement Algorithm

In order to enable the YOLO algorithm to identify the positions of targets in an image and determine the category of the objects at those positions, it is necessary to incorporate activation functions, loss functions, and Intersection over Union (IoU) loss functions into the network architecture [7]. The former plays a crucial role in the forward propagation of the entire training process, while the latter two significantly influence the weights and biases during the backward propagation in the training and validation process of the YOLO algorithm [8].

In the final prediction layer of YOLOv5, the Sigmoid activation function is chosen, and its mathematical expression is:

$$f(x) = \frac{1}{1 + e^{-x}} \tag{1}$$

However, the activation function in the neck's middle layer is Leaky ReLU, and its mathematical expression is:

$$f(x) = \begin{cases} x, x \geq 0 \\ \alpha x, x < 0 \end{cases} \tag{2}$$

where α is a small positive constant.

The purpose of the loss function is to compare the parameter vector obtained from each iteration with the initial parameter vector, thereby determining whether the model-selected relevant weight coefficients and biases can ensure the stability and robustness of the network algorithm model during actual production use. The loss function of the YOLO algorithm is the sum of the anchor box loss function (L_{box}), the object detection loss function (L_{obj}), and the object classification loss function (L_{class}), and its expression is:

$$L_{all} = L_{box} + L_{obj} + L_{class} \tag{3}$$

$$L_{box} = L_{xy} + L_{wh} \tag{4}$$

where L_{xy} and L_{wh} are both the mean squared loss values for the predicted bounding boxes of the object detection network.

Since the coordinates of the top-left corner of the bounding box and the width and height of the bounding box are not independent in the algorithm, the IoU loss [9] function is introduced, and its expression is:

$$L_{IoU} = 1 - \frac{\left|B \cap B^{gt}\right|}{|B \cup B^{gt}|} \tag{5}$$

where B is the size or rectangular set of the predicted bounding box, and B^{gt} is the size or rectangular set of the actual bounding box.

The geometric distance measurement with a monocular camera involves the transformation of three-dimensional real objects into two-dimensional pixel information through the camera imaging device. Therefore, before obtaining the distance measurement results, it is essential to perform the coordinate transformation and inverse transformation processes. The Fig. 1 illustrates the coordinate transformation process during the imaging of an industrial camera.

The first coordinate system transformation in imaging is the conversion from the world coordinate system to the camera coordinate system. Its essence involves

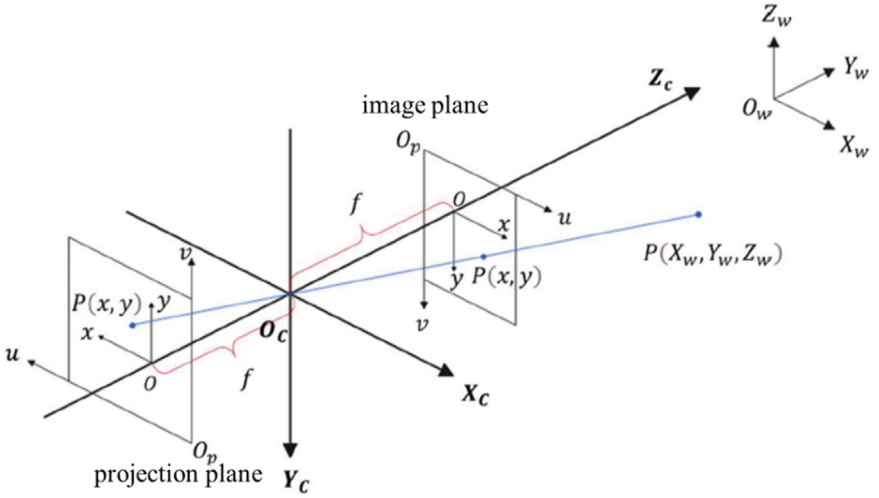


Fig. 1 The coordinate transformation process for the industrial camera

rotation and translation of the coordinate system, and its expression is:

$$\begin{bmatrix} X_c & Y_c & Z_c \end{bmatrix}^T = R\begin{bmatrix} X_w & Y_w & Z_w \end{bmatrix}^T + T \tag{6}$$

where R is the orthogonal rotation parameter matrix, and T is the translation parameter matrix.

The camera coordinate system is transformed into a 2D plane coordinate system through projection and dimensionality reduction, and its expression is:

$$Z_c\begin{bmatrix} x & y & 1 \end{bmatrix}^T = \begin{bmatrix}\begin{bmatrix} f & 0 & 0 & 0 \\ 0 & f & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}\end{bmatrix}\begin{bmatrix} X_c & Y_c & Z_c & 1 \end{bmatrix}^T \tag{7}$$

where f is the focal length of the camera imaging.

Then the conversion between the optical image coordinate system and the digital pixel image coordinate system involves only scaling of the coordinate system and translation of the coordinate system origin:

$$\begin{bmatrix} u & v & 1 \end{bmatrix}^T = \begin{bmatrix} \frac{1}{dx} & 0 & u_0 \\ 0 & \frac{1}{dy} & v_0 \\ 0 & 0 & 1 \end{bmatrix}\begin{bmatrix} x & y & 1 \end{bmatrix}^T \tag{8}$$

Finally, the intrinsic and extrinsic parameter matrices of the camera can then be determined, and the expression is:

Table 1 Experimental environment

Operating system	Windows 11 22H2
Camera	A3B00CG000
Pytorch	1.12.1
GPU	NNDIA GeForce RTX 3060
CPU	Intel i7-12700H
CUDA	11.3

$$M_1 = \begin{bmatrix} R & T \\ 0 & E \end{bmatrix}^{-1} \tag{9}$$

$$M_2 = \begin{bmatrix} \frac{1}{dx} & 0 & u_0 \\ 0 & \frac{1}{dy} & v_0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} f & 0 & 0 & 0 \\ 0 & f & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \tag{10}$$

where M_2 represents the intrinsic parameters of the camera, determined by the camera's own hardware and software conditions, and M_1 represents the extrinsic parameters of the camera, determined by the camera's coordinate system and its position in the actual space. Both parameters can be obtained through camera calibration methods.

3 The Experiment and Results Analysis

3.1 Experimental Environment

This paper builds the deep learning model based on PyTorch 1.7.1, and all subsequent experiments are conducted in the hardware and software environment listed in Table 1.

The network training uses the SGD optimizer, with an initial learning rate and final learning rate of 0.01; the image input size is 640 × 640, the batch size is 8, the maximum number of training epochs is 300, and early stopping is set to 25.

3.2 Dataset

In this paper, the industrial camera selected in Sect. 3.1 is used to independently collect dataset images in the experimental site. By setting the acquisition parameters of the industrial camera, 27 photos are taken in succession. The collected photos are then calibrated, as shown in the Fig. 2., categorizing all the cross points in the images into three types: the cross points projected by the robot itself (target cross),

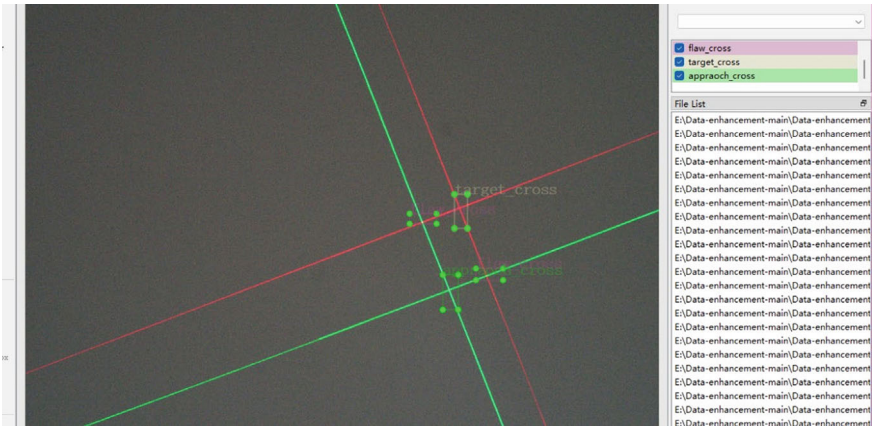


Fig. 2 Labeling by labelimg

the positioning cross points that the robot needs to approach (approach cross), and the interference cross points formed by the intersection of multiple rays (flaw cross).

To ensure that the model trained on the dataset has broad applicability, data augmentation techniques were employed to further process the images. An algorithm was designed to randomly generate noise for a single image or to perform random operations such as rotation, cropping, translation, and brightness adjustment, while also automatically labeling the images. Each original sample image in the collected dataset underwent ten different random operations, expanding the entire image dataset to 270 images. After processing, the training set contains 243 images, and the validation set contains 27 images.

3.3 Process

By training and validating the dataset, the corresponding evaluation metric curves for the algorithm can be obtained. The Fig. 3 shows the confusion matrix at the yolo5s weights, the Fig. 4 shows the precision-confidence curve, and the Fig. 5 shows the recall-confidence curve.

The confusion matrix indicate that all predicted true values are consistent with the actual true values throughout the entire training process, and there was no influence from background noise. To further determine whether the model is overfitting, the trends of the loss function for the training set and the validation set with respect to the number of epochs are analyzed, as shown in the Fig. 6. The analysis results indicate that the decreasing trends of the loss function values for both the training set and the validation set are nearly consistent, proving that the model is not overfitting. When the IoU is at 50%, the area under the precision-recall curve shows that the model's mean Average Precision (mAP) is close to 1. The visualization of the detection

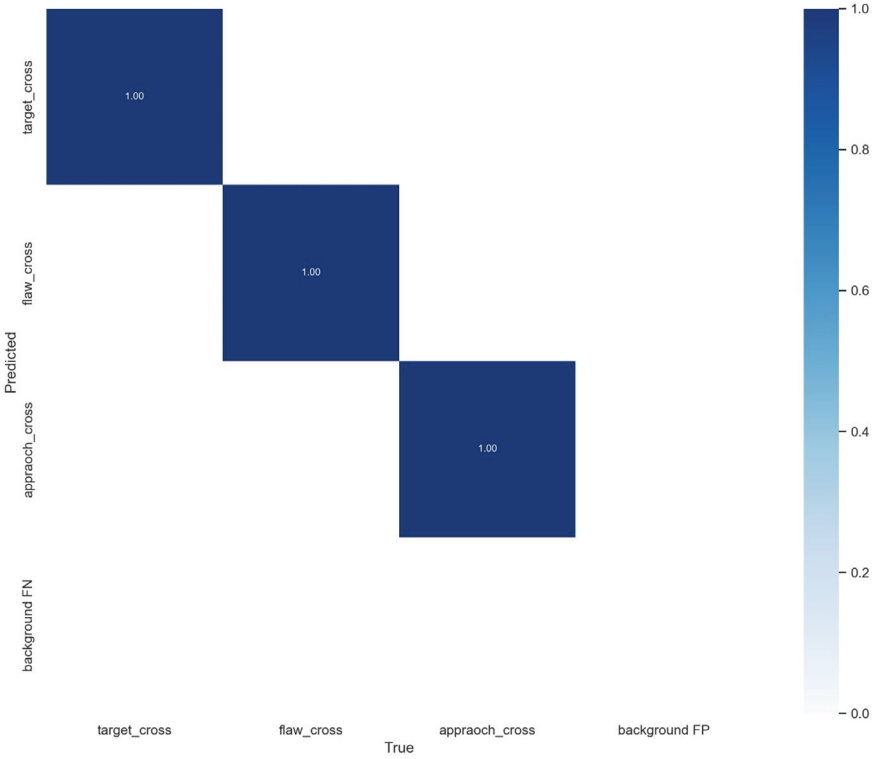


Fig. 3 Confusion matrix

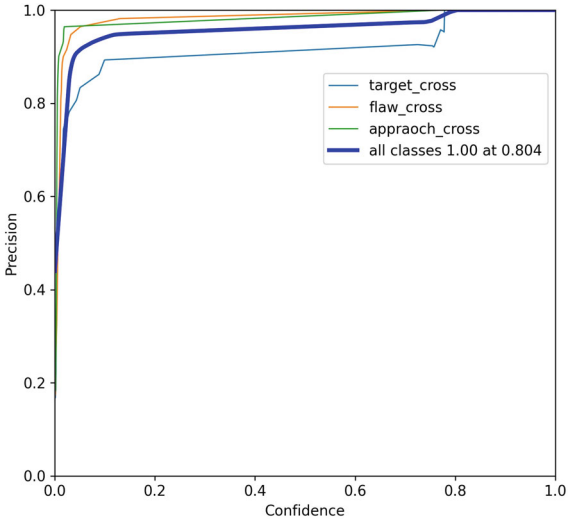


Fig. 4 Precision-confidence curve

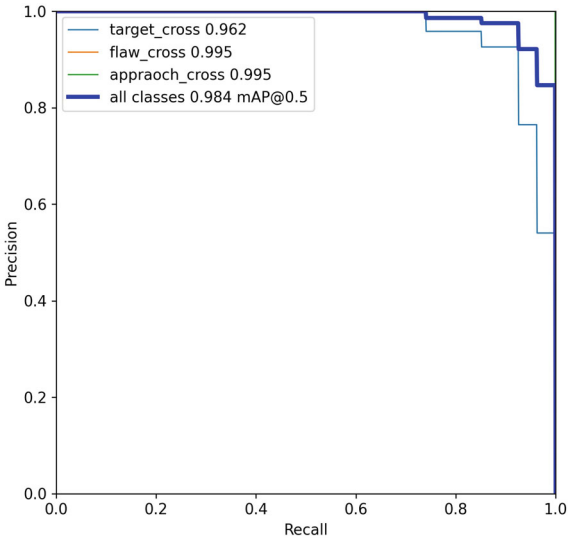


Fig. 5 Recall-confidence curve

results for the validation set is shown in the Fig. 7. The results indicate that the pre-trained model can accurately recognize different types of cross points, with high confidence for each bounding box. Therefore, this model can be used for subsequent experiments.

Before conducting the distance measurement experiment, it is necessary to calibrate the monocular vision industrial camera. The Zhang's calibration method is used to obtain the intrinsic and extrinsic parameter matrices of the camera, which in turn provides the scale factor. In this experiment, it is 9.78.

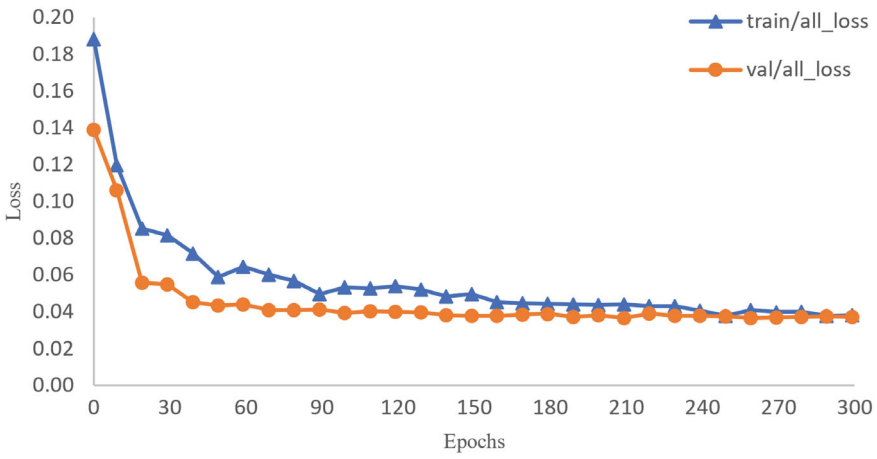


Fig. 6 Training loss and validation loss with number of epochs



Fig. 7 The visualization result of the objection detection

The essence of distance measurement and localization in this method is to measure the distance between the centers of two types of cross points. The Canny edge detector is introduced to perform edge detection on the anchored regions of interest.

To obtain a more accurate edge image, optimization processes such as dilation and erosion are applied to the regions of interest. As shown in the Figs. 8 and 9, the optimized lines are more pronounced, and the number of white noise pixels around the target lines is significantly reduced. The processed image is then subjected to edge detection and convex hull fitting, resulting in the corresponding line and intersection point location information, as illustrated in the Fig. 10.

Finally, by developing the soft trigger program for the industrial camera, the camera can be automatically controlled by the host computer to conduct monocular vision distance measurement experiments.

3.4 Analysis of the Results

Multiple experiments were conducted in the same hardware and software, as shown in Table 2. The visualization example of the experiment results is illustrated in Fig. 11.

Comparison of the values derived from the subsequent manual measurements with the values calculated from the coordinates of the intersection points obtained by edge detection, it can be observed that the values are nearly equal. This indicates

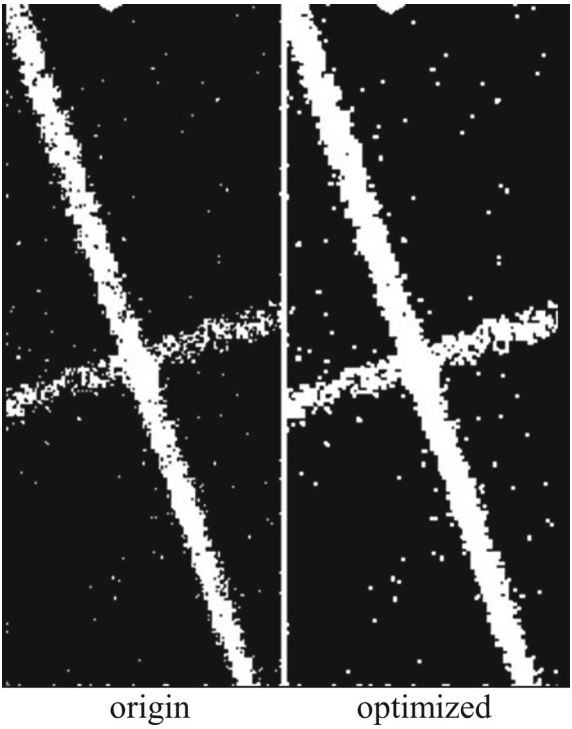


Fig. 8 Threshold segmentation before and after optimization

that the intersection points coordinates derived from edge detection are reasonable and usable.

Furthermore, as shown in the table, there is a certain error between the identified intersection positions and the actual intersection positions, but the error is only in the millimeter range, achieving an accuracy of 96%, which meets the positioning accuracy requirements for indoor construction in large buildings.

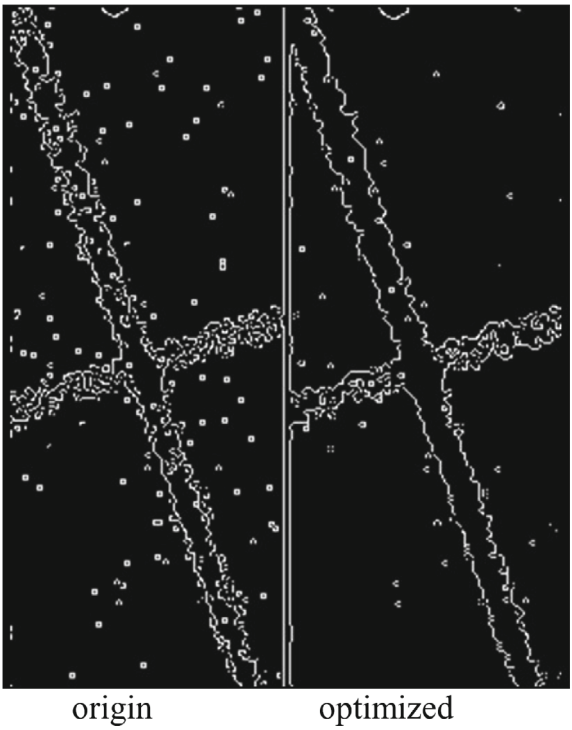


Fig. 9 Edge detection before and after optimization

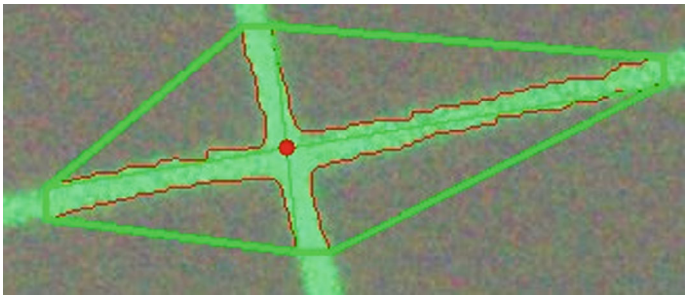


Fig. 10 The corresponding line and intersection point

Table 2 Analysis of results for monocular vision object distance measurement

No.	Distance on x axis (mm)	Distance on y axis (mm)	Central distance (mm)	Real distance (mm)	Error (mm)
1	165.76	107.26	197.44	205.35	7.91
2	187.45	110.34	217.51	213.47	4.04
3	207.45	109.53	234.59	229.39	5.20
4	209.53	130.45	246.82	246.84	0.02
5	230.90	142.20	271.17	274.34	3.17
6	246.18	149.67	288.11	286.98	1.13
7	250.45	156.53	295.34	293.17	2.17

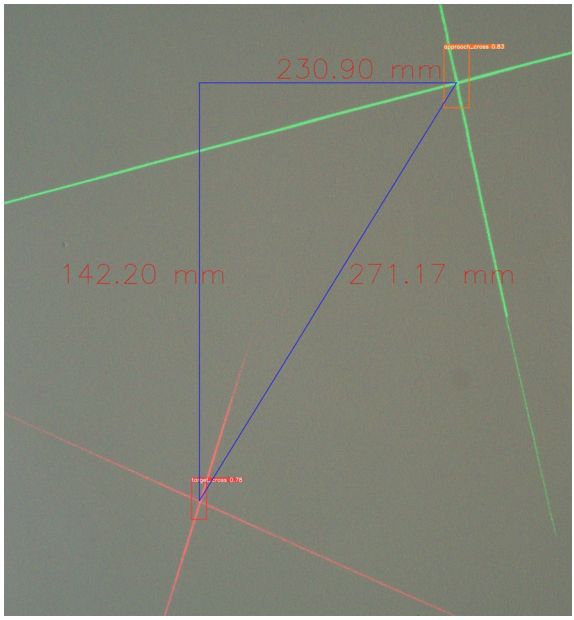


Fig. 11 Visualization example of the experiment results (No. 5)

4 Conclusion

This paper addresses the positioning problem of robots in large indoor construction environments by proposing a monocular vision distance measurement method based on YOLOv5. The results indicate that this algorithm performs well in laser cross point-assisted positioning scenarios, validating the effectiveness of the algorithm. However, due to the limitations of the experimental environment and hardware, as well as the relatively small size of the self-constructed dataset compared to commonly used construction datasets, issues such as missed detections and false detections of nearby cross point targets occur in some images, especially in cases of insufficient

background exposure. In future research, the focus will mainly be on expanding the dataset or designing better image optimization methods to address this issue. Additionally, since the training experimental environment is cost-effective and easy to set up, it is expected that the model can be transferred and improved for edge-moving platforms, making it lighter and easier to deploy.

Acknowledgements This work was supported by the National Key R&D Program of China (Grant No. 2023YFF0613200).

References

1. Duy CV, Duc HL (2023) A review and performance comparison of visual servoing controls. Int J Intell Robot Appl 7(1):10–11
2. Rosdi MHAB, Abdul Ghani AS (2022) Investigation on accuracy of sensors in sensor fusion for object detection of autonomous vehicle based on 2D lidar and ultrasonic sensor. In: Ab Nasir AF, Ibrahim AN, Ishak I, Mat Yahya N, Zakaria MA, Abdul Majeed A (eds) Recent trends in mechatronics towards industry 4.0. Lecture notes in electrical engineering, vol 730. Springer, Singapore
3. Redmon J, Divvala S, Girshick R, Farhadi A (2016) You Only Look once: unified, real-time object detection. In: 2016 IEEE conference on computer vision and pattern recognition (CVPR).Las Vegas, NV, USA, pp 779–788
4. Matsuda Y, Tagami N, Sugi T et al (2019) Teleoperation system for a mobile robot arm with visual servomechanism based on turning radius determination using angle information of image. Artif Life Robotics 24(2019):106–113
5. Wu H, Lou L, Chen CC, Hirche S, Kühnlenz K (2010)A framework of networked visual servo control system with distributed computation. IN: 2010 11th international conference on control automation robotics & vision. Singapore, pp 1466–1471
6. Sangeetha D, Deepa P (2019) FPGA implementation of cost-effective robust Canny edge detection algorithm. R-Time Image Process 2019(16):957–970
7. Tijtgat N, Van Ranst W, Volckaert B, Goedeme T, De Turck F (2017) Embedded real-time object detection for a UAV warning system. In: 2017 IEEE international conference on computer vision workshop (ICCVW) .Venice, Italy, pp 2110–2118
8. Fan S, Liang X, Huang W et al (2022) Real-time defects detection for apple sorting using NIR cameras with pruning-based YOLOV4 network. Comput Electron Agric 193:106715
9. Yao L, Yaoyao Q (2020) Insulator detection dased on GIOU-YOLOv3.2020 Chinese Automation Congress (CAC).Shanghai, China, pp 5066–5071

Open Access This chapter is licensed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License (http://creativecommons.org/licenses/by-nc-nd/4.0/), which permits any noncommercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if you modified the licensed material. You do not have permission under this license to share adapted material derived from this chapter or parts of it.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.



Ergonomic Simulation Analysis of Forklift Drivers Based on JACK



Pingkai Wang, Dingxuan Zhao, Ruixuan Zhu, and Zeren Wang

Abstract As an important equipment of logistics mechanization system, forklift plays an important role in realizing the mechanization of handling, improving work efficiency and reducing labor intensity. But there are thousands of industrial injuries and casualties caused by forklifts every year, which also brings important safety problems. The study found that the poor visual field environment of the driver is an important factor leading to the accident. This paper applies Jack virtual simulation technology to analyze the working state of forklift drivers. By analyzing the relationship between the driver's body posture and the change of the visual area when driving forward and reversing, it is proposed to change the layout area of the cab seat and mast of the forklift to form a position shift, so as to achieve the purpose of improving the driver's visibility, alleviating fatigue and reducing the accident rate.

Keywords Jack · Virtual simulation technology · Forklift · Reverse driving · Field of vision

1 Introduction

In today's era, the logistics industry is developing rapidly, and the traditional manual handling has been unable to meet the needs of enterprises. As an important equipment of logistics mechanization system, forklift is more and more widely used in stations, ports, mines, airports, freight yards, factories, workshops, warehouses, circulation and distribution centers and other places forklift driving is a special operation with

P. Wang · D. Zhao (✉) · R. Zhu
School of Mechanical and Aerospace Engineering, Jilin University, Changchun, China
e-mail: zdx-yw@ysu.edu.cn

P. Wang · Z. Wang
Engineering Training Centre, Changchun University of Technology, Changchun, China

D. Zhao
Hebei Provincial Key Laboratory of Special Delivery Equipment, Yanshan University, Qinhuangdao, China

© The Author(s) 2026

H. Saman et al. (eds.), *The 9th International Conference on Advances in Construction Machinery and Vehicle Engineering*, Lecture Notes in Mechanical Engineering,
https://doi.org/10.1007/978-981-96-5527-4_55

long operation time, high intensity and great difficulty. Long time of high load operation will lead to forklift driver fatigue, which increases the probability of safety accidents to a certain extent [1]. According to the National Safety Council, forklifts were the source of 79 work-related deaths and 8140 non-fatal injuries involving days away from work (DAFW) in 2019. The Bureau of Labor Statistics calculates that forklifts were involved in 9050 non-fatal workplace injuries in 2017, resulting in workers taking a median of 13 DAFW [2]. The common safety accidents of forklifts mainly include collision safety hazards, extrusion and shear safety hazards, rollover safety hazards, object falling safety hazards, sight blocking safety hazards and other hidden dangers [3]. A study investigating forklift accidents concluded that many hazards encountered objects during loading and unloading and had the highest risk of collision during forward or backward driving. And most accidents were also associated with obstructed vision in the drivers.

When the forklift is working, due to the low visibility caused by the driver's sight block, the forklift has repeatedly crashed into other objects or personnel, causing the accident. In the 'Industrial trucks-Safety rules for the application, Operation and maintenance' Chinese national standard GB36507-2018, it has clearly stipulated that "If there are other vehicles or pedestrians nearby, and in intersections and other occasions where the line of sight is blocked, the operator should send a sound warning signal, slow down/stop, and confirm the safety of the slow through" [4].

In recent years, many researchers have conducted diversified studies on forklift design and its safe driving. Daniel C. and Priyadarshini D. G., Southeast Louisiana University, used ergonomics to establish a forklift training program to improve forklift driving safety [5]. Jing Ou takes the operator's active driving behavior as the research object, optimizes the design of forklift cab through behavior analysis, and provides guidance for the modeling design of forklift cab [6]. Tianqi Wang uses TRIZ theory to establish a functional model of forklift and proposed a design method of high position forklift to improve safety [7]. Hongmin Yang, designs the operating table of the engineering vehicle based on ergonomics to improve the comfort of operation [8]. Shijiang Hou takes the perceptual image as the basis of the forklift modeling design, and uses the visual dynamic analysis to design the forklift appearance modeling [9].

2 Virtual Simulation of the Forklift Cab

Jack is a human body modeling and simulation ergonomics analysis software, it has a rich and comprehensive digital human body models. The digital maniquin contains 69 sites (Segment), 68 joints (Joint), and 135 degrees of freedom. When creating the digital manikin, we can customize the various size data of the body of the digital manikin, which makes the man–machine efficacy analysis more targeted and accurate.



Fig. 1 Forklift 3D model

2.1 Create the Model

(1) Establish a 3-dimensional model of the forklift

A 1:1 forklift model was created according to the actual forklift size in the 3 D design software. Lay the foundation for the subsequent ergonomics simulation (see Fig. 1).

(2) Establish a human body simulation model

Jack integrates multiple national anthropometric databases, provides relatively accurate human biological models, and provides the main body size for Chinese adults [10]. In addition, Advanced Scaling is used to accurately modify the individual or all 26 size parameters of the human model, and then create a digital human model closer to real people [11]. According to the regional criteria in anthropometric data, the 5% to 95% percentile of Chinese adult males provided by GB10000-1988 [12]. Based on the principle of averaging in this paper, the 50th percentile of Chinese adult males were selected to generate the driver 3D model simulation in Jack.

2.2 Create a Virtual Work Behavior

The working posture of the operator is mainly divided into normal driving state and reversing driving state. Complete the simulation of the model within Jack, (see Figs. 2 and 3).

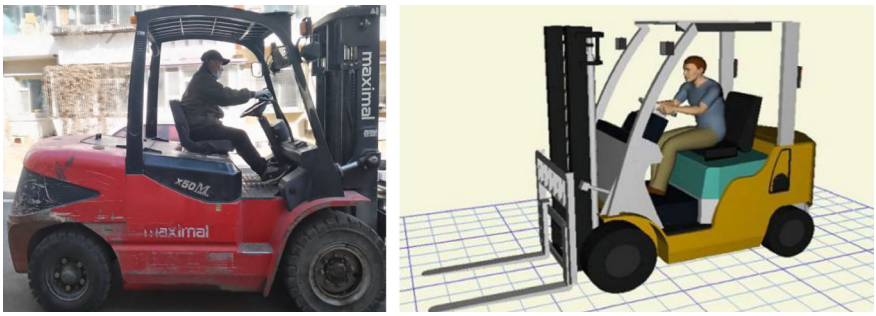


Fig. 2 Normal driving status

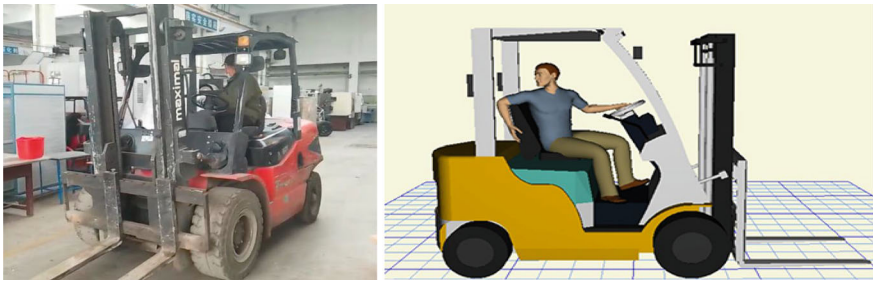


Fig. 3 Reverse driving state

Normal driving requires the operator to sit in the seat in the cab, put his hands on the steering wheel, and put his feet on the clutch and accelerator. When working, the body position forward, hands, eyes, feet forward coordinated movement.

The reverse driving state requires the operator to sit on the seat, one hand on the steering wheel, the upper body turns backward to enable the view to see the rear things of the vehicle, the other arm is on the back of the seat, and the feet step on the clutch and accelerator to control the movement of the forklift.

3 The Ergonomic Simulation Analysis of Drivers

3.1 Comfort Analysis

Comfort analysis is to compare the comfort value of 'digital human' with the reference value range under a specific posture to judge the comfort level. The Kris module in the Jack software belongs to the multi-joint comfort analysis. By studying the driving posture of the operator, the regression formula is used to calculate the relationship between the measured value and the height of the leg, the position of the

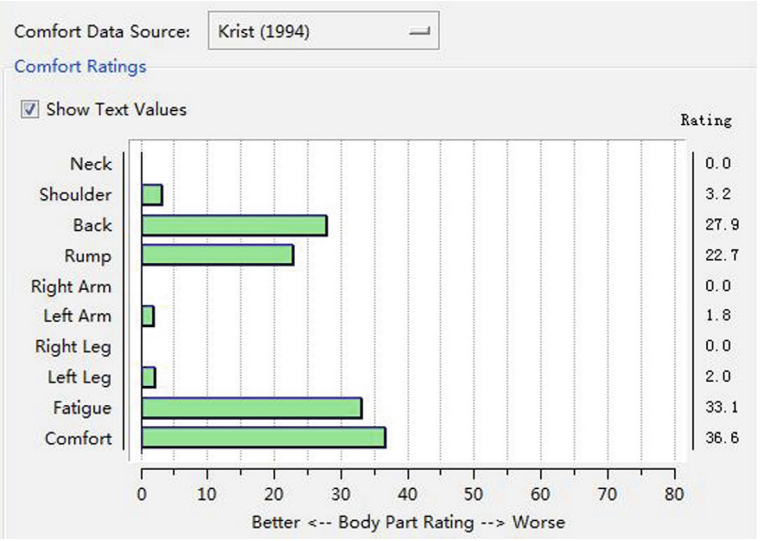


Fig. 4 Analysis of Kris in normal driving state

console and the Angle, and the test data are obtained. The Kris data gives a numerical scoring criterion from 0 to 80, with lower values representing higher comfort. The body parts involved in the data included the neck, shoulder, back, hip, right arm, left arm, right leg and left leg. The Fatigue score indicates the overall fatigue of the current posture, and the comfort score is the overall assessment of comfort for all parts mentioned above. When the comfort value exceeds "40", it indicates that the human body obviously feels uncomfortable. The comfort level of reverse driving state is lower than the normal driving state (see Figs. 4 and 5), and the neck and shoulders of drivers have obvious discomfort in the reverse driving state. This conclusion can provide reference for the follow-up work scheduling arrangement of forklift drivers.

3.2 Visual Domain Analysis

Use the Visual Fields tool to analyze the visibility of the forklift, (see Figs. 6 and Fig. 7).

Forklift drivers in the normal driving position, the front perspective and the left and right limit perspective of the visual range (see Fig. 6); The visual range of the driver's eyes in the reversing driving state (see Fig. 7). It can be seen that the forklift in no-load and normal driving condition, the driver's visual range is better; However, in the reverse driving state, the visual range of the forklift driver is limited.

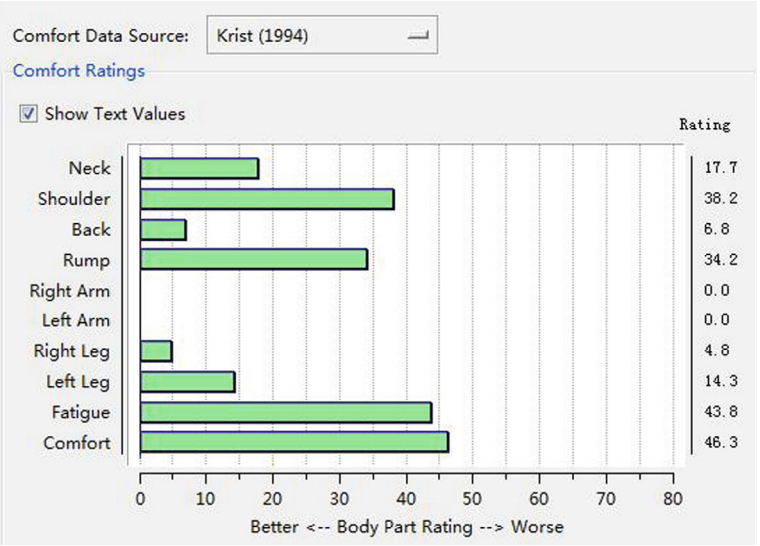
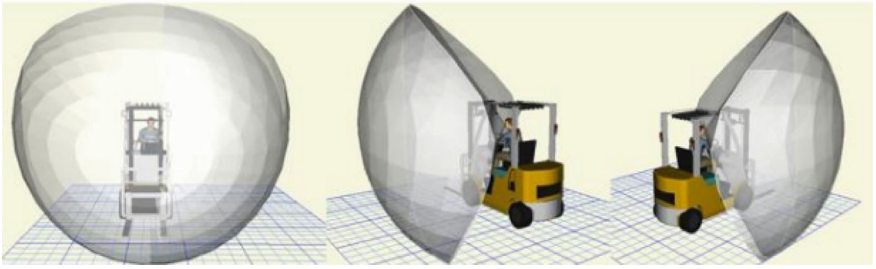


Fig. 5 Analysis of Kris reverse driving state



(a) Front view (b) Left extreme view (c) Right limit view

Fig. 6 Vision analysis of drivers under normal driving state

3.3 Cover Domain Analysis

The Obscuration Zones tool is used to analyze the shaded domain of the driver when the forklift has cargo condition, and the gray area is the shaded part of the vision. The visual field range distance was set to 200 m, as shown in (see Fig. 8) and (see Fig. 9). When there are goods on the fork rack, the vision in front of the operator is blocked under the normal driving condition, and the safety frame behind the forklift partially blocks the vision of the driver.

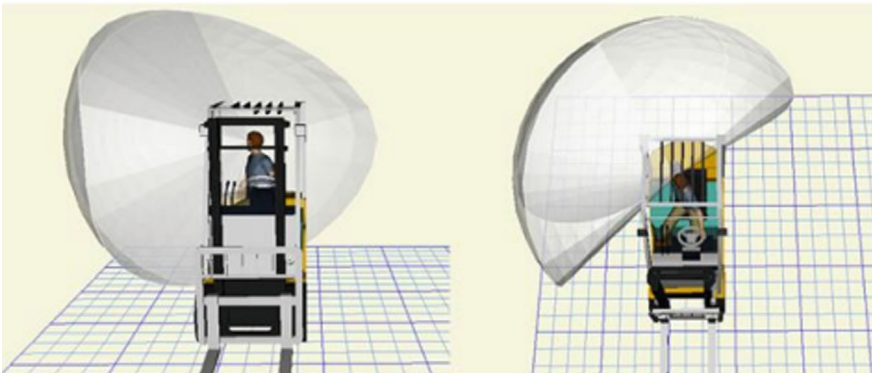


Fig. 7 Vision analysis of drivers in reverse driving state

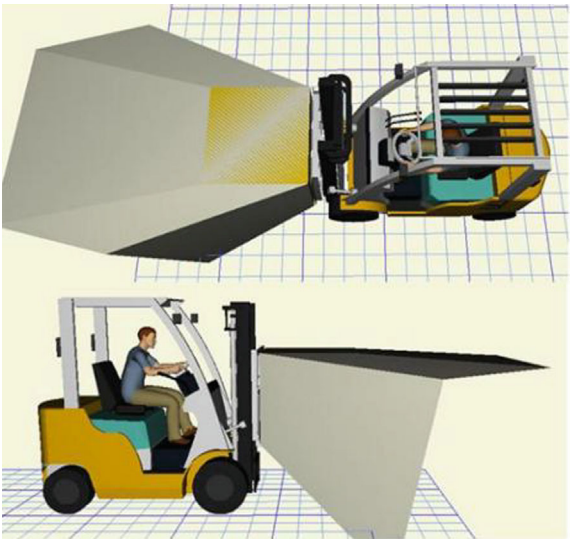


Fig. 8 Analysis of shelter domain in normal driving state

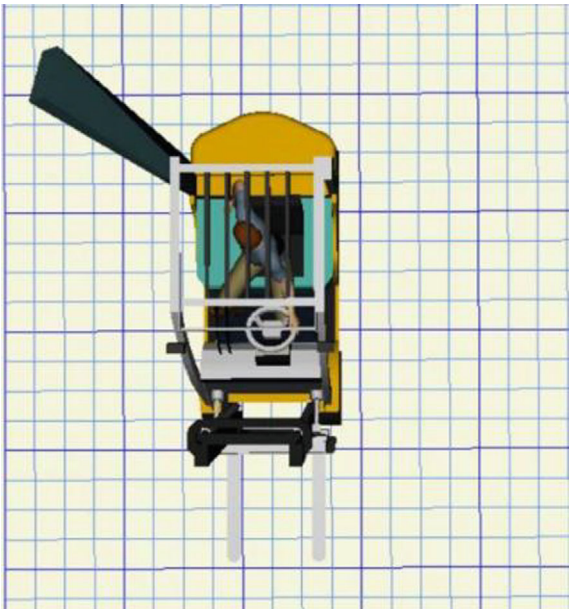


Fig. 9 Analysis of shelter domain in reverse driving state

4 Analysis of Physiological Signals of Drivers

4.1 Physiological Signal Detection Experiment

In the analysis of the driver's comfort, it can be seen that compared with the normal driving posture, the driver's neck and shoulder will have obvious discomfort and fatigue; Based on this, through the interview experiment method, to find the driver muscle fatigue feeling obvious parts. The paper uses the Biopac physiological signal detection system to detect the cervical trapezius muscle (see Fig. 10), the deltoid muscle of the shoulder (see Fig. 11), and the lumbar erector spinae muscle (see Fig. 12).

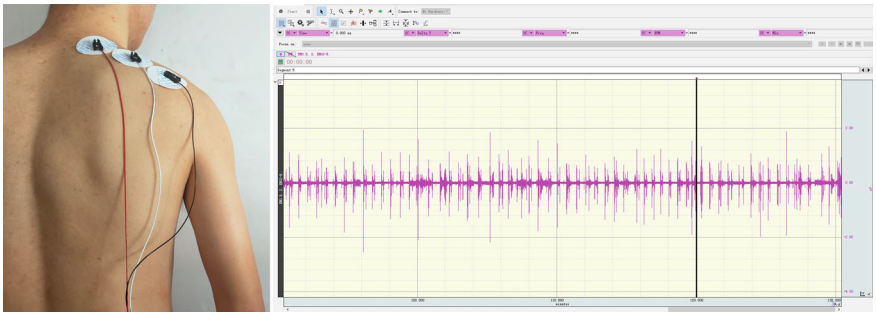


Fig. 10 Detects the trapezius muscle

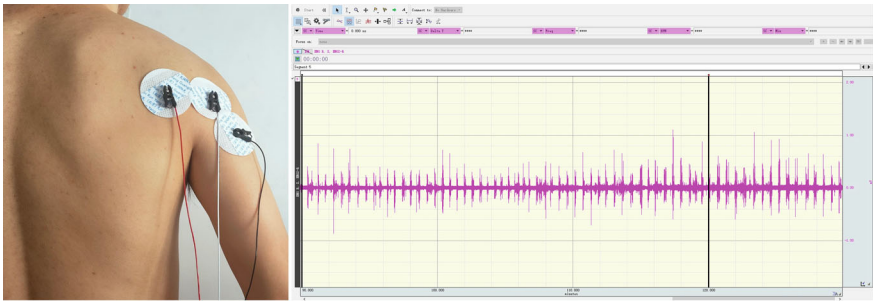


Fig. 11 Detects the deltoid muscle

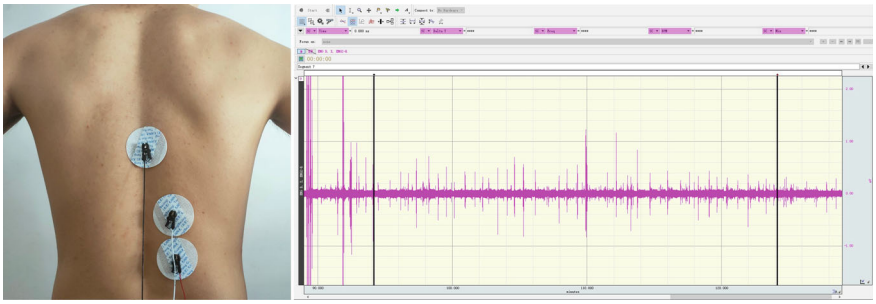


Fig. 12 Detects the erector spinae

4.2 Physiological Signal Analysis

The physiological data of each subject acquired in the experiment were divided into 13 groups in units of 10 min. After the time-domain analysis and the frequency-domain analysis processing of the overall data, the median frequency values of the EMG signals were obtained in the AcqKnowledge software.

$$iEMG = \int_{t}^{t+T} |EMG(t)|dt \tag{1}$$

$$RMS = \left(\frac{1}{T} \int_{t}^{t+T} EMG^2(t)dt \right)^{\frac{1}{2}} \tag{2}$$

$$MF = \int_{-\infty}^{+\infty} p(f)df = \frac{1}{2} \int_{-\infty}^{+\infty} p(f)df \tag{3}$$

T is unit time; f is frequency; and $p(f)$ is the density function of the power spectrum curve.

All three muscle MF data showed a downward trend (see Fig. 13), indicating that all the three muscles of the driver were tired with the work progress and time accumulation. Where the slope of the trapezius muscle decreases faster. It shows that in the three muscles, the trapezius muscle produces a higher fatigue degree, which also indicates that the reversing driving posture has a greater impact on the fatigue of forklift drivers.

5 Ergonomics Improvement and Simulation Results

5.1 Optimization and Improvement Direction Analysis

According to the above analysis, it can be seen that the forward driving and reversing driving positions are acceptable to most people. There are safety problems caused by environmental changes in the factory. In addition to the design problems in the forklift cab, the vision of drivers at work is also one of the main reasons for the accident.

The design is carried out based on the principles of safety and easy operation of forklift design. After further sorting out the operation process of the forklift truck, the specific design point was determined. Analysis for the perspective of forklift driver, and optimize the cab position [14].

5.2 Optimization of Forklift Truck Cab

Forklift operators in driving, especially in the state of loading goods to block the view, seriously affect the driving safety. But it is normal to load the goods on the forklift truck. Therefore, by changing the position of the cab seat and the mast of

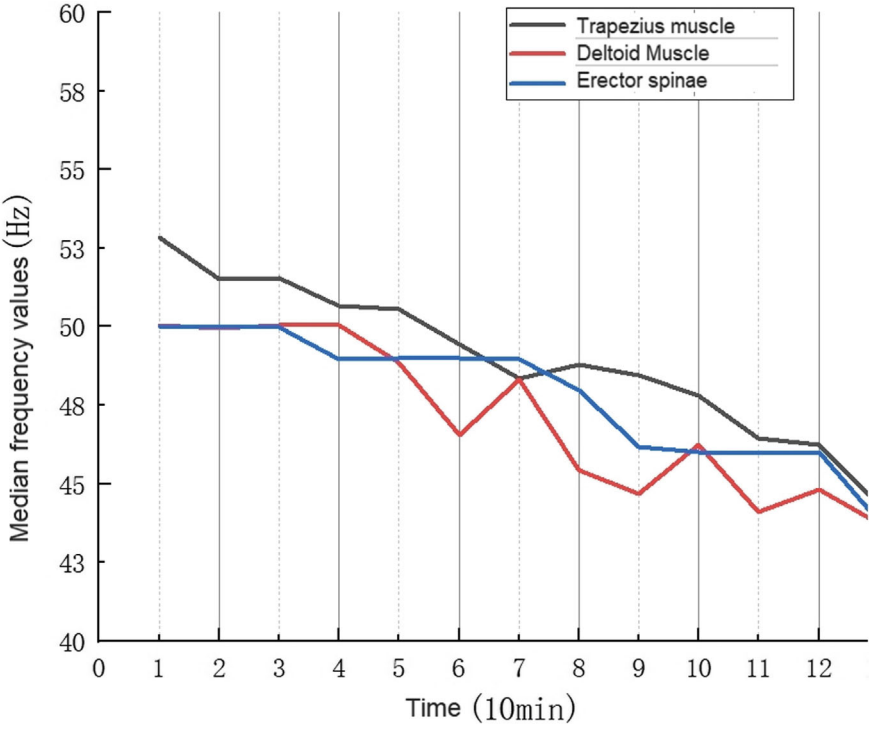


Fig. 13 Data trends of the MF values of the three muscles

the forklift to form a position deviation, so as to improve the blocking of the driver's vision when transporting goods (see Figs. 14 and 15).The improved forklift truck needs to redistribute the counterweight to achieve balance and stable operation.

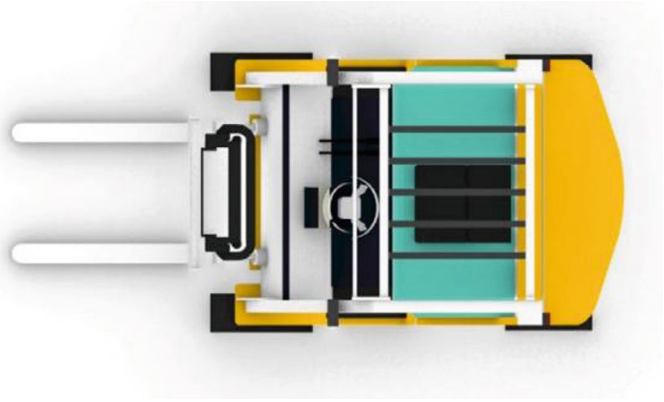


Fig. 14 Position before the cab optimization

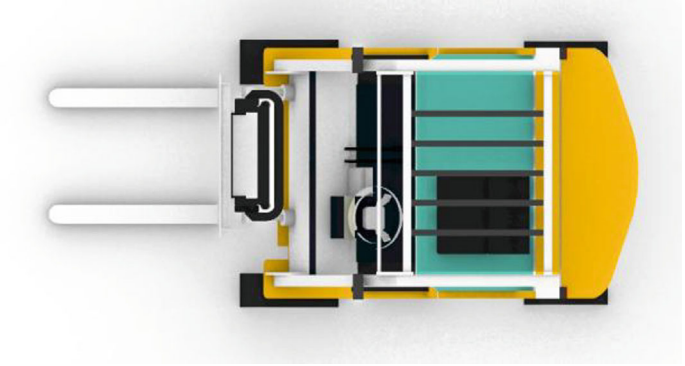


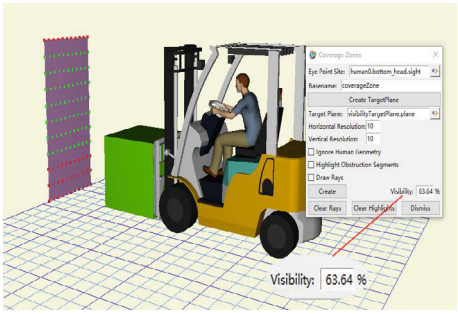
Fig. 15 Optimized position of the cab

5.3 Comparison of the Simulation Results

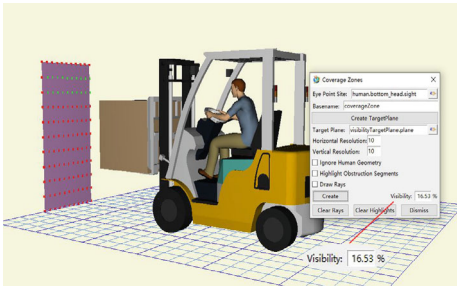
Using the Coverage Zones tool in Jack to analyze the operator view in the forklift cab. Coverage domain analysis is the study of those perspectives that are hindered by observed objects. The line of sight is projected directly from the eye onto a two-dimensional plane plate, and then analyzes whether the line of sight is blocked by an object when it reaches the plane. If the line of sight is not obstructed, the corresponding spot on the plate appears in green, otherwise in red. In this paper, the cargo lifting height is set to 700 mm; the number of test points in the horizontal and vertical direction is set to 10 * 10.During normal driving, the distance between the target plate and the forklift body is 2600 mm; the height of the plate is the same as the forklift height is 2195 mm; the width of the plate is 1255 mm. When reversing, the distance between the target plate and the forklift body is 1350 mm, and the height and width of the target panel are set the same as in the front.

Before the optimization, the forklift operator visually covered different areas in three different driving positions (see Fig. 16). At this point, when the goods don't lifted, the visibility of the driver's vision is 63.64%, (see Fig. 16a); When the cargo are lifted, the visual visibility of the forklift operator is 16.53% (see Fig. 16b); The visibility of the operator's view in reverse driving is 76.86% (see Fig. 16c).

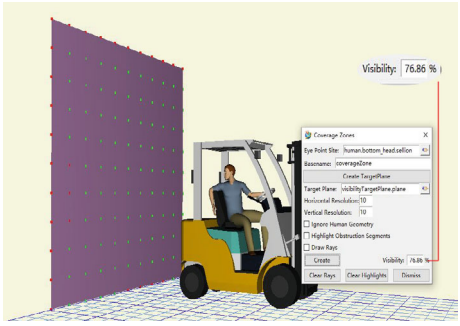
After the system improvement, the coverage domain analysis of the three cases (see Fig. 17). When the goods aren't lifted, the visibility of the driver's vision is 65.29% (see Fig. 17a), 1.65% higher than before the improvement; When the goods are lifted, the visibility of the forklift driver's view is 40.50% (see Fig. 17b), 23.97% higher than before the improvement; The visibility of the driver's view in reverse driving is 94.21%, (see Fig. 17c)17.35% higher than before the improvement [13].



(a) The goods don't raised

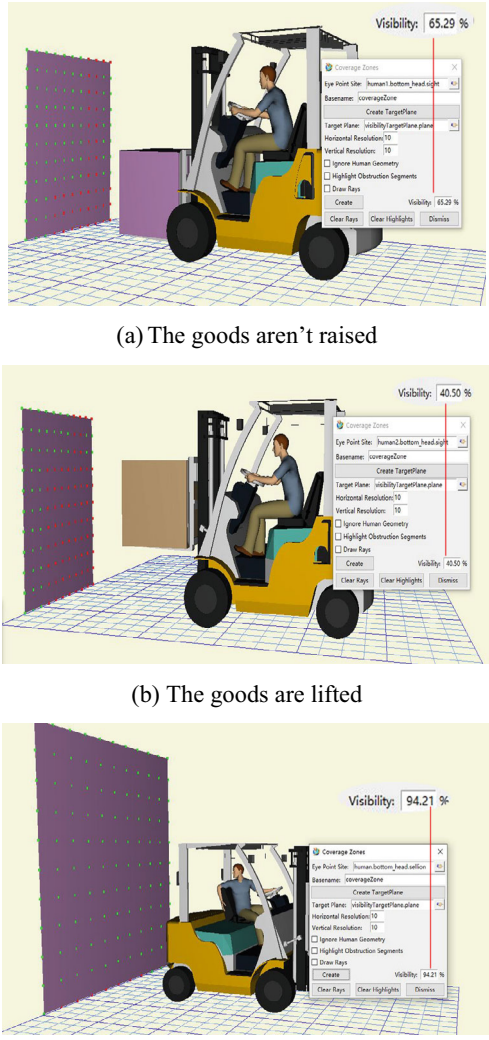


(b) The goods are lifted



(c) Driver reversing state

Fig. 16 Forklift drivers visibility coverage domain before system optimization



(a) The goods aren't raised

(b) The goods are lifted

(c) Driver reversing state

Fig. 17 Forklift drivers visibility coverage domain after system optimization

6 Conclusion

This paper mainly analyzes the ergonomic characteristics of forklift operators in normal driving and reversing driving. By changing the cab seat layout, the view coverage of the forklift driver in the carriage of goods is improved. Through optimization, the visibility of the driver is increased by 1.65% when the goods are not lifted; When the cargo are lifted, the driver's visibility increased by 23.97%; When

reverse driving state, the driver's vision visibility increased by 17.35%. The improvement of visual field visibility effectively reduces the probability of accidents among forklift drivers, and ensures the safety of personnel and property. Jack virtual simulation technology is used to improve the safety problems and vision problems of forklift drivers, improve the operation efficiency of driving operations and the safety of drivers, and provide valuable solutions for humanized design.

This paper mainly studies the upright position of the driver, and the situation of the operator in other positions is still insufficient. Future research will focus on the visual field of operators when their bodies are tilted left and right and their heads are rotated, as well as multi-muscle fatigue detection and prediction models.

References

1. Horberry T, Larsson TJ, Johnston I, Lambert J (2004) Forklift safety, traffic engineering and intelligent transport systems: a case study. Appl Ergon 35:575–581
2. Sharotry A, Jimenez JA, Méndez Mediavilla FA (2023) Forklift operator safety & productivity: a review of current research and future directions. In: Progress in material handling research, 16th Proceedings. Dresden Germany
3. Guo Z (2019) Study on the common safety hazards of forklifts in the factory. Equipment supervision 08:32–33
4. Industrial trucks-Safety rules for the application (2018) Operation and maintenance, Chinese national standard GB36507-2018
5. Daniel C, Priyadarshini DG (2019) An ergonomic forklift training program. Ergon Int J 3(4):000212
6. Ou J, Zheng Y, Zhao J (2019) Driving behavior analytics—based forklift cab appearance design. Packag Eng 40(14):128–134
7. Wang T, Wang B, Xu Y (2020) Design of raised table lift truck based on TRIZ theory. Pioneer Sci Technol Mon 33(02):79–82
8. Yang H, Liu Z (2021) Study on the optimal design of engineering vehicle operating console based on ergonomics. Internal Combust Engine Part 04:211–212
9. Hou S, Wang J, Liu J (2021) Visual dynamic analysis of forklift modeling and its image bionic design. Packag Eng 42(22):108–114
10. Niu J, Zhang L (2012) JACK human cause engineering foundation and application examples. Electronic Industry Press
11. Chen W,Yang G, Xu W (2020) Analysis of subway cab machine comfort based on JACK. South Agric Mach 51(21):189-191
12. Human size for Chinese adults (1988) The state bureau of quality and technical supervision, GB/T1000-1988
13. Zhu R (2022) Study on fatigue analysis of forklift operators under human machine system. Changchun University of Technology, Changchun
14. Miao Y (2020) Research on optimization design of engineering vehicle cab comfort. Yanshan University, Qinhuangdao

Open Access This chapter is licensed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License (http://creativecommons.org/licenses/by-nc-nd/4.0/), which permits any noncommercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if you modified the licensed material. You do not have permission under this license to share adapted material derived from this chapter or parts of it.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.



Modeling and Analysis of Mechanical-Braking Thermomechanical Coupling for Direct-Drive Hub-Motor



Kaikun Pei, Lijun Zhang, Dejian Meng, Wei Tian, and Zhuang Zhang

Abstract With the transformation of the automotive industry towards electrification, hub motor technology has become a hot research topic, especially direct drive hub motors, which have received widespread attention. When the electric wheel undergoes mechanical braking, a large amount of heat is generated, and the impact of this heat on the hub motor has not been fully understood. Therefore, this paper selects direct drive hub motors as the research object and focuses on exploring the temperature distribution and thermal deformation characteristics of electric wheels under mechanical braking conditions. By geometric modeling and parameterization of key components of the electric wheel, ABAQUS software is used to simulate temperature and deformation under emergency braking conditions. The simulation results show that the temperature rise of the brake disc is most significant, reaching up to 75 °C, while the temperature rise in the motor part is relatively small and can be ignored. In terms of deformation, the maximum deformation of the brake disc is 0.18 mm, while that in the motor part is smaller with a maximum rotor deformation of only 0.058 mm. This provides some reference for heat dissipation design for hub motors.

Keywords Braking thermomechanical coupling · ABAQUS · Temperature field · Deformation field · Finite element analysis

1 Introduction

In recent years, with the support and attention of the state, new energy vehicles, especially pure electric vehicles, have developed rapidly [1–5], and electric vehicles will be an important means of transportation for people in the future. Hub motor as one of the most core components of electric vehicles, its biggest feature is the drive,

K. Pei · L. Zhang · D. Meng · W. Tian (✉) · Z. Zhang
School of Automotive Studies, Tongji University, Shanghai, China
e-mail: 971013314@qq.com

© The Author(s) 2026

H. Saman et al. (eds.), *The 9th International Conference on Advances in Construction Machinery and Vehicle Engineering*, Lecture Notes in Mechanical Engineering,
https://doi.org/10.1007/978-981-96-5527-4_56

transmission and braking devices integrated into the hub, omit the clutch, transmission, transmission shaft, differential, transfer case and other traditional transmission components, due to the relatively high degree of integration, heat dissipation has become a key issue of the electric wheel.

Due to the high temperature of the working environment of the hub motor, especially in the pure mechanical braking, its internal temperature management becomes particularly important. Tang Chong's study used PrePoMax to build a finite element model and perform a heat transfer analysis to study the effect of a "ring disk" configuration at different speeds on the temperature of a permanent magnet [6]. In addition, Dong Haoxuan analyzed the influencing factors of energy recovery in the regenerative braking system of electric vehicles driven by four-wheel hub motors, and proposed an optimized parameter matching method to reduce the temperature rise during braking [7]. In Zhu Junyu's research, the deformation problem of a new fan-shaped axial magnetic hub motor which can retain the original braking system of the vehicle under high load is studied, and the loss suppression method is proposed [8]. Guo Lixiang proposed a cooling device which can be used for the effective heat dissipation of the motor inside the electric wheel, and verified that the device can ensure the stable operating temperature of the motor inside the electric wheel through simulation analysis [9]. In addition, Wang Zengquan simulated the temperature field distribution of disc brake through finite element software, which provided important basic data for the thermal management design of electric wheel [10]. In BoZhen Xie's research, a joint simulation of braking energy recovery including a supercapacitor model was built to analyze the energy recovery rate under different electrical/mechanical braking distributions, and the deformation problem during braking was discussed [11]. In addition, Jasper Nonneman evaluated various permanent magnet motor rotor cooling techniques by establishing a centralized parameter model and found that rotor jet cooling is the most effective method to reduce the temperature of permanent magnets, meeting the demand for high power density in electric vehicles [12]. Wei Liang studied the heat dissipation issues of high-torque hub motors used for heavy-duty truck wheel-end drive. By designing a water channel cooling structure, the study analyzed the impact of water channel structure forms, coolant flow rate, and coolant medium on motor cooling effectiveness and pressure drop using CFD-based methods [13]. Muhammad Hasan Albana researched the thermal characteristics of hub motors for electric scooters and discussed the effects of different speeds (250, 500, 750 rpm) on motor temperature [14].

It can be seen that at present, although some scholars have studied the brake temperature of hub motors [15–19], there is a lack of research on the effect of pure mechanical braking on electric wheels. In the mechanical braking condition, the electric wheel not only has to bear the thermal load caused by the braking, but also has to deal with the thermal deformation caused by it. Therefore, this paper takes direct-drive hub motor as the research object and carries out modeling and analysis of temperature-thermal deformation coupling (Thermomechanical Coupling). Firstly, a geometric model of direct drive hub motor is established, and then based on ABAQUS finite element analysis software, the influence of heat generated by

brake discs under pure mechanical braking conditions on key components of electric wheels is analyzed.

2 Heat Transfer Analysis

In the microscopic case, if there is a temperature difference between various points inside the object, heat transfer will occur; In the macro case, if there is a temperature difference between the object and the object, heat transfer will also occur, which is heat transfer, there are three main ways to transfer heat, respectively, heat conduction, heat convection and heat radiation.

Heat transfer occurs between the internal molecules of the object with a temperature difference or between two objects in contact with each other with a temperature difference. The essence of heat transfer is the thermal motion of microscopic particles. However, heat conduction can only transfer heat from the high-temperature part inside the object to the low-temperature part by the thermal motion of molecules, or transfer heat from the high-temperature object to the low-temperature object between two objects in contact with each other. The following formula is the basic formula of Fourier's heat conduction law.

$$Q = -\lambda S \frac{dT}{dn} \tag{1}$$

where Q is the heat (W) passing through a certain surface in unit time; S is the area of heat passing through the surface (mm^2); λ is the thermal conductivity of the surface material through which the heat passes (mW/mm K); $\frac{dT}{dn}$ is the temperature gradient from high to low temperature (/mm) °C.

Convective heat transfer generally occurs between the fluid and the fluid or between the fluid and the solid, when the relative displacement occurs between the parts of the fluid, the heat transfer process caused by the mixing of cold and hot fluids is called convective heat transfer. In this paper, it is the convective heat transfer between the solid and the fluid, and when the convective heat transfer occurs, it is usually accompanied by heat conduction. The basic formula for calculating convective heat transfer is Newton's cooling formula, as shown below.

$$Q = Ah(t_s - t_f) \tag{2}$$

where Q is heat transfer (mJ); A is the convective heat transfer area (mm^2); h is the convective heat transfer coefficient (mW/(mm^2 k)); t_s is temperature on the outer surface of the solid (K); t_f is the temperature of the convective heat transfer gas with the solid surface (K).

Radiation is a way of transmitting energy to the outside world through the emission of electromagnetic waves, and thermal radiation is a way of transmitting energy to the outside world through the emission of electromagnetic waves because of heat,

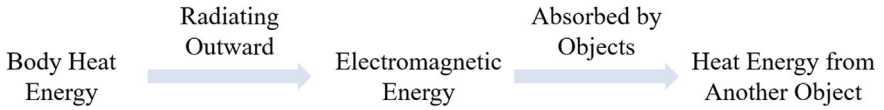


Fig. 1 Heat radiation

as shown in Fig. 1. The general consideration of thermal radiation is to consider the thermal radiation between two objects, and the exchange of heat between two objects in the form of radiation.

$$Q = \varepsilon A \sigma F_{12}\left(T_1^4 - T_2^4\right) \tag{3}$$

where Q is heat emission (mJ), ε is the radiant rate of the object, which is related to the type of the object and the surface shape, σ is the Stefan-Boltzmann constant, F_{12} is the form factor from face 1 to face 2, T_1 is the temperature of face 1(K), T_2 is temperature of face 2(K).

In this article, make the following assumptions about the electric wheel model:

(1) The influence of the polar arc coefficient on the temperature field of the electric wheel is ignored.
(2) Assume that the loss of the motor is evenly distributed in all parts, the iron loss is mainly distributed on the stator core, the copper loss is distributed on the copper wire winding, and the permanent magnet loss is distributed on the magnetic pole of the permanent magnet.
(3) Ignoring the motion of the electric wheel, the electric wheel is static equivalent in this paper.
(4) Ignore the mechanical loss of the electric wheel, only consider the stator iron loss, permanent magnet loss, copper wire winding copper loss, and the resulting loss is converted into heat.
(5) Because the simulation time is relatively short, ignore the influence of thermal radiation on the temperature field of the electric wheel.

3 Build the Finite Element Model

3.1 Convective Heat Transfer Coefficient

In this paper, the temperature field and thermal deformation of the global model of the electric wheel are simulated. There are no adiabatic boundary conditions, only convective heat transfer boundary conditions and heat conduction boundary conditions. The main materials of electric wheel parts in this paper include: stator core material silicon steel sheet DW315, permanent magnet pole material NdFeB (NdFeb), brake disc solid gray cast iron, hub motor shell material aluminum alloy,

Table 1 Thermodynamic parameters of electric wheel materials

	Density ρ (t/mm^3)	Specific heat c (mJ/t K)	Thermal conductivity λ(mW/mm K)	Coefficient of thermal expansion a(1/k)	Modulus of elasticity (Gpa)	Poisson's ratio
Brake disc	7.31E−009	700,000,000	43	1.15E−005	135,000	0.23
Housing	2.78E−009	880,000,000	25.2	2.1E−007	69,000	0.35
Rotor	7.44E−009	518,000,000	40	1.37E−007	200,000	0.26
Permanent magnet	7.4E−009	501,000,000	7.7	4E−006	160,000	0.24
Coil	8.954E−009	383,100,000	386	1.77E−007	110,000	0.34
Stator core	7.44E−009	518,000,000	40	1.37E−007	200,000	0.26

hub motor outer rotor material silicon steel, and hub motor coil winding material Cu (copper), the specific physical properties are shown in Table 1.

Each cooling surface of the electric wheel model in this paper is: the outer surface of the brake disc, the outer surface of the shell, the end face of the shell, the end face of the rotor, the inner surface of the rotor, the end face of the permanent magnet, the end face of the air gap, the end face of the winding, the end face of the stator core and other surfaces. The convective heat dissipation coefficient of these surfaces can be calculated by the related theory of heat transfer. The results are as Fig. 2.

3.2 Calculation of Heat Flux

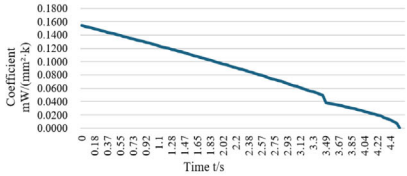
The heat flux of the electric wheel in the braking process can be calculated according to the following formula, according to the speed, deceleration, braking pressure and other parameters under different working conditions, the braking time of the car and the heat flux of the friction surface of the input brake disc block under different working conditions.

$$I \cdot \frac{a}{R_c} = M = 2 \cdot P \cdot S_P \cdot \mu \cdot R_d \tag{4}$$

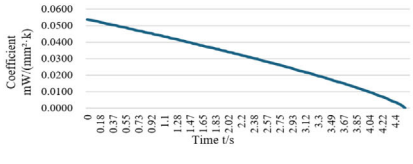
where I is 1/4 electric vehicle moment of inertia (kg m^2), A is vehicle deceleration (m/s^2), R_c is tire equivalent radius (mm), M is 1/4 electric vehicle braking torque (Nm), P is braking pressure Mpa, S_P is piston area (mm^2), μ is friction coefficient of brake disc block friction pair, R_d is effective radius of brake disc (mm).

The model parameters related to heat flux calculation in this paper are shown in Table 2.

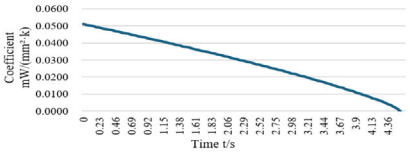
In this experiment, ABAQUS software was used to simulate the temperature field distribution and deformation of the key part of the electric wheel during the braking



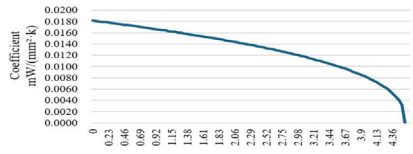
(a) Surface of Brake Disc



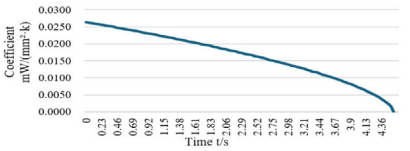
(b) Surface of Brake Motor



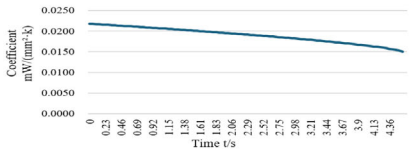
(c) Surface of the Outer Rotor of the Motor



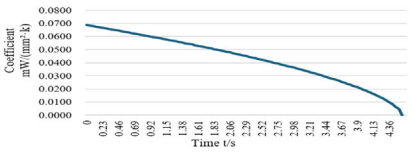
(d) Permanent Magnet



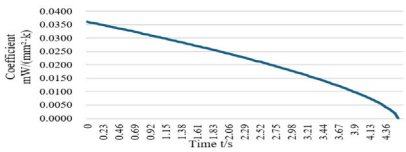
(e) Outer Surface of the Winding



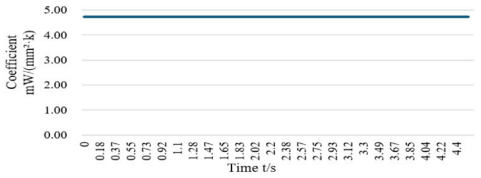
(f) Inner Surface of Winding



(g) Outer Surface of the Winding



(h) Inner Surface of Wind-ing



(i) Cooling Water

Fig. 2 Convective heat transfer coefficient

Table 2 Geometric parameters of the model

Model geometric parameters	Parameter values
Piston acting area (mm^2)	4247
Brake friction coefficient	0.32
Effective radius of brake disc (mm)	160
Tire equivalent radius (mm)	318
1/4 car moment of inertia (kg m^2)	75
Area of friction pair of single disc block (mm^2)	4536.35

Table 3 Parameters of primary braking conditions

Name of working condition	Operating condition parameter value
Initial speed (km/h)	100
Initial speed (r/min)	834
Final speed (km/h)	0
Final speed (r/min)	0
Brake pressure (MPa)	3.3
Deceleration (m/s^2)	6.12
Braking time (s)	4.54

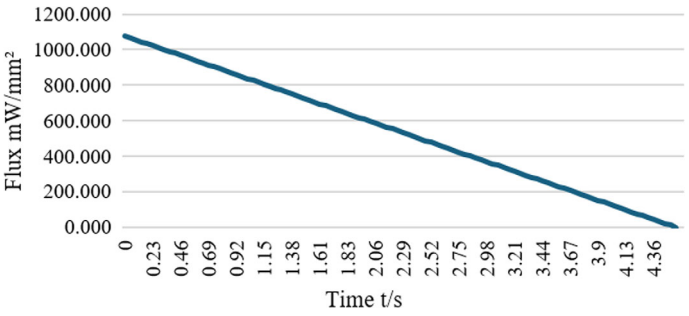


Fig. 3 Heat flux of mechanical brake

process of single brake under a single brake condition. The specific parameters of the single brake condition are shown in Table 3.

As can be seen from the braking parameters, the braking pressure is constant, so the braking deceleration is constant, the electric wheel does uniform deceleration, the electric wheel speed changes linearly with time, and the heat flux can also change linearly with time from the formula. According to the formula, it can be calculated that the value of the heat flux of the single brake disc input brake disc disk changes with time in a single braking condition, and the calculation results are shown in Fig. 3.

4 Model Results and Analysis

The following will analyze the simulation results of single friction braking conditions from two aspects of temperature and thermal deformation.

4.1 Temperature Field Analysis

Figure 4 shows the cloud picture of temperature field distribution under single friction braking condition. The key parts of the electric wheel, such as brake disc, motor shell, motor rotor, motor permanent magnet, winding and stator core, are analyzed in combination with the cloud image.

Figure 5 shows the temperature field cloud diagram of the key components of the electric wheel system, from which it can be seen that the temperature of the disc surface of the brake disc is relatively high, and the temperature from the disc surface to the disc surface gradually decreases, and the highest temperature of the disc reaches 75 °C. The overall temperature rise of the shell is not high, and the maximum temperature rise of the highest temperature rise node is only 2.4 °C, mainly because the braking time is relatively short, only 4.54 s, and the heat generated by braking is too late to transfer from the brake disc to the motor housing. It can be concluded that the influence of the heat of the brake disc on the temperature rise of the motor housing can be ignored during short braking time. The temperature of the rotor in the part connected with the brake disc is relatively high, and the temperature decreases rapidly along the axial direction. Moreover, because the braking time is relatively short, the temperature rise of the rotor is not very high. With the increase of the axial distance, the temperature rise decreases rapidly, mainly because the braking time is relatively short, a large amount of heat on the brake disc, and the heat transferred to the rotor is a small part. It can be concluded that part of the heat can be transmitted to the rotor and the connecting part of the brake disc in a short time, which has a great influence on the temperature rise of the rotor. The temperature of the permanent magnet rises below 1 °C, because the distance from the brake disc is farther, the heat of the brake disc is transmitted, and it needs to pass through bolts, rotors and other components. The reason for the relatively high temperature in the figure is that

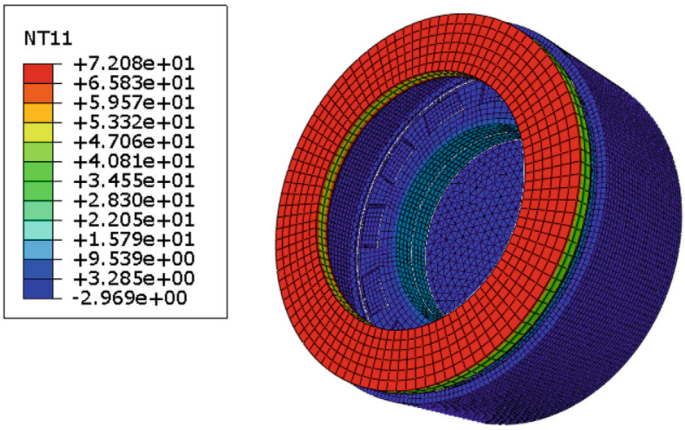


Fig. 4 Temperature field cloud image of electric wheel

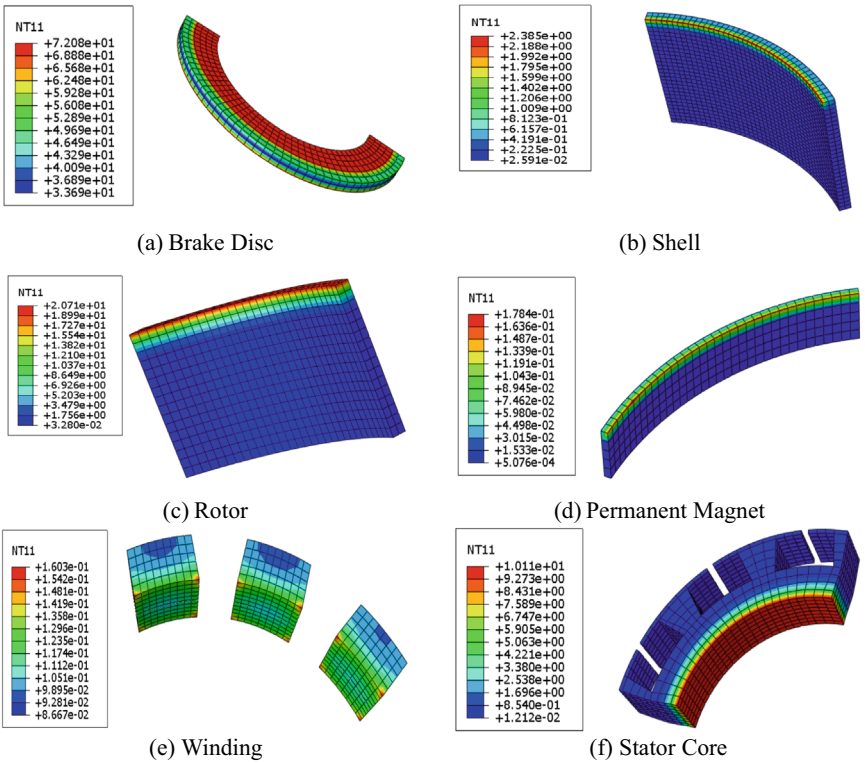


Fig. 5 Temperature field cloud image of key components of electric wheel

the permanent magnet is in contact with the air gap, and the temperature of the air gap rises faster at the ambient temperature of 20 °C, so the temperature of the part connected with the permanent magnet and the air gap is higher than that of the part connected with the rotor. The temperature rise of the winding has basically no effect, and the temperature rise is very low. The temperature rise in the stator core water jacket is higher, and the temperature rise here is not because the heat of the brake disc is transmitted to the stator core, but because the convective heat transfer intensity is relatively high here, which is affected by the environment of 20 °C. And the stator core from the farthest heat is difficult to transfer in 4.54 s, so it can be considered that the brake disc heat has no effect on the stator core.

4.2 Thermal Deformation Analysis

Through the analysis of the temperature field of the electric wheel, it can be seen that the temperature rise is only 20.3 °C in the connection part of the brake disc

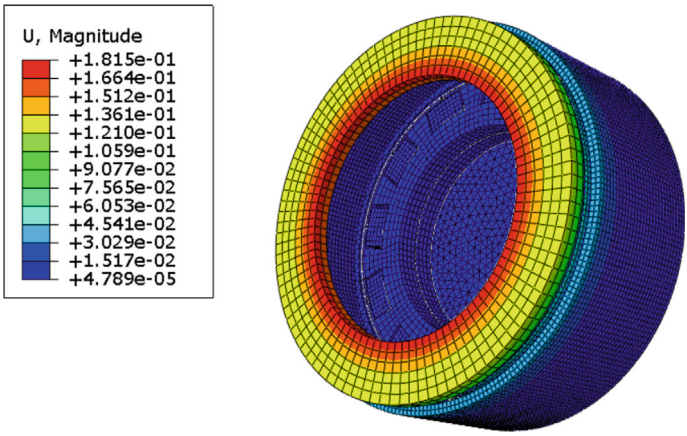
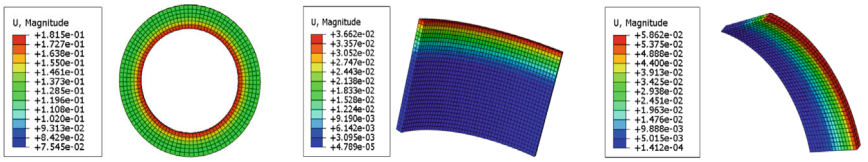


Fig. 6 Cloud image of thermal deformation field of electric wheel

and the outer rotor, and the temperature rise range is relatively small. Therefore, the deformation of the electric wheel studied in this paper during the single friction braking process mainly occurs in the brake disc, rotor and shell parts. Figure 6 is the cloud image of the thermal deformation of the electric wheel.

Figure 7 shows the thermal deformation cloud map of key components of the electric wheel system, from which it can be seen that the brake disc produces different deformations at different radial radii after a single friction braking. The smaller the radius, the larger the shape variable, the maximum deformation of the inner surface of the brake disc is 0.15 mm, and the maximum shape variable of the outer surface of the brake disc is 0.18 mm. The deformation of the inner surface of the shell gradually decreases and tends to 0 mm indefinitely. The maximum deformation occurs at the end face of the shell near the brake disc. The main reason is that the heat generated by the friction of the brake disc is transmitted to the end face of the shell, resulting in a large deformation of the end face of the shell, and the maximum deformation is 0.036 mm. With the increase of the distance between the rotor and the brake disc, the axial shape variable of the rotor gradually decreases. The place where the rotor deformation is the largest is connected to the brake disc, and the maximum deformation is 0.058 mm, which is smaller than that of the brake disc. Permanent magnet, winding, stator core and brake disc, rotor and other parts of the deformation is relatively small, can be ignored.



(a) Brake Disc (b) Shell (c) Rotor

Fig. 7 Thermal deformation cloud image of key components of electric wheel

5 Conclusion

This paper determines the convective heat transfer coefficient of each convective heat transfer surface and the heat flux of the disc under mechanical braking conditions, and does the simulation analysis under single friction braking conditions, and can draw the following conclusions: the temperature rise of the electric wheel brake disc is the highest, 75 °C, followed by the rotor and housing connected with the brake disc, the temperature rise of 20.3 and 2.4 °C. The temperature rise of the motor part is relatively low and can be ignored. The largest part of the deformation is 0.18 mm at the minimum radius of the brake disc, the deformation of the motor part is relatively small, the largest deformation is the rotor connected to the brake disc, the maximum deformation is 0.058 mm, it can be considered that under the single friction braking condition, the motor deformation is not affected.

Acknowledgements Project is supported by the Natural Science Foundation of Chongqing [CSTB2023NSCQ-MSX0063].

References

1. Zhou Y, Pan Y (2019) Financial subsidies and tax relief: an analysis of new energy vehicle industry policy from the perspective of transaction costs. Manag World 35(10):133–149. https://doi.org/10.19744/j.carol carroll nki. 11-1235/f 2019.0137
2. Ma J, Liu X, Chen Y et al (2018) China new energy vehicle industry and technology development status and countermeasures. China J Highw Transp 31(08):1–19
3. Minggao O (2016) Research and development and prospect of china's new energy vehicles. Sci Technol Rev 34(06):13–20
4. Yahong Z, Yulu P, Shiyi C et al (2015) Government support and development of new industries: a case study of new energy. Econ Res J 50(06):147–161
5. Ouyang M (2006) China's energy saving and new energy vehicle development strategy and countermeasures. J Automob Eng (4):317–321. https://doi.org/10.19562/j.carol carroll hinasae QCGC. 2006.04.001
6. Shang C (2023) Based on a wheel hub motor braking heat transfer study PrePoMax. J Mech Des 40(4):99–104. https://doi.org/10.13841/j.carol carroll nki JXSJ. 2023.04.015
7. Dong H (2018) Research on Regenerative braking system of electric vehicle driven by four-wheel hub motor. Chang'an University

8. Zhu J (2023) Research on multi-physical field analysis and loss suppression of integrated wheel motor with drive and brake. Yantai University. 10.27437/, dc nki. Gytdu. 2023.000399
9. Guo L (2020) Integration of electric cars, electric wheel design optimization and temperature field analysis. Shenyang University of Technology. 10.27323/dc nki. Gsgyc. 2020.000342
10. Wang Z, Shen Y, Jin C et al (2012) Articulated electric wheel brake continuous braking temperature field analysis. J Coal Mine Mach 2012(01):110–112. https://doi.org/10.13436/j.m. KJX 2012.01.030
11. Xie B (2016) Simulation and experimental research on brake energy recovery of wheel motor. Zhejiang University, 2017. Uikey R, Lone HR, Agarwal A (2024) Indian traffic sign detection and classification through a unified framework. IEEE Trans Intell Transp Syst 25(10):14866–14875. Author F, Author S (2016) Title of a proceedings paper. In: Editor F, Editor S (eds) Conference 2016, LNCS, vol 9999. Springer, Heidelberg, pp 1–13
12. Nonneman J, Van der Sijpe B, T'Jollyn I et al (2021) Evaluation of high performance rotor cooling techniques for permanent magnet electric motors. In: 2021 IEEE international electric machines and drives conference (IEMDC). IEEE, pp 1–7
13. Liang W, Mei L, Deng J et al (2023) Heat dissipation investigation of a high torque hub motor. In: 2023 26th international conference on electrical machines and systems (ICEMS). IEEE, pp 1061–1066
14. Albana MH, Guntur HL, Putra ABK (2023) The effect of rotational speed on the thermal characteristics of a wheel hub motor on an electric scooter. In: 2023 international seminar on intelligent technology and its applications (ISITIA). IEEE, pp 749–754
15. Li Y, Wang Y, Zhang T et al (2023) Parameter identification approach using improved teaching and learning based optimization for hub motor considering temperature rise. Metrol Measurement Syst 30(1)
16. Li G, Bai Q, Shi W (2024) Energy recovery of four-wheel hub motor driven vehicle in steering-braking condition. In: Proceedings of the institution of mechanical engineers, Part D: J Automob Eng:09544070241256711
17. Wang D, Yang G, Luo Y et al (2023) Optimal design and stability control of an automotive magnetorheological brake considering the temperature effect. Smart Mater Struct 32(2):025020
18. Zhu X, Yang Z, Wu J et al (2024) Multiphysics dimension-reduced optimization of a PM hub motor based on electromagnetic-thermal mesh-shared network model. IEEE Trans Transp Electrification
19. Wang Q, Li R, Zhao Z et al (2023) Temperature field analysis and cooling structure optimization for integrated permanent magnet in-wheel motor based on electromagnetic-thermal coupling. Energies 16(3):1527

Open Access This chapter is licensed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License (http://creativecommons.org/licenses/by-nc-nd/4.0/), which permits any noncommercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if you modified the licensed material. You do not have permission under this license to share adapted material derived from this chapter or parts of it.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.



Dynamics Modeling and Analysis of a Single Flexible Link-Flexible Joint Manipulator



Jianyuan Wang and Shanying Jiang

Abstract This paper presents a dynamic model of a single-link flexible link-flexible joint manipulator system. The Lagrange method and the assumed mode method are combined in the modeling process, and multiple factors affecting the motion of the flexible manipulator system are considered, such as the coupling effect of link flexibility and joint flexibility, gravity, input torque, and the influence of the mass and inertia of the load and rotor on the system. Subsequently, the derived model is numerically simulated by MATLAB using the fourth-order Runge–Kutta method.

Keywords Manipulator · Nonlinear-vibration · Dynamic modeling

1 Introduction

As the scale of social production continues to expand, various industries have put forward higher requirements for work efficiency, precision, higher work intensity and adaptability to complex environments. In order to meet these needs, automation continues to improve, and manipulators are widely used in many fields.

For a long time in the past, rigid manipulators have occupied the mainstream position. This type of manipulator is often made of rigid and heavy materials, but it also leads to defects such as large mass, large volume, high cost, low load-to-weight ratio, slow speed, and difficulty in operating in a large space, which has become a bottleneck in its development and gradually cannot meet the needs of use. Compared with rigid manipulators, flexible manipulators have the characteristics of light weight, fast response, high load-to-weight ratio, low cost, low energy consumption, and suitability for operation in large spaces. They have been widely used: in the aerospace

J. Wang · S. Jiang (✉)
Faculty of Intelligence Technology, Shanghai Institute of Technology, Shanghai, China
e-mail: jiangshanying@sit.edu.cn

J. Wang
e-mail: 226091174@mail.sit.edu.cn

© The Author(s) 2026

H. Saman et al. (eds.), *The 9th International Conference on Advances in Construction Machinery and Vehicle Engineering*, Lecture Notes in Mechanical Engineering,
https://doi.org/10.1007/978-981-96-5527-4_57

field, flexible robotic manipulators are widely used for docking, maintenance, experiments, transportation, and astronaut extravehicular activities [1, 2]; in the medical field, flexible manipulators are used for detection and surgery [3, 4]; in the logistics field, flexible manipulators are also used for sorting [5], In the military field, flexible manipulator systems are used to control weapon systems [6].

However, due to the flexible characteristics of the flexible manipulator, various problems may arise during operation: in reciprocating motion, elastic deformation seriously affects the accuracy of the manipulator's trajectory tracking; in rapid motion, elastic deformation will cause nonlinear vibration. Nonlinear vibration not only affects the accuracy and stability of the flexible manipulator system, but also aggravates structural fatigue, causes safety problems, and affects work efficiency. For example, the Expedition 1 launched by the United States in 1958 suffered from energy loss in the system due to the nonlinear vibration of the four whip antennas, resulting in attitude instability [7]; the manipulator used in the space station needs nearly ten hours to wait for the residual vibration to disappear before the manipulator can be used to perform the next task [8]. Therefore, in theoretical research, constructing a dynamic model that can better reflect the actual motion situation and designing a more effective control strategy have become two research focuses. However, due to the high coupling and infinite-dimensional characteristics of the dynamic model of the flexible manipulator, it is quite challenging to establish an accurate dynamic model and design a controller.

The flexible manipulator is an infinite-dimensional system. It is difficult to accurately reflect the actual situation of the flexible robotic arm system dynamics modeling. Therefore, it needs to be simplified to facilitate subsequent calculations.

Common modeling methods include Lagrange method, Hamilton principle, Newton–Euler method, etc. The first two methods are modeled from the perspective of energy, and the latter is based on the principle of force system balance.

Generally, in research, discretization methods are also used to simplify infinite dimensions to finite dimensions to reduce the difficulty of modeling and calculation. Commonly used discretization methods mainly include assumed mode method and finite element method etc. Assumed mode method is based on the Rayleigh–Ritz method. Through the modal truncation technology, the sub-modes of each structure in the flexible manipulator system are integrated into the mode of the entire system; the finite element method splits the infinite-dimensional system into a finite-dimensional unit set, which is represented by elastic units, rigid nodes, stiffness and other units.

In addition, due to the complexity of the flexible manipulator system, a simplified model is usually used in modeling. For the link part, there are several beam theories used to describe the deformation of flexible manipulators, such as Timoshenko beam and Euler–Bernoulli beam. Different simplified models also differ in the deformation factors considered, which are generally determined according to the characteristics of the research object. For flexible joints, the simplified model proposed by Spong [9] is often used to transform flexible joints into inertia-free linear springs.

Due to the convenience of calculation and analysis, the combination of Lagrange method and assumed modal method has been widely used in past studies and has developed into a complete and reliable modeling method. This paper also uses the

Lagrange-AMM method for modeling, and uses Euler–Bernoulli beam theory and equivalent spring assumption to simplify the model. Compared with previous studies, this paper not only considers the flexibility of the link and the flexibility of the joint, but also takes gravity into consideration. At the same time, in order to fit the actual movement of the flexible manipulator as much as possible, the influence of input torque, load, and the mass and inertia of the rotor are also considered in the modeling.

2 Modeling of Flexible Link-Flexible Joint Manipulator

2.1 Model Introduction

The structure of a single-link flexible link-flexible joint manipulator is shown in Fig. 1, which mainly includes a flexible link and a rotor. The link adopts a simplified Euler–Bernoulli beam model, so it is assumed that the link is a slender link, only considering lateral deformation, and has a constant cross-sectional area and uniform material properties, and satisfies the small deflection assumption. The mass of the link is m_1, the length is L, the linear density is ρ, and the bending stiffness is EI. A load m_P regarded as a point mass is hung at the end of the link, and the moment of inertia is I_P. The rotor has a mass of m_2 and a moment of inertia of I_m. It is driven by the input torque τ. The kinetic energy of the rotor is mainly generated by the rotation of the rotor, so the moment of inertia of the rotor is symmetrical about its rotation axis. The flexible joint is equivalent to a massless linear spring with a stiffness coefficient of k, which connects the link to the rotor. The damping coefficient of the rotor to the link is C_1, and the damping coefficient of the support structure to the rotor is C_2.

X_0OY_0 and X_1OY_1 are the inertial coordinates and the tracking coordinates respectively, $\theta_1(t)$ and $\theta_2(t)$ are the angular positions of the link and the rotor respectively, and $D(x_1, t)$ and $D^*(L, t)$ are the deflections at any point P on the link and at the end point P* of the link respectively.

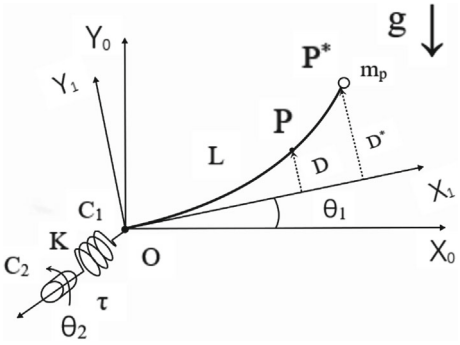


Fig. 1 Structure of single-link flexible link-flexible joint manipulator

2.2 Dynamics Modeling

By using the assumed modal method, the deflection can be simplified to the product of the length-dependent modal vibration shape function and the time-dependent modal displacement function. We write the deflection as

$$D(x_1, t) = \sum_{n=1}^{m} W_n(x_1) T_{\mathrm{n}}(t), \tag{1}$$

$W_n(x_1)$ is the nth order modal vibration function, $T_n(t)$ is the nth order modal displacement function, m is the total number of modes, and for the convenience of calculation, m = 2 in this paper.

Euler–Bernoulli beam can be expressed as [10]

$$EI \frac{\partial^4 D(x_1, t)}{\partial x_1^4} + \rho \frac{\partial^2 D(x_1, t)}{\partial t^2} = 0, \tag{2}$$

Through boundary conditions and orthogonality principle, the expression of modal vibration function can be written as

$$W_n(x_1) = N_n\{\cos(\beta_n x_1) - \cosh(\beta_n x_1) + r_n[\sin(\beta_n x_1) - \sinh(\beta_n x_1)]\}, \tag{3}$$

N_n is the normalization constant, N_n = m_1, and r_n is given by the following equation

$$r_n = \frac{\sin(\beta_n x_1) - \sinh(\beta_n x_1) + \frac{m_p \beta_n}{\rho}(\cos(\beta_n x_1) - cosh(\beta_n x_1))}{\cos(\beta_n x_1) + \cosh(\beta_n x_1) + \frac{m_p \beta_n}{\rho}(\sinh(\beta_n x_1) - sin(\beta_n x_1))}, \tag{4}$$

The parameter β_n is obtained by the following equation

$$\begin{aligned} &1 + \cosh(\beta_n x_1)\cos(\beta_n x_1) \\ &+ \frac{m_p \beta_n}{\rho}(\cos(\beta_n x_1)\sinh(\beta_n x_1) - \cosh(\beta_n x_1)\sin(\beta_n x_1)) \\ &- \frac{I_p \beta_n^2}{\rho}(\cosh(\beta_n x_1)\sin(\beta_n x_1) + \cos(\beta_n x_1)\sinh(\beta_n x_1)) \\ &- \frac{m_p I_p \beta_n^4}{\rho^2}(\cosh(\beta_n x_1)\cos(\beta_n x_1) - 1) = 0. \end{aligned} \tag{5}$$

In addition, the nth order natural frequency ω_n and β_n have the following relationship

$$\omega_n = \beta_n^2 \sqrt{\frac{EI}{\rho}}, \tag{6}$$

As shown in Fig. 1, the position vector of point P in the inertial coordinate system can be expressed as the product of the rigid transformation matrix T and the position vector r of P in the moving coordinate system, expressed as

$$P = Tr = \begin{bmatrix} \cos\theta_1 & -\sin\theta_1 \\ \sin\theta_1 & \cos\theta_1 \end{bmatrix} \begin{bmatrix} x_1 \\ D \end{bmatrix}, \tag{7}$$

Similarly, the position vector of the end point P can be expressed as

$$P^* = Tr^* = \begin{bmatrix} \cos\theta_1 & -\sin\theta_1 \\ \sin\theta_1 & \cos\theta_1 \end{bmatrix} \begin{bmatrix} L \\ D^* \end{bmatrix}, \tag{8}$$

Dynamic modeling requires calculating the kinetic energy and potential energy of the flexible link-flexible joint manipulator system. The kinetic energy of the system can be expressed as

$$T = T_{rotor} + T_{link} + T_{load}, \tag{9}$$

T_{rotor}, T_{link}, and T_{load} represent the kinetic energy of the rotor, the kinetic energy of the link, and the kinetic energy of the load, respectively, and they can be expressed as

$$T_{rotor} = \frac{1}{2} I_m \dot{\theta}_2^2, \tag{10}$$

$$T_{link} = \frac{1}{2} \rho \int_0^L \dot{P}^T \dot{P} dx_1, \tag{11}$$

$$T_{load} = \frac{1}{2} m_p \int_0^L \dot{P}^{*T} \dot{P}^* dx_1 + \frac{1}{2} I_p \dot{\theta}_1^2. \tag{12}$$

Similarly, the potential energy of the flexible link-flexible joint manipulator system can be expressed as

$$U = U_{joint} + U_{link} + U_{load}, \tag{13}$$

U_{joint}, U_{link}, and U_{load} represent the potential energy of the joint, the potential energy of the link, and the potential energy of the load, respectively, and they can be expressed as

$$U_{joint} = \frac{1}{2}k(\theta_2 - \theta_1)^2, \tag{14}$$

$$U_{link} = \frac{1}{2}EI \int_0^L \left(\frac{\partial^2 D}{\partial x_1^2}\right)^2 dx_1 + \rho \int_0^L G^T P dx_1, \tag{15}$$

$$U_{load} = m_p G^T P^*. \tag{16}$$

where $G^T = [0\ g]$, which is the transpose of the vector form of the gravitational acceleration.

The dynamic equations of motion of a single link flexible link—flexible joint manipulator can be derived utilizing the Lagrange's equations the equation can be obtained as

$$L = T - U, \tag{17}$$

$$\frac{\partial}{\partial t}\frac{\partial L}{\partial \dot{Q}} - \frac{\partial L}{\partial Q} + \frac{\partial E_D}{\partial \dot{Q}} = F. \tag{18}$$

where Q is the generalized coordinate and F is the generalized torque or generalized force. In this model, the generalized coordinates are the link angle θ_1, the rotor angle θ_2, and the link mode displacement T_1 and T_2. The F can be written as

$$F = \left[-C_1\left(\dot{\theta}_2 - \dot{\theta}_1\right) - C_1\dot{\theta}_2 - \tau\ C_1\left(\dot{\theta}_2 - \dot{\theta}_1\right)\ 0\ 0\right]^T, \tag{19}$$

The first term of the vector is the generalized torque of the rotor, the second term is the generalized torque of the link, and the third and fourth terms are the generalized torques corresponding to the mode displacement.

Substituting (9) and (13) into (18), after simplification and merging, the dynamic equation of the single-link flexible link flexible joint manipulator can be expressed as

$$I_m\ddot{\theta}_2 + k(\theta_2 - \theta_1) = -C_1\left(\dot{\theta}_2 - \dot{\theta}_1\right) - C_2\dot{\theta}_2 - \tau, \tag{20}$$

$$\boldsymbol{M}_{3*3}\begin{bmatrix}\ddot{\theta}_1\\ \ddot{T}_1\\ \ddot{T}_2\end{bmatrix}+\begin{bmatrix}a_1\\ a_2\\ a_3\end{bmatrix}\cos(\theta_1)$$
$$+\begin{bmatrix}f_1\left(\dot{\theta}_1,T_1,T_2,\dot{T}_1,\dot{T}_2\right)\\ f_2\left(\dot{\theta}_1,T_1\right)\\ f_3\left(\dot{\theta}_1,T_2\right)\end{bmatrix}+\begin{bmatrix}g_1(T_1,T_2)\\ 0\\ 0\end{bmatrix}\sin(\theta_1) \tag{21}$$
$$-\begin{bmatrix}k(\theta_2-\theta_1)\\ 0\\ 0\end{bmatrix}-\begin{bmatrix}C_1\left(\dot{\theta}_2-\dot{\theta}_1\right)\\ 0\\ 0\end{bmatrix}=0$$

where M is the inertia matrix, a_1, a_2, a_3 are the products of parameters including the gravity terms of the link and the rotor, f_1, f_2, f_3 are the functions of the interaction between the angular velocity of the link and the mode displacement, and g_1 is the function related to the gravity and modal displacement of the link and the load.

3 Simulation and Discussion

Based on the fourth-order Runge–Kutta method, the dynamic model was numerically simulated using MATLAB, and the response diagrams of the link angular position, rotor angular position, and link modal displacement without any interference were plotted. The specific parameter data used are shown in Table 1.

Figures 2, 3, 4 and 5 shows the response of the link angle, rotor angle, and link modal displacement. It can be observed that the amplitude of the rotor angle θ_2 is

Table 1 Flexible link-flexible joint manipulator parameters

Symbol	Parameter	Value	Unit
m_1	Mass of link	0.1	kg
L	Length of link	0.5	m
ρ	Mass density	0.2	kg/m
EI	Flexural rigidity	1	Nm^2
m_p	Mass of load	0.5	kg
Ip	Inertia of load	0.0025	kgm^2
m_2	Mass of rotor	1	kg
I_m	Inertia of rotor	0.08	kgm^2
C_1	Damping coefficient of link	0.016	Ns/m
C_2	Damping coefficient of rotor	0.008	Ns/m
g	Gravity	9.8	m/s^2
k	Stiffness constant	1	N/deg
τ	Input torque	0.3	Nm

significantly greater than the link angle θ_1, and the positions after entering the steady state are also different, which reflects the influence of joint flexibility on the motion of the robot system; the amplitude of the first-order modal displacement is much larger than the second-order mode, so it dominates the nonlinear vibration of the link.

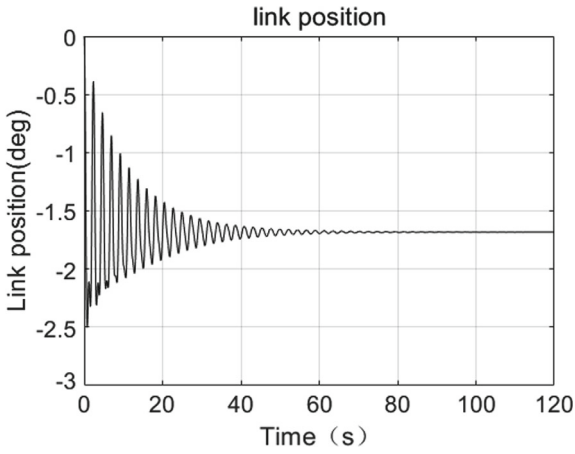


Fig. 2 Angular position of the link

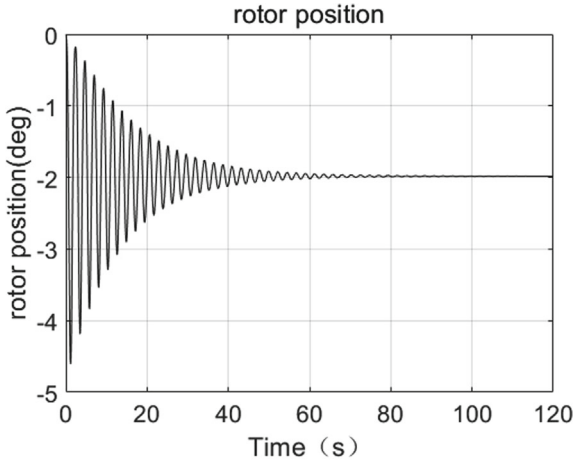


Fig. 3 Angular position of the rotor

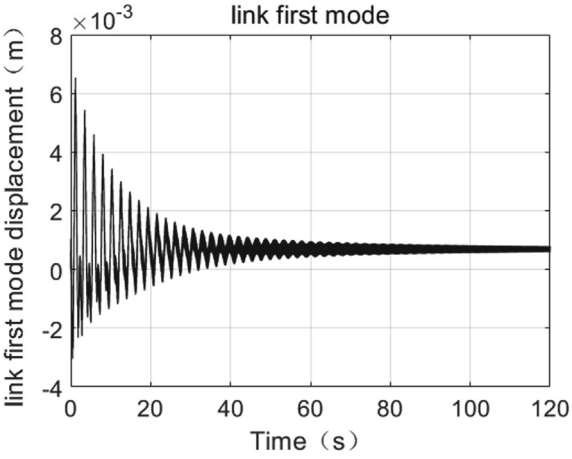


Fig. 4 Response of the first modal displacement

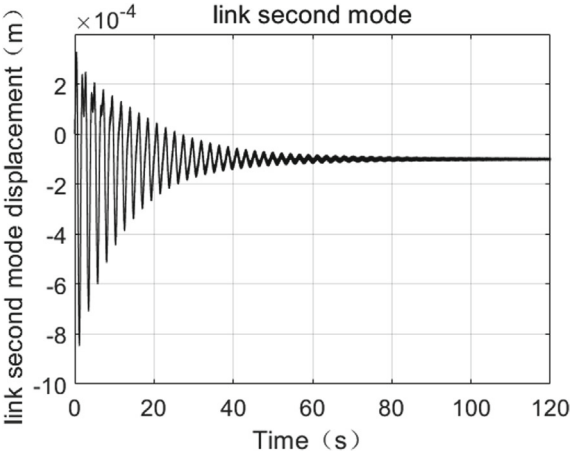


Fig. 5 Response of the second modal displacement

4 Conclusion

This paper uses the Lagrange-AMM method to dynamically model the manipulator system of a single-link flexible link-flexible joint. In the modeling process, the coupling effect of link flexibility and joint flexibility, gravity, and factors such as the mass and inertia of the load and rotor are considered. The fourth-order Runge–Kutta method is used to numerically simulate the response of the system. The simulation proves that the flexible manipulator system will produce nonlinear vibrations during movement, and the flexibility factor has a significant impact on the movement. The simulation results have a certain reference value for designing a single-link flexible manipulator system and designing a control scheme.

References

1. Shang D, Li X, Yin M, Li F (2022) Dynamic modeling and fuzzy adaptive control strategy for space flexible robotic arm considering joint flexibility based on improved sliding mode controller. Adv Space Res 70(11):3520–3539
2. Kraïem S, Rognant M, Biannic JM, Brière Y (2023) Dynamics and robust control of a space manipulator with flexible appendages for on-orbit servicing. CEAS Space J 15(5):681–700
3. Jayaswal K, Palwalia DK, Kumar S (2021) Analysis of robust control method for the flexible manipulator in reliable operation of medical robots during COVID-19 pandemic. Microsyst Technol 27(5):2103–2116
4. Kim J, de Mathelin M, Ikuta K, Kwon DS (2022) Advancement of flexible robot technologies for endoluminal surgeries. Proc IEEE 110(7):909–931
5. Cai S (2023) Robust adaptive control of flexible manipulator for express delivery sorting. https://doi.org/10.27040/d.cnki.ggzdu.2023.001142
6. Uyulan Ç (2021) Design and stability analysis of a robust-adaptive sliding mode control applied on a robot arm with flexible links. Vibration 5(1):1–19
7. Jiang D (2010) Nonlinear vibration research of the flexible manipulator in out-space environment. https://kns.cnki.net/kcms2/article/abstract?v=r9IaLYgXogU1Cc6xWNAQBKOmPhg44REeVhpV-wIfjn_9zrqBhE_qBeiQAOi6zwxrjt64uI_ND76UcyAOryjOYCFKNujVFmxvKWxuo5d_IYZF4b-XOq_4iluuo4PYBpvWL_pYs_21h8Vn7z0uHTNJO0yPwHHFkbQ4m8N4kZWEYY0YoqEjTtghOKJbVxA6MhHbSfY7NJCiZ0=&uniplatform=NZKPT&language=CHS
8. Alberts TE, Xia H, Chen Y (1992) Dynamic analysis to evaluate viscoelastic passive damping augmentation for the space shuttle remote manipulator system. J Dyn Syst Meas Control 114(3):468–475
9. Spong MW (1989) Modeling and control of elastic joint robots. Math Comput Model 12(7):912
10. Subudhi B, Morris AS (2002) Dynamic modelling, simulation and control of a manipulator with flexible links and joints. Robot Auton Syst 41(4):257–270

Open Access This chapter is licensed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License (http://creativecommons.org/licenses/by-nc-nd/4.0/), which permits any noncommercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if you modified the licensed material. You do not have permission under this license to share adapted material derived from this chapter or parts of it.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.



Design and Structural Optimisation of Vehicle and Its Equipment

ANSYS Contact Simulation Analysis Between Shed Tunnel Structure and Ultra-High Voltage Wire



Yue Wang, Xiangxiang Wang, and Hongsheng Zhang

Abstract In order to explore the contact safety problems of broken iced ultra-high voltage wires on the construction of shed tunnels, a project design sketch is first established based on actual engineering data and the research technical route is determined. Secondly, the iced wires, shed tunnels, and their contact parts are regarded as the research objects, and the APDL modeling method is used to establish a finite element analysis model. The SOLID185 element is utilized to establish the finite element analysis model of the contacted shed tunnel structure, and the BEAM189 element is applied to establish the ANSYS models of six iced lines. Arrange contact element CONTA175 on the iced wire and target element TARGE169 on the foam concrete roof of the shed tunnel which will be contacted by wires, and set the contact pair properties according to the actual working conditions. Contact between the wire and the shed tunnel is performed. Finally, results are obtained from different conditions and they are analyzed, which is beneficial to determine the dangerous contact area on the roof of the shed tunnel.

Keywords Broken wire · Shed tunnel · Fine and sparse mesh · ANSYS contact simulation · Contact force

1 Introduction

The three ultra-high voltage transmission lines above Shang-He-Hang Railway, namely Hu'an I and II lines, Jinsu line, and Fufeng line, were successively built and put into operation from 2010 to 2013. It is crucial to establish a protective concrete shed tunnel above Shang-He-Hang Railway to prevent wire breakage from affecting railway traffic. This concrete shed tunnel can reduce unnecessary economic losses caused by broken transmission line wires and ensure the safety of operating railway lines [1, 2]. Due to the different ground heights of different phase wires, it

Y. Wang · X. Wang · H. Zhang (✉)
School of Mechatronics Engineering, Harbin Institute of Technology, Harbin, Heilongjiang, China
e-mail: zhanghs@hit.edu.cn

© The Author(s) 2026

H. Saman et al. (eds.), *The 9th International Conference on Advances in Construction Machinery and Vehicle Engineering*, Lecture Notes in Mechanical Engineering,
https://doi.org/10.1007/978-981-96-5527-4_58

is necessary to specifically consider the contact of wire breakage on the underlying buildings in the case of three-phase transmission wires. Therefore, how to determine the contact force and contact area of the iced wire fracture on the shed tunnel above the railway has become the core issue in the design and construction process of the shed tunnel [3–5].

With the continuous development of CAE software, the use of advanced computer software has achieved significant results in simulation calculations of parts, components and even the entire equipment. ANSYS, as a powerful software, can analyze contact problems, which have been widely used in engineering. Especially in bridges, cars, ships [6] and so on in research on contacts with large obstacles [7–10], cases of using ANSYS for contact force simulation calculations are not uncommon. These literature mainly focus on solving the problems of complete machine discretization and stress concentration, and some of them also propose ANSYS numerical simulation of iced wires [8, 9]. However, limited research on the simulation focus on contact between shed tunnels and wires. Therefore, further exploration and improvement are needed on how to use ANSYS for simulation analysis of this problem.

The contact between high-voltage power lines and shed tunnels in practical engineering applications is regarded as the research object. Concerning the contact problem of iced wires, different elements will be selected. At the same time, simplify the structure of the shed and select key points. Based on this, according to the technical route in Fig. 1, ANSYS contact analysis is performed.

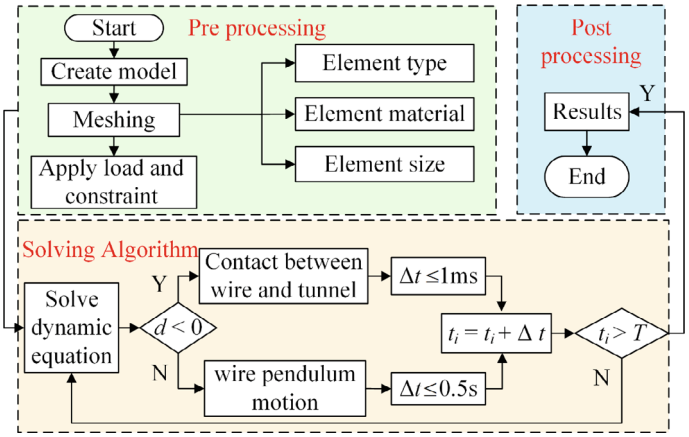


Fig. 1 Technical route

2 Establishment of Finite Element Model

2.1 Establishment of Geometric Models for Sheds and Wires

Construction of ANSYS Geometric Model for Shed Tunnel. As is well known, the accuracy of finite element calculations based on ANSYS is strongly correlated with the quality of the mesh. To ensure mesh high-quality, the geometric model of the shed pillar is simplified. The sectional diagram of geometric model of shed tunnel is shown in Fig. 2. Compared with the drawing model, ANSYS geometric model simplifies some details of the stormwater outlet and drainage ditch. Besides, in ANSYS analysis of 3D problems, hexahedral mesh can provide accurate geometric representation and high computational efficiency, so shed tunnels are divided by hexahedral mapping mesh. To facilitate the subsequent hexahedral mesh division, the shed tunnel geometry is modeled according to the black dividing line shown in Fig. 2.

When building the shed model, the shed model near the starting point of the shed is built firstly, and the rest of the shed tunnel structure is built by copy-paste. Because the structure of the shed is symmetrical, it is possible to build half of the structure of the shed during modeling, as shown in the left of Fig. 3. After that, the modeling work of the shed tunnel is completed by mirror copy, as shown in the right of Fig. 3.

Construction of ANSYS geometric model for wire. In the three lines, the ultra-high voltage transmission line is not only composed by transmission wires, but also includes insulator strings and spacer bars. In addition, each line contains several transmission wires rather than one transmission wire. The number of wires is determined by the number of wire splits in the spacer bar. Taking the Hu'an line as an example, the spacer bar model of the transmission wire is FJZ-840/630, and its shape is shown in Fig. 4a. The spacer bar model is simplified into ANSYS simplified model form as shown in Fig. 4b. The model is radial and has 8 splits, so the Hu'an line contains 8 wires and it shown in Fig. 5.

Because the contact force generated by the transmission wire under ice and wind load on the roof of the shed is greater than that without ice and wind load, this research

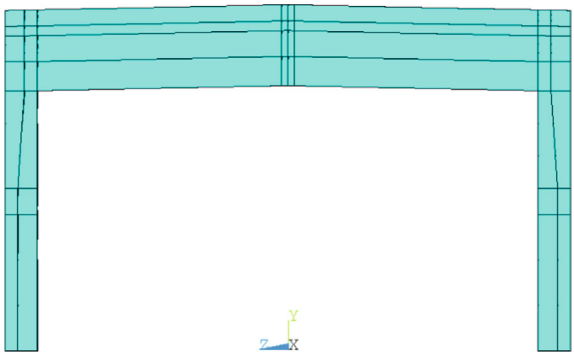


Fig. 2 Cross-section view of the shed tunnel geometry model in ANSYS

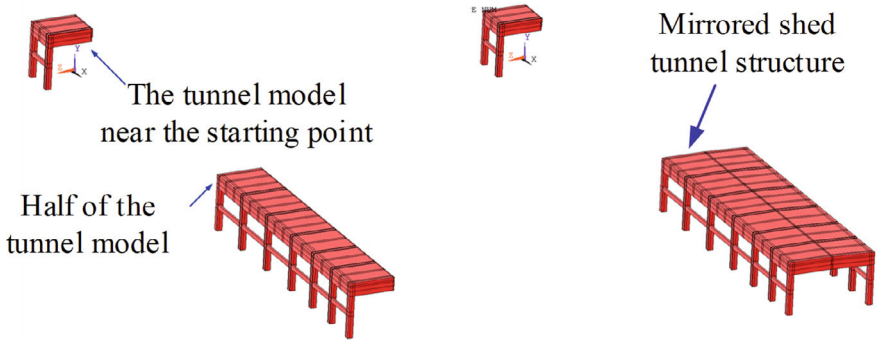
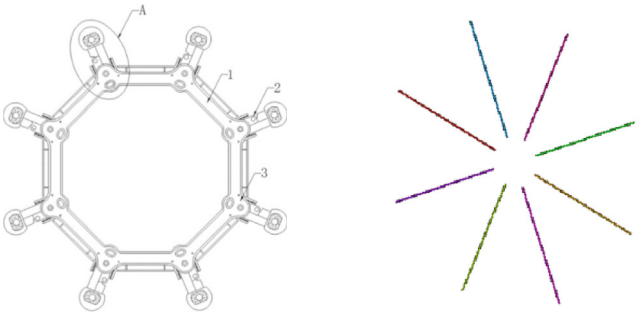


Fig. 3 ANSYS geometric model of shed tunnel



(a) Product shape diagram (b) ANSYS simplified model

Fig. 4 Spacer bar model product shape diagram and ANSYS simplified model

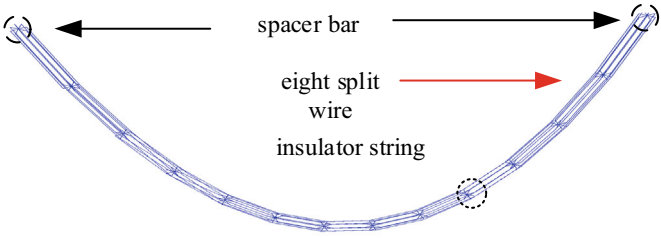


Fig. 5 ANSYS model diagram of Hu'an line

only models and analyzes the transmission wire under ice and wind load. Table 1 gives the geometric model parameters of three high voltage transmission lines.

Table 1 Geometric parameters of three wires

Name of lines	Hu'an line	Jinsu line	Fufeng line
Name of wires	LGJ-630/45	JL/G3A-90 0/40	ACSR-720/45
Wind yaw angle (°)	3.132	2.549	2.767
Number of wire splits	8	6	6
Wire spacing (m)	33	22	22
Span (m)	553	589	519
Pylon practical height (m)	105.4	66	55.8
Sag (m)	20.8	24.95	19.65
Circular section radius (mm)	33.6	16.9	15.7

2.2 Mesh Division

Mesh division of the shed tunnel model. The tunnel diagram, transition mesh, partial magnification of front view and partial magnification of side view are given in Fig. 6.For the shed tunnel model, solid element SOLID185 is selected to discretize it. The foam concrete roof of the shed tunnel is made of C30 foam concrete. For other components including columns, side beams, beams, main longitudinal beams, auxiliary longitudinal beams and cement roof are made of concrete. The foam concrete at the top of the shed tunnel has direct contact with the transmission wire, and the contact stress of the foam concrete is the key content concerned in this research. Therefore, the foam concrete roof needs to be discretized by a finer mesh to enhance the accuracy of the computational analysis. The foam concrete shed roof is divided by hexahedron and mapping mesh, and the maximum side length of the element is controlled below 0.25 m. For section made of concrete materials, the stress caused by the contact of falling transmission wires is not concerned in this study. The element mesh division of the above four structures can be sparse. The four-part structure is also divided into hexahedra by mapping mesh, and its maximum side length is controlled below 2 m. The auxiliary longitudinal beams and the cement roof are transitioned with a free mesh.

The schematic diagram in Fig. 6 shows the structure of the entire roof of the shed tunnel and the beam columns. The partial enlarged diagram of the transition mesh clearly shows the transition between the foam concrete roof and the concrete structure, highlighting the connection between the finer mesh of the roof and the sparser mesh of the beam. Combined with the side view local magnification and the front view local magnification, the mesh distribution characteristics of the shed structure are reflected.

Mesh division of transmission wires. In the mesh division of the transmission wire model, the BEAM189 element is selected to discretize wires. In order to better characterize the flexibility of the wire, the section moment of inertia of the BEAM189 element is reduced. In addition, the BEAM189 element is also used to discretize spacer bars and insulator strings in the ultra-high voltage transmission wires. The

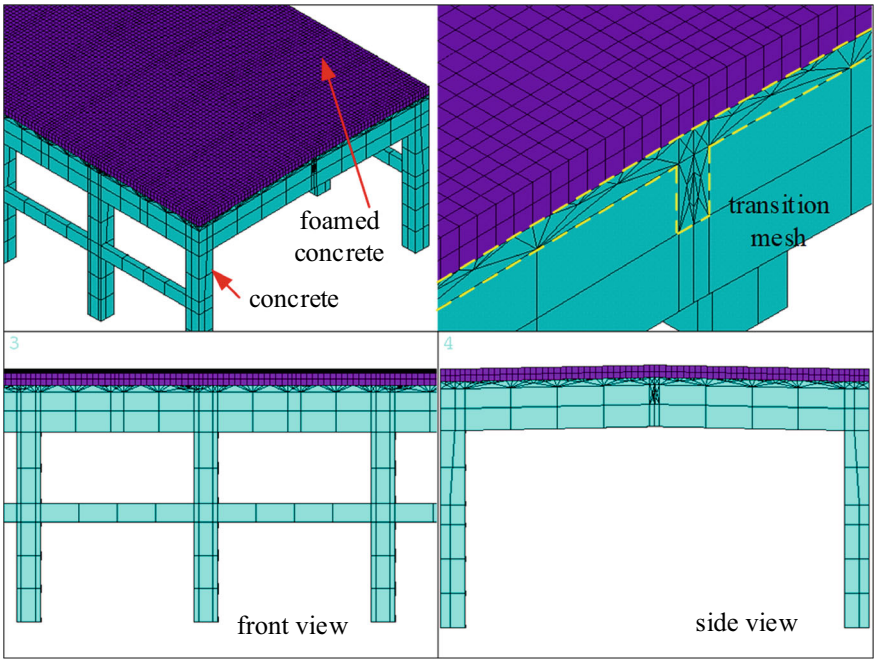


Fig. 6 ANSYS model mesh division effect of shed tunnels

material characteristic parameters used in tunnel wires and other constructions are shown in Table 2.

Table 2 Parameter of shed hangar tunnel's and wire's accessories

Component name	Material or type	Density (kg/m^3)	Poisson's ratio	Elastic modulus (MPa)	Dynamic elastic modulus (MPa)
Shed tunnel	Concrete	2500	0.2	–	37,800
	Foam concrete	700		–	480
Iced wire	LGJ-630/45	1458	0.3	63,000	–
	JL/G3A-90 0/40	1785		63,700	–
	ACSR-720/45	1720		63,700	–
Spacer bar	FJZ-840/630	7850	0.3	210,000	–
Insulator string	XS300H	7500	0.32	76,500	–
	X3-2A				–

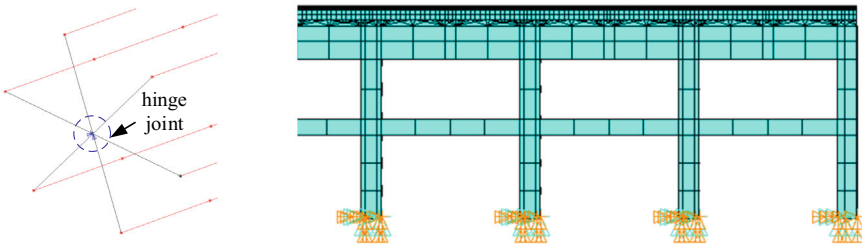
3 Contact Model Settings

When performing contact analysis in ANSYS, contact pairs need to be set up [11] to determine contact surface and target surface of the contact between shed tunnel's roof and the specified wire respectively. Wherein, the roof of the shed tunnel is regarded as the target surface, and the flexible target surface property is selected. Select the wire as the contact surface and select the contact mode opposite the node. Select TARGE170 to discretize the shed tunnel's roof as the target surface, and use CONTA175 element to discretize the transmission wire as the contact surface.

The properties and contact parameters of the contact pair need to be set to realize the contact simulation between the broken line and the shed tunnel roof. In terms of contact algorithms, Augmented Lagrange can be chosen to deal with complex contact conditions [12]. Besides, it's necessary to specify the direction of the contact normal and the direction towards the target plane is chosen. The behavior of the contact surface can choose the standard behavior [13]. The friction between the wire and the roof of the shed tunnel is ignored. Since the wire is mainly bent during the contact with the shed tunnel, the normal penalty stiffness (FKN) coefficient is set as a smaller value as 0.01. In addition, the maximum allowed insertion depth is set as the thickness of the foam concrete roof 0.5 m. This ensures wires will not become too inserted into the roof.

3.1 Imposition of Boundary Conditions

In the contact analysis, the boundary conditions need to consider the force boundary and displacement boundary, and also need to determine the boundary conditions in the contact area. Figure 7 shows the constraint simulation diagram. When the wire breaks, one of the suspension points will be released, while the other suspension point is still subject to hinged constraints. The constraint of the wire on the suspension point can only have zero displacement in the X, Y, and Z directions, as shown in Fig. 7a for constraint simulation diagram. When the wire breaks, one of the suspension points will be released, while the other suspension point is still subject to hinged constraints. The constraint of the wire on the suspension point can only have zero displacement in the X, Y, and Z directions, as shown in Fig. 7a. For the shed tunnel structure, the shed tunnel is always limited by the embedding constraint of the ground facing the bottom of the column before and after the contact. The freedom of all nodes on the bottom of the column is zero, as shown in Fig. 7b.



(a) Wire constraint diagram (b) Shed tunnel constraint diagram

Fig. 7 Model constraint simulation diagram

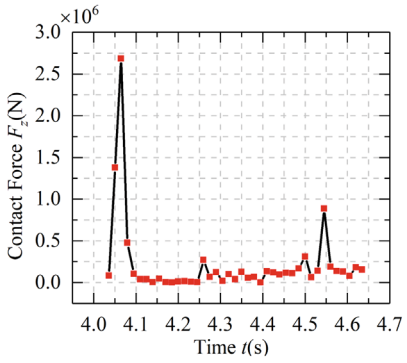
3.2 Solving Settings

In this research, the contact problem between the wire and the shed tunnel contains strong contact nonlinearity. Besides, the large deformation of the wire after the contact makes the geometric shape of the transmission wire have strong nonlinearity. The above two reasons make the dynamic equation in this research have strong nonlinearity, and its computational convergence and stability are poor. Therefore, different iterative steps are set for different stages of the motion process. In the period when the transmission wire does not contact with the roof of the shed tunnel, the calculation step can be set to a larger value to improve the calculation efficiency, and the maximum calculation step is set to $\Delta t = 0.5$ s; when the wire is in contact with the shed, the dynamic equation has contact nonlinearity and large deformation, so a small calculation step is required. The maximum calculation step is set to $\Delta t = 1$ ms.

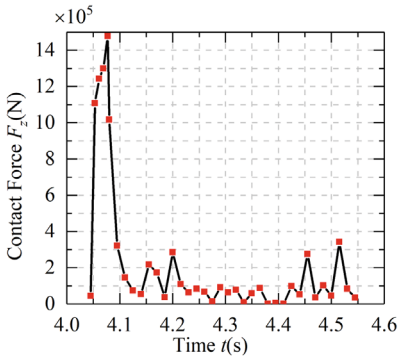
4 ANSYS Contact Calculation Results

Through the above analysis steps, the vertical contact force results of Hu'an Line, Jinsu Line, Fufeng Line and shed tunnel roof can be obtained, time history diagram of the vertical contact force are shown in Fig. 8. Combined with these curves, the contact force calculation results of each ultra-high voltage transmission wires are analyzed:

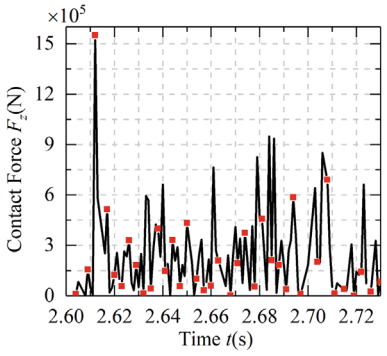
The moment when the maximum vertical contact force resultant force is generated between the wire I of Hu'an Line and the shed tunnel is 4.065 s, and the resultant contact force on the wire is 2684941 N. (2) The maximum vertical resultant contact force between Hu'an wire II and the shed tunnel generates in 4.0772 s, and the resultant contact force is 1477570.82 N. (3) The moment when the contact between Jinsu wire 1 and the shed tunnel produces the maximum contact force in the vertical direction is 2.9616 s, and the force is 1477570.82 N. (4) The moment when the collision between Jinsu wire 2 and the shed tunnel produces the maximum contact



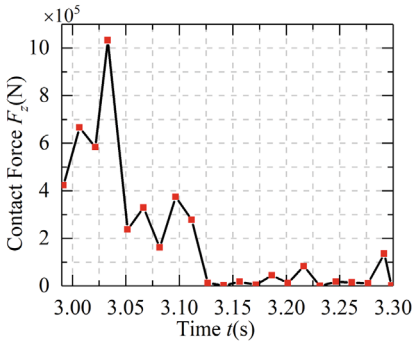
(a) Contact force curve of Hu'an wire 1



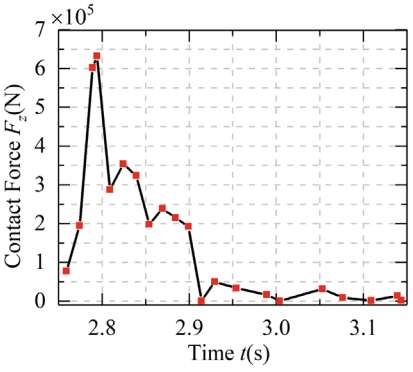
(b) Contact force curve of Hu'an wire 2



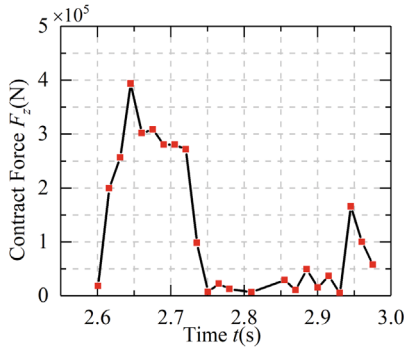
(c) Contact force curve of Jinsu wire 1



(d) Contact force curve of Jinsu wire 2



(e) Contact force curve of Fufeng wire 1



(f) Contact force curve of Fufeng wire 2

Fig. 8 Curves of vertical contact force between transmission lines and shed tunnel roof

force in the vertical direction is 3.0330 s, the force is 1033267.7539 N. (5) The moment when the contact between Fufeng wire 1 and the shed tunnel produces the maximum contact force in the vertical direction is 2.7941 s, and the force is 632271.5549 N. (6) The maximum contact force between Fufeng wire 2 and the shed tunnel in the vertical direction generates in 2.6454 s, and the force is 393535.1183 N.

In summary, we get the contact force of each wire and the roof of the shed tunnel through the simulation results. These results can provide important reference for the design and improvement of the shed tunnel roof.

5 Conclusion

In order to explore the protective effect of concrete shed tunnels under the three groups of ultra-high voltage power wires on Shang-He-Hang Railway, based on working conditions, this paper presents a simulation analysis based on ANSYS for the dynamic problems of concrete shed tunnels under the impact of broken iced wires. The study calculates and analyzes the contact process between high-voltage wires and concrete sheds after different iced wires break and fall, providing important references for the design, manufacturing, and improvement of shed tunnels.

When conducting contact dynamics simulation analysis based on ANSYS, according to the construction drawings and technical parameter requirements, geometric modeling is carried out at first, including concrete shed tunnels, Hu'an Line, Jinsu Line, and Fufeng Line that are covered with ice and subjected to wind load. Secondly, SOLID and BEAM elements are selected to discretize the concrete shed tunnel and iced wires respectively, and mesh of different sizes are used to discretize the concrete shed tunnel to balance computational efficiency and accuracy. Then, the contact model between iced wire and foam concrete roof is set up, and parameters of the contact model are introduced in detail. Afterwards, different calculation steps are set for different motion states of wires to improve the computational efficiency of contact analysis. Finally, curves of the vertical contact force of the shed roof over time are presented. This article comprehensively elaborates on the contact problem between falling iced wires and concrete shed tunnels, providing theoretical and practical guidance for the construction of protective concrete shed tunnels.

References

1. Zhou G, Guo Q, Ning L et al (2022) Highway crossing existing railway protection shed frame device on freeway, has vertical rod whose bottom is connected with protective shed frame base through pipe frame lower support. Sinohydro Eng Bureau 14 Corp Ltd (Cnpw-C).CN217949823-U

2. Liu C, Gao M, Yang S et al (2023) Replaceable rockfall protection sheds for railway bridges, has replaceable rockfall protection sheds for railway bridges along bridge direction include multiple independent shed sections, and connecting plate extends below wing plate. In: China Railway First Survey & Design Inst (Crcc-C) China Railway First Survey and Design Inst (Crcc-C).CN111270620-A; CN212000678-U
3. Lan P, Li K, Yu Z (2019) Computer implementation of piecewise cable element based on the absolute nodal coordinate formulation and its application in wire modeling. Acta Mech 230(3):1145–1158
4. Gu Y, Lan P, Cui Y et al (2021) Dynamic interaction between the transmission wire and cross-frame. Mech Mach Theory 155
5. Park Y, Lee K (2012) Implement of dynamic performance measurement system between pantograph and contact wire in tunnel. J Electr Soc 61(11):1732–1736
6. Tabri K, Varsta P, Matusiak J (2010) Numerical and experimental motion simulations of nonsymmetric ship collisions. J Mar Sci Technol 15(1):87–101
7. Yan B, Zhang Z, Zhang H et al (2024) Dynamic contact characteristics analysis of heavy-duty track rollers. Int J Automot Mech Eng 21(1):11155–11165
8. Sun JF, Wang GL, Zhang HO (2005) Elasto-plastic contact problem of laying wire rope using FE analysis. Int J Adv Manuf Technol 26(1–2):17–22
9. Liu C, Chen Q, Zhang E et al (2012) A calculation model and characteristics analysis of radio wave propagation in rectangular shed tunnel. In: 2012 10th international symposium on antennas, propagation and Em theory (Isape). Xi'an, Peoples R China, pp 535–539
10. Sun RY, Cho H, Kim S-H (2016) Damage evaluation of flexible concrete mattress considering steel reinforcement modeling and collision angle of anchor. J Ocean Eng Technol 30(2):109–116
11. Zhang X, Hu J, Zhou Y et al (2023) Analysis of bolt drop strength of parking brake skid based on ANSYS/LS-DYNA. J Vib Shock (in Chinese) 42(7):217–224
12. Jin J, Chen Y, Wu X (2008) Application of contact algorithm in complicated structure dynamic simulation. In: Proceedings of first international conference of modelling and simulation, modelling and simulation in science and technology, vol 1. Nanjing, Peoples R China, pp 516–519
13. Wang Y, Cui H, Fang Y et al (2016) FEA of the rotated casting machine based on hypermesh and ANSYS. In: 2016 3rd international conference on mechanical, industrial, and manufacturing engineering (Mime 2016). Los Angeles, California, USA, pp 128–131

Open Access This chapter is licensed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License (http://creativecommons.org/licenses/by-nc-nd/4.0/), which permits any noncommercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if you modified the licensed material. You do not have permission under this license to share adapted material derived from this chapter or parts of it.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.



Research on the Aerodynamic Performance of a Suspended Monorail Train Designed like a Whale



Wenhui Ma, Yazhuo Zou, Jiachun Zou, and Jingyi Wang

Abstract In order to improve the operational efficiency of suspended intelligent monorail trains, a whale is used as a biomimetic object to design the appearance of the train and study its aerodynamic performance. Innovatively design the head shape of a suspended intelligent monorail train using whale bionics. According to the special morphology of different whale parts, the upper surface, side surface, and lower surface of the train were designed separately. The aerodynamic performance of the three innovative design trains was analyzed through turbulence models. Under exponential wind load, the air resistance of the train after innovative design was significantly reduced, with the highest resistance being around 3500 N. At different wind speeds and wind direction angles, the overturning moment and lateral force of the train were lower than the corresponding limit values, ensuring the stability and safety of train operation. Compared to traditional models, the design of whale like trains has lower resistance and higher operating efficiency. Its form design and aerodynamic performance research have certain value and significance for the development of suspended intelligent monorail trains.

Keyword Whale bionics · Pneumatic performance · Turbulence model · Air resistance

1 Introduction

Traditional transportation methods are no longer able to meet the modern urban demand for efficient, convenient, and environmentally friendly travel [1]. As an innovative form of rail transit, the suspended intelligent monorail train has brought new possibilities for modern urban transportation with its unique design concept and advanced technological applications [2]. This innovative design also reflects

W. Ma · Y. Zou (✉) · J. Zou · J. Wang
Dalian Jiaotong University, Dalian, China
e-mail: 276613628@qq.com

© The Author(s) 2026

H. Saman et al. (eds.), *The 9th International Conference on Advances in Construction Machinery and Vehicle Engineering*, Lecture Notes in Mechanical Engineering,
https://doi.org/10.1007/978-981-96-5527-4_59

a profound practice of environmental protection concepts, making positive contributions to the sustainable development of cities by reducing ground contact and energy consumption. However, the innovative design of suspended intelligent monorail trains is not limited to the technical level, but rather a profound reflection and reshaping of modern urban lifestyles and cultural concepts. For example, Huang Xiaocui et al. established a mathematical model for the lateral stiffness of suspension systems in theoretical analysis, and analyzed the influence of different stiffness values on the stability and comfort of vehicle operation [3]. Yu Rensheng et al. proposed a new lightweight material application scheme for structural design, aiming to reduce train weight and energy consumption. In terms of power and control systems, advanced electric drive technology and intelligent control strategies have been introduced [4]. However, the above methods did not fully consider the factors of train air resistance. In response to the above issues, the article proposes an innovative design of a suspended intelligent monorail train based on whale bionics, and conducts aerodynamic performance analysis of the train to verify its feasibility.

2 Theoretical Foundations

2.1 Biomimetic Design

Biomimetic design is established on the common foundation of bionics and design, including various disciplines such as dynamics, materials science, color science, biology, etc.; its wide audience is also the reason for the complexity and diversity of its research content [5]. Biomimetic design is based on all living organisms in nature, by studying their forms, structures, colors, functions, and other related elements, extracting certain features for design, and combining interdisciplinary knowledge to form new ideas and obtain new design methods [6]. The rapid development of bionics has strongly promoted the development of different disciplines, indicating the potential for significant impact on new technologies [7].

2.2 Turbulence Model

In theory, the turbulence calculation area is unbounded, but in numerical simulation analysis of the aerodynamic performance of a suspended intelligent monorail train, it needs to be calculated in a finite space. Therefore, based on the principle of not affecting the air flow near the suspended intelligent monorail lightning car, the calculation area of the turbulence model for aerodynamic performance analysis is selected [8]. When the distance between the suspended intelligent monorail train and the boundary of the calculation area is far, the airflow disturbance formed during

the operation of the suspended intelligent monorail train will not affect the air flow in the calculation area [9].

When using turbulence models to analyze the resistance of suspended intelligent monorail trains, the calculation formula is as follows:

$$Z = \frac{2Q}{\rho V^2 A} \tag{1}$$

Among them, the air resistance of the suspended intelligent monorail train is Q; The driving speed is V; The maximum cross-sectional area is A; The air density is ρ.

The aerodynamic lateral force is:

$$Z_S = \frac{Q_S}{0.5\rho V^2 A} \tag{2}$$

Among them, the aerodynamic lateral force is Z_S.

The overturning moment is:

$$Z_M = \frac{Q_M}{0.5\rho V^2 A G} \tag{3}$$

Among them, the reference height for the center of gravity of the suspended intelligent monorail train is G; Overturning moment is Q_M.

The pressure at the nose tip of the suspended intelligent monorail train is:

Among them, the stationed pressure is P.

3 Innovative Design of the Appearance of Suspended Intelligent Monorail Trains

3.1 Design Steps

Biomimicry is a discipline that draws inspiration from nature and applies it to engineering design and manufacturing. Whales, as aquatic mammals, have undergone long-term natural selection and evolution of their body shape, possessing excellent fluid dynamics characteristics that can reduce air resistance and improve the efficiency and stability of train operation. The design elements of biomimetic whales bring unique appearance and performance advantages to the suspended intelligent monorail train.

The specific steps for innovatively designing the front end shape of a suspended intelligent monorail train based on the shape of a whale's head are as follows:

Step 1: Extract features of whale heads.

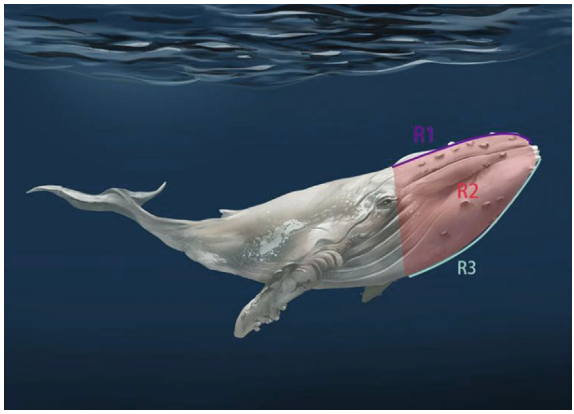


Fig. 1 Whale characteristic curve

Step 2: Assign different weights to each feature based on the importance and impact of the extracted whale head features, and then determine the primary and secondary features by comparing the weights. Ignore secondary features, preserve primary features, and extract feature lines of primary features.

Step 3: Process the main characteristic lines according to the control lines and body contour lines of the front end of the suspended intelligent monorail train, and use them in the exterior design of the front end of the train to complete the exterior design of the front end of the suspended intelligent monorail train.

3.2 Whale Feature Extraction

As shown in Fig. 1, extract the top curve of the whale's head as R1, the side curve of the whale's head as R2, and the bottom curve of the whale's head as R3, corresponding to the base, side, and top of the train locomotive, respectively.

3.3 Design Scheme

Taking the suspended intelligent monorail train of a certain passenger sightseeing route as the experimental object, this paper innovatively designs the suspended intelligent monorail train of the passenger sightseeing route using the method proposed in this paper, reducing the running resistance of the suspended intelligent monorail train and improving passenger comfort.

The innovative design of the suspended intelligent monorail train for this passenger sightseeing route was carried out using the method described in this article. The innovative design results of the suspended intelligent monorail train's appearance are shown in Fig. 2.



Fig. 2 Innovative design results of suspended intelligent monorail train appearance

4 Aerodynamic Performance Analysis of Trains

4.1 Analysis of Resistance Based on Turbulence Model

Aerodynamic performance analysis of an innovative designed suspended intelligent monorail train using turbulence models. To improve the accuracy of aerodynamic performance analysis of suspended intelligent monorail trains, an area of 1000 m in length and 100 m in width is taken as the calculation area. Within this calculation area, the top of the suspended intelligent monorail train is 80 m above the ground, the distance between the nose tip of the train and the entrance is 400 m, and the distance between the nose tip of the train tail and the exit is 600 m.

The boundary conditions of the turbulence model are:

(1) Calculate the entrance of the area and use the speed of the suspended intelligent monorail train as the boundary condition for the speed entrance [10].
(2) Calculate the outlet of the calculation area, with a static pressure of 0 as the boundary condition for the pressure outlet, to avoid backflow.
(3) Calculate the upper and both sides of the region, with the boundary normal velocity being 0 and the remaining gradient being 0, as symmetric boundary conditions.
(4) Calculate the lower part of the calculation area, using the running speed of the suspended intelligent monorail train as the sliding speed [11], and design the sliding wall boundary conditions.

By unstructured meshing, the calculation area of the turbulence model for the suspended intelligent monorail train is divided [12]. The turbulence model of the suspended intelligent monorail train innovatively designed by biomimetic whales has around 14 million internal grids.

The suspended intelligent monorail train innovatively designed in this article consists of four parts, namely the head car, middle car 1, middle car 2, and tail car. The resistance situation at different train positions is analyzed when the train runs smoothly at a speed of 50 km/h. The analysis results are shown in Fig. 3.

Analyze the distribution of resistance at different positions of the train in Fig. 3. The resistance of the leading car is the highest, about 2500 N, followed by the resistance of the trailing car, about 2000 N, while the resistance of the middle cars 1

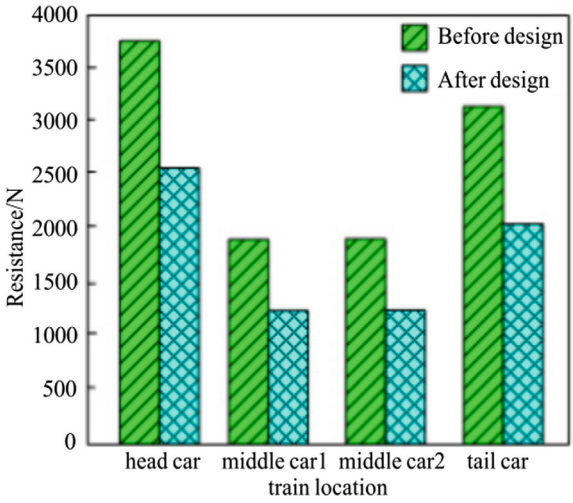


Fig. 3 Resistance analysis results at different train positions

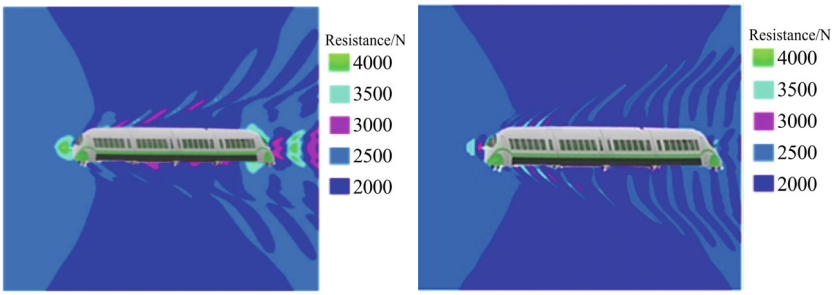
and 2 is basically the same and relatively small. This resistance distribution pattern reflects the different effects of air resistance on various parts of the train during operation. Firstly, as the leading part of the train, the shape and size of the locomotive have the most direct and significant interference with air flow. Due to the fact that the airflow faced by the leading car is the original airflow that has not been disturbed by other parts of the train, its resistance is relatively high. The tail car, on the other hand, has relatively less resistance due to the influence of the wake formed by the head car and the middle car. Due to being located in the middle of the train, the shape and size of intermediate cars 1 and 2 have relatively less interference with the airflow, resulting in lower resistance. Secondly, comparing the resistance data before and after the innovative design, it was found that the resistance of the suspended intelligent monorail train designed using the method proposed in this paper was significantly reduced during smooth operation. The aerodynamic performance of the entire train has been improved.

In the turbulence model, when analyzing the exponential wind load, this paper innovatively designs the resistance changes of the suspended intelligent monorail train before and after the bottom of the train based on the R1 curve design method of the whale head. The analysis results are shown in Fig. 4.

This article innovatively designs the resistance changes of the suspended intelligent monorail train on the front and rear sides of the train based on the surface design method of the R2 whale head. The analysis results are shown in Fig. 5.

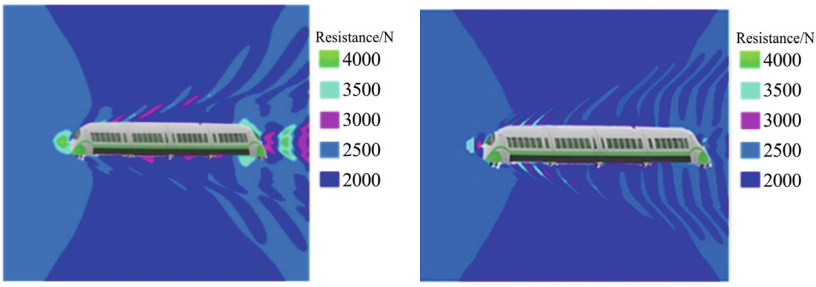
This article innovatively designs the resistance variation of the suspended intelligent monorail train at the top of the train based on the R3 curve design method of the whale head. The analysis results are shown in Fig. 6.

According to the data shown in Figs. 4a, 5a, and 6a, as well as Figs. 4b, 5b, and 6b, it can be concluded that the resistance performance of the suspended intelligent monorail train has undergone significant changes under exponential wind loads.



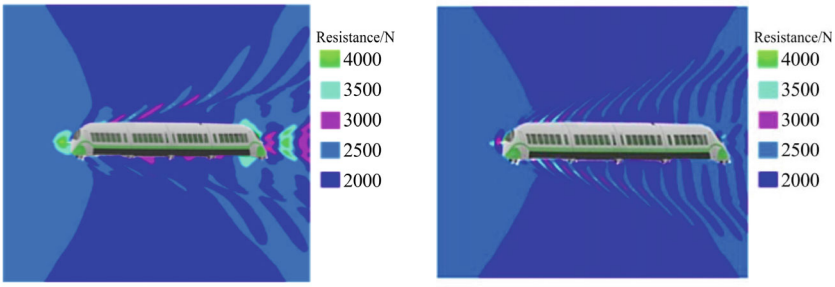
(a) Resistance changes before innovative design (b) Resistance changes after innovative design

Fig. 4 Resistance changes before and after innovative design (Scenario 1)



(a) Resistance changes before innovative design (b) Resistance changes after innovative design

Fig. 5 Resistance changes before and after innovative design (Scenario 2)



(a) Resistance changes before innovative design (b) Resistance changes after innovative design

Fig. 6 Resistance changes before and after innovative design (Scenario 3)

Before innovative design, the maximum resistance of the train was approximately 4000 N. After innovative design, the maximum resistance of the train was reduced to about 3500 N, which clearly indicates the positive impact of innovative design on



(a) Cloud map of windward pressure under uniform wind load



(b) Wind pressure cloud map on the windward side under turbulent wind load

Fig. 7 Windward side pressure cloud map of the front end of the suspended intelligent monorail train

train resistance. This article innovatively designs the train head by mimicking the whale head, which can effectively reduce train resistance.

When analyzing different crosswind loads, the pressure cloud map on the windward side of the suspended intelligent monorail train innovatively designed in this paper is shown in Fig. 7.

According to the data shown in Fig. 7a, b, it can be seen that the trend of surface pressure changes at the front end of the train is similar under two different types of crosswinds: uniform wind and turbulent wind. Whether it is uniform wind or turbulent wind, the pressure at the front end of the car is the highest, and the maximum pressure value is maintained at a level of about 2000 Pa. When the air flow reaches the nose tip of the front end of the train, a large pressure is generated at the nose tip, called nose tip stagnation pressure. When the nose tip stagnation pressure of the suspended intelligent monorail train exceeds 5789 Pa, the train head will deform. According to the analysis of Figs. 4a, b, 5a, b and 6a, b, it can be seen that the pressure at the nasal tip stationary point does not exceed 5789 Pa. This observation indicates that the suspension intelligent monorail train innovatively designed in this paper can maintain a safe and stable front structure without deformation when dealing with uniform and turbulent wind loads.

Table 1 Changes in overturning moment and lateral force of trains under different wind speeds

Wind direction angle/°	Overturning moment/Nm	Lateral force/N	Overturning moment limit/Nm	Lateral force limitation/N
15	960	569	≤ 9000	≤ 4500
30	1112	664	≤ 8000	≤ 4000
45	2201	772	≤ 7000	≤ 3500
60	2422	889	≤ 6000	≤ 3000
75	3108	1099	≤ 5000	≤ 2500
90	3329	1493	≤ 4000	≤ 2000

4.2 Wind Speed Impact Analysis

Analyze the changes in overturning moment and lateral force of the suspended intelligent monorail train innovatively designed in this paper under different wind speeds, when running at a speed of 50 km/h. The analysis results are shown in Table 1.

According to Table 1, as the wind speed increases, the overturning moment and lateral force of the suspended intelligent monorail train innovatively designed in this paper show an upward trend. This indicates that the increase in wind speed has indeed had an impact on the stability and safety of trains. However, it can also be seen from the data that although the overturning moment and lateral force increase with the increase of wind speed, they are always below the corresponding limit values, indicating that the train can maintain stable operation at different wind speeds. This innovative train design has excellent dynamic performance and stability, and can effectively cope with external factors such as wind speed, providing passengers with a safe and comfortable riding experience.

4.3 Analysis of Wind Direction Angle Influence

Analyze the changes in overturning moment and lateral force of the suspended intelligent monorail train innovatively designed in this paper under different wind direction angles, when running at a speed of 50 km/h. The analysis results are shown in Table 2.

By analyzing the data in Table 2, we can observe that as the wind direction angle increases, the overturning moment and lateral force of the suspended intelligent monorail train innovatively designed in this paper show an upward trend. This is because the change in wind direction angle directly affects the direction and magnitude of the wind force that the train is subjected to, thereby affecting the overturning moment and lateral force of the train. However, although the overturning moment and lateral force increase with the increase of wind direction angle, it is worth noting that they are always below the corresponding limit values. This indicates that the suspended intelligent monorail train innovatively designed in this article has good

Table 2 Changes in overturning moment and lateral force of trains under different wind direction angles

Wind speed (m/s)	Overturning moment/Nm	Lateral force/N	Overturning moment limit/Nm	Lateral force limitation/N
4.5	1016	267	$\leq$ 10,000	$\leq$ 5000
6.5	1038	387	$\leq$ 8000	$\leq$ 4000
8.5	1059	357	$\leq$ 8000	$\leq$ 4000
10.5	1098	479	$\leq$ 6000	$\leq$ 3000
12.5	1125	489	$\leq$ 6000	$\leq$ 3000
14.5	1247	546	$\leq$ 6000	$\leq$ 3000
16.5	1307	668	$\leq$ 4000	$\leq$ 2000
18.5	1419	765	$\leq$ 4000	$\leq$ 2000
20.5	1478	880	$\leq$ 2000	$\leq$ 1000
22.5	1699	986	$\leq$ 2000	$\leq$ 1000
24.5	1817	998	$\leq$ 2000	$\leq$ 1000

stability and safety under different wind direction angles. This design can effectively reduce the impact of wind on trains under various wind direction and angle conditions, thereby maintaining the smooth operation of trains.

5 Conclusion

This article uses finite element simulation to analyze the aerodynamic energy and confirms that it can effectively design the shape of a suspended intelligent monorail train in the biomimetic form of a whale. The front end of the train is designed in the shape of a whale's head, which can effectively reduce air resistance. This is because streamlined design can effectively guide the airflow, reduce the vortices and turbulence generated by the airflow on the surface of the train, thereby reducing air resistance. For suspended intelligent monorail trains, reducing air resistance can not only improve the operational efficiency of the train, but also reduce energy consumption and lower operating costs.

Project Fund Humanities and Social Sciences Research Project of Liaoning Provincial Department of Education (LJKMR20220851).

Research on Cultural Gene Expression and Genealogy Evolution Method in the Design of Rail Transit Vehicles.

References

1. Zhu E, Qin Z, Bai Z (2023) Calculation and analysis of torsional performance of bottom opening suspension monorail transit track beams.Urban Rail Transit Res 26(07):20–25+37
2. Min Z, Bing Z, Zhen Z et al (2022) Wind vehicle bridge coupling dynamic analysis of suspended monorail small radius curved bridge. Railway Standard Design 66(12):66–71
3. Xiaocui H, Xiaoping J, Maohai F et al (2022) Optimization design and verification of lateral stiffness of suspension system for suspended monorail vehicles. Locomot Electr Transm (05):9–16
4. Yu S, Shi Y, Bai L (2022) Design points and optimization directions of network control system for new energy suspended monorail trains. Urban Rail Transit Res 25(09):45–48+52
5. Zhao Q (2019) A brief analysis of biomimetic design in industrial products. Ind Des (07):66–67
6. Hairong Z (2020) Research on biomimetic design and application in modern industrial design. Sci Inf Technol 17:70–76
7. Hu N, Feng P, Dai G (2013) The gift from nature: bio inspired strategy for developing innovative bridges. J Biomim Eng Engl Ed 04:10
8. Wang L, Gao G, Wu Y et al (2022) Aerodynamic design of high speed train body friction brake plates. J Central South Univ (Natural Science Edition) 53(05):1655–1667
9. Li J, Liu S, Zhu B et al (2022) Simulation study on coupled vibration of suspended monorail vehicle bridge system under wind force. Urban Rapid Transit 35(05):104–110
10. Yao Y, Li Q, Li Z (2023) Design and key technology research of intelligent operation and maintenance system for suspended monorail turnouts. Railw Stand Des 67(10):101–107
11. Yuwen T, Kui C, Li Y et al (2023) Fault analysis and optimization design of pantograph head support plate for monorail vehicles. Mech Des Res 39(05):225–227
12. Li F (2023) Design method for sustainable operation of high speed trains. Mech Des Res 39(05):210–218+224

Open Access This chapter is licensed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License (http://creativecommons.org/licenses/by-nc-nd/4.0/), which permits any noncommercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if you modified the licensed material. You do not have permission under this license to share adapted material derived from this chapter or parts of it.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.



Experimental Study on Electro-Hydraulic Cooperative Braking by Wire Safety Performance of Mining Electric Vehicles



Zhi-jun Hao

Abstract The dynamic performance of explosion-proof brake by wire affects the safety of the electric vehicles. Analyze the composition of wet brake by wire system. Research the key technologies which affects the safety of braking system as response time and distribution ratio between electric brake and hydraulic brake. Through design the simulation model of the brake by wire system, modify and test the electro-hydraulic system in the prototype. The results shows that the electro-hydraulic cooperative braking can effectively reduce the rear axle braking response time of mining electric vehicles by wire, solve the braking time difference between the front and rear axles, shorten the overall response time by 0.41 s, improve the safety of mining electric vehicles. The study provides a fast and effective solution for the practical engineering application of wire controlled braking of mining electric vehicles, establish a foundation for accelerating the intelligent development of mining electric vehicles.

Keywords Brake by wire · Electro-hydraulic distribution · Simulation testing · Response time

1 Introduction

Promote the integrated development of intelligent technology and the coal industry, enhance the level of coal mine intelligence, the China National Development and Reform Commission and the National Energy Administration jointly issued the

Z. Hao (✉)
Shanxi Tiandi Coal Mining Machinery Co., Ltd., Taiyuan, China
e-mail: mkyhzj@163.com

China Coal Technology and Engineering Group Taiyuan Research Institute Co., Ltd., Taiyuan, China

China National Engineering Laboratory for Coal Mining Machinery, Taiyuan, China

Shanxi Key Laboratory of Comprehensive Drivage and Shortwall Mining Equipment for Coal Mine, Taiyuan, China

© The Author(s) 2026

H. Saman et al. (eds.), *The 9th International Conference on Advances in Construction Machinery and Vehicle Engineering*, Lecture Notes in Mechanical Engineering,
https://doi.org/10.1007/978-981-96-5527-4_60

"Guiding Opinions on Accelerating the Development of Coal Mine Intelligence" with other eight ministries in 2020. The opinions pointed out that large coal mines whose production capacity over 1 million tons should be intelligentization by 2025, key positions in underground should be operated by robots, achieves intelligent continuous operation and unmanned transportation [1]. After several years, large coal mines such as Shaanxi Coal Group and CE Shendong Group have conducted pilot tests on explosion-proof unmanned driving vehicles for mining, but none of them have been widely used. On the one hand, the safe operation of unmanned vehicles for auxiliary transportation equipment is concerned; On the other hand, there are many uncertain factors in the underground, such as uneven road surfaces, potholes, water accumulation, and low illumination. The manner vehicles do not respond timely, which brings unsafe risks.

So the manner vehicles must have the ability to monitor and process information independently, and can process information timely and operate reliably [2, 3]. At present, explosion-proof manner vehicles mainly adopts wire hydraulic braking. The maximum operating speed of underground vehicles can reach 40 km/h, the explosion-proof underground vehicles are not equipped with ABS antilock braking systems, the front and rear wheels with different loads cannot reach the braking lock state at the same time, which may cause the vehicle loses steering ability or sideslips, affecting the safety of unmanned vehicle in underground coal mine.

2 Key Technologies of Wet Brake by Wire

Different from the braking system in ground vehicles, according to Chinese standards MT/T 1199 "General Safety Technical Conditions for Explosion proof Diesel Engine Trackless Rubber Wheel Vehicles Used in Coal Mines", wet brakes should be used for the trackless rubber wheel vehicles in coal mines. So the brake by wire system for trackless rubber tyred vehicles used in coal mines is based on a wet brake wire control system.

2.1 System Composition

As usual, the vehicles in coal mine uses a hydraulic brake pedal valve to control the flow of brake fluid, then pushes the brake friction plate to generate braking force. While wet brake by wire replaces the mechanical brake pedal valve with an electronic components, is an electro-hydraulic combined system. The system converts the brake pedal position information into an electrical signal and sends it to the explosion-proof electro-hydraulic valve. At the same time, the CAN bus is used to control the opening of the explosion-proof electro-hydraulic valve, omitting the pipeline connection between the pedal valve and the diversion valve, reducing the braking response time, greatly improving the braking performance. To ensure the safety of

underground vehicles, the wet brake by wire system also retains a brake pedal valve. In emergency, drivers can take over the brake system to increase the vehicle safety redundancy [4].

The explosion-proof wet brake by wire system mainly includes oil pump motor, explosion-proof electro-hydraulic valve, diversion valve, pressure reducing valve, brakes in front and rear wheels, and pressure sensor. The system also includes a vehicle controller, a drive motor controller, an explosion-proof lithium battery pack, a brake pedal signal acquisition module, etc., among which the various controllers exchange information through CAN bus. The system also has the driving motor feedback brake, between the explosion-proof lithium battery pack, the driving motor controller, and the driving motor with an electric current, as shown in Fig. 1.While the vehicle controller sends out the braking signal, which is transmitted to the explosion-proof electro-hydraulic valve through the brake pedal signal acquisition module, also transmitted to the drive motor controller to control the motor feeds back brake force. The motor feedback braking and hydraulic braking simultaneously affect the wet brake by wire efficiency. So the explosion-proof electro-hydraulic valves, the layout of hydraulic pipelines, and motor feedback braking affect the brake response time of the vehicle [5].

When the explosion-proof vehicles brake, high-pressure oil is injected into the brake cylinder to push the dynamic brake pads to combine with the static brake pads. Because of the wet brake piston displacement is much larger than a caliper piston brake in ground vehicles. To ensure force balance, the wet brake is more than 3 cylinders even distribution around the brake ring. The required brake fluid increases, and the high-pressure brake fluid flow through the explosion-proof electro-hydraulic valve increases. Therefore, the wet brake often has obvious braking delay problems. The explosion-proof electro-hydraulic valve as the key of the wire by brake system affects the speed of the braking response time. For this reason, multiple universities and research institutes in China have gradually overcome this key technology, and the response time of the electro-hydraulic valve can reach 3.5 ms [6].

2.2 Motor Feedback Braking

The wet brake by wire system includes hydraulic braking and motor feedback braking. The motor feedback braking as an auxiliary braking system, can convert some of the dynamic energy during the vehicle's driving process into electrical energy stored in the explosion-proof lithium battery, can save energy and improve energy utilization efficiency, extend the vehicle's driving distance, increase the life of hydraulic braking pads and braking intensity, reduce frictional heat generation and improve vehicle braking safety. Motor feedback braking is an electric braking system that relies on the motor for reverse traction, and its response speed is significantly faster than hydraulic braking, which has advantages in improving the response of wire controlled hydraulic braking systems [7].

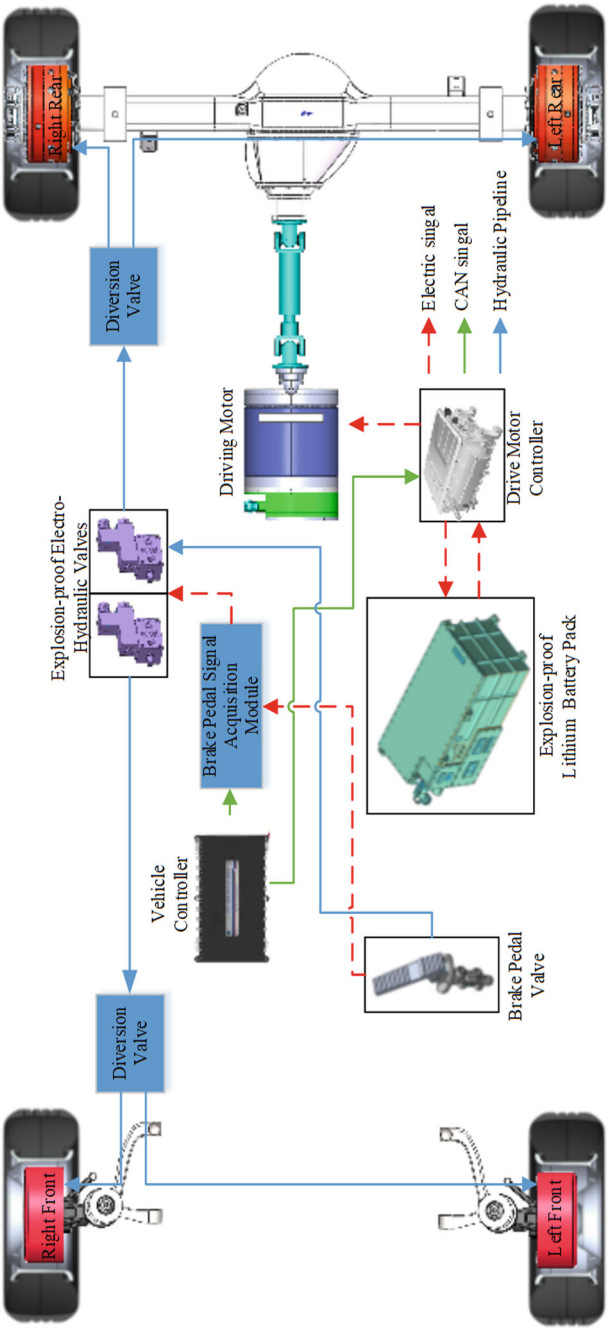


Fig. 1 Structure and composition of wet brake by wire for rear drive trackless rubber wheeled vehicles. Explosion-proof electro-hydraulic valve

The motor feedback braking is influenced by both the motor power generation and the battery charging characteristics. For the motor, during the braking process, the output torque is closely related to the vehicle speed, and the power generation efficiency of the motor will also changes with its speed and torque; For battery, the allowable charging conditions during the braking energy recovery process mainly depend on the battery's state of charge (SOC), allowable charging current, and battery temperature. In practical, mining electric rubber tyred vehicles face special application conditions of long-distance downhill after fully charged, the motor will not perform feedback braking when the vehicle's SOC is greater than or equal to 90%.

2.3 Electric Brake and Hydraulic Brake Distribution Strategy

Due to the fast response time of motor feedback braking, there are also adverse effects such as excessive feedback intensity increased vehicle shaking. How to allocate hydraulic braking and motor feedback braking is the key to improve response speed and braking safety.

For rear drive mining electric vehicles, designed a controller which integrates the mechanical characteristics of the motor, the charging and discharging characteristics of the battery pack, and the braking safety, etc. The total braking torque requirement is obtained by identifying the driving braking intention through the brake pedal and signal acquisition module. The motor feedback braking torque is determined by the braking strength Z. Then the motor feedback braking torque and the hydraulic braking torque provide the whole braking force [8] (Fig. 2).

3 Design Simulation and Prototype Testing

3.1 Simulation

Model and simulate the structure of the wet brake by wire system in MATLAB [9, 10]. The model with two input signals control the on/off of the front and rear solenoid valves, the pressure and stroke of the left front, right front, left rear, and right rear brakes as monitoring targets. The simulation model is established as Fig. 3.

When all four brake wheels only use hydraulic braking, and motor feedback braking is zero. The front and rear explosion-proof electro-hydraulic valves receives a step signal at 0.9 s simultaneously. The pressure variation curves of the four brakes are shown in the Fig. 4. Because of the longer length of the rear axle brake pipeline than the front and the larger volume, the peak pressure of the front and rear brakes is reached at different times. The full response time of the four brakes is 0.6 s, and the brakes on both sides of the rear axle lag behind the front axle brake by about 0.1 s to reach the peak pressure. The four wheels did not reach the locking state at the same

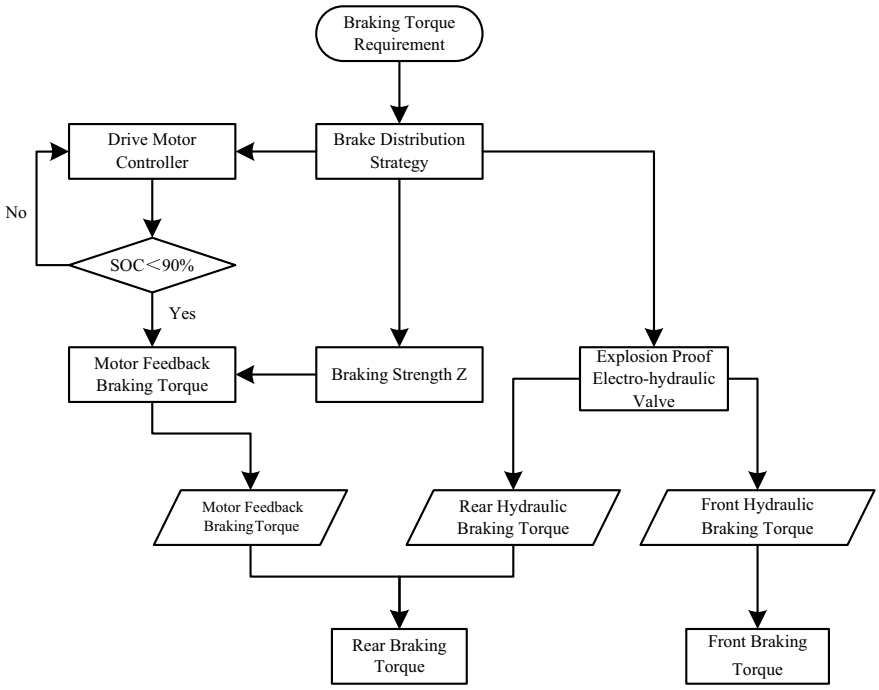


Fig. 2 Braking force distribution control strategy

time, the front wheels reach peak brake pressure before the rear wheels, indicating that the vehicle is expected to lose steering ability in a brief moment when braking.

3.2 Prototype Modification and Testing

Install a wet brake by wire system on a 5-seater explosion-proof electric vehicle, which curb weight of 3200 kg, front and rear axle load ratio of 40%:60%, rated brake pressure of 3.6 MPa, front braking torque of 1800 N m, and a rear braking torque of 2400 N m. Two sets of explosion-proof electro-hydraulic valves and pressure control valves as shown in Fig. 5 are used to control the hydraulic flow and pressure of the front and rear brakes, and four pressure sensors as shown in Fig. 6 are used to monitor the changes in brake pressure at the inlet during the braking process. Auxiliary module such as data acquisition module, CAN signal acquisition module, 24 V power supply and sensor cables shown in Fig. 7.

The pressures change in the four brakes inlet ports monitored by sensors are shown in Fig. 8. At 0 s, a electrical signal is input to the explosion-proof electro-hydraulic valves, means to the vehicle begins to brake. High-pressure oil inlets brakes on both sides of the front and rear axles continuously. The front and rear brakes reached

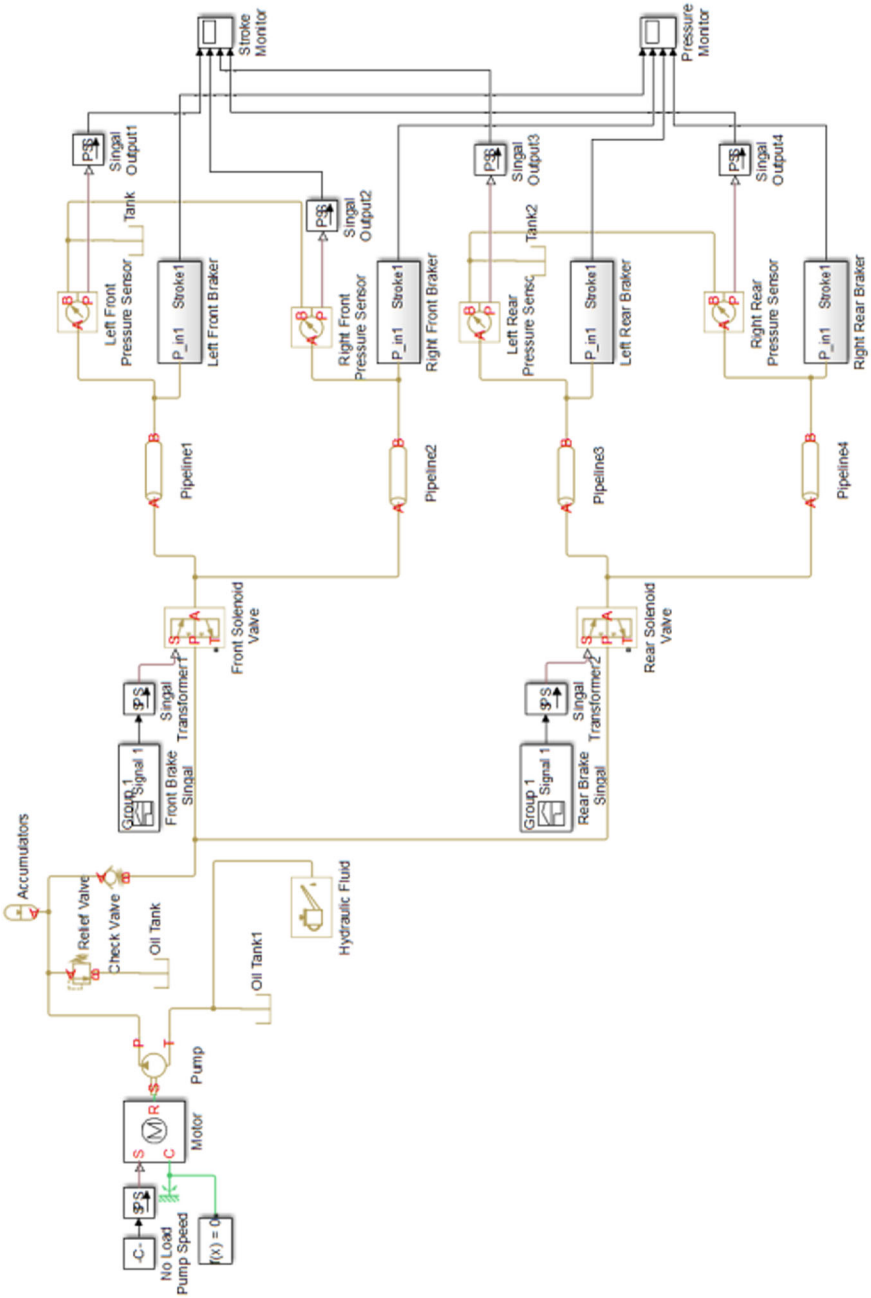


Fig. 3 Simulation model of a wet brake by wire model in Simulink

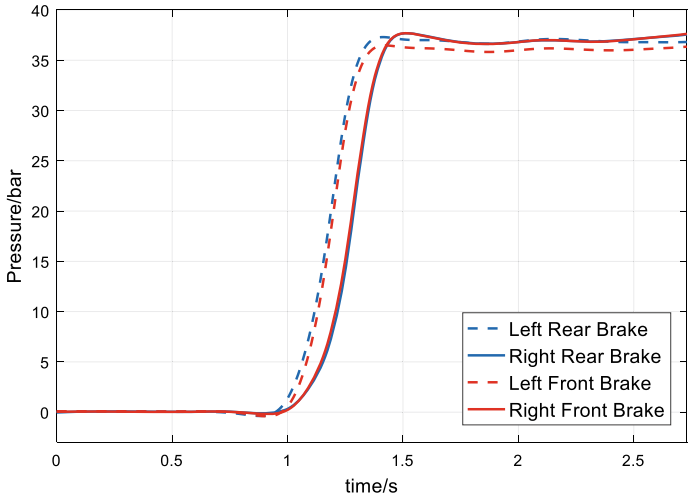


Fig. 4 Pressure variation curves of the four brakes when hydraulic braking



Fig. 5 Explosion-proof electro-hydraulic valves



Fig. 6 Pressure sensor

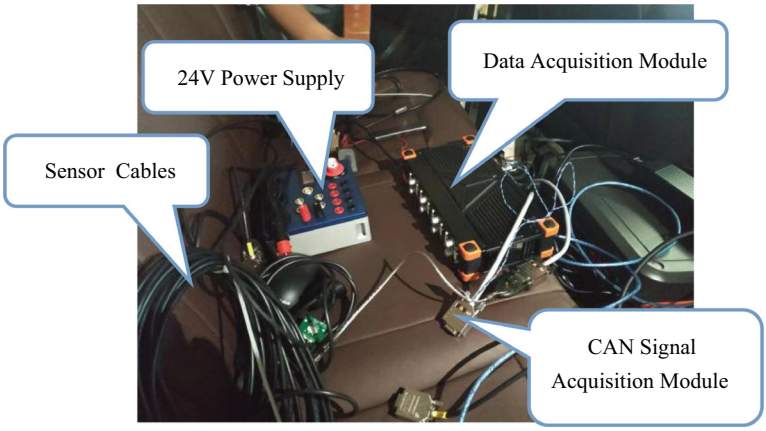


Fig. 7 Testing equipment's

design pressure at 0.4 s and 0.5 s respectively, at which point the front and rear axles were locked. The measured and simulated pressure curves responses of the vehicle are in good agreement with each other.

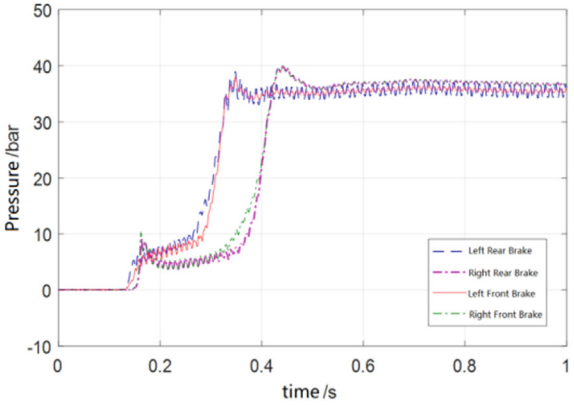


Fig. 8 Pressure curves at the four brake inlet ports without motor feedback braking

In order to test motor feedback braking distribution strategy, a linear proportional control strategy is adopted with the brake pedal stroke. When the braking strength $Z \leq 0.5$, the motor feedback braking torque increases with the increase of braking intensity. When the braking strength $Z > 0.5$, the motor feedback braking is set to 0.3 times of the motor maximum torque.

Through adding motor feedback braking into hydraulic braking, the pressure curves on rear axle sides shown in Fig. 9 are smoother than Fig. 8, while the motor feedback braking acts on the rear axle braking, the whole braking response time is shortened. The response time of brake by wire is 0.41 s to reach the design value, and the overall response time is shortened by 0.09 s. The front and rear wheels are basically synchronized and reach a locked state, and the time for the front wheels to lose steering is significantly reduced by 90%, improving the vehicle safety (Fig. 10).

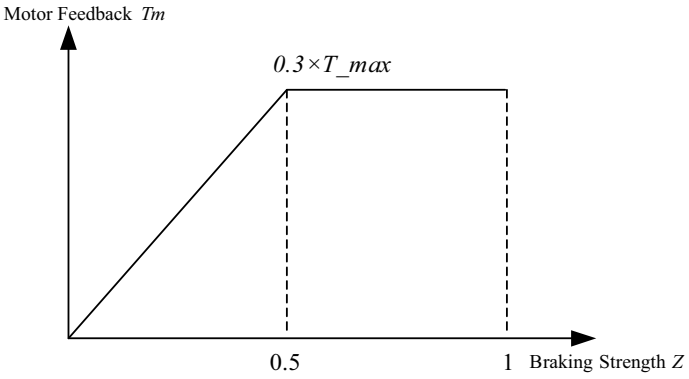


Fig. 9 Feedback brake curve

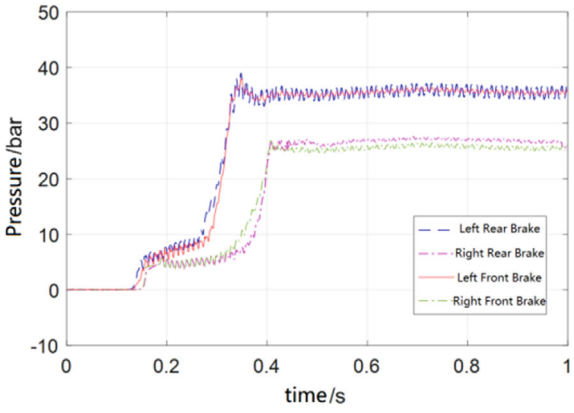


Fig. 10 Pressure curves at the four brake inlet ports within motor feedback braking

4 Conclusion

The article studies wet brake by wire system of mining electric vehicles, analyzes braking response time and motor feedback braking distribution, proposes a strategy to improve the braking performance by controlling the motor feedback braking ratio. Through simulation and testing the wet brake by wire of 5-seater explosion-proof electric rubber tyred vehicle, the following conclusions are drawn:

(1) When the vehicle's are hydraulic braking without motor feedback braking, simulation analysis and experimental testing all show that the complete braking response time is about 0.5 s when the front and rear axle wheels all brake, and there is a time difference when the front and rear axles reach the maximum pressure, that is, the front and rear axles cannot lock at the same time, and there is a risk of the front wheels losing steering during braking;
(2) By designing a motor feedback braking strategy, the rear axle braking adopts motor feedback braking and hydraulic braking. The complete response time is shortened to 0.41 s, the time difference of the front and rear axles reach the maximum pressure at the same time is significantly reduced. The front and rear wheels of the vehicle are almost simultaneously locked, and the braking safety of the vehicle is improved;

At present, mining electric trackless rubber tyred vehicles are not equipped with ABS because of explosion-proof problem. Further research is needed on brake by wire with ABS to solve the braking problem of mining electric vehicles, the safety performance can be further improved. This will promote the rapid development of explosion-proof unmanned vehicles and enhance the intelligence level of coal mines.

Acknowledgements Sponsored by Shanxi Province Basic Research Program Jointly Project (202104021301067).

References

1. Chen Y, Wei J (2022) Study on automatic steer-by-wire system of mine explosion-proof auxiliary vehicles. METAL MINE 555(9):174–179
2. Chen X (2022) Current status and development trend of intelligent technology of underground coal mine transportation system. J Mine Autom 48(6):6–14
3. Zhou L (2022) Research on unmanned driving system of underground trackless rubber-tyred vehicle in coal mine. J Mine Autom 48(6):37–48
4. Huang J (2015) Research on regenerative braking charging safety and braking stability for battery electric vehicle. Chongqing University doctor paper, chongqing
5. Wachter E, Ngu TQ, Alirand M (2019) Virtual simulation of an electro-hydraulic braking system. ATZ Worldwide 121(07–08):54–59
6. Li B, Li D, Ge W et al (2021) Precision control of hydraulic pressure in fast-response brake-by-wire system based on direct-drive valve. China J Highw Transp 34(09):121–132
7. Ma X, Tao X, Wen L et al (2023) Optimization of smoothness for electro-hydraulic composite braking of mining explosion-proof electrical rubber-tired vehicle. Coal Eng 55(11):248–251
8. Vignati M, Belloni M, Sabbioni E, Tarsitano D (2021) A regenerative braking strategy for independently driven electric wheel accounting for contemporary use of electric and hydraulic brakes. In: The IAVSD international symposium on dynamics of vehicles on roads and tracks, pp 1256–1268
9. Yang W (2018) Simulation and experiment research on dynamic characteristic of automobile EHB hydraulic system. Hunan University doctor paper, Changsha
10. Dehua G (2020) Simulation and experimental study of electro-hydraulic combined braking system for heavy-duty vehicles. Jilin University doctor paper, Changchun

Open Access This chapter is licensed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License (http://creativecommons.org/licenses/by-nc-nd/4.0/), which permits any noncommercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if you modified the licensed material. You do not have permission under this license to share adapted material derived from this chapter or parts of it.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.



Parameter Optimization Design of Air Springs for Subway Vehicles



Xianhui Su

Abstract By constructing a dynamic model of subway vehicles that includes air springs, the dynamic data of the air springs were selected. The main optimized parameters are the lateral stiffness and vertical stiffness of the air spring, and the main evaluation indicators include the vehicle's motion stability, running smoothness, and running safety (curve passing performance). The research results indicate that an increase in vertical stiffness will lower the critical speed of the vehicle. An increase in the horizontal stiffness of the air spring will have a slight impact on the lateral stability index, while an increase in vertical stiffness will significantly affect the vertical stability index, but will reduce the lateral vibration of the frame. Considering all factors, it is recommended that the lateral stiffness of the secondary air spring be 0.15 MN/m and the vertical stiffness be 0.35 MN/m.

Keyword Air springs · Dynamics · Stiffness · Subway vehicles

1 Introduction

Air spring has excellent dynamic performance, therefore, it is widely used in the secondary suspension of subway vehicles [1]. Air springs can not only withstand vertical vibrations, but also utilize their excellent lateral stiffness characteristics to withstand lateral forces. As one of the key technologies of bogies, the performance of air springs determines the stability, safety, and stationarity of vehicles. Therefore, the design and optimization of air springs have become a focus of research for scholars both domestically and internationally.

Garcia et al. [2] summarized and compared the equivalent models, thermodynamic models, analytical models, and finite element models of railway vehicle air springs over the past decade. They pointed out that there is currently limited research on lateral models, and further research should be conducted on the nonlinearity and fluid

X. Su (✉)
Hefei Rail Transit Group Operation Co., Ltd. Hefei, Anhui, China
e-mail: suxianhui@hfgdjt.com

© The Author(s) 2026

H. Saman et al. (eds.), *The 9th International Conference on Advances in Construction Machinery and Vehicle Engineering*, Lecture Notes in Mechanical Engineering,
https://doi.org/10.1007/978-981-96-5527-4_61

solid coupling of rubber airbags. Zhang [3] established a high-speed train dynamics model that includes an air spring system. Based on vertical vibration transmission test of air spring, the accuracy of the nonlinear dynamic model was verified. Numerical simulation was used to analyze the vertical dynamic stiffness characteristics and vibration transmission laws of the air spring. Qi [4] derived the air equation inside the air spring based on fluid dynamics theory, established a fluid dynamics model of the air spring based on dynamic mesh theory, and the accuracy of the model was verified by bench tests. Explored the calculation method of turbulence inside the air spring system, the stiffness characteristics was studied under different loads. Feng [5] used ABAQUS software to simulate the rubber air spring system, and proved that the air spring has good vibration reduction effect.

In this article, a nonlinear dynamic vehicle model containing air springs is optimized using Simpack software to design the parameters of the air springs. The main optimized parameters are the horizontal and vertical stiffness of the air spring, and the main analysis indicators include the vehicle's motion stability, running smoothness, and running safety (curve passing performance).

2 Dynamic Performance Evaluation Criteria

For the evaluation of the dynamic performance of subway vehicles, the standard adopted is the national standard GB/T5599-1985 "Code for Evaluation and Testing of Dynamic Performance of Railway Vehicles", and the calculated dynamic performance indicators must meet the following requirements.

2.1 Stability

The design speed of the subway vehicle studied in this article is 80 km/h, which requires the critical speed for instability to be higher than this speed. And there should be sufficient safety margin.

2.2 The Safety of Passing Through Curves

Safety is evaluated through three indicators: derailment coefficient, wheel load reduction rate, and lateral wheel rail force.

Derailment coefficient assessed by wheel rail force-Q/P.

$$\frac{\mathrm{Q}}{P} < 0.8 \tag{1}$$

Table 1 Limit value of lateral wheel rail force

Working condition	Wheel load/kN	Lateral wheel rail force/kN
Motor car (AW0)	44.78	24.93
Trailer (AW0)	42.37	24.12
Motor car (AW3)	76.38	35.46
Trailer (AW3)	74.19	34.73

Q is the lateral force between the wheel-rail; P is the vertical force between the wheel-rail.

Wheel load reduction rate:

$$\frac{\Delta \mathrm{P}}{P} < 0.6 \tag{2}$$

△P is a wheel load reduction capacity; P is the average wheel weight for increasing and decreasing the load.

The allowable limit of lateral force between wheel-rail adopts the following standards.

$$H_{\lim} = \beta * (10 + P_0/3) \tag{3}$$

$H_{\lim}$ is limit value of lateral force between wheel and rail. P_0 is empty axle load. For locomotives, high-speed trains and other vehicles, $\beta = 1$.

According to the formula (3), the following result can be obtained, as shown in Table 1.

2.3 Stability of Operation

According to the national standard GB/T5599-1985, the stability of operation is evaluated based on the stability index and maximum acceleration.

Stability index W, W < 2.5 is excellent level, 2.5 < W < 2.75 is medium level.

Maximum acceleration: The maximum vertical and lateral acceleration of the car body should be less than 0.25 g.

The stability index of operation is calculated according to the following formula.

$$W = 7.08^{10}\sqrt{\frac{A^3}{f}F(f)} \tag{4}$$

In the formula, W is stability index, A is acceleration of the vehicle body, f is vibration frequency, F(f) is frequency correction coefficient.

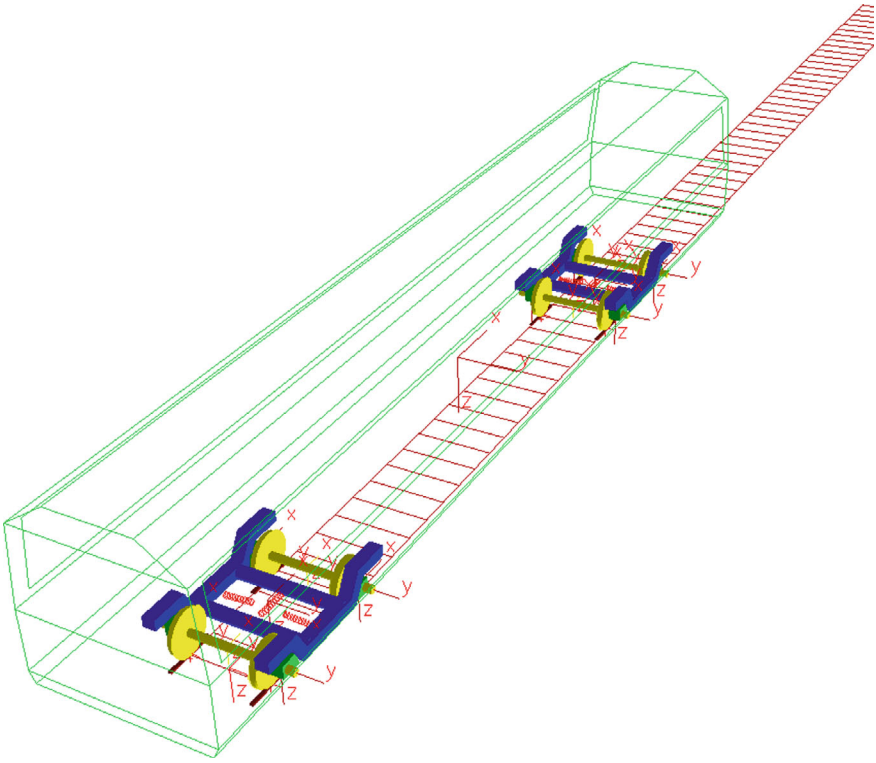


Fig. 1 Dynamic model

3 Dynamic Simulation Model

The nonlinear factors considered in the dynamic model include the geometric and interactive relationships between wheels and rails, and nonlinear force of suspension system. The dynamic model is shown in Fig. 1.

4 Parameter Optimization

In this chapter, the suspension parameters of the vehicle in an empty state are optimized, and the running speed selected for the stability calculation is 80 km/h; The working condition of a curved line is to pass through a bend with a radius of 800 m at a speed of 60 km/h.

The lateral stiffness range of the air spring is 0.09–0.24 MN/m, and the vertical stiffness range is 0.15–0.55 MN/m. The lateral stiffness and vertical stiffness are cross optimized.

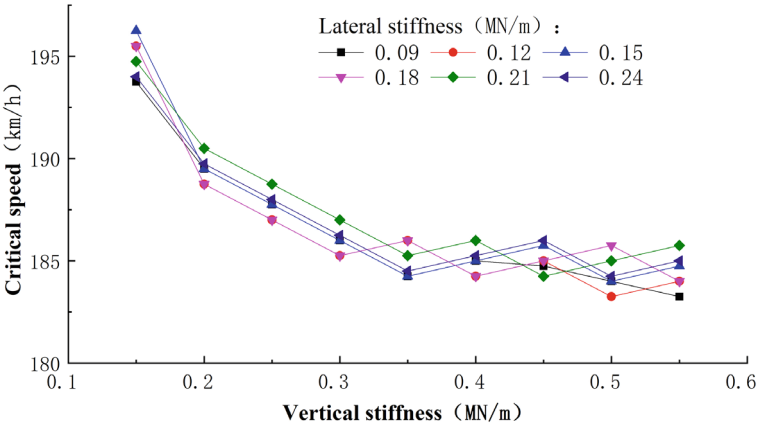


Fig. 2 The influence of air spring stiffness on the critical speed of vehicles

4.1 Stability

As shown in Fig. 2, according to the analysis results, as the lateral stiffness of the air spring increases, the change in the critical speed of the vehicle during operation is not significant, and an increase in vertical stiffness will lower the critical speed of the vehicle.

4.2 The Safety of Passing Through Curves

An increase in the lateral stiffness of the air spring will cause an increase in the vehicle's lateral stability index, but will reduce the lateral acceleration of the frame. As the vertical stiffness of the air spring increases, the vertical stability index of the vehicle first slowly increases, and then the trend of change increases, and the vertical vibration of the frame decreases to some extent. These data are shown in Figs. 3, 4, 5, 6, 7 and 8.

4.3 Stability of Operation

Under curved working conditions, as the vertical stiffness of the air spring increases, the lateral force between the wheel and rail first decreases. After the vertical stiffness reaches 0.25 MN/m, the lateral force between the wheel and rail changes very little. Under curved lines, the lateral stiffness has little effect on other safety indicators. As the vertical stiffness increases, safety parameters such as derailment coefficient,

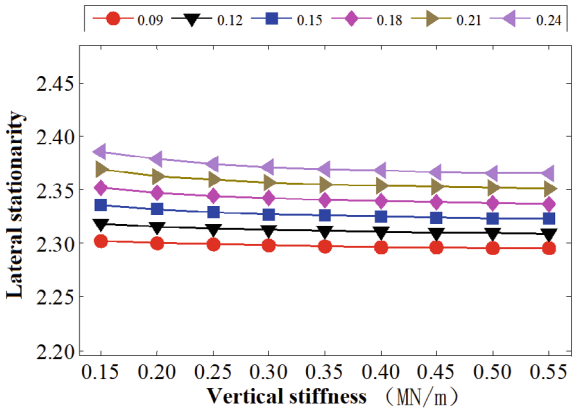


Fig. 3 The influence of stiffness on lateral stability

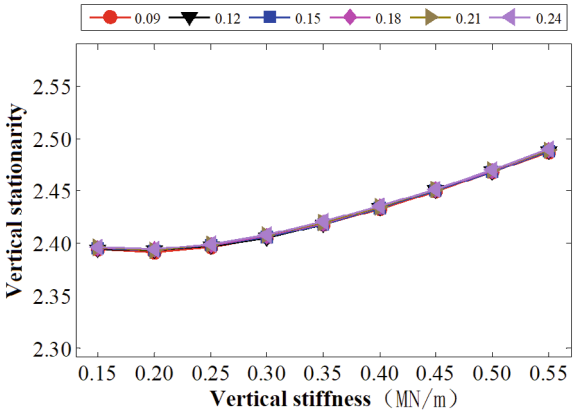


Fig. 4 The influence of stiffness on vertical stability

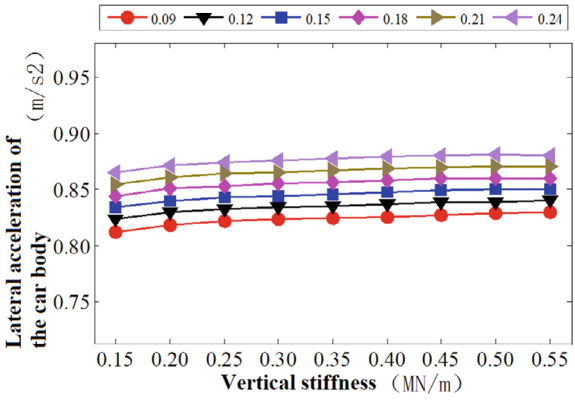


Fig. 5 The influence of stiffness on lateral acceleration of car body

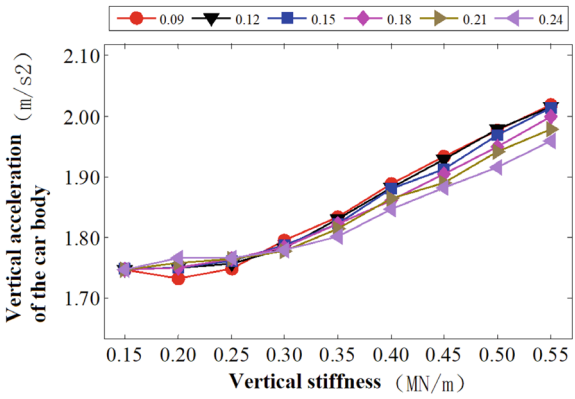


Fig. 6 The influence of stiffness on vertical acceleration of car body

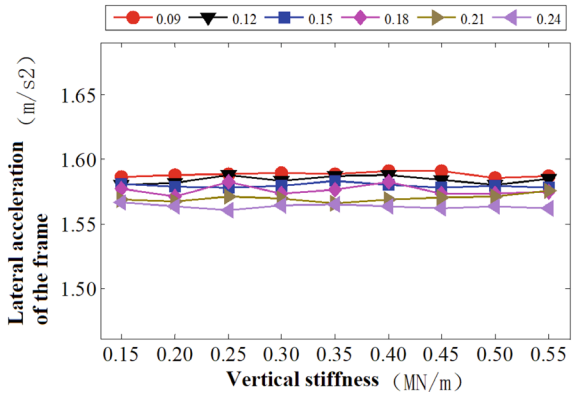


Fig. 7 The influence of stiffness on lateral acceleration of frame

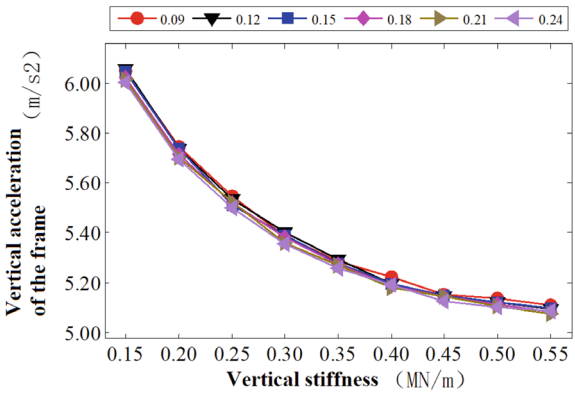


Fig. 8 The influence of stiffness on vertical acceleration of frame

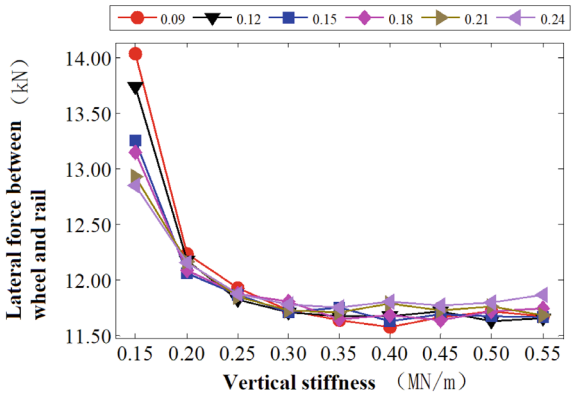


Fig. 9 The influence of stiffness on lateral force

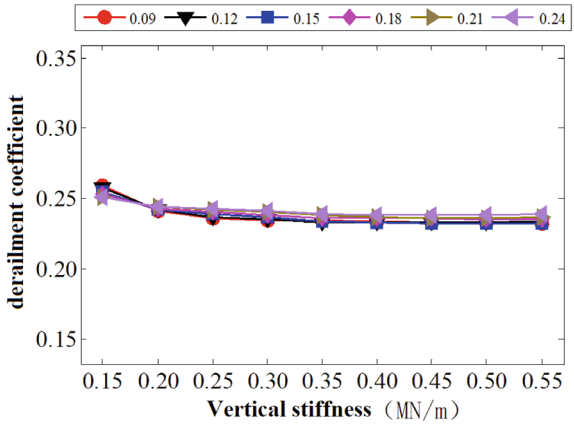


Fig. 10 The influence of stiffness on derailment coefficient

wear power metal gradually decrease. When the vertical stiffness exceeds 0.35 MN/m, the changes are small. These data are shown in Figs. 9, 10, 11, 12 and 13.

5 Conclusions

1. As the lateral stiffness of the air spring increases, the change in the critical speed of the vehicle during operation is not significant, and an increase in vertical stiffness will lower the critical speed of the vehicle.
2. An increase in the lateral stiffness of the air spring will cause an increase in the vehicle's lateral stability index, but will reduce the lateral acceleration of the frame. As the vertical stiffness of the air spring increases, the vertical stability index of the vehicle first slowly increases, and then the trend of change increases, and the vertical vibration of the frame decreases to some extent.

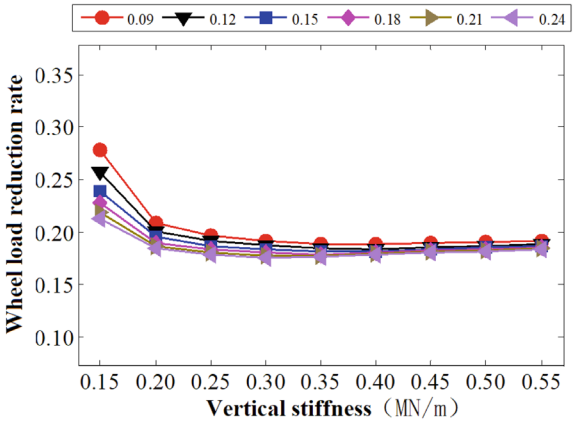


Fig. 11 The influence of stiffness on wheel load reduction rate

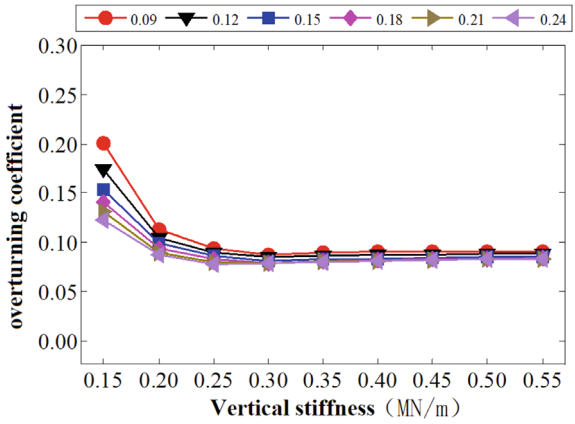


Fig. 12 The influence of stiffness on overturning coefficient

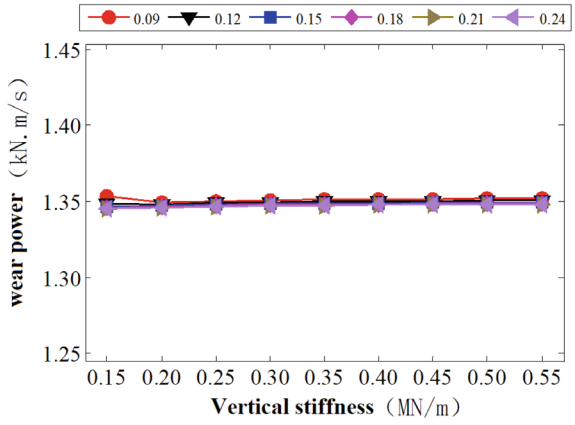


Fig. 13 The influence of stiffness on wear power

3. Under curved working conditions, as the vertical stiffness of the air spring increases, the lateral force between the wheel and rail first decreases. After the vertical stiffness reaches 0.25 MN/m, the lateral force between the wheel and rail changes very little. Under curved lines, the lateral stiffness has little effect on other safety indicators. As the vertical stiffness increases, safety parameters such as derailment coefficient, wear power metal gradually decrease. When the vertical stiffness exceeds 0.35 MN/m, the changes are small.

Considering all factors, it is recommended that the lateral stiffness of the air spring be 0.15 MN/m and the vertical stiffness be 0.35 MN/m.

References

1. Garcia I, Gil N, Pradera A et al (2022) A survey on the modelling of air springs secondary suspension in railway vehicles. Veh Syst Dyn 60(3):835–864
2. Zhang C, Luo R (2024) Research on vertical vibration of secondary suspension with air spring of high-speed train. Mach Build Autom 53(2):214–219
3. Qi Z, Jiang L, Liu P et al (2024) Analysis of internal flow field of air spring in high-speed EMUs and its effects on air spring dynamic characteristics. J Vib Shock 43(19):19–27, 36
4. Fan F, Jincao Z (2022) Finite element analysis of vibration characteristics of automobile rubber air spring based on ABAQUS. Adhesion 49(10):145–148
5. Li Y, Xiao S, Xie J et al (2023) Nonlinear dynamic mechanical characteristics of air springs based on a fluid–solid coupling simulation method. Appl Sci 13(23):12677

Open Access This chapter is licensed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License (http://creativecommons.org/licenses/by-nc-nd/4.0/), which permits any noncommercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if you modified the licensed material. You do not have permission under this license to share adapted material derived from this chapter or parts of it.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.



Numerical Simulation Analysis of Vehicle Equipment Safety Fastening System Design Scheme



Jingqiao Liu, Xin Jiang, Guo Wang, Mengxiang Li, Kun Liu, and Shiye Liu

Abstract In order to ensure the stability and safety of various types of vehicle equipment in the process of mothership transportation, the numerical simulation technology of vehicle equipment fastening system is studied based on nonlinear multi-body dynamics software. Taking typical vehicle equipment as the research object and numerical simulation modeling technology, inertial load application and model test analysis technology as the starting point, this paper analyzes the relative deck movement of vehicle equipment under various working conditions and the load time relationship of each lashing rigging in the safety fastening system of each vehicle equipment, and verifies and evaluates the safety of each lashing rigging.

Keywords Numerical simulation analysis · Vehicle equipment · Ships

1 Introduction

When transporting vehicle equipment on ships, it is necessary to secure the loaded vehicles to prevent them from moving or overturning during transportation, which could lead to damage to the ship and vehicles. As the importance of shipborne transportation becomes increasingly evident, this issue has also become an urgent problem that needs to be addressed [1]. Therefore, it is necessary to conduct safety studies on the fastening systems of shipborne vehicle equipment, establish relevant numerical analysis methods, and improve the fastening design level of large transport ships.

To ensure the stability and safety of various types of equipment during the transportation process on the mother ship, they need to be secured to the loading deck. However, the complexity of the marine environment and the large size and diversity

J. Liu · X. Jiang · G. Wang (✉)
Marine Design and Research Institute of China, Shanghai, China
e-mail: 917715625@qq.com

M. Li · K. Liu · S. Liu
Jiangsu University of Science and Technology, Zhenjiang, China

© The Author(s) 2026

H. Saman et al. (eds.), *The 9th International Conference on Advances in Construction Machinery and Vehicle Engineering*, Lecture Notes in Mechanical Engineering,
https://doi.org/10.1007/978-981-96-5527-4_62

of the loaded vehicles pose certain challenges in determining the fastening scheme. Therefore, it is necessary to consider different types of fastening objects, different fastening positions, different marine environments of the mother ship's sailing areas, and fastening costs comprehensively to establish a scientific and reasonable fastening scheme to ensure the stability and safety of the secured equipment.

Scholars at home and abroad have conducted a large number of studies on fastening issues in shipping. Turnbull et al. [2] simplified the suspension, tires, and brackets of semi-trailers, established a numerical model for the fastening and securing of semi-trailers on the deck of a ro-ro ship, and analyzed in detail the forces on the vehicle fastening equipment as it follows the overall rolling and pitching motion of the ro-ro ship. They studied the impact of different rolling and pitching motion periods, different initial phases of motion, and different positions of vehicles on the deck on the maximum load of the fastening equipment. Jia [3, 4] conducted experimental research on the frictional forces on the tires of cars on ro-ro ships and found that the safety of vehicle transportation mainly considers sliding. At the same time, under certain weather conditions and on some routes with good sea conditions, vehicles can be fixed on ships without being tied down. Surendran [5] established a numerical model of ship motion to ensure safety during navigation, and conducted a comprehensive study on the rolling problem of ro-ro ships in waves and how to improve the stability of ships in resonance with waves. Zhao [6] combined the multibody dynamics analysis software ADAMS to establish a shipborne helicopter fastening model, simulated and calculated the tension of the shipborne helicopter fastening equipment, and systematically analyzed the factors affecting the tension of the shipborne helicopter fastening equipment, which can be used to guide the subsequent fastening research of shipborne equipment.

This paper will focus on the safety issues of fastened vehicles during the mother ship's voyage, establish a coupled motion simulation analysis technology for the mother ship-fastened equipment-fastening equipment, and apply it to the verification and evaluation of vehicle fastening methods. Based on this, it will clarify the impact patterns of different fastening parameters, carry out research on the safety fastening scheme of vehicle equipment, and form related fastening design guidance documents to support the improvement of equipment fastening design level.

2 Multibody Dynamics Analysis Theory

In recent years, multibody dynamics analysis methods have gradually come into the public eye and have become one of the hot and difficult points in current dynamic simulation research. The emergence of ADAMS software has made the solution and analysis of complex multibody systems simpler. Using this software for multibody dynamics modeling, calculation, and analysis has the advantages of wide applicability, the ability to calculate large displacement motion, and high model precision [7].

In the ADAMS software, the centroid Cartesian coordinates of rigid body B and the Euler angles that reflect the orientation of the rigid body are used as generalized coordinates, that is to say: $q = [x, y, z, \psi, \theta, \varphi]^T$, let $R = [x, y, z]^T$, $\gamma = [\psi, \theta, \phi]^T$, $q = \left[R^T, \gamma^T\right]^T$. The coordinate transformation matrix between the centroid reference coordinate system of the component and the ground coordinate system is:

$$A^{gi} = \begin{bmatrix} \cos\psi\cos\phi - \sin\psi\cos\theta\sin\phi & -\cos\psi\sin\phi - \sin\psi\cos\theta\cos\phi & \sin\psi\sin\theta \\ \sin\psi\cos\phi + \cos\psi\cos\theta\sin\phi & -\sin\psi\sin\phi + \cos\psi\cos\theta\cos\phi & -\cos\psi\sin\theta \\ \sin\theta\sin\phi & \sin\theta\cos\phi & \cos\theta \end{bmatrix} \tag{1}$$

Define an Euler axis coordinate system, in which the three unit vectors are the axes of the aforementioned three Euler rotations. Since these three axes are not mutually perpendicular, the coordinate transformation matrix from this coordinate system to the component's centroid coordinate system can be written as:

$$B = \begin{bmatrix} \sin\theta\sin\phi & \cos\theta & 0 \\ \sin\theta\sin\phi & -\sin\theta & 0 \\ \cos\theta & 0 & 1 \end{bmatrix} \tag{2}$$

Additionally, the angular velocity of the component can be expressed as:

$$\omega = B\dot{\gamma} \tag{3}$$

In ADAMS, the variable ω_e is introduced as the component of angular velocity in the Euler axis coordinate system:

$$\omega_e = \dot{\gamma} \tag{4}$$

In ADAMS software, considering the constraint equations, the energy form of the first kind of Lagrange's equation with Lagrange multipliers can be used to solve the following equation:

$$\frac{d}{dt}\left(\frac{\partial T}{\partial \dot{q}_j}\right) - \frac{\partial T}{\partial q_j} = Q_j \sum_{i=1}^{n} \lambda_i \frac{\partial \Phi}{\partial q_j} \tag{5}$$

In Eq. (5), T represents the kinetic energy of the system with respect to the generalized coordinates, q_j are the generalized coordinates, Q_j represents the generalized force in the direction of generalized coordinate q_j, and the last term involves the constraint equations and Lagrange multipliers, representing the constraint reaction force in the direction of generalized coordinate q_j. In the ADAMS software, the expression for generalized momentum is further introduced:

$$P_j = {}^{\partial T}\big/_{\partial \dot{q}_j} \tag{6}$$

Simplify the expression of the constraint reaction force as:

$$C_j = \sum_{i=1}^{n} \lambda_i \frac{\partial \Phi}{\partial q_j} \tag{7}$$

Simplifying Eq. (5), we get:

$$\dot{P}_j - \frac{\partial T}{\partial q_j} = Q_j - C_j \tag{8}$$

Additionally, the expression for kinetic energy can be further simplified to:

$$T = \frac{1}{2}\dot{R}^T M \dot{R} + \frac{1}{2}\dot{\gamma}^T B^T J B \dot{\gamma} \tag{9}$$

In Eq. (9), MM represents the mass matrix of the component, and JJ represents the inertia matrix of the component in the centroid coordinate system. Transforming Eq. (8) into the structural form of translation and rotation directions is as follows:

$$\dot{P}_R - \frac{\partial T}{\partial q_R} = Q_R - C_R \tag{10}$$

$$\dot{P}_\gamma - \frac{\partial T}{\partial q_\gamma} = Q_\gamma - C_\gamma \tag{11}$$

thereinto $\dot{P}_R = \frac{d}{dt}\left(\partial T / \partial \dot{q}_R\right) = \frac{d}{dt}(M\dot{R}) = M\dot{V}$, $\frac{\partial T}{\partial q_R} = 0$.

Equation (10) can be simplified to:

$$M\dot{V} = Q_R - C_R \tag{12}$$

Thereinto, $P_\gamma = \left(\frac{\partial T}{\partial \dot{q}_\gamma}\right) = B^T J B \dot{\gamma}$, Since B$B$ contains Euler angles, no further derivation is made in the ADAMS software. To simplify the derivation process, it is treated as a variable to be solved. Therefore, in the software, each component has the following 15 variables and 15 equations, not 12. The specific variables and equations are as follows:

$$\begin{cases} V = \left[V_x, V_y, V_z\right]^T \\ R = [x, y, z]^T \\ P_\gamma = \left[P_\psi, P_\theta, P_\phi\right]^T \\ \omega_e = \left[\omega_\psi, \omega_\theta, \omega_\phi\right]^T \\ \gamma = [\psi, \theta, \phi]^T \end{cases} \tag{13}$$

$$\begin{cases} M\dot{V} = Q_R - C_R \\ V = \dot{R} \\ \dot{P}_\gamma - \frac{\partial T}{\partial q_\gamma} = Q_\gamma - C_\gamma \\ P_\gamma = B^T JB\omega_e \\ \omega_e = \dot{\gamma} \end{cases} \tag{14}$$

The ADAMS software can automatically establish the dynamic equations of the entire system (differential–algebraic equations):

$$\begin{cases} \dot{P} - {}^{\partial T}\big/_{\partial q} + \Phi_q^T \lambda + H^T F = 0 \\ P = {}^{\partial T}\big/_{\partial \dot{q}} \\ u = \dot{q} \\ \Phi(q, t) = 0 \\ F = f(u, q, t) \end{cases} \tag{15}$$

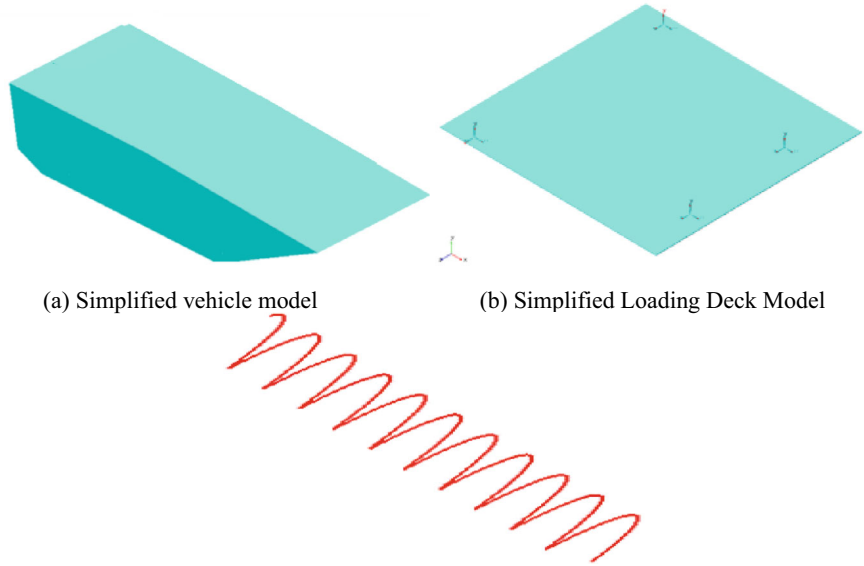
In the equation, P represents the generalized momentum of the entire system, andH is the coordinate transformation matrix for external forces.

3 Establishment of Numerical Simulation Models

In the vehicle securing technology on ship decks, it includes the contact between vehicles and ship decks, the connection between securing equipment and vehicles, the connection between securing equipment and ship decks, and the settings of connecting joints and motion joints during the ship's motion. At the same time, there are complex contact algorithms involved, which makes the calculation work extremely large. During the vehicle securing motion process, the main focus is on the dynamic response of the securing equipment, while the force and deformation of the deck and vehicles are basically not considered. When performing the analysis and calculation, it is necessary to appropriately and reasonably simplify the scene of vehicles loaded on the deck of a sailing ship and arrange the vehicle securing scheme.

When conducting research on the performance of deck vehicle securing, the focus is mainly on the load-bearing capacity of the securing equipment. Therefore, the vehicle and ship deck model structures can be simplified to be rigid, and simplified vehicle (Fig. 1a) and loading deck (Fig. 1b) structural models can be established respectively. Securing equipment securing points (marker points) are reserved on the vehicle and deck structures according to the specific securing form. By defining the securing equipment (Fig. 1c), the predefined securing points on the vehicle and deck are connected to form an overall numerical simulation model.

In the ADAMS software, spline curves can be utilized to define the stiffness of securing equipment. During the definition of the spline curve, the deformation and



(a) Simplified vehicle model (b) Simplified Loading Deck Model

(c) Simplified Securing Equipment Model

Fig. 1 Simulation calculation model

force variation of the securing equipment are used to represent its stiffness curve, and the required stiffness can be adjusted according to actual needs. Additionally, the pretension force of the securing equipment can be added on demand during the establishment of the equipment using the preload option. In the ADAMS software modeling process, when defining the contact between the vehicle deck, the threshold values for the static and dynamic friction coefficients can be defined in the contact settings. Based on the actual motion conditions of rolling, pitching, and heaving, parametric definitions are made, and point-driven methods are used to establish the motion of the overall model structure. Parametric equations for motion can be input in the drive settings.

4 Numerical Simulation Verification

To verify the validity of the numerical model, this paper relies on the three-degree-of-freedom sway test bench of Jiangsu University of Science and Technology to carry out model scale tests. The numerical simulation model of the vehicle equipment is shown in Fig. 1, and its securing scheme is shown in Fig. 2. The main dimensions of the vehicle scale model in length, width, height, and weight are 13,240 mm, 3370 mm, 3600 mm, and 224 kg, respectively. The other ship motion conditions

and main parameters of the vehicle equipment are shown in Tables 1 and 2, and the results of the numerical simulation and model test are shown in Figs. 3 and 4.

From the above results, it can be observed that under the 6th level sea condition, the load of the securing equipment for each single degree of freedom motion changes with time following a basically consistent pattern, all showing a sinusoidal/cosecant periodic change trend. The period is consistent with the swaying period of the swing table. The periodicity and change trend of the securing load time history curves obtained from the experiment and simulation for the vehicle equipment are basically consistent, with only a certain error existing in the peak load. Under the scenario of the 6th level sea condition, the peak values of the securing loads of the securing equipment under the action of ship rolling are 271.3 and 290.7 N for the experiment and simulation, respectively, with a difference of 7.15%. Under the action of pitching, the peak values of the securing loads of the securing equipment are 203.1 and 197.6 N for the experiment and simulation, respectively, with a difference of 2.71%. It can be seen that the error during pitching motion is relatively small, and the error during rolling motion is relatively large, but still within the range of 10%. This indicates that the established numerical simulation model can relatively accurately predict the load characteristics of the securing equipment under ship motion scenarios.

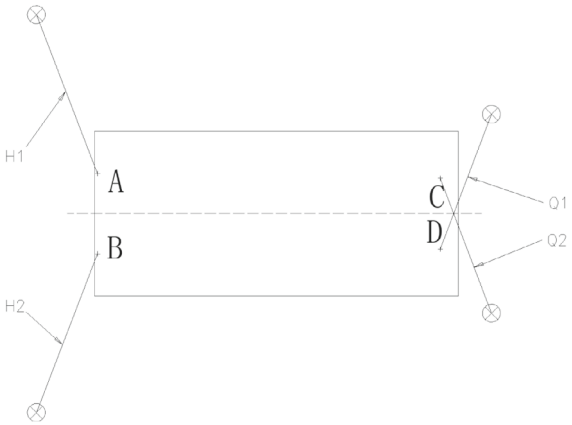


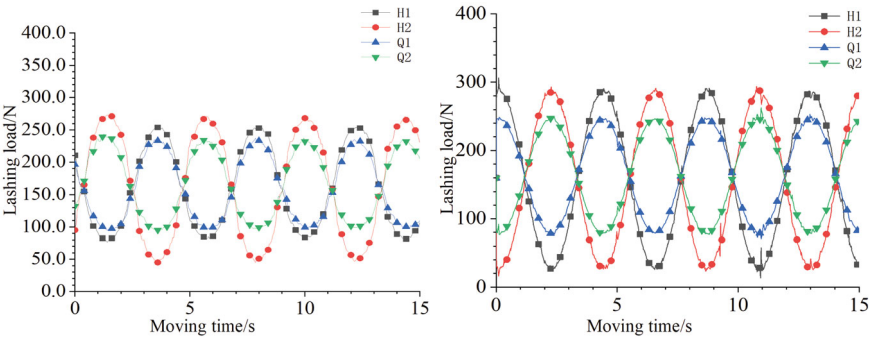
Fig. 2 Model vehicle tethering scheme

Table 1 Ship movements corresponding to class 6 sea conditions

Motion parameters		Level 6
Rolling	Maximum roll angle (°)	25.4
	Rolling period (s)	9.6
Pitching	Maximum pitching angle (°)	4.4
	Pitching period (s)	9.6

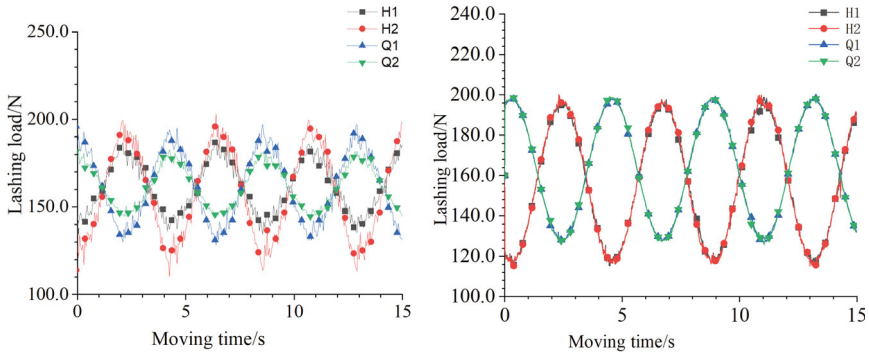
Table 2 Basic information of vehicle equipment

Project	Numerical
Vehicle weight (t)	55.000
Front left and right hook height (Lower)	0.780
Rear left and right hook height (Lower)	0.840
Vertical distance from vehicle center of gravity to deck Zc	1.437
Horizontal distance from vehicle center of gravity to side support point LG transverse	1.658
Horizontal distance from vehicle center of gravity to side support point LG longitudinal (Distance from rear)	2.070
Horizontal distance from vehicle center of gravity to side support point LG longitudinal (Distance from front)	2.761



(a) Experimental Results (b) Simulation Results

Fig. 3 Relation curve of vehicle tethered rigging load with time under the action of class 6 sea state sag



(a) Experimental Results (b) Simulation Results

Fig. 4 Time variation curve of vehicle tethered rigging load under class 6 sea state pitching

Table 3 Corresponding motion parameters of mother ship

	Level 6 sea state	Level 9 sea state
Maximum hull roll angle ø(°)	25.400	30.000
Maximum hull pitch angle φ(°)	4.400	10.000
Maximum hull heave value (m)	2.8	7
Waterline width B (m) corresponding to full load displacement	34.000	34.000
Length between perpendiculars Lpp (m)	240.000	240.000
Initial stability height h (m) corresponding to full load displacement	3.938	3.938
Roll period Tr (s)	9.600	13.000
Pitch period Tp (s)	9.600	13.000
Heave period Th (s)	9.75	12.8
Vertical distance Zg (m) from the center of gravity of the hull to the baseline at full load displacement	14.4	
Vertical distance h (m) from the deck to the baseline of the hull	12.800	
Longitudinal position of the center of gravity of the hull corresponding to full load displacement (Negative for aft of amidships) (m)	− 1.319	
Preload force of the securing equipment (N)	20,000	

5 Analysis of Typical Vehicle Equipment Safety Securing Scheme Design

5.1 Different Motion Scenarios of the Mother Ship

This paper primarily focuses on the issues of rolling, pitching, and heaving motions during ship navigation. Based on the different motion conditions of the hull being studied, the motion scenarios of the hull in the simulation calculations are mainly divided into: two-degree-of-freedom coupled motion and three-degree-of-freedom coupled motion. The corresponding motion parameters of the mother ship are listed in Table 3.

5.2 Analysis of Vehicle Equipment Securing Scenarios

Due to the variety of equipment securing schemes, in order to more clearly and conveniently represent the position of the securing equipment when analyzing the load-bearing capacity of the securing equipment, a unified numbering system is adopted for the securing equipment. We use “Q” to denote the securing equipment on the front side of the vehicle and “H” for the rear side. Facing the direction of the vehicle’s front, the numbering sequence is arranged from left to right. For example,

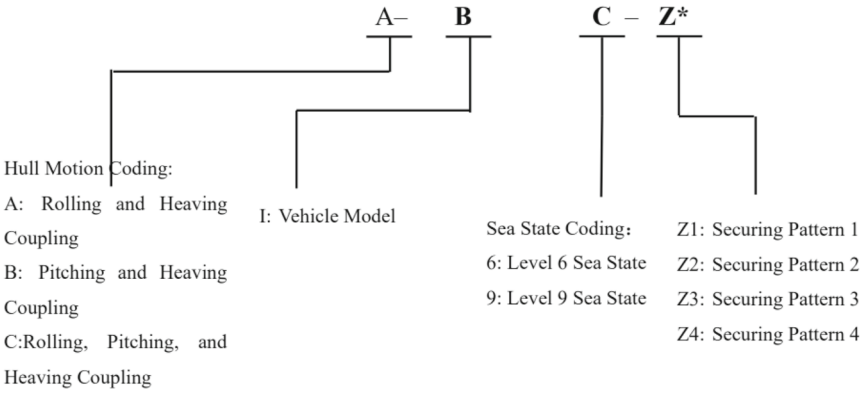


Fig. 5 Calculation conditions naming rules

the first securing equipment on the left front side of the vehicle is denoted as Q1, and the second one as Q2.

For clarity and convenience, a standardized naming convention for all calculation scenarios is established as shown in Fig. 5.

Taking A-I-6-Z1 as an example: it indicates the calculation scenario where the vehicle equipment is secured in way 1 under the coupled rolling and heaving conditions of a ship in a 6th-level sea state.

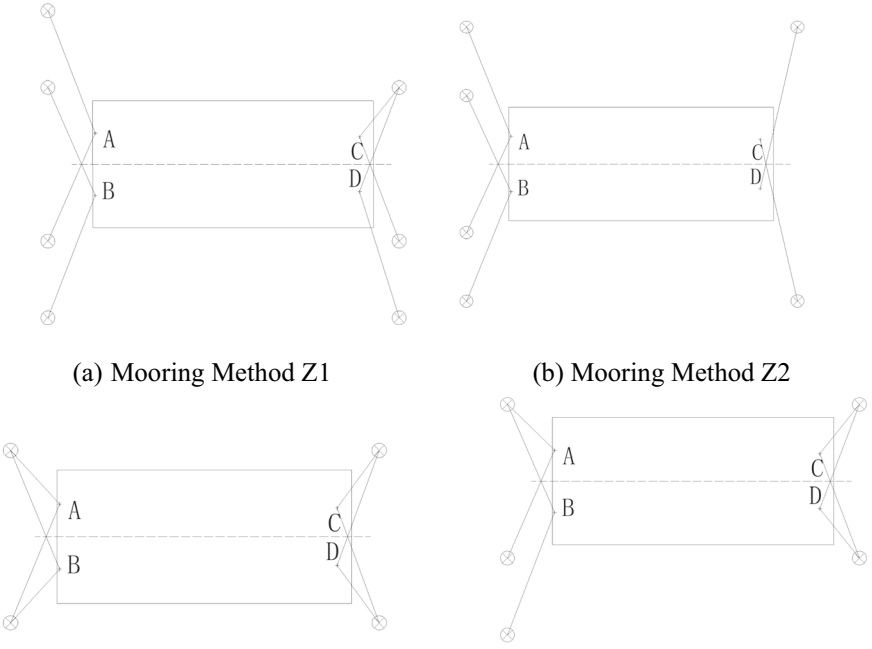
This paper focuses on several typical securing scenarios for equipment I and uses ADAMS to establish the corresponding securing simulation models as shown in Fig. 6.

5.3 Analysis of Computational Results for Typical Securing Scenarios of Vehicle Equipment

To ensure the safety of equipment securing on the ship's deck under typical securing scenarios, it is necessary to analyze the force conditions of the securing equipment under the influence of different ship motion scenarios, securing forms, and the spatial positions of the equipment securing. This section will conduct a comprehensive analysis and calculation of the aforementioned parameters, thereby revealing the variation rules of the securing force of the vehicle securing equipment, and providing a reference for the design of safe securing schemes for vehicle equipment on the ship's deck.

Level 6 Sea State

(1) *Rolling and Heaving Coupling.*



(a) Mooring Method Z1 (b) Mooring Method Z2

(c) Mooring Method Z3 (d) Mooring Method Z4

Fig. 6 Typical tethered scenario model

Figure 7 shows the securing loads of the securing equipment under the securing methods Z1 to Z4 in a 6th-level sea condition. As shown in Fig. 7b, during the simulation, the securing equipment on the left and right sides of the vehicle are symmetric about the centerline in the direction of the vehicle's length, and when the rolling center is on the vehicle's centerline, the force conditions on the left and right sides of the securing equipment are also basically symmetric throughout the motion process. However, due to the presence of 2 securing equipment on the front of the vehicle and 4 on the rear, the maximum tension of 39.6 kN appears on the front securing equipment Q1 and Q2 ropes, while the maximum securing force of 25.9 kN on the rear securing equipment appears on the H4 securing equipment. Then, as shown in Figs. 7a–c, when the vehicle is in a special position, such as near the bulkhead, causing the vehicle's securing scheme to be asymmetric on the left and right, or when the rolling center is not on the vehicle's centerline, there will be a significant difference in the force on the securing equipment. From Figs. 2, 3, 4, 5, 6 and 7, it can be seen that since the securing equipment Q1 and Q2 are on the same side, and Q3 and Q4 are on the same side, the securing forces of Q1 and Q2 are in phase, and the securing forces of Q3 and Q4 are also in phase, with a half-cycle difference between the two, which is the same as the actual swinging situation. At this time, the maximum securing force of the securing equipment appears on Q3, and the amplitude reaches 35.6 kN. From Fig. 7d, it can be seen that the maximum

securing force of the securing equipment appears on Q3, and the amplitude reaches 32.7 kN.

(2) *Pitching and Heaving Coupling.*

Under 6th-level sea conditions, as shown in Fig. 8, the amplitudes of the securing equipment's securing forces appear on Q1, Q2, Q4, and Q1, respectively, and are approximately 26.6, 29.6, 26.6, and 25.0 kN. During the simulation, the forces on the securing equipment during ship pitching differ significantly from those during ship rolling. When the ship rolls, the securing equipment on the left or right side of the vehicle is subjected to force simultaneously. For the ship pitching scenario, the securing equipment on the front or rear side of the vehicle is subjected to force simultaneously. When the securing equipment on the front or rear side of the vehicle is symmetric about the centerline of the vehicle's length, the force conditions on the front or rear securing equipment are also basically the same throughout the coupled pitching and heaving motion process. If the vehicle's securing scheme is not symmetrical on the left and right sides of the front and rear, or if the vehicle layout is not symmetrical along the length of the ship, there will be a significant difference in the force on the securing equipment.

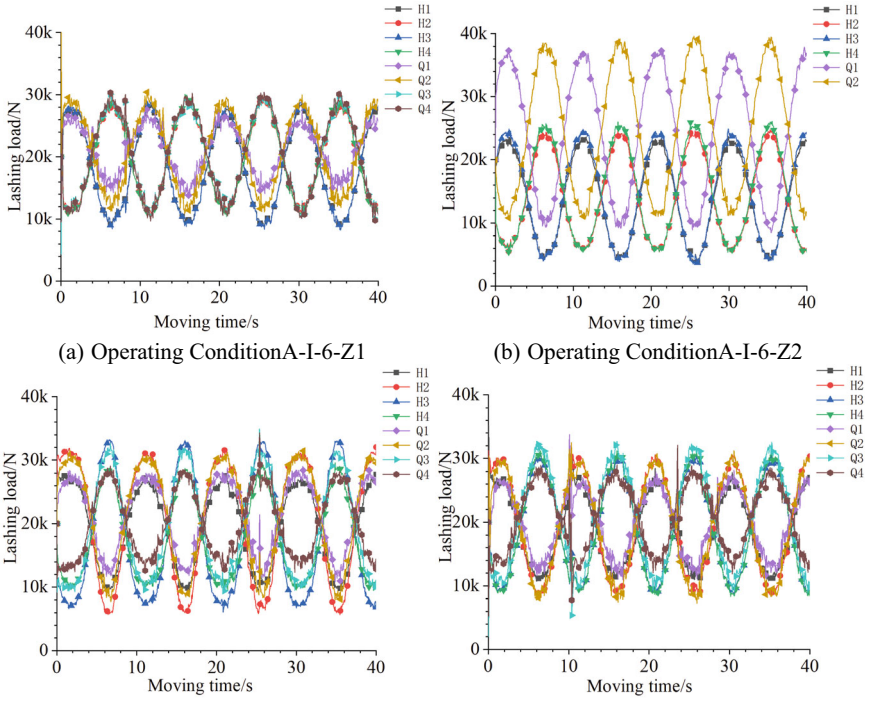
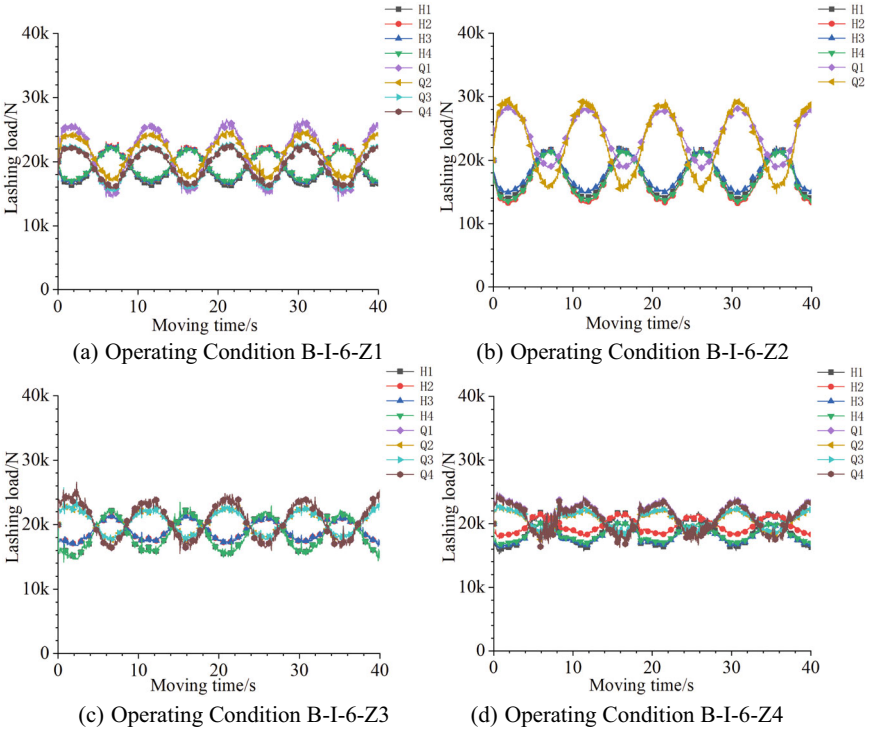


Fig. 7 Stress time history curve of tethered rigging in the coupled rolling and heave scenario under six level sea conditions



(a) Operating Condition B-I-6-Z1 (b) Operating Condition B-I-6-Z2

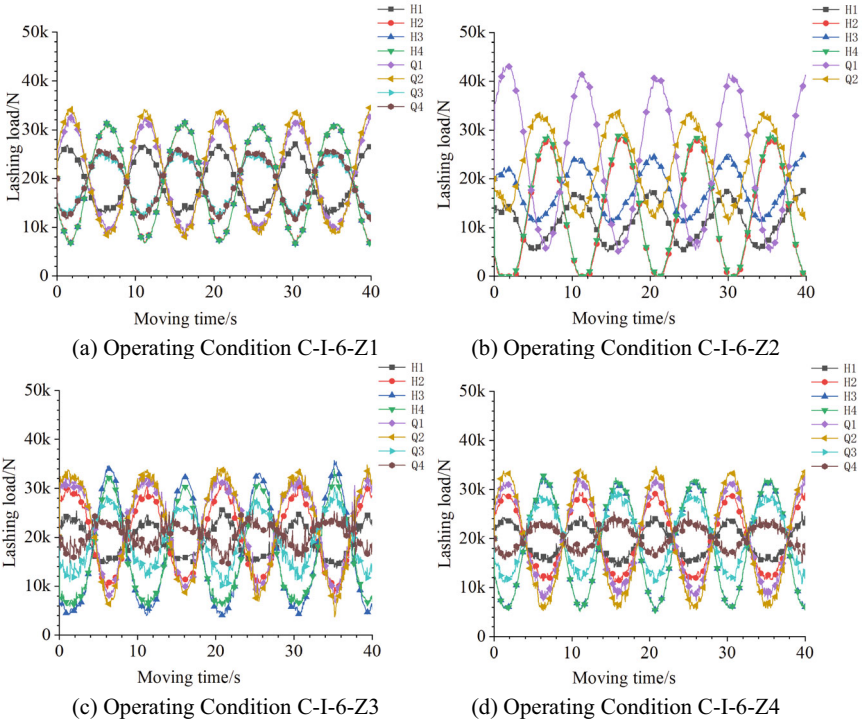
(c) Operating Condition B-I-6-Z3 (d) Operating Condition B-I-6-Z4

Fig. 8 Stress history curve of tethered rigging under coupling scenario of pitch and heave in six-level sea conditions

(3) *Rolling, Pitching, and Heaving Coupling.*

Under the combined pitching and heaving conditions with a sea state of level six, Fig. 9a shows that due to the influence of the vehicle's center of gravity and securing method, the maximum securing force of the securing equipment appears on Q2, with an amplitude reaching 34.8 kN. Figure 9b indicates that the maximum securing force of the securing equipment appears on Q1, with an amplitude reaching 43.5 kN. Figure 9c reveals that the maximum securing force of the securing equipment appears on H3, with an amplitude reaching 35.7 kN. Figure 9d shows that the maximum securing force of the securing equipment appears on Q2, with an amplitude reaching 34.8 kN.

Under the sea state of level six, it can be seen from the figures that when rolling, pitching, and heaving in three degrees of freedom are coupled, the rolling motion has a dominant influence, causing the force–time curve of the securing equipment to be characterized by the motion response features under the action of rolling alone. Pitching and heaving only play an auxiliary role in the entire motion process and cannot change the form of the securing equipment's load curve.



(a) Operating Condition C-I-6-Z1 (b) Operating Condition C-I-6-Z2

(c) Operating Condition C-I-6-Z3 (d) Operating Condition C-I-6-Z4

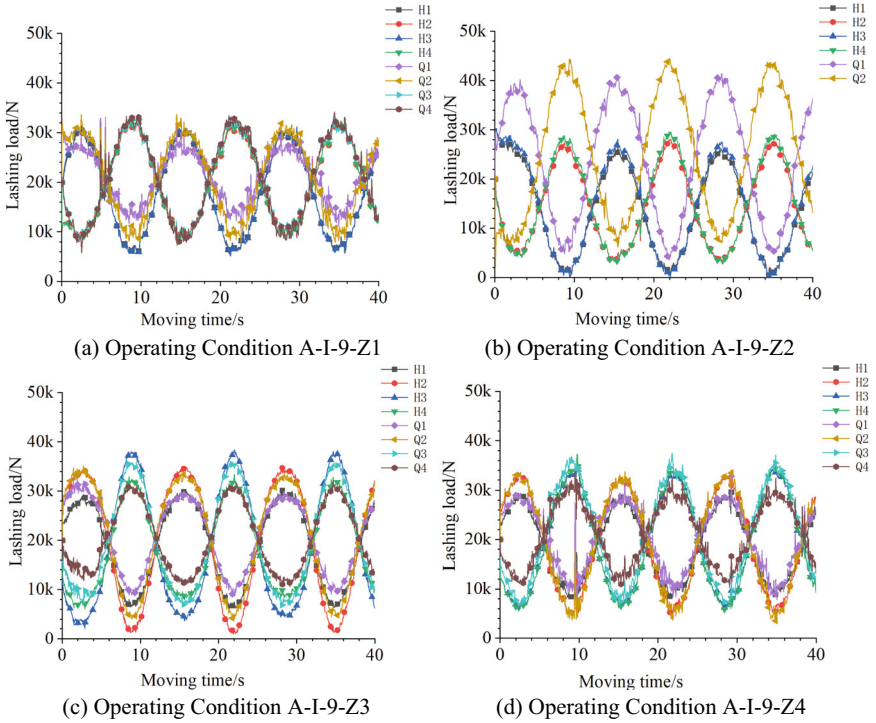
Fig. 9 Stress history curve of tether rigging under rolling, pitching and heave coupling scenarios in six level sea conditions

9-level Sea Conditions

(1) *Coupling of Roll and Heave.*

Under the conditions of a level 9 sea state, for equipment I, the influence of different sea conditions on the securing force of the equipment's securing equipment is studied under the same spatial arrangement. Analysis shows that the securing force of the securing equipment increases with the increase of the maximum rolling angle, which is easy to understand; the securing force of the securing equipment decreases with the increase of the rolling period, mainly because the increase of the rolling period indicates a slower swing, reducing the inertial force effect on the securing equipment; the securing force of the securing equipment increases with the increase of the heave amplitude and decreases with the increase of the heave period.

Under the combined rolling and heaving conditions with a sea state of level 9, Fig. 10a shows that due to the influence of the vehicle's center of gravity and securing method, the maximum securing force of the securing equipment appears on Q4, with an amplitude reaching 34.6 kN. Figure 10b indicates that the maximum securing force of the securing equipment appears on Q2, with an amplitude reaching 44.7 kN. Figure 10c reveals that the maximum securing force of the securing equipment



(a) Operating Condition A-I-9-Z1

(b) Operating Condition A-I-9-Z2

(c) Operating Condition A-I-9-Z3

(d) Operating Condition A-I-9-Z4

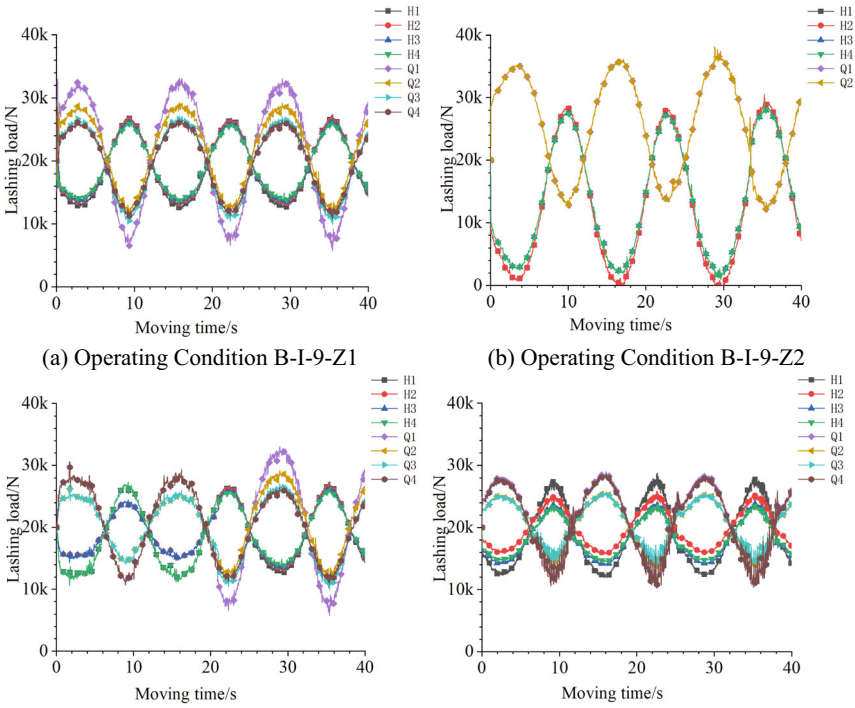
Fig. 10 Stress history curve of tethered rigging in the coupled rolling and heave scenario under class 9 sea conditions

appears on H3, with an amplitude reaching 38.5 kN. Figure 10d shows that the maximum securing force of the securing equipment appears on Q3, with an amplitude reaching 37.4 kN.

(2) *Coupling of Pitch and Heave.*

Under the conditions of a level 9 sea state, as shown in Fig. 11, due to the presence of pretension and a relatively small pitching amplitude of 4.4°, the force on the securing equipment is relatively smaller compared to the rolling. If the securing equipment at the front and back of the vehicle is roughly symmetrical, the force curve of the securing equipment will fluctuate around the force at the equilibrium position, which is the pretension force of 20 kN. Compared to the level 6 sea state, the motion in the level 9 sea state is more turbulent, and the force amplitude of the securing equipment under each securing scheme in the level 9 sea state is also larger than that under the same scheme in the level 6 sea state.

The amplitudes of the securing forces of the securing equipment all appear on the front securing equipment. Among them, the maximum value of the Z1 securing scheme appears on Q1, approximately around 33.9 kN, the maximum value of the Z2 securing scheme appears on Q2, approximately around 38.1 kN, the maximum



(a) Operating Condition B-I-9-Z1

(b) Operating Condition B-I-9-Z2

(c) Operating Condition B-I-9-Z3

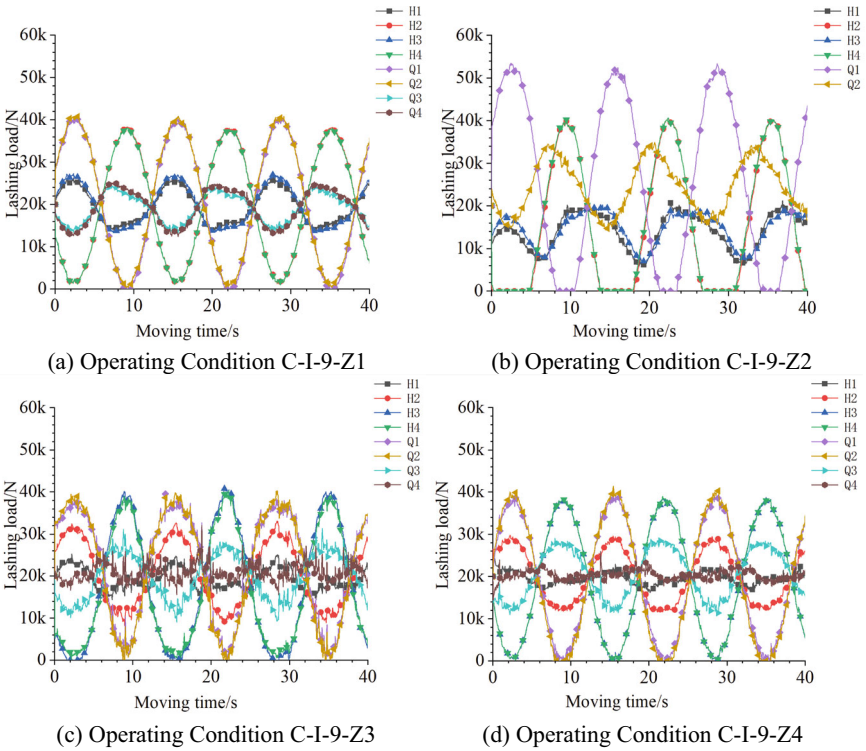
(d) Operating Condition B-I-9-Z4

Fig. 11 Stress history curve of tethered rigging under coupling scenario of pitch and heave in class 9 sea conditions

value of the Z3 securing scheme appears on Q1, approximately around 33.9 kN, and the maximum value of the Z4 securing scheme appears on Q1, approximately around 29.2 kN.

(3) *Coupling of Roll, Pitch, and Heave.*

Under the conditions of a level 9 sea state, as shown in Fig. 12, the maximum loads of the securing equipment for scenarios Z1, Z2, Z3, and Z4 are approximately 40.7 kN, 53.5 kN, 40.7 kN, and 41.4 kN, respectively, occurring on the Q2, Q1, H3, and Q2 securing equipment. When the motions of rolling, pitching, and heaving in the three degrees of freedom are coupled, the rolling motion has a dominant influence, causing the force–time curve of the securing equipment to be characterized by the motion response features under the action of rolling alone. Pitching and heaving only play an auxiliary role in the entire motion process and cannot change the form of the securing equipment's load curve. Under the coupled motion conditions of the three degrees of freedom, the load-time curve of the securing equipment exhibits a periodic change trend consistent with that under the action of a single rolling motion, which conforms to the actual force characteristics of the securing equipment and the overall motion of the ship. Therefore, the forces on the securing equipment on the same side



(a) Operating Condition C-I-9-Z1

(b) Operating Condition C-I-9-Z2

(c) Operating Condition C-I-9-Z3

(d) Operating Condition C-I-9-Z4

Fig. 12 Stress history curve of tethered rigging under rolling, pitching and heave coupling scenarios in class 9 sea conditions

but different front and back sides of the vehicle are more consistent compared to other securing equipment.

6 Conclusion

This paper simplifies the securing model structure based on the ship deck vehicle securing scenario, analyzes the securing situation of the securing equipment under different navigation conditions, securing forms, and the spatial position of the equipment securing, and establishes a numerical simulation model using the multibody dynamics software ADAMS. By comparing and analyzing the results with the model test, the effectiveness of the numerical model is verified. The numerical simulation analysis technology is used to quantitatively analyze the securing performance of the securing equipment under the influence of different securing parameters. The main conclusions are as follows:

(1) The motion scenarios of vehicles secured on ship decks during navigation are analyzed, and the securing forces of the securing equipment are arranged from largest to smallest as follows: considering the coupled effects of rolling, pitching, and heaving > rolling and heaving coupled > pitching and heaving coupled. At the same time, the peak values of the securing forces of the securing equipment in all scenarios under level 9 sea conditions are all greater than those under level 6 sea conditions. Therefore, to check the safety of the securing equipment, the safety of the securing scheme under the scenario considering the coupled effects of rolling, pitching, and heaving in level 9 sea conditions can be mainly focused on.
(2) The established multibody dynamics modeling technology has high computational accuracy and can accurately reflect the changes in the securing forces of the securing equipment under different securing schemes. Under certain circumstances, as the number of securing equipment increases, the tensile force distributed to each piece of securing equipment is continuously decreasing, which can effectively reduce the maximum tensile force of the securing equipment.
(3) By analyzing the typical vehicle equipment securing design scheme, the calculation layout scenario is determined. It can be found through the numerical calculation of the multibody dynamics model that under level 6 and level 9 wave sea conditions, the securing loads of the vehicle securing scheme meet the usage requirements.

References

1. Shi L (2006) On Cargo lashing to ensure safety of navigation. In: Ship maneuvering in rough seas (Book I), vol 5
2. Turnbull SR, Dawson D (1997) The securing of rigid semi-trailers on roll-on/roll-off ships. Int J Mech Sci 39:1–14
3. Jia J, Ulfvarson A (2006) Dynamic analysis of vehicle—deck interactions. Ocean Eng 33:1765–1795
4. Jia J (2007) Investigations of vehicle securing without lashings for ro-ro ships. J Mar Sci Technol 12:43–57
5. Surendran S, Lee SK, Reddy JVR, Lee G (2005) Non-linear roll dynamics of a ro-ro ship in waves. Ocean Eng 32:1818–1828
6. Xiaobing Z, Cheng Q (2018) Simulation analysis of shipborne helicopter fastening system based on ADAMS. Mechatron Dev Innov 31(01):48–50
7. Hua Z (2020) Performance analysis and optimization of air suspension vehicle system based on multi-body dynamics and collaborative optimization theory. Jiangsu University of Science and Technology

Open Access This chapter is licensed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License (http://creativecommons.org/licenses/by-nc-nd/4.0/), which permits any noncommercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if you modified the licensed material. You do not have permission under this license to share adapted material derived from this chapter or parts of it.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.



Mechanical Structure Design of a Continuously Rotating Flight Test Device



Shaozun Hong, Deng He, Daiyong Mao, and Xiaodong Zheng

Abstract In order to expand wind tunnel test capabilities and improve wind tunnel test efficiency, a test device with continuous rotation during the test process is designed through analysis of design requirements. The design and simulation results indicate that the new device has stronger torque resistance and remote adjustment ability, and can be well applied in future experiments.

Keyword Test device · Wind tunnel · Mechanical design

1 Introduction

Wind tunnel test is widely used in aerodynamics research, aircraft development, and other fields related to airflow [1]. Especially in the development of aircraft, wind tunnel test has become an important part of the aircraft development process due to its advantages of high accuracy, convenient operation, high efficiency and low test costs. For different test models, test conditions, sensor installation methods, etc., it is necessary to design appropriate test devices in advance to fix or support the models before conducting wind tunnel tests. Currently, with the continuous development of large transport aircraft, civil aircraft, drones, and other new concept aircraft, there is a more diverse demand for wind tunnel test technology [2–6]. Taking a high aspect ratio aircraft as an example. On the one hand, adopting a high aspect ratio during flight is beneficial for increasing lift and improving flight efficiency. On the other hand, high aspect ratio aircraft require higher torque resistance from wind tunnel test devices, which makes existing test devices unable to meet their requirements. In addition, some specific aircraft require research on their external flow field characteristics under continuous rotation. However, at the beginning of the design of the existing

S. Hong · D. He · D. Mao (✉) · X. Zheng
High Speed Institute, China Aerodynamics Research and Development Center, CARDC, Mianyang, China
e-mail: 275079137@qq.com

© The Author(s) 2026

H. Saman et al. (eds.), *The 9th International Conference on Advances in Construction Machinery and Vehicle Engineering*, Lecture Notes in Mechanical Engineering,
https://doi.org/10.1007/978-981-96-5527-4_63

support device for the 2.4 m transonic wind tunnel, these test conditions were not considered, and it can no longer meet the updated requirements.

The support device of the existing 2.4 m transonic wind tunnel is shown in Fig. 1. According to the coordinate system shown in Fig. 1, it has degrees of freedom around the X, Y, and Z coordinate axes, which are used to adjust the α, β and γ angles of the test model. Among them, the angle adjustment of α and β is carried out through the hydraulic drive system of the wind tunnel, which can continuously change during the test. While the angle adjustment of γ is manually adjusted using a worm gear and assisted by an expansion device for locking. It is necessary to loosen the expansion device before each test, adjust the angle to the desired position, and lock the expansion device before conducting the test. In the current use process, it has been found that its torque resistance is no longer sufficient to meet the requirements of certain high aspect ratio model tests. In addition, due to the manual adjustment of γ, on the one hand, it cannot meet the requirement of continuous rotation during the test, and on the other hand, it greatly limits the efficiency of the test.

Therefore, it is necessary to redesign the test support device of the wind tunnel to meet the current test requirements and effectively improve test efficiency.

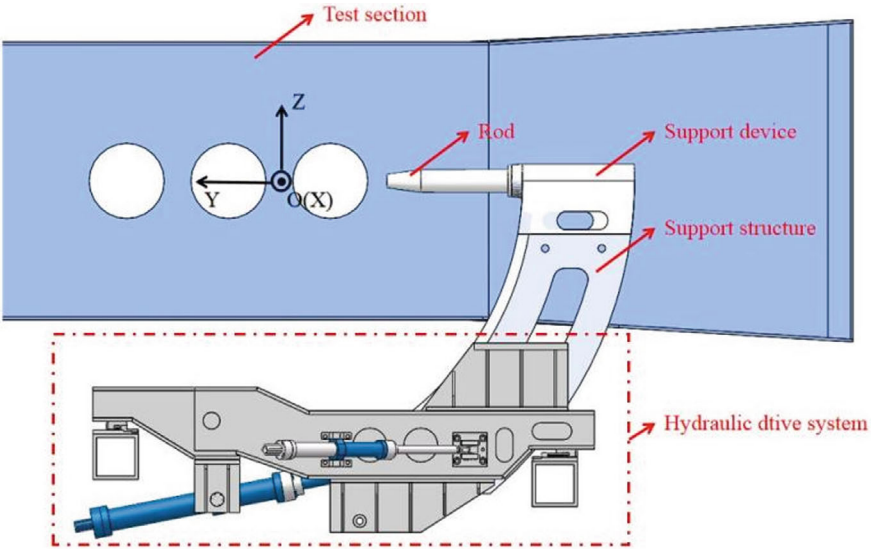


Fig. 1 Schematic diagram of the existing support device

2 Mechanical Structure Design

2.1 Design Requirement Analysis

The design requirements are analysed based on the application background of the supporting device from two aspects: functional indicators and performance indicators.

Firstly, in terms of functional indicators. Through long-term test experience accumulation and analysis of future application requirements, the functional indicators and purposes that the device needs to achieve are shown in Table 1.

Secondly, in terms of performance indicators. According to the coordinate system established in Fig. 1, the performance indicators of the original support device are shown in Table 2.

In order to meet the new experimental requirements, the new support device needs to have a M_y of no less than 2700 N m, an accuracy of no less than 0.02° and a rotation speed of no less than 10°/s during the continuous rolling change test. Meanwhile, it should be ensured that other indicators are equivalent to the original support device. When conducting a fixed angle test, the new support device must have a M_y of no less than 12,000 N m through remote locking.

Table 1 Functional indicators

No.	Function	Purpose
1	Can achieve continuous changes in γ	Meet the needs of future test applications
2	Can achieve remote locking	Meet experimental requirements and improve test efficiency
3	The shaft should be a hollow structure	Meet sensor wiring requirements
4	The interface remains consistent with the original equipment	Meet the demand for interchangeable use
5	The external dimensions should not be too large	Meet the requirements of wind tunnel test section size and blockage degree

Table 2 Performance indicators of the original support device

Name (Unit)	X (N)	Y (N)	Z (N)
Indicator	10,000	25,000	10,000
Name (Unit)	M_x (N m)	M_y (N m)	M_z (N m)
Indicator	2000	2500	2000

2.2 Detail Design

The mechanical structure of the new support device is designed based on the functional indicators and performance indicators analyzed above.

The first is the transmission part. In order to achieve continuous rolling of the shaft, the original support device used a manual worm gear structure. The worm gear structure has the advantages of compact structure and high reduction ratio, which can provide huge torque output under limited structural size. However, on the one hand, due to the use of manual control, it cannot meet the needs of remote continuous rolling changes in the test. On the other hand, due to the inherent clearance and low transmission efficiency of the worm gear, it cannot guarantee the accuracy and rotational speed required for the test. In order to compensate for these shortcomings and considering various factors, it has been decided to design the transmission of the support device as "motor + bevel gear + reducer", as shown in Fig. 2.

The purpose of selecting bevel gears is, on the one hand, to change the direction of motor torque transmission, making the overall structure more compact, and on the other hand, to further increase the output torque by setting the reduction ratio of gear input and output to 2:1. Meanwhile, the bevel gear connected to the transmission shaft will be designed as a hollow structure to ensure the passage of internal sensor cables. In addition, adopting this method, which means the motor is not directly connected to the reducer, can make it more convenient for the motor to be repaired or replaced. Harmonic reducers have the advantages of compact structure, high transmission accuracy and high transmission efficiency. Therefore, a C55 harmonic reducer with a rated output torque of 2705 N m and a speed of 6.7 rpm is selected as the reducer, which can meet the requirements for continuous rotation test [7].

The next is the locking part. The original support device adopts a mechanical expansion sleeve in terms of locking. By installing a mechanical expansion sleeve

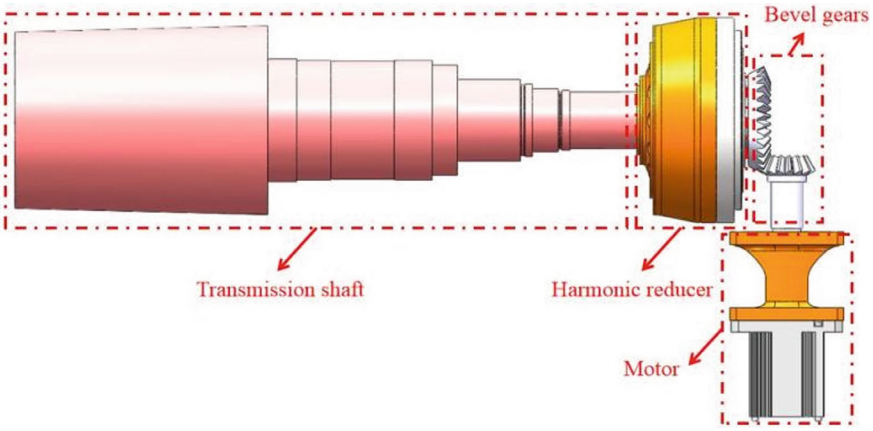


Fig. 2 Schematic diagram of the transmission part



Fig. 3 The ETP-EXPRESS-90 hydraulic expansion sleeve

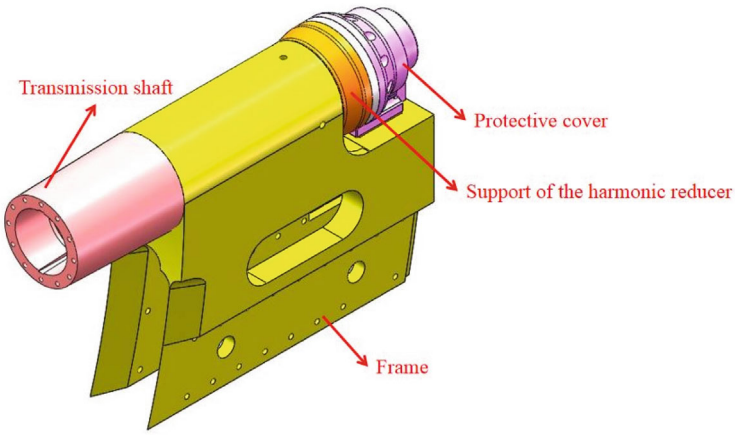
at the rear end of the transmission shaft, when locking is required, the screw on the expansion sleeve is turned. After the expansion sleeve expands, it is locked using the friction force between it and the transmission shaft. When adjusting the angle each time, it is necessary to manually loosen and lock the expansion sleeve, which mainly brings three disadvantages: first, remote control cannot be achieved; second, low efficiency; and third, the torque value of each locking cannot be accurately controlled and judged.

In order to compensate for the above shortcomings, considering multiple factors, it has been decided to use the ETP-EXPRESS-90 hydraulic expansion sleeve shown in Fig. 3 and make improvements. The principle of this hydraulic expansion sleeve is that the part used for locking is a hollow structure. Hydraulic oil is stored in the structure using screws, and the volume of the internal hollow structure is changed by tightening the screws, compressing the hydraulic oil and generating huge pressure. This pressure will cause the part used for locking to expand and deform, resulting in close contact with the locked structure and accompanied by huge pressure, achieving the purpose of locking.

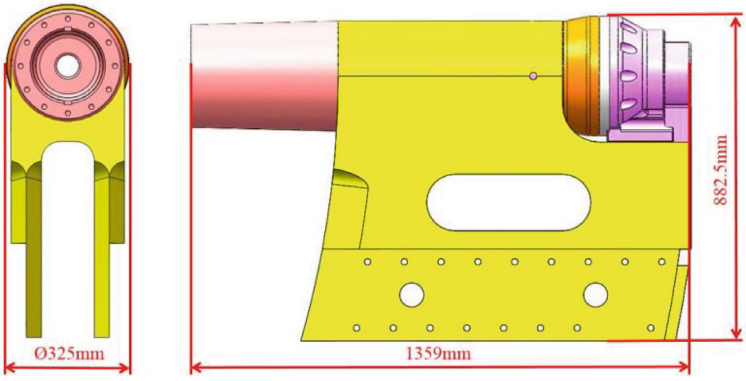
However, the ETP-EXPRESS-90 hydraulic expansion sleeve cannot be operated remotely as well. In order to achieve the purpose of remote operation, based on its working principle, improvements have been made. Specifically, the original screws have been removed and replaced with high-pressure oil pipe interfaces. The hydraulic transmission system in the wind tunnel is connected through a high-pressure oil circuit, and the pressure of the hydraulic oil is used to replace the pressure applied by the screws. Remote locking and loosening are achieved by controlling the hydraulic oil pressure.

Finally, there are other aspects of the design. The structure or device inside the transmission shaft, like its axis, adopts a hollow design to ensure the passage of internal sensor cables. The combination of sliding copper sleeve and rolling bearing is used to ensure that the transmission shaft can still rotate smoothly under high load conditions. In addition, an encoder is installed on the transmission shaft to achieve precise measurement of the actual output angle and form a closed-loop control with the motor control system, thereby achieving a rotational accuracy of no less than 0.02°. At the same time, strictly ensure that all types of structural dimensions

are consistent with the original installation. In terms of appearance, it meets the requirements of wind tunnel test section size and blockage degree. The final design is shown in Figs. 4 and 5.



(a)



(b)

Fig. 4 Schematic diagram of the continuously rotating flight test device **a** overall outline diagram, **b** outline dimension diagram

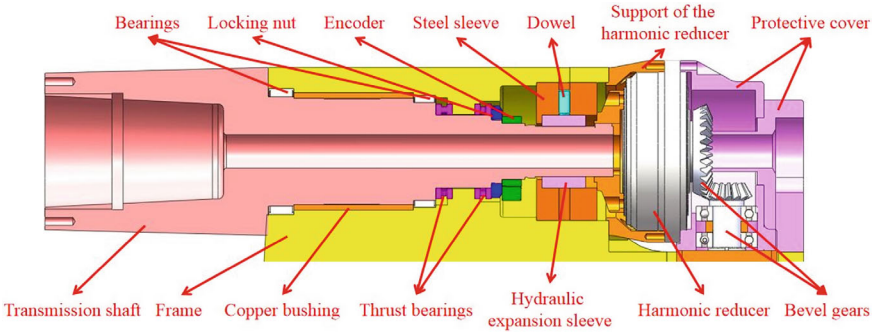


Fig. 5 Schematic diagram of the internal structure of the continuously rotating flight test device

3 Key Structure Simulation and Analysis

Analysis of the stress conditions of various structural components of the continuously rotating flight test device designed above during actual use reveals that the transmission shaft, as the main component supporting the model, transmitting torque and resisting impact, needs to be simulated and analyzed in advance to ensure the rationality of the design and the reliability during subsequent use.

3.1 Selection of Material

Considering the comprehensive mechanical properties of the material, the material for the transmission shaft is selected as 30CrMnSiA in the quenched and high-temperature tempered state. 30CrMnSiA is the most widely used quenched and tempered steel in the aircraft manufacturing industry with good machinability. After heat treatment, it has good impact toughness and fatigue strength, and can operate stably for a long time in harsh environments such as high temperature, high pressure and high speed. This steel is mainly used for manufacturing important aviation forgings, machined parts, sheet metal parts and welded parts [8–10]. Its specific performance indicators are shown in Table 3.

In the subsequent simulation process, the lower limits of the performance parameters in the table will be set, such as a yield strength of 835 MPa and a tensile strength of 1080 MPa.

Table 3 Performance parameters of 30CrMnSiA

Chemical composition			
Name (Unit)	C	Si	Mn
Indicator	0.27–0.34	0.90–1.20	0.80–1.10
Name (Unit)	P	S	Cr
Indicator	≤ 0.035	≤ 0.035	0.80–1.10
Mechanical property			
Name (Unit)	Yield strength σ_s (MPa)	Tensile strength σ_b (MPa)	Percentage elongation δ_5 (%)
Indicator	≥ 835	≥ 1080	≤ 10
Name (Unit)	Reduction of area (%)	Impact toughness KV2 (J)	Hardness (HB)
Indicator	≥ 45	≥ 39	≤ 229

3.2 Settings of Simulation Conditions

According to the operating conditions faced by the test device during use, when conducting continuous rolling change tests, M_y should not be less than 2700 N m, and when conducting fixed angle tests, M_y should not be less than 12,000 N m. By analyzing the two operating conditions, it is evident that the conditions required for conducting fixed angle tests are more severe. Therefore, it is considered to simulate the transmission shaft under more severe conditions, where M_y is set to 12000 N m and the forces and torques in other directions are set to the indicator values in Table 2. Removing irrelevant components, the support method of the transmission shaft is shown in Fig. 6.

The transmission shaft mainly provides radial support through two rolling bearings and a sliding bearing, ensuring that the axis of the transmission shaft remains constant. By using two thrust bearings and a locking nut to limit its axial freedom, it ensures that the transmission shaft only has rotational freedom around the y-axis

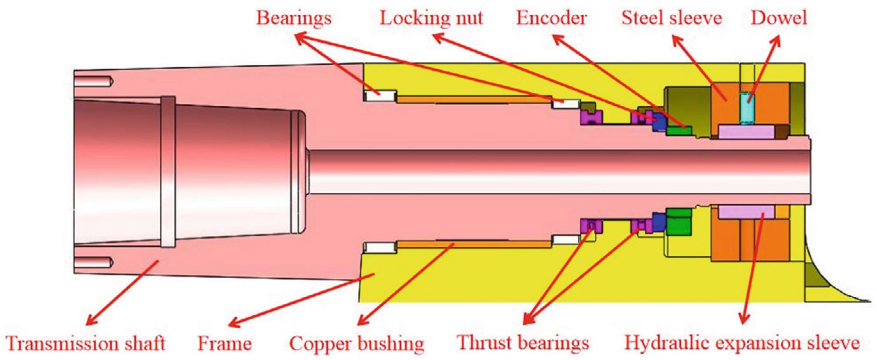


Fig. 6 The support method of transmission shaft during fixed angle test

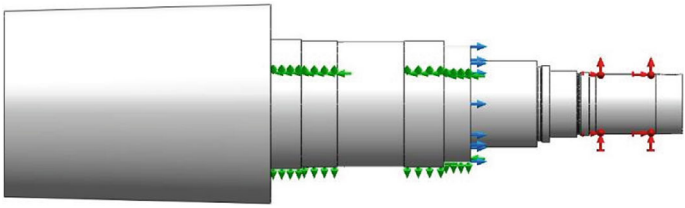


Fig. 7 Simulation condition setting diagram of the transmission shaft

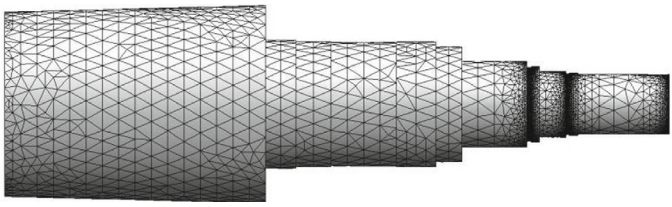


Fig. 8 Grid division diagram of the transmission shaft

direction. When conducting a fixed angle test, the hydraulic expansion sleeve at the end of the transmission shaft restricts the rotational freedom around the y-axis direction, thereby causing the transmission shaft to lock. During the test, the model is connected to the transmission shaft through a support rod, which fits with the transmission shaft through a conical surface and is loosened using screws. Therefore, it can be assumed that the torque generated by the model will ultimately be transmitted to the inner cone surface at the front end of the drive shaft. According to the support method of the transmission shaft mentioned above and the force and torque received by the transmission shaft in various directions, the simulation condition is set in SOLIDWORKS 2016 [11, 12], as shown in Fig. 7.

Through the grid division module in the software, the transmission shaft was meshed and optimized locally. Finally, the transmission shaft was divided into 126,880 grid units, as shown in Fig. 8.

3.3 Results and Analysis of Simulation

After simulation, the stress distribution and displacement of the transmission shaft under the above conditions are shown in Figs. 9 and 10.

From the simulation results, it can be seen that the maximum stress distribution of the transmission shaft is at the front end of the locking thread, which is 297.2 Mpa. And the cumulative maximum deformation is at the front end of the transmission shaft, which is 0.436 mm. Compared with the yield limit of 30CrMnSiA, the simulation results have a safety factor of nearly three times, and the deformation meets

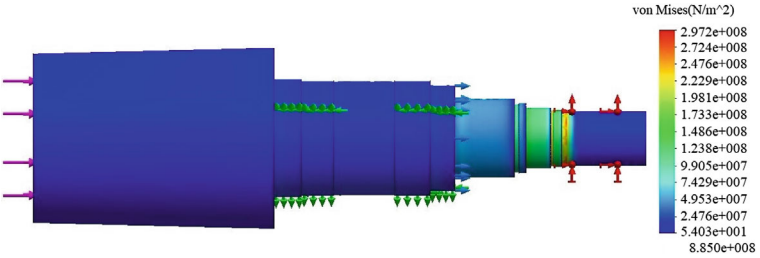


Fig. 9 The stress distribution figure

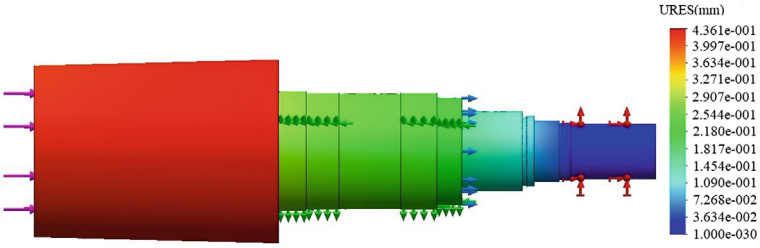


Fig. 10 The displacement figure

the requirements of coaxiality, which can meet the usage requirements. In the actual manufacturing process, a fillet will be added at the front end of the locking thread for transition, further improving the situation of stress concentration.

4 Conclusions

This article provides a detailed analysis of the functional and performance indicators of the support device from its application background. Through reasonable design, simulation, and analysis, a continuously rotating flight test device that meets the requirements of use is finally designed. The device achieves the indicators of anti torsion torque not less than 12000 N m during fixed angle test, anti torsion torque not less than 2700 N m during continuous rolling change test, rotational accuracy not less than 0.02°, and rotational speed not less than 10°/s. Compared with the original device, its M_y has increased by more than 4 times during the fixed angle test, fully meeting the expected design effect. In addition, the designed device also has good passability for sensor wiring which is lacking in the original device. At the same time, the interface and external dimensions are basically consistent with the original equipment, which is good in interchangeability. With the help of the corresponding control system, the device can be remotely controlled, greatly improving the efficiency of the experiment. In the future, further optimization of design will be carried

out in processing and manufacturing, and the device will be put into use as soon as possible.

References

1. Wu JQ, Cui XC, Wei Z, Du WT (2023) Research status and development tendency of large aircraft wind tunnel technologies in China. Acta Aerodynamica Sinica 41(12):83–94. https://doi.org/10.7638/kqdlxxb-2023.0147
2. Zhan PG, Zhao X (2011) New Development of model techniques for wind tunnel test. Aeronaut Sci Technol 5:8–11
3. Wang HF, Zhan JX, Chen K, Chen X, Chen XJ (2022) Development and application of aerodynamic research technologies for fighters at high angle of attack. Acta Aerodynamica Sinica 40(1):1–25. https://doi.org/10.7638/kqdlxxb-2021.0306
4. Zhou J, Zhang J, Chen Q, Wei W, Liu L, Qian DD (2023) Wind tunnel test technique of continuous varying Mach number for air-breathing vehicle. J Exp Fluid Mech 37(6):76–85
5. Gao YW, Wei BB, Liang D (2021) Research status of airfoil wind tunnel experimental technology. Acta Aerodynamica Sinica 39(6):85–100. https://doi.org/10.7638/kqdlxxb-2021.0381
6. Luo PY, Sun ZX, Sun HY, Chen XL, Li WJ (2021) Overview of the development of hypersonic wind tunnel test capability in the United States. Aerodyn Missile J 6:33–31. https://doi.org/10.16338/j.issn.1009-1319.20210826
7. Sheng YD (2021) Simulation and experimental research on mechanical performance of hollow harmonic flexible wheel. Kunming University of Science and Technology
8. Pang XT, Gong QF, Wang ZJ, Li ZG, Yao CW (2020) Microstructures and mechanical properties of 30CrMnSiA and 30CrMnSiNi2A high-strength steels after laser-cladding repair. Chinese J Lasers 47:11:1–11. https://doi.org/10.3788/CJL202047.1102002
9. Li ZH, Wang CJ, Luo XY (2019) Research on the final heat treatment press of 30CrMnSiA steel. New Technol New Process 40(4):1–2. https://doi.org/10.16635/j.cnki.1003-5311.2017.10.001
10. Chao BX, Wang BL, Jiang KQ (2019) Research on control and prevention technology for temper brittleness of 30CrMnSiA steel. Heat Treat Technol Equip 40(4):15–19. https://doi.org/10.19382/j.cnki.1673-4971.2019.04.004
11. Zhu C, Fu XL, Wang M, Zhang HX (2020) Analysis and design of gear transmission structure based on solid works. Jiangsu Agric Sci 48(16):241–249. https://doi.org/10.15889/j.issn.1002-1302.2020.16.047
12. Ning TZ, Liu YK (2019) Parametric modeling of shafts based on solid works 2016 and visual basic 6.0. J Yancheng Inst Technol 32(1):19–26. https://doi.org/10.16018/j.cnki.cn32-1650/n.201901004

Open Access This chapter is licensed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License (http://creativecommons.org/licenses/by-nc-nd/4.0/), which permits any noncommercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if you modified the licensed material. You do not have permission under this license to share adapted material derived from this chapter or parts of it.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.



Airworthiness Compliance Verification and Analysis of Turbofan Engine Bird Ingestion



Qing Zhang, Xigang Shen, Kun Niu, and Haiyang Zhang

Abstract The verification of airworthiness compliance for aero-engines ingesting birds is a comprehensive test of engine performance and safety. Due to inexperience, issues with automatic engine shutdown arose during these tests. To address this, cause analysis and improvement measures were implemented. By examining key test parameters, changes before and after bird ingestion, test recordings, engine inspections, and simulation analyses, the cause of the shutdown is identified. It is found that the high-pressure compressor surge caused by the first bird ingestion led to abnormal casing deformation near the booster stator, resulting in axial locking and excessive rotor friction. This friction caused the rotor to gradually slow down and eventually stop. Improvements include extending the casing's positioning surface by 4 mm in the turbocharging stage to prevent abnormal rubbing. Additionally, the bird-ingestion requirement from the 1988 CCAR 33.77 clause is deemed the applicable standard for this engine type.

Keywords Bird ingestion · Airworthiness compliance · Test and verification · Engine

1 Introduction

With the rapid development of aviation, the harm of birds to flight is becoming more and more prominent [1, 2]. For a high-by-pass ratio engine with the increase of the inlet area, the probability of bird swallowing event increases, and so does the harm of bird swallowing. After being hit by a bird, it may cause structural damage at the inlet of the engine, such as the fan blade and the intake fairing [3], and the bird entering the core will cause the compressor surge and cause the engine thrust loss [4].

All the major aviation countries in the world have formulated corresponding airworthiness regulations and requirements for the problem of engine swallowing

Q. Zhang (✉) · X. Shen · K. Niu · H. Zhang
AECC Shenyang Engine Research Institute, Shengyang, China
e-mail: zq252098052@163.com

© The Author(s) 2026

H. Saman et al. (eds.), *The 9th International Conference on Advances in Construction Machinery and Vehicle Engineering*, Lecture Notes in Mechanical Engineering,
https://doi.org/10.1007/978-981-96-5527-4_64

birds [5], which stipulates the basic safety and working ability of the engine after being hit by a bird. The three major foreign engine companies had the ability to test the airworthiness compliance of swallowing birds in the 1980s, and carried out related research on the bird swallowing test technology of the whole engine. In China, Luo et al. [6] carried out the analysis method of key parameters of aero-engine airworthiness verification, and began the research on the parameters of airworthiness verification of swallowing birds, which provided a basis for the analysis of aircraft test parameters. However, the above research still stays in the verification method and laboratory verification and has not really been used to guide the actual verification of an engine.

It is the first time in China to carry out the verification of airworthiness compliance of turbofan engine swallowing birds in accordance with the CCAR33.76 bird-swallowing terms, but it is inexperienced. Aiming at the problem of the automatic stop of the engine after swallowing birds in the test, this paper analyzes the direct cause and mechanism, and puts forward the corresponding improvement measures to provide support for the design of airworthiness compliance and technical ability of verification.

2 General Situation of Test

2.1 Test Parameters

According to the CCAR33.76 bird swallowing clause, according to the simulation analysis of the engine inlet roar area, the key parameters of the swallowing bird flock test are established as shown in Table 1, with the domestic chickens used as the experimental birds. In Table 1, one chicken weighed 1.16 kg and five chickens weighed 0.7, 0.71, 0.7, 0.7, and 0.73 kg are used as the experimental birds. These birds are adopted to evaluate the impact of different bird weights on engine suction. Among them, one chicken weighing 1.16 kg is launched to the position of the fan blade corresponding to the flow channel of the core machine, and the other five chickens weighing 0.7–0.73 kg are thrown to the shoulder of the fan blade and the tip of the fan blade, in which three chickens weighing 0.7 kg are evenly distributed according to the inlet area, as shown in Fig. 1.

2.2 Test Performance

According to the test procedure, the physical speed of the low pressure rotor reaches the test state after the engine runs for 10 min. The 1 min 10 s runs stably in the test state, triggering the bird gun launch button. 6 chickens are fired at the specified position within 1 s, and the engine tail vents are seen from the video surveillance.

Table 1 Test parameters of swallowing birds by an engine

	Test requirement	Test parameter
Bird weight (kg)	1.15 for one bird and 0.7 for the other five birds	1.16 for one bird and 0.7,0.71, 0.7, 0.7, 0.73 for the other five birds
Impact velocity of bird (m/s)	100	98–105
Engine speed (r/min)	Maximum speed	Maximum speed + 35
Impact position	Core	1.16

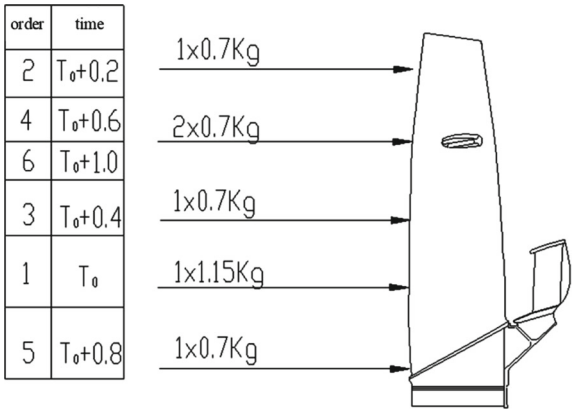


Fig. 1 Impact position parameters

The low-pressure speed drops sharply and drops to 1141 r/min 4 s after the button is triggered, and there is no sign of recovery. It is pulled to the slow state, the low-pressure speed continues to decline, and the low pressure rotor stops rotating after 21 s.

2.3 Structural Damage

When the engine is decomposed after the test, it is found that the inlet edge of 6 fan rotor blades is deformed, the shoulder is lapped, 12 fan stator blades fall off, the rotor blades of all levels of the high pressure compressor are seriously rubbed with the casing, and the tip of the blade is worn in a circular arc. The inner ring of the stator of the last stage of the supercharging stage is deformed, the large deformation area is detached from the positioning surface of the intermediate casing, and the metal at the front edge of the inner ring is axially rubbed against the top of the drum of the booster stage. There are 32 final stator blades protruding from the inner ring. The damage is shown in Fig. 2, and the rubbing is shown in Fig. 3.

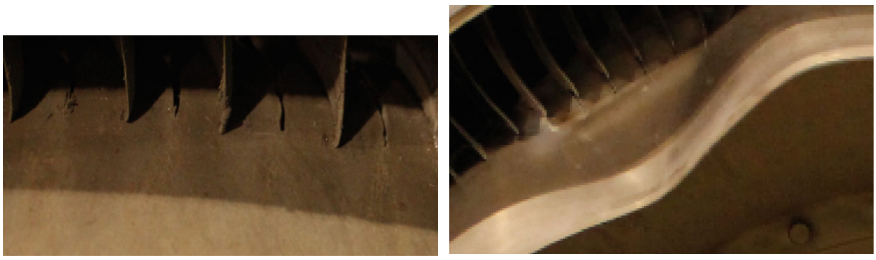


Fig. 2 Last stage stator damage of supercharging stage

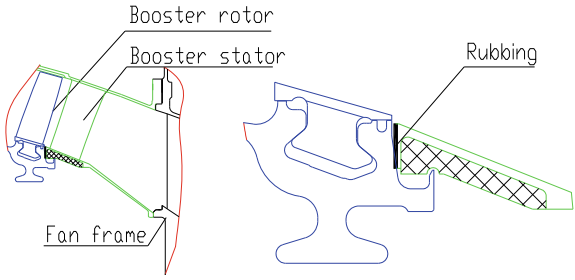


Fig. 3 Indication of rubbing position

Under the condition of the whole machine before decomposition, the low-pressure rotor cannot rotate. In the process of decomposition, when the booster stage is disconnected from the low-pressure rotor, the rotation capacity of the low-pressure rotor is restored. Therefore, it is preliminarily judged that the abnormal axial rubbing between the leading edge of the inner ring of the stator and the trailing edge of the rotor at the last stage of the booster is the direct cause of the failure of the test.

3 General Situation of Test

Through the combination of test parameters and deformation simulation, this paper comprehensively analyzes the causes of the abnormal axial friction between the stator inner ring and the rotor at the last stage of the booster stage, which is the direct reason for the failure of the bird swallowing test.

3.1 Analysis of Test Parameters

Through the test process, the image recording parameters and measuring parameters are analyzed to analyze the change in engine state after swallowing birds. From the measured changes of the fan rear pulsating pressure measuring point P12m, the high

pressure compressor front fluctuating pressure measuring point P25m and the high pressure compressor rear fluctuating pressure measuring point P31m, it can be seen that all the parameters of the engine are normal before the test. The measuring point P12m is used to monitor changes in the gas pressure of the fan rear, the measuring point P25m is employed to monitor the instantaneous change of the gas pressure at the inlet of the high pressure compressor, and the measuring point P31m is utilized to monitor momentary changes in high pressure compressor outlet gas pressure. When the first 1.15 kg bird hits the fan blade and enters the main runner, the inner flow path is affected by extrusion shock and blockage, and the internal pressure of the high-pressure compressor increases sharply, which exceeds the surge boundary after 11 ms. The pressure increases successively in the three places, and the high-pressure ratio reaches the maximum in 11 ms, and then decreases rapidly. The peak value of P25m reaches 0.61 MPA and the peak value of P31m reaches 6.57 MPa, as shown in Fig. 4. At the same time, the image recording results of the process in the engine inlet region show that Mars appears in the outer culvert region now, as shown in Fig. 5. The results show that the 11 ms high pressure compressor surges after the first bird impact.

After the high pressure compressor surge, the low pressure rotor speed, the total inlet pressure P25 and the static pressure PS3 at the exit of the high pressure compressor continue to decrease, while the inlet temperature of the low pressure turbine two-stage deflector T495 increases. The measuring point P25 is used to monitor high pressure compressor inlet gas pressure, and the measuring point PS3 is employed to monitor high pressure compressor outlet gas pressure. The control system is controlled according to the control law but limited by other parameters. The fuel supply continues to decrease, and the engine stops after 4 s of bird launch. The specific process is shown in Fig. 6. The results show that after the high pressure compressor surge, the engine control law is normal, and tries to exit the surge, but

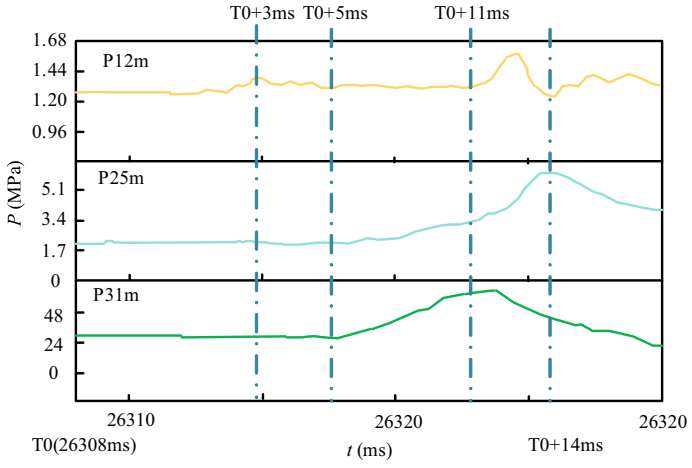


Fig. 4 Variation of pulsation parameters



Fig. 5 The fire light of the outer culvert

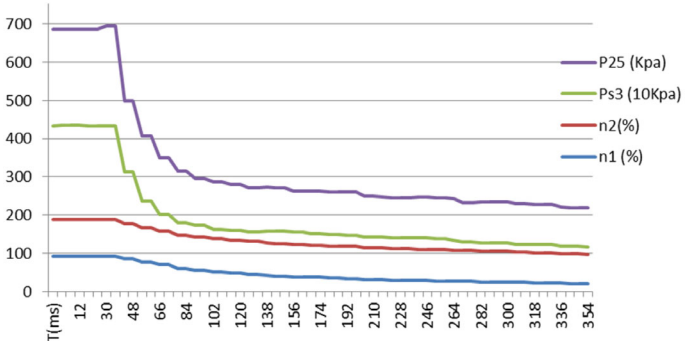


Fig. 6 Variation curves of experimental parameters

due to the limitation that the low pressure rotor speed cannot be recovered, it cannot recover.

Therefore, the surge of the high-pressure compressor after swallowing the first chicken, which leads to the normal rubbing of the low-pressure rotor, is the direct cause of automatic parking.

3.2 Analysis of Rubbing Process

Under the normal working conditions of the engine, the rear end face of the booster stage rotor and the front end face of the last stage stator, that is, the rotor in the rubbing position, are in a clearance state, the design value is 3.25 mm, and no abnormal rubbing has occurred.

In the bird swallowing test, when the fan blade is impacted by the 1.15 kg bird body, the fan rotor is additionally subjected to the huge axial impact load caused by the bird impact, and the stator of the last stage of the booster stage is subjected to the huge surge load caused by the surge of the high pressure compressor. Therefore, taking the actual bird speed, fan speed and impact position parameters as inputs,

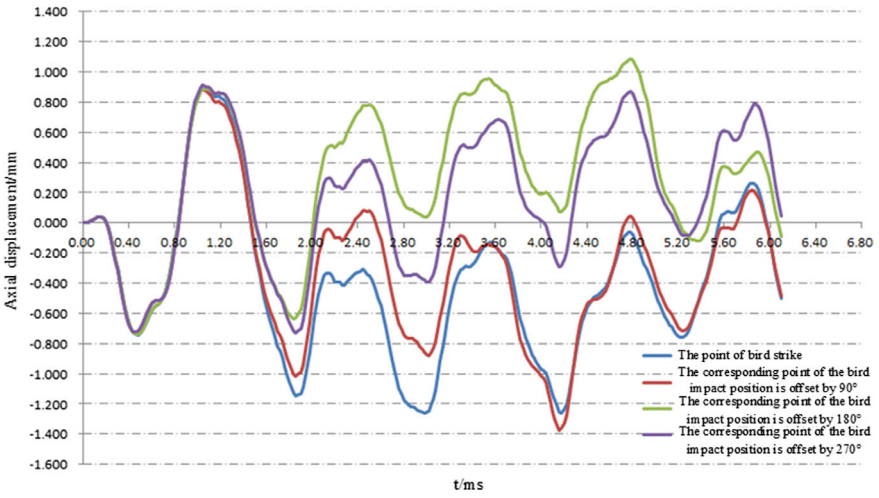


Fig. 7 Axial displacement of the back end of the booster rotor during bird impact

the simulation analysis of booster stage rotation and stator deformation in the bird impact process under the above two loads is carried out to determine the occurrence process of booster stage rotation and stator abnormal rubbing.

Analysis of Deformation Caused by Impact Loads on a Bird Strike. The impact load caused by the impact of the first bird is simulated and analyzed. The analysis results show that from the collision between the bird body and the blade to the complete end of the impact, the axial impact force on the fan rotor varies with the impact time, the range is – 130 to 105 kN (positive backward), and the axial deformation of the rear end of the pressurized stage rotor is − 1.4 to 1.1 mm (positive backward), as shown in Fig. 7. When the axial load is 32 kN, the back end of the booster rotor is deformed 1.1 mm backward, that is, the bird impact causes the minimum gap between the two to become 2.15 mm, and the rubbing will not occur.

Surge Load

A finite element model of the inner ring and blades in the pressurization stage is established. In the modeling process, the pressure applied on the surface of the inner ring flow channel is 0.6 MPa (which can be obtained through actual measurement), and the working pressure inside the cavity applied on the inner surface of the inner ring is 0.21 MPa.

As shown in Fig. 8, the inner ring of the fourth stage of the pressurized stage exhibits significant forward deformation under surge loads. The deformation ranges from – 2.8 mm to 0 (positive values denote rearward deformation), which exceeds the displacement caused by the bird strike load. At maximum surge loads, the minimum clearance between the aft end of the booster stage drum and the leading edge of the

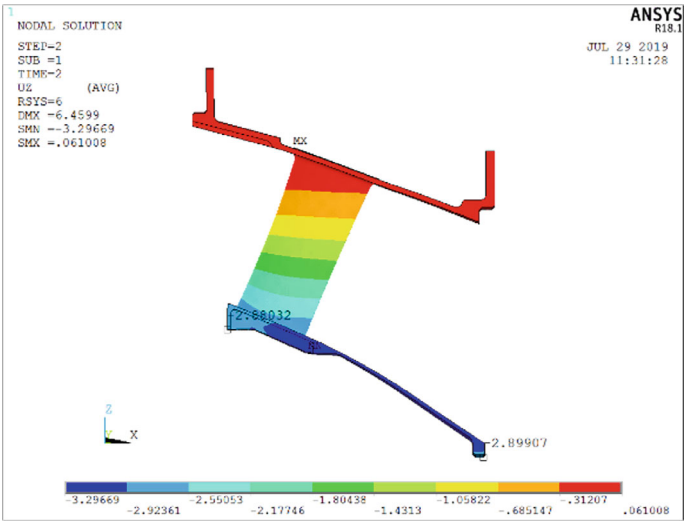


Fig. 8 Axial displacement of the inner ring of the pressurised stage stator under surge loading

booster stage fourth stage stator inner ring is 0.45 mm. Although no actual bumping has yet occurred, this tiny gap is in a critical state, and the gap is very close to the threshold at which a bumping occurs. At the same time, the maximum deformation of 2.8 mm exceeded the length of the radial locating surface on the intermediate casing by 2.65 mm, indicating that the upper stage stator of the pressurised stage detaches from its radial locating surface.

Buckling Analysis of the Inner Ring of the Last Stator

The inner ring of the stator of the last stage of the supercharging stage falls off from the radial positioning surface under the surge loads, which would cause its original constraints to disappear. In addition, considering that the component is a large-diameter, thin wall circle structure, to assess its structural stability under extreme conditions, a buckling analysis for the maximum surge loading condition is needed.

First, a finite element model of the pressurized stage inner ring and blades is built, as shown in Fig. 9. The inner ring runner surface is subjected to the same analytical load as the test-measured surge pressure, which is set to 0.6 MPa. The inner surface of the inner ring is applied with a working pressure of 0.21 MPa. Then, the boundary of the model is set to the same constraint state as one of the test pieces. Furthermore, the rear end of the inner ring disengaged from the intermediary magazine stop and the 32 blades at the front end of the inner ring at the position corresponding to the disengagement are set as free boundaries, and axial constraints are applied at the remaining positions.

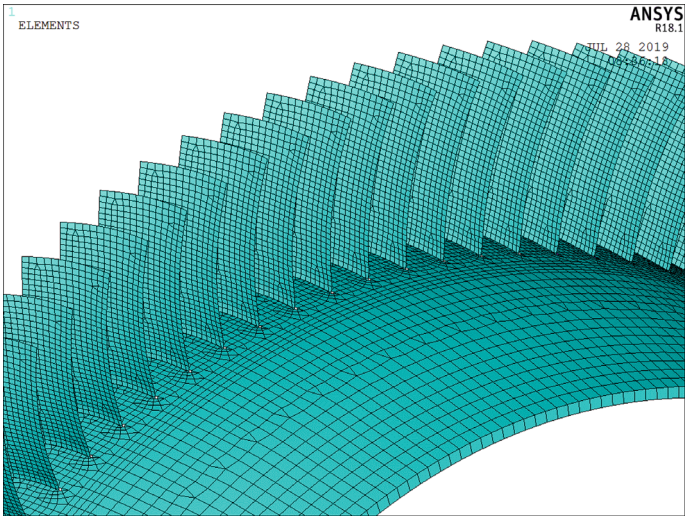


Fig. 9 Buckling analysis model of upper stage stator inner ring of pressurised stage

As shown in Fig. 10, the buckling instability reserve of the structure can be obtained based on simulation analysis results with $\lambda_1 = 0.09674$, $\lambda_2 = 0.105444$, and $\lambda_3 = 0.233803$. These results show that the stator inner ring will have buckling deformation at the corresponding position after the rear end of the stator inner ring escapes from the intermediate casing lip and some blades detach from the inner ring. The deformation pattern and amount of deformation of the inner ring in the simulation analysis are similar to those of the test workpiece.

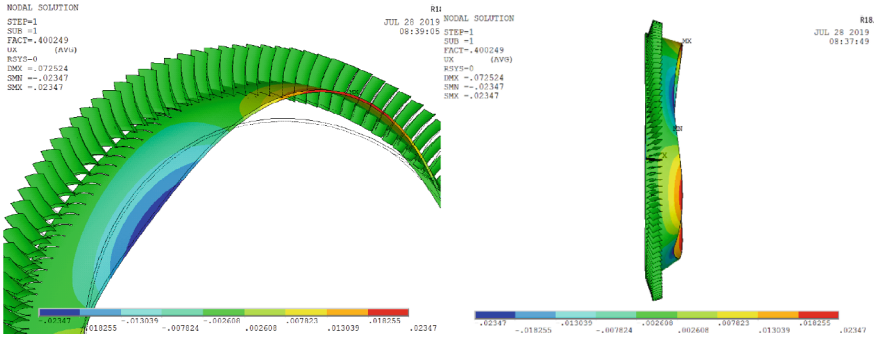
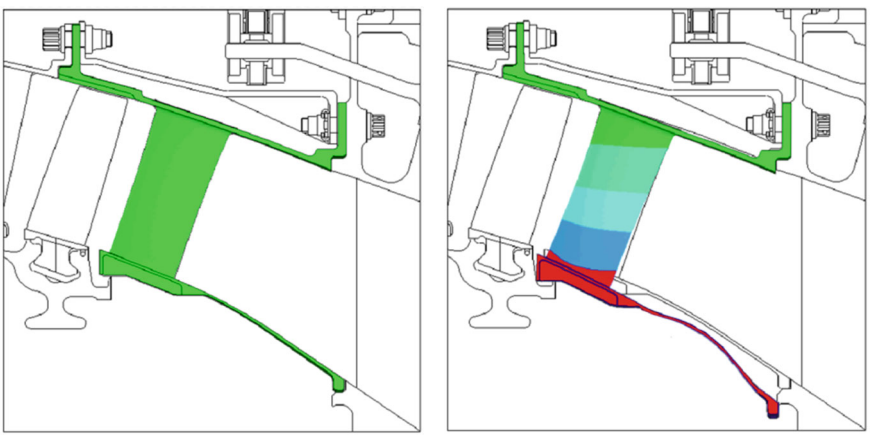


Fig. 10 First order buckling mode

3.3 Analysis of Rubbing Process

Causes of Collision. The simulation results show that the surge occurs after swallowing birds. Under the action of surge load, the rear end of the inner ring of the last stage stator of the booster stage moves forward 2.8 mm, which is detached from the radial positioning structure of the intermediary box, and the gap between the leading edge of the last stage stator and the trailing edge of the rotor becomes smaller. After detaching, the inner ring of the stator of the last stage of the booster stage is buckled under the action of surge pressure, resulting in a local depression at the back end of the inner ring, axial locking with the radial positioning structure on the intermediate casing, and local axial deformation of the leading edge of the last stage stator. Friction with the axial end face of the rotor occurs, resulting in a rapid decrease in low-pressure speed, as shown in Fig. 11.

Improvement Measures. In the bird impact test, the initial bird strike caused the high-pressure compressor to surge, leading to abnormal axial deformation of the inner ring of the booster stage stator. This deformation resulted in axial locking and excessive rotor friction against the compressor casing, ultimately causing the rotor to slow down and stop gradually. To solve this problem, it is proposed to increase the length by 4 mm at the positioning surface of the inner ring of the booster stage stator, increasing it from 2.65 mm to 6.65 mm to prevent abnormal friction, as shown in Fig. 12. By increasing the length of the positioning surface, the axial positioning capability of the inner ring of the booster stage stator can be enhanced, improving its ability to resist abnormal loads. Additionally, the overlap length between the rear end of the inner ring of the booster stage stator and the middle positioning surface has been increased from 1.42 to 2.65 to 3.92 to 5.38 mm. This increase is more significant



a) Normal state b) After swallowing birds

Fig. 11 Schematic diagram of rubbing

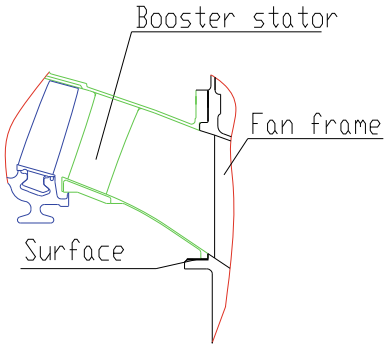


Fig. 12 Location surface growth

than the displacement of the rear end of the inner ring caused by the surge, thereby preventing friction between the front edge of the stator inner ring and the rear end of the rotor.

4 Analysis of the Compatibility Between Airworthiness Regulations and Development Technology

4.1 Characteristics of the Evolution of Airworthiness Regulations

The first edition of American aero-engine airworthiness regulation FAR33 "Aero-engine Airworthiness Standard" was released in 1965, and it has been revised 34 times in 35 editions so far. According to incomplete statistics, about 15 series 170 turbofan engines have completed airworthiness certification according to this regulation. The continuous revision and improvement of the regulations are verified and assessed by many models, and are the embodiment of the technical capability and accumulation of aero-engine technology in the United States for many years, which have complemented each other with the mature development of the aviation industry technology in the same period and coordinated discussions with aero-engine manufacturers.

With the development and application of new industrial technology, it became possible to popularize and implement the content of FAR 33.76. It was not until the 1990s, with the development and application of new industrial technology, that the bird-swallowing clause of FAR 33.76 began to be studied in 1976. At the same time, based on the success of a large number of aero-engine companies, after 24 years, the special 33.76 bird-swallowing clause was formally promulgated on December 13, 2000. The application of the FAA bird-swallowing clause is closely related to the development and maturity of advanced engine technology in the world.

China's engine airworthiness regulations were formulated with reference to the FAR33 of the United States. In 1988, the first edition of aero-engine airworthiness

regulation CCAR-33 was issued, and then the R1 version and R2 version CCAR-33 were released respectively. Up to now, a total of three versions have been released, which have been used for the model approval review of foreign civil aero engines imported from various types, but there is not yet a self-developed large culvert ratio turbofan engine to complete the airworthiness certification according to this regulation. In the process of regulation formulation, the method of directly referring to the airworthiness regulations of the United States in the same period is adopted, and the process of continuous integration and practice of airworthiness regulations, corresponding design capabilities and product technology is not experienced. Therefore, the formulation of regulations is not based on the technological basis of China's aviation industry, but more follows the pace of international advanced technology.

4.2 Evolution of Typical Bird-Swallowing Clauses

The assessment requirement of bird-swallowing ability in FAR33 originated from AC33-1 in 1965, and was formally incorporated into the airworthiness rules with FAR 33.77 foreign matter suction clause in 1974, which became the airworthiness clause. Bird swallowing requirements were put forward together with water, hail, ice, "sand-sand" mixture, lost blades, tire fragments and other foreign objects. Up to now, the content of the clause has been amended seven times and evolved into the current effective FAR 33.76 bird-swallowing clause.

In the course of the evolution of the seven provisions, taking the version released on December 13th, 2000 as the dividing line, which is quite different from the previous version, the relevant contents of the original "FAR 33.77 Foreign object inhalation" were cancelled, and the bird swallowing request was put forward as an independent "33.76 bird devouring" clause. The change in the content of the clause is mainly focused on the verification of the middle bird flock, including the weight of the bird, the projection position, the qualification criterion and so on. Compared with the original FAR 33.77 clause content, the FAR 33.76 clause is much more stringent. Specific changes include:

(1) Bird weight: the original standard requirement is 0.68 kg, but the new standard adds the weight requirement of 1.15 kg.
(2) Projection position: the original standard requires projection in the key area, but the new standard specifies that the largest bird is projected to the main runner of the core machine.
(3) Qualification criterion: the original standard requires that the thrust loss is not more than 25%, the engine cannot stop within 5 min after bird suction, and the potentially dangerous situation cannot be caused. The new standard adds the requirement to run 20 min according to the specified procedure after bird ingestion.

The contents of the two articles are representative to a certain extent, corresponding to different engine design technologies, which are called typical bird-swallowing clauses.

4.3 Engine Bird-Resistant Design Technology

Provisions Against the Need for Bird Technology. The typical bird-swallowing clause has evolved in the direction of stringent requirements for engine bird-swallowing and improved bird-swallowing ability. From the evolution of the two swallowing birds clause, we can see that its essential change mainly lies in the addition of three capacity requirements to the engine. The main results are as follows:

(1) By increasing the weight grade requirements of the middle birds, the anti-bird impact strength requirements of the engine inlet structure are improved.
(2) By increasing the position requirement that the first bird body must be aligned with the core machine, the requirements of the core machine's bird-swallowing ability are increased, and the stability requirements of the core machine after swallowing foreign objects are improved, that is, the aerodynamic design requirements of the core machine are improved.
(3) By increasing the requirements of a stable working state and working time after swallowing birds, the requirements of structural stability and aerodynamic stability of the whole engine after swallowing birds are improved, that is, the requirement of stable running ability of the engine.

Anti-Bird Design Technology. In the early anti-bird design of the engine, the main design goal is the safety of the engine after being hit by a bird. The design technology usually minimizes the structural damage on the engine inlet section caused by swallowing, which mainly includes:

(1) The material selection and structural design of the imported rotor blade, intake casing and stator blade make it have sufficient strength to reduce the damage, serious deformation, tearing and fracture after impact.
(2) The front stage rotor and stator with protruding shoulder and inner ring.
(3) Compressor (or fan) without intake guide vane.
(4) The rotor blade profile of the bird can be effectively chopped.
(5) The main bearing structure and bearing support of the engine are strong enough to provide a strength margin when a huge unbalanced force representing the damaged rotor blade occurs, so as to safely pull and stop the engine and rotate to the low-speed windmill.
(6) A casing against perforation and tearing or another shield sufficient to contain damaged rotor blades and stator blades.
(7) During the period of airflow disturbance caused by external swallowing, the engine has a rich stall margin, and the combustion chamber has good stability and the ability to reignite quickly.

At present, with the improvement of engine structure design technology, especially the continuous development and breakthrough of large diameter fan blade design and manufacturing technology, the difficulty of engine bird swallowing is no longer the anti-bird impact design of engine inlet structure. Technical research is being carried out with the goal of aerodynamic stability of the core engine after swallowing birds and reducing the surge amplitude of the high-pressure compressor. Considering the development of high-pressure compressor design technology, the bird-resistant design technology currently adopted by airlines around the world mainly focuses on reducing the number of birds entering high-pressure compressors, which mainly include. The main results are as follows:

(1) The design of the hat cover surface adopts the hat cover flow path design of the front cone and the back ellipse, which can improve the throwing ability of the external object and make the foreign object be introduced into the outer culvert.
(2) The axial clearance between the compressor first stage rotor blade (fan) and the inlet guide blade is increased, so that more foreign objects can be introduced into the outer culvert.
(3) The wide string fan blade increases the axial distance between the fan inlet and the booster stage inlet, prompting more foreign objects to be introduced into the outer culvert.
(4) The leading edge of the inlet guide vane is designed to tilt to introduce the foreign body into the outer culvert.
(5) The radial height of the inlet of the booster stage avoids the trajectory of the external object in the root flow path of the fan blade and reduces the total amount of foreign matter entering the connotation.
(6) The design of the flow path of the transition section before the inlet of the high-pressure compressor promotes more bird fragments to be thrown out and reduces the amount of entry into the core machine.
(7) The corresponding emission structure is designed before the inlet of the high-pressure compressor to encourage more bird fragments to be thrown out and reduce the amount of entry into the core machine.

In addition, the latest anti-bird technology summarized and analyzed by GE Company is shown in Fig. 13. The core content of the technology is to design the internal and external culvert flow path of the engine inlet to make the bird body fly out of the outer culvert after impact, reducing the total amount of foreign matter entering the core machine, and thus reducing the surge amplitude of the high-pressure compressor.

Technical Characteristics of Bird Resistance Design. Through the above anti-bird design technology, we can see that all the design technologies are essentially to reduce the number of birds entering the core machine, and the specific technology can be subdivided into two directions: reducing the import of foreign objects into the engine into the connotation and reducing the content of external objects into the core machine. The main results are as follows:

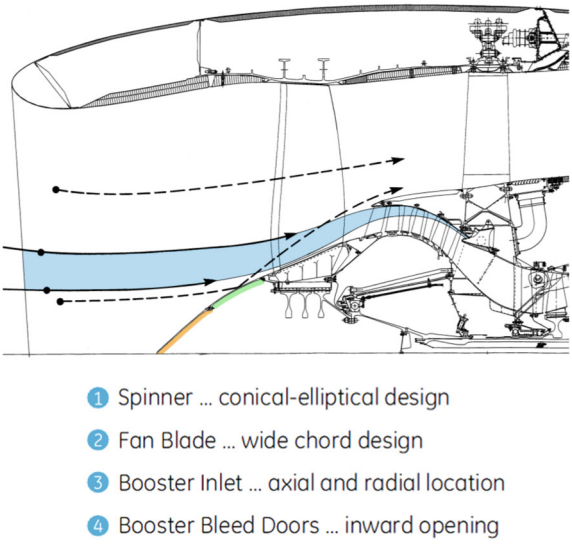


Fig. 13 The latest anti-bird technology of GE

(1) Reduction of foreign objects entering the connotative flow. For example, through the design of the intake cap, fan and the inlet flow path of the booster stage to the profile, the external objects entering the engine inlet can be thrown out of the outer culvert before entering the booster stage, thus reducing the amount of foreign matter entering the connotative flow path. In addition, through the design of the axial distance between the wide-string fan blade, the trailing edge of the fan blade and the inlet of the booster stage, the flow distance of the external object before entering the booster stage is increased. Combined with the design of the inlet flow path, the external object is thrown out of the outer culvert before entering the booster stage, thus reducing the amount of foreign matter entering the connotative flow path.
(2) Reduction of the external objects of the connotative flow entering the core machine. If an inward opening bleed valve is arranged at the front end of the inlet of the high pressure compressor, the external objects entering the booster stage are provided with a throwing channel before entering the core machine, so as to further reduce the number of external objects entering the core machine.

4.4 Application of Typical Bird-Swallowing Clause

Taking the typical bird-swallowing provisions of American airworthiness regulations as an example, the application of typical bird-swallowing provisions in representative turbofan engines with a large culvert ratio and the anti-bird technology applied to the corresponding products are analyzed, as shown in Table 2. It can be seen that the application of the latest anti-bird technology is the basis for the promulgation and

Table 2 Application of typical bird-swallowing clauses

Typical clause	Application product	Anti-bird design technology
FAR 33.77	CF6-80E1	None (Shoulder narrow chord fan blade)
	CF6-80C2	
	CFM56-5	
	PW4000	
	V2500	Wide chord fan blade
FAR 33.76	CFM56-7B	Wide chord fan blade Conical-elliptical spinner Axial and radial flow path of booster inlet
	Leap-X	Wide chord fan blade, Conical-elliptical spinner Axial and radial flow path of booster inlet Inward opening booster bleed doors
	GE90-92B	
	GE90-115B	
	Trent1000	

application of FAR 33.76 provisions. These technologies at least include wide-string fan blades, front cone and rear elliptical caps, and so on. However, in the stage of breakthrough and application of these technologies, the FAR 33.77 clause is still used to verify the airworthiness compliance of the engine with a large culvert ratio.

4.5 Essential Cause Analysis

The engine tested is a conventional dual-rotor turbofan engine, which adopts the design technology of front cone rear elliptical cap and shoulder narrow chord fan blade, but in the design of imported internal and external culvert flow path, it is not designed to make more birds fly out of the outer culvert after impact in order to reduce the total amount of foreign objects entering the core engine.

The anti-bird design technology adopted in the engine is more similar to that of CF6-80E1, CF6-80C2, CFM56-5 and PW4000, but does not have the bird-swallowing technology required by the FAR 33.76 clause. Therefore, the mechanism of the engine's automatic stop in the test is that the airworthiness verification of the swallowing bird is based on the mismatch between the FAR 33.76 clause and the design technical capability of the engine.

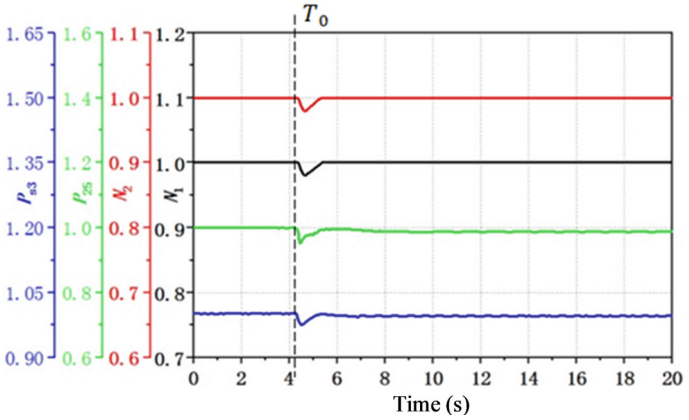


Fig. 14 The curves of test parameters over time

5 Verification of Airworthiness Compliance

The FAR 33.77 bird swallowing requirement is selected as the test basis, the radial positioning surface length structure of the upper stage stator of the pressurized stage on the intermediate casing is adopted, and the airworthiness compliance verification of a large bypass ratio engine swallowing bird swarm is conducted.

The relationship of test parameters and time is shown in Fig. 14, where N_1 and N_2 represent the speeds of low pressure and high pressure rotors, respectively; P_{25} and P_{s3} represent the high pressure compressor inlet total pressure and exit stator pressure, respectively, and all of these parameters are dimensionless. It can be seen that the high pressure compressor inlet total pressure and high pressure compressor outlet static pressure remain stable before the engine ingests the bird. The bird ingestion at time T_0, the total pressure at the inlet of the high-pressure compressor, and the static pressure at the outlet of the high-pressure compressor show a trend of first decreasing and then increasing, and then returning to a stable state, which is essentially the same as the test parameters before the bird is ingested by the engine. The test workpiece after the test is not found to have an unusual axial rub between the leading edge of the upper stage stator inner ring of the pressurized stage and the trailing edge of the rotor, which gives a solid validation of the effectiveness of the improvement measures.

6 Conclusion

In this paper, the verification of airworthiness of turbofan engines is analyzed, the direct cause and mechanism of engine stopping in the test are analyzed, and the corresponding improvement measures are put forward. Also, through the subsequent airworthiness compliance verification, the effectiveness of the measures is proved,

which improves the engine bird airworthiness compliance design and verification ability, and also promotes the improvement of domestic aero-engine airworthiness compliance verification technology. In addition, based on the analysis results, the following conclusions are highlighted:

(1) Based on the engine decomposition inspection, test parameter analysis, process image data analysis and simulation analysis under real parameters, it is determined that the surge of the high pressure compressor is caused by the impact of the first bird. The surge load causes the buckling deformation of the inner ring of the stator at the last stage of the booster stage after detaching from the positioning surface. This leads to the locking phenomenon axially at the rear end of the booster rotor, which is the direct cause of parking.
(2) Based on the evolution of airworthiness regulations of swallowing birds, the technical analysis of engine anti-bird design and the matching analysis between the connotation of regulations and technical development, it is determined that the mismatch between the FAR 33.76 clause and the technical capability of engine design is the essential cause of test parking.
(3) Based on the results of direct cause and mechanism analysis, the measures of increasing the length of positioning surface and selecting the CCAR 33.77 clause as verification basis are put forward, which are verified in the verification test of airworthiness compliance of swallowing birds.
(4) The improvement measures proposed in this paper can improve the bird strike resistance of the engine, such as avoiding the automatic shutdown caused by surge of the high-pressure compressor, enhancing the axial positioning capability of the inner ring of the booster stage stator, and improving its ability to resist abnormal loads, thus avoiding structural deformation and friction caused by surge loads.

References

1. Van GH, Krijgsveld KL, Klauke N et al (2019) Aeroecology meets aviation safety: early warning systems in Europe and the Middle East prevent collisions between birds and aircraft. Ecography 42(5):899–911
2. Metz IC, Ellerbroek J, Mühlhausen T et al (2020) The bird strike challenge. Aerospace 7(3):26
3. Wu B, Lin J, Xie A et al (2022) Flocking bird strikes on engine fan blades and their effect on rotor system: a numerical simulation. Aerospace 9(2):90
4. Chiavegatto R, Li YG (2024) Thrust rebalance to extend engine time on-wing with consideration of engine degradation and creep life consumption. J Eng Gas Turbines Powers 146(9)
5. Yang P, Yue W, Li J et al (2022) Review of damage mechanism and protection of aero-engine blades based on impact properties. Eng Fail Anal 140:106570
6. Luo G, Zhang HY, Wu CB et al (2019) General airworthiness analysis method of aeroengine bird ingestion and aircraft bird impact. Aeroengine 45(6):90–96

Open Access This chapter is licensed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License (http://creativecommons.org/licenses/by-nc-nd/4.0/), which permits any noncommercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if you modified the licensed material. You do not have permission under this license to share adapted material derived from this chapter or parts of it.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.



Aerodynamic Optimization Design of Propellers for Multi-rotor Unmanned Aerial Vehicles Based on NSGA-II



Ming Gao, Wei Guo, Jiaqiang Sun, Yunlong Huang, and Yong Cai

Abstract The aerodynamic optimization design of a propeller for unmanned aerial vehicles (UAVs) is a complex and critical process that directly affects the performance and flight efficiency of the UAVs. This paper aims to improve the UAVs' performance and efficiency, ensuring their stability and safety under various flight conditions. With thrust and hover efficiency of its propellers as the two optimization objectives, the propeller blade chord length and twist angle distribution as design variables, and the absorbed power of the propeller used as a constraint, the strip theory combined with the second-generation Non-dominated Sorting Genetic Algorithm (NSGA-II) are used in this paper for design optimization of the propeller. Through aerodynamic optimization and propulsion performance analysis, the hover efficiency of the propeller is enhanced. Results show that the optimized propeller increases its thrust and hover efficiency by 2% and 3%, respectively, providing significant support for the widespread application and development of multi-rotor UAVs.

Keywords Multi-rotor · UAVs · Propeller · Aerodynamic optimization · NSGA-II

1 Introduction

The optimization design of a propeller for unmanned aerial vehicles (UAVs) is a key element in improving the performance of UAVs. As the only lift component of multi-rotor UAVs, the performance parameters of a propeller directly determine the flight performance of UAVs [1–3]. Therefore, the design and optimization of a

M. Gao · J. Sun · Y. Huang (✉)
Quzhou Power Supply Company of State Grid Zhejiang Electric Power Co., Ltd., Quzhou, Zhejiang, China
e-mail: 576321120@qq.com

W. Guo
Institute of Unmanned System, Beihang University, Beijing, China

Y. Cai
Quzhou Guangming Electric Power Investment Group Co. Ltd, Quzhou, China

© The Author(s) 2026

H. Saman et al. (eds.), *The 9th International Conference on Advances in Construction Machinery and Vehicle Engineering*, Lecture Notes in Mechanical Engineering,
https://doi.org/10.1007/978-981-96-5527-4_65

propeller are crucial for enhancing the overall performance of the UAVs. The design process involves a comprehensive search and trade-off to ensure both operational requirements and other flight conditions are met. Traditional design methods rely on experience, characterized by preliminarily selecting the number of blades, propeller diameter, and the average chord length of the blades, and then choosing the radial distribution of blade thickness and an appropriate airfoil. However, this method is labor-intensive due to the various chord lengths and twist angle distributions. Hence, developing rapid design optimization methods to achieve the best performance of propellers under specific flight conditions is a significant approach for enhancing the performance of UAVs.

Since the advent of propellers up to the 1960s, various theories relevant to propellers have been developed, including momentum theory, blade element theory, vortex theory, and strip theory [4–6]. Among them, momentum theory [7] was first introduced in the nineteenth century, with Rankine and Froude presenting the axial momentum theory for propellers in 1865 and 1889, respectively. Although initially used for marine propellers, this theory is also applicable to air propellers and is thereby considered the beginning of theoretical research on propellers. Momentum theory simplifies the propeller to a thin circular disk with a uniform distribution of pressure in front of and behind it, thus generating a forward thrust for the propeller. The presence of pressure in front of and behind the disc contributes to the acceleration of the airflow through the disc. However, the momentum theory does not account for the physical mechanism of how the propeller generates thrust while consuming power, making it a simplistic theory. The propeller propulsion efficiency can be calculated based on such theory. Nevertheless, since only the energy carried away by the axial acceleration in the slipstream is considered in this theory, while ignoring other forms of energy loss, it is also referred to as ideal efficiency. Subsequent results of plenty of experiments have shown that the actual efficiency differs from the ideal efficiency by a certain ratio. As a result, this theory is generally used for preliminary predictions of propeller propulsion efficiency.

To delve deeper into the aerodynamic performance of propellers in detail, William [8] first introduced the blade element theory in 1878, dividing the propeller blade into infinitesimal elements along the radius. Each element provides thrust and consumes torque to a certain extent. Drzewiecki [9] provided a systematic discussion and solution methods for blade element theory in 1885, which was often referred to as Drzewiecki's theory. Subsequently, Lanchester refined Drzewiecki's work in 1907. It is worth noting that simple blade element theory assumes that disturbances caused by each blade element at any position along the radius are two-dimensional, without taking any induced factors into account. The aerodynamic performance of propellers is calculated using the aerodynamic data of airfoils based on the speed triangle formed by the axial flight speed and the rotational speed. Despite its limitations, such as a 28% discrepancy between theoretical and experimental thrust efficiency values, blade element theory has guided the development of subsequent theories. By comparing the computational data and experimental results of over 80 rotors in the follow-up research, Durand and William [10] found significant limitations in simple blade element theory, with a discrepancy of up to 28% between computational

and experimental values of propulsion efficiency, but the introduction of the blade element concept in strip theory has been instrumental in guiding the proposition and development of subsequent theories.

Given the limitations of simple blade element theory, various propeller theories have been proposed, including the Prandtl vortex theory and strip theory. Among them, the Prandtl vortex theory was established by Prandtl [4] through four papers between 1912 and 1915. It involves assumptions: first, the size of attached vortices along the radius is constant; second, there is an infinite number of propeller blades, leading to the formation of a propeller vortex structure, comprising a central vortex belt, a vortex column profile, and an attached vortex base. However, these assumptions have significant deficiencies. With the development of Prandtl's lift line theory and the introduction of attached and free vortices, H. Glauert applied Prandtl's finite wing theory to the propeller vortex model and proposed the strip theory [4, 11, 12] in 1926. According to the finite wing theory, when airflow passes over a lifting finite wing, the direction of the airflow changes, resulting in a downward flow known as downwash. The downwash angle depends on the lift produced, the tangent direction of the airfoil's trailing edge, and the wing span. This downwash angle reduces the attack angle of the blade element relative to the airflow, and in subsequent propeller performance calculations, this downwash angle is known as the interference angle. Although these methods provide a better understanding of propellers and significant advances have been made in propeller analysis, they involve complex calculations that are difficult for most designers to master. In 1936, H. Glauert's method, combining momentum theory and blade element theory, provided a good approach to the theoretical analysis of propellers [13]. Even though this method is based on the principle of 'independence on blade elements,' it has subsequently been proven to have a very high computational accuracy. Meanwhile, it simplifies propeller calculations dramatically and is thereby widely used.

The objective of aerodynamic optimization for air propellers is to improve propeller efficiency. Parameter optimization includes airfoil optimization, optimization of chord length and twist angle along radial distribution, and optimization of diameter and blade number. The most common optimization variables are the propeller chord length and twist angle distribution, which determine the blade shape and significantly impact propeller efficiency. There are many methods for optimizing air propellers to enumerate, among which the genetic algorithm stands out as the most prevalent. The genetic algorithm is a widely recognized and employed intelligent optimization technique.

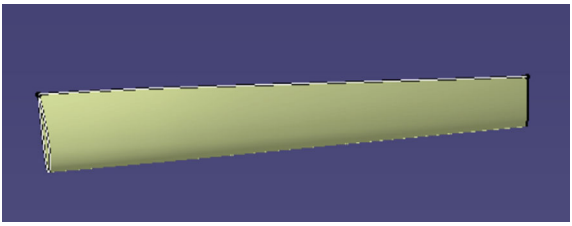


Fig. 1 CAD drawing of the propeller

2 Settings of Optimization

2.1 Propeller Model

The propeller used in the optimization is a motor-driven six-rotor, fixed-pitch, and two-blade propeller. The installation angle is adjusted at 0.7 R, indicating the angle between the geometric chord line of the airfoil and the plane of rotation. The propeller diameter (D) is 1.9 m, and Fig. 1 shows the propeller's schematic diagram.

2.2 Calculation Method for Aerodynamic Force

The speed of optimization relates to the calculation speed of the objective function. The establishment of the objective function in propeller optimization necessitates the propeller's efficiency, thrust, and power. Various methods, such as strip theory and computational fluid dynamics (CFD) simulation, can be used to calculate these parameters. Strip theory, the most common method for aerodynamic calculation of propellers, combines momentum theory and vortex theory to calculate the aerodynamic interference of blade elements. This method only requires the lift and drag coefficients of the blade elements to determine thrust, power and efficiency, and it is relatively fast in calculation.

The forces on a blade element are shown in Fig. 2, of which the thrust of the blade element is calculated as

$$dF = \frac{1}{2}\rho V_0^2 \frac{C_L b(1+a)^2}{\sin^2\phi\cos\gamma}\sin(\phi+\gamma)dr \tag{1}$$

The circumferential force of the blade element is given by:

$$dF = \frac{1}{2}\rho V_0^2 \frac{C_L b(1+a)^2}{\sin^2\phi\cos\gamma}\sin(\phi+\gamma)dr \tag{2}$$

The torque of the blade element is given by

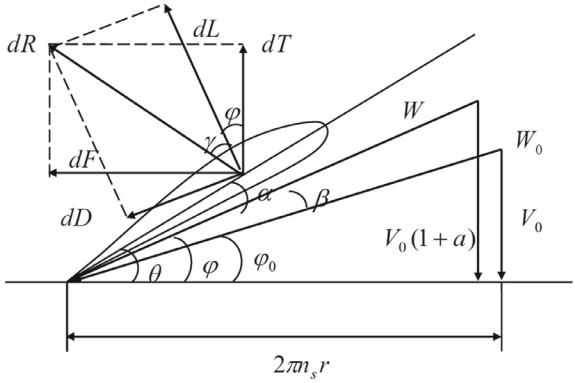


Fig. 2 Blade element stress diagram

$$dM = rdF = \frac{1}{2}\rho V_0^2 \frac{C_L b(1+a)^2}{\sin^2\phi\cos\gamma}\sin(\phi+\gamma)rdr \tag{3}$$

In the solution process, the crucial step lies in the determination of the interference angle. To calculate the interference angle β, by equating the thrust and torque obtained from the momentum equation with those from blade element analysis, we can solve for the interference angle.

$$C_L\sigma = \frac{4\sin(\phi_0+\beta)tg\beta}{1-tg\gamma tg\beta} \tag{4}$$

Finally, the interference angle is obtained through the Newton–Raphson iteration method, which subsequently enables the derivation of other parameters regarding the blade element.

2.3 Optimization Algorithm

The multi-objective genetic algorithm is used to optimize the aerodynamic performance of the propeller. As an evolutionary algorithm used to analyze and solve multi-objective optimization problems, the core issue is to coordinate the relationships between multiple objective functions to find an optimal solution set that maximizes (or minimizes) all objective functions. Among various multi-objective genetic algorithms, NSGA-II is the most influential and widely applied algorithm [14]. It has become a fundamental algorithm in multi-objective optimization due to its simplicity, effectiveness, and obvious advantages.

The basic idea of the NSGA-II is as follows: First, a random initial population of size N is generated, and the first-generation offspring population is obtained through

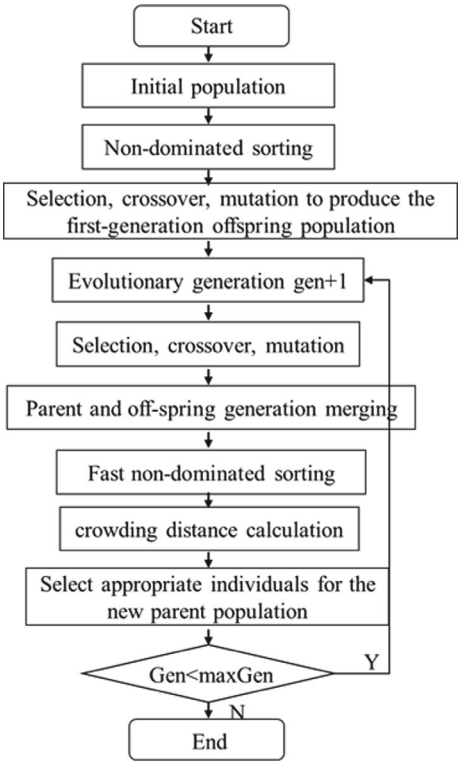


Fig. 3 Flowchart of NSGA-II

non-dominated sorting and genetic operations such as selection, crossover, and mutation. Second, starting from the second generation, the parent and offspring populations are merged for fast non-dominated sorting. At the same time, the crowding distance of individuals in each non-dominated level is calculated to select appropriate individuals for the new parent population based on the non-dominated relation and the crowding distance of individuals. Finally, the fundamental operations through genetic algorithm produce new offspring populations, proceeding in this manner until the termination condition of the program is met. The flowchart of the algorithm is shown in Fig. 3.

2.4 Design Variables and Constraints

The parameterization of the propeller blade shape is the basis of the design optimization. In this project, the NURBS (Non-uniform Rational B-spline) curves are used to discretize the twist angle and chord length distribution. The complex surface of the propeller blade requires high surface smoothness, so Non-Uniform Rational B-Spline (NURBS) curves are used for description. The mathematical definition is

as follows:

$$P(u) = \frac{\sum_{i=0}^{m} N_{i,m}(u) R_i P_i}{\sum_{i=0}^{m} N_{i,m}(u) R_i} \tag{5}$$

The design optimization of the propeller blade in this paper selects a propeller with a geometric shape as a reference propeller based on design experience and assumes a hover state with a propeller speed of 600 RPM/s and a flight altitude of 0 km. The optimization uses the airfoil twist angle and chord length of the blade at different spanwise positions as design variables. Nine design variables are considered, including the scaling factors and twist angles at each section.

During optimization, the power in the hover state is constrained, using the absorbed power under such a state as a constraint function. The optimization aims to maximize thrust and hover efficiency for the blade shape in the hover state. This leads to two optimization objectives, which are propeller thrust and hover efficiency. The mathematical model for the optimization design is as follows:

$$\begin{aligned} y_{objective} &= \min[f_1(x), f_2(x)] \\ & T \geq T_0 \\ & P \leq P_0 \\ s.t. \quad & \eta_{hover} \geq \eta_0 \\ \eta_{hover} &= \frac{T^{1.5}}{P \cdot \sqrt{2\rho\pi r^2}} \end{aligned} \tag{6}$$

where $f_1(x)$ and $f_2(x)$ are two sub-objectives for optimization, constructed from the weighted propulsion efficiency of the propeller under cruising and climbing conditions in both plain and high-altitude environments. They are defined as:

$$f_1(x) = -T \tag{7}$$

$$f_2(x) = -\eta_{hover} \tag{8}$$

2.5 Setup for Optimization Process

The iSIGHT platform is used for optimization, involving iterative calculations through geometric parameterization programs and aerodynamic performance assessment programs, and the NSGA-II, based on the strip theory engineering prediction model, is used for optimization search, as shown in Fig. 4.

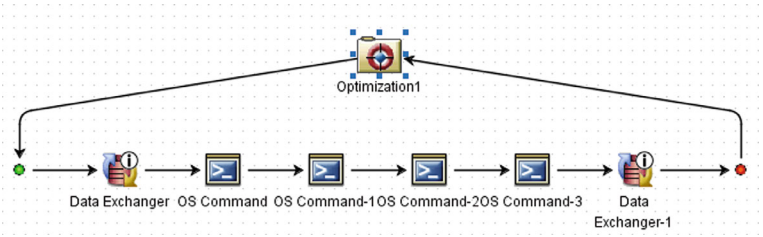


Fig. 4 Propeller optimization platform

3 Analysis of Optimization Results

The initial population size for the NSGA-II is set to 300, with a maximum of 2000 iterations. After convergence of the optimization, sample points and individuals recalculated from each generation are combined into a new population. This population is subjected to non-dominated sorting to determine the Pareto front, as shown in Fig. 5. The most aerodynamically optimal individuals are selected from the Pareto front.

The geometry, width, and twist angle distribution of the optimized propeller are shown in Fig. 6. It can be observed that, compared with the original propeller, the optimized propeller has a smaller chord length in the blade root region (less than 25%) and a larger chord length in the blade middle and tip regions. Additionally, the installation angle of the optimized propeller is larger for blade sections below 62% and smaller for sections above 62%.

Through calculations using strip theory, the optimized propeller shows an increase in thrust and hover efficiency by 2% and 2.1%, respectively, compared to the original propeller.

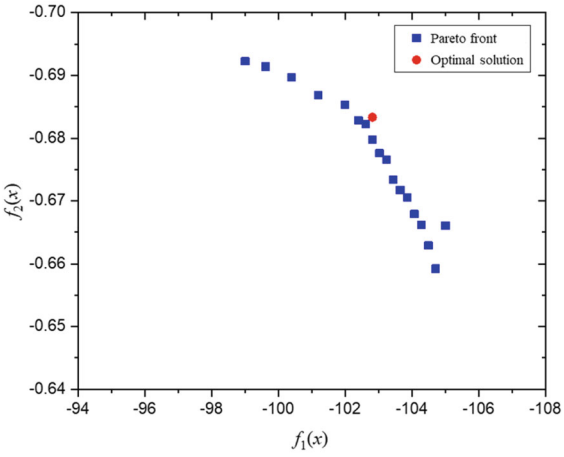
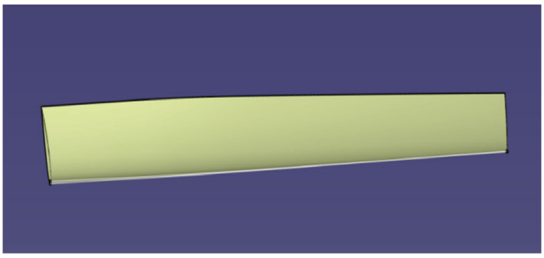
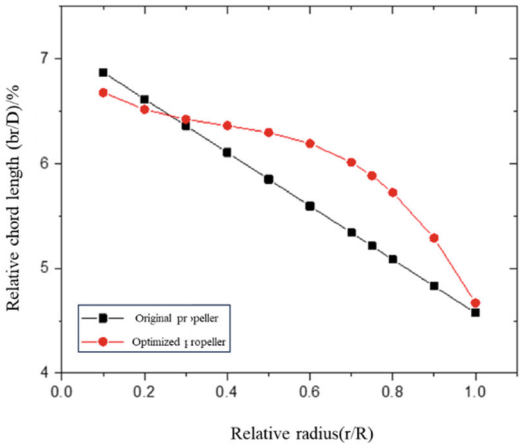


Fig. 5 Pareto front distribution map

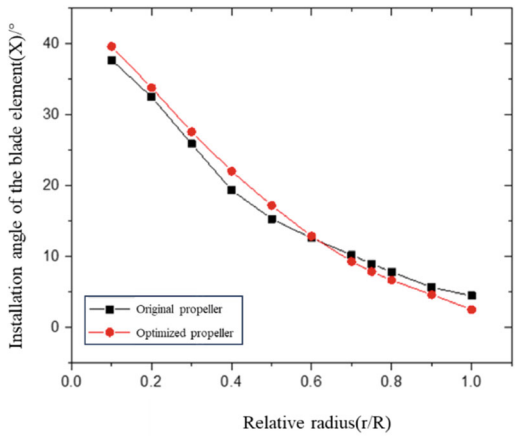
Fig. 6 The geometry parameters of the optimized propeller



(a) Shape of the propeller blade



(b) Width distribution of the propeller blade



(c) Propeller blade installation angle distribution

4 Conclusions

Propellers for multi-rotor UAVs are typically fixed-pitch propellers. To ensure performance, the chord length of the propeller is distributed non-linearly along the radius. NURBS curves are used to fit the distribution of chord length and airfoil twist angle, and appropriate parameters for chord length distribution are provided based on experience to obtain an initial propeller airfoil chord length distribution. Preliminary optimal angles of attack for each airfoil are calculated, and strip theory is used to determine the propeller's design dimensions and further calculate performance parameters of the propeller, such as axial induced velocity, thrust, torque, and power. Combining these with a multi-objective genetic algorithm, a Pareto optimal set for thrust and hover efficiency is obtained. The optimized propeller effectively enhances the thrust and hover efficiency of propellers for multi-rotor UAVs.

References

1. Weigang AN, Lin T, Wang S (2024) Structural design for a certain near-space composite propeller of airship using adaptive region division blending model. Chin J Aeronaut 37(5):301–316. https://doi.org/10.1016/j.cja.2023.11.018
2. Cruzatty C, Sarmiento E, Valencia E et al (2022) Design methodology of a UAV propeller implemented in monitoring activities
3. Asher PK, Deviparameswari K, Antonitta BF et al (2021) Comparative aerodynamic performance analysis on modified UAV's propeller by using CFD. https://doi.org/10.1007/978-981-16-0119-4_78
4. Liu PQ (2006) Theory and application of air propellers. Beijing University of Aeronautics and Astronautics Press
5. Wang X, Chen C, Liu Y et al (2024) A review of propel configurations and aerodynamic interaction between propellers and components of UAVs. In: Advances in aeronautical science and engineering, pp 1–16
6. Chen W, Yu Z, Zhong B et al (2023) Comprehensive design method for tilting ducted propeller. J Aerosp Power
7. Dahms J, Bardenhagen A (2019) Propulsion model for (hybrid) unmanned aircraft systems (UAS). Aircr Eng Aerosp Technol 91(2):373–380. https://doi.org/10.1108/AEAT-01-2018-0033
8. Froude W (1920) On the elementary relation between pitch, slip, and propulsive efficiency. Tech Rep Arch Image Libr
9. Carroll TJ, Carroll TR (2015) Wright Brothers' invention of 1903 propeller and genesis of modern propeller theory. J Aircr 42(1):218–223. https://doi.org/10.2514/1.6361
10. Durand WF, Lesley EP (1952) Comparison of model propeller tests with airfoil theory. Tech Rep Arch Image Libr
11. Lee DH, Kim CJ, Lee SH (2021) Development of unified high-fidelity flight dynamic modeling technique for unmanned compound aircraft. Hindawi. https://doi.org/10.1155/2021/5513337
12. Delaurier JD (1993) An aerodynamic model for flapping-wing flight. Aeronaut J 97(964):125–130
13. Rubin RL, Zhao D (2020) New development of classical actuator disk model for propellers at incidence. AIAA J 59(792):1–15. https://doi.org/10.2514/1.J059734
14. Srinivas N, Deb K (1995) Multi-objective function optimization using non-dominated sorting genetic algorithm. Evol Comput

Open Access This chapter is licensed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License (http://creativecommons.org/licenses/by-nc-nd/4.0/), which permits any noncommercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if you modified the licensed material. You do not have permission under this license to share adapted material derived from this chapter or parts of it.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.



A Quantitative Energy Efficiency Index for Heat Transfer and Energy Conservation of Plate Heat Exchanger in Commercial Vehicle



Haidong Wang, Yunjia Liang, Zhiao Wang, Jiaxin Liu, Lingqin Meng, and Weimin Liu

Abstract Heat transfer and energy conservation are crucial for plate heat exchangers (PHEs). To set up a criterion covering heat transfer and energy conservation for PHEs, a numerical analysis on the PHEs from commercial electric vehicle was performed and verified by its corresponding experiment. Then, via the derived entropy generation and *JF* factor, a quantitative energy efficiency index (EEI) was proposed in terms of *j* and *f* factors. This work may somehow offer a reference for the PHE comprehensive optimization.

Keywords Plate heat exchanger · Entropy generation · Energy efficiency index

1 Introduction

Thermal performance and energy conservation of plate heat exchangers (PHEs), a vital part of battery heat management system, flourish in the respective professional field since its birth. Considering both in the application of PHE, the integration of those two criteria into a quantitative energy efficiency index (EEI) is often involved as an essential problem in research and industrial fields.

Traditionally, many evaluation criteria have been investigated and developed in professional fields by published literature and research, such as *JF* factor ($j/f^{1/3}$, $j/f^{1/2}$, and j/f) in heat transfer and entropy generation in Thermodynamics.

JF factor is the ratio of the Colburn factor to the friction factor. A higher value of the *JF* factor typically implies enhanced comprehensive performance. Due to its simple and quantitative characteristics, *JF* factor has been widely applied for the optimization and design of PHE. Tang et al. [1] conducted experiments to assess capability of the thermal hydraulic of five distinct fin-and-tube heat exchangers (FATHEs)

H. Wang · Y. Liang · Z. Wang · J. Liu (✉) · L. Meng · W. Liu
College of Mechanical Engineeing, North China University of Science and Technology, Tangshan, China
e-mail: liujiaxin@ncst.edu.cn

© The Author(s) 2026

H. Saman et al. (eds.), *The 9th International Conference on Advances in Construction Machinery and Vehicle Engineering*, Lecture Notes in Mechanical Engineering,
https://doi.org/10.1007/978-981-96-5527-4_66

on the air sides. They optimized the FATHE with a genetic algorithm, utilizing the *JF* factor as the evaluation standard. Then the comparison of the numerical results confirmed that optimized fins of vortex-generator could supply a better heat transfer capability than slit fin. Wen et al. [2] presented a new type of corrugated-sawtooth fin, a combination of corrugated and sawtooth fins, and its Colburn factor was 19.97–38.22% higher than that of corrugated fin and 25.09–33.83% higher than that of serrated fin. Ma et al. [3] presented the optimal comprehensive performance of three mixing wavy tube at phase shift angle of 324° after investigating the distributions of their mean Nu, f, and NPEC. Tepe [4] applied a new punched ramp vortex generator with triangular shape to FATHEs and numerically described the ideal standardized height of ramp of 0.161 and ramp angle of 20° for the most feasible geometric design according to NPEC. Wang et al. [5] found that the little impact of spiral flow on the heat transfer efficiency and significant reduction in flow resistance inside the multi-tube HEs with helically corrugated tubes or transverse corrugated, and NPEC heat exchanger of the multi-tube and staggered helically corrugated tubes was higher than those of the others. Liu et al. [6] employed numerical modelling to estimate the capability of FATHEs, subsequently verifying the predictions through experimentation. After comparison, the maximum error of the heat transfer coefficient was 3.62%, while the maximum error in pressure loss was 4.51%. The fitted correlation is capable of predicting over 88.9% of j with an average error of less than 3.10%, and over 96.30% of f with an average error of less than 2.33%.

Ovando-Chacon et al. [7] showed that the performance of HE increased with more than 1 Richardson number *Ri*. For a novel geometry suggested by Shamsdanesh et al. [8], heat transfer rate, heat transfer improvement number, and effectiveness of the PHE increased by 46%, 10%, and 57%, respectively.

In this work, the thermal–hydraulic capability of a PHE for commercial electric vehicle was numerically studied and verified it with a corresponding experiment. A proposal was put forth that an EEI serve as a mutual criterion in the PHE for the conservation of energy and heat transfer. The following objects are the focus of this work: (a) The precision of the numerical simulation could be validated by a comparison of the heat transfer coefficient and pressure loss between the experimental and simulation results. (b) A comprehensive evaluation EEI $= \Delta S/\alpha_{jf}$ expressed with j and f factor would be derived from JF factor and entropy generation. (c) The proposed EEI could be a comprehensive criterion for the further structural optimization under heat transfer and energy conservation.

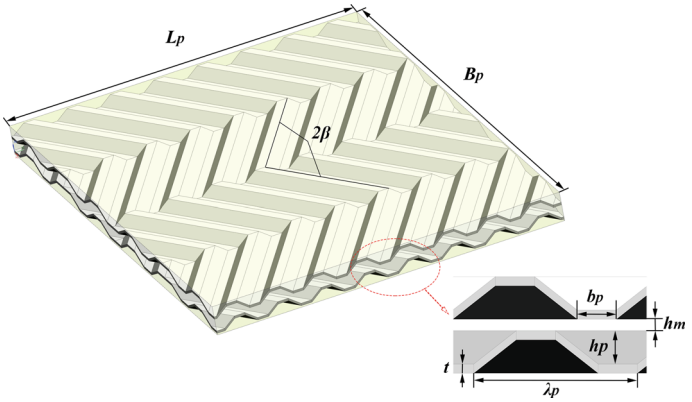


Fig. 1 Structural diagram of PHE

Table 1 Structural parameters of the PHE

Structure parameters (Unit)	Variable	Value	Structure parameters (Unit)	Variable	Value
Corrugated pitch	λ_p	4 mm	Spacing of plate	h_m	0.3 mm
Length of plate	L_p	58 mm	Thickness of plate	t	0.3 mm
Height of plate	h_p	1 mm	Chevron angle	β	30°
Width of plate	B_p	40 mm	Peak width	b_p	1 mm

2 Experimental Validation on Numerical Investigation

2.1 3D Model of PHE

An elementary unit of the PHE for commercial electric vehicle was confirmed as the analysis object. The model and its structural diagram are presented in Fig. 1. The detailed structural parameters were given as Table 1.

2.2 Meshing Configuration

GAMBIT 2.4.6 was applied to the grid generation for higher heat transfer accuracy, robust grid independence, and less resource consumption. To balance veracity and cost, the mesh independence experiment was conducted, and 8 configurations of 0.5, 0.96, 2.56, 3, 3.46, 4.8, 5.4, and 6.12 million grids were analyzed under the same conditions. Among those results, when grids increased from 4.8 to 6.12 million, the deviation was less than 0.61% for pressure loss and 1.58% for heat transfer coefficient, respectively. Thus, the number of grids was restrained to 4.8–5.4 million

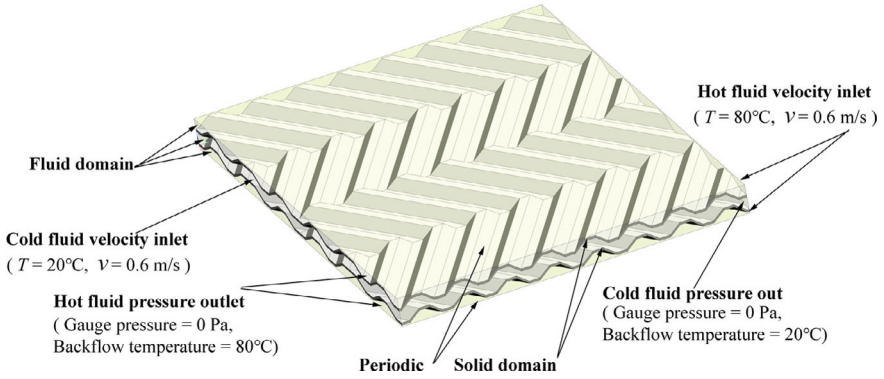


Fig. 2 Boundary conditions

in this work. In the implementation, the grids of fluid and solid domains were set with different sizes, 0.1 mm and 0.25 mm in detail, respectively.

2.3 Boundary Condition

The boundary conditions are set Fig. 2. More details are listed in Tables 2 and 3.

2.4 Introduction to Experimental Rig

The specific arrangement of the experimental rig is shown in Fig. 3. All experimental apparatuses were checked and calibrated before testing. To reproduce the heat transfer of the PHE, the test employed the use of coolant and film heaters. Throughout the test, the temperature and flow rate of the coolant at the entrance to the test section were maintained at constant values. The flow rate of fluid could be visually controlled by a peristaltic pump. The temperature of the plate surface was guarded by a thermal image. Sensors were used to obtain the inlet and outlet temperatures and pressures, respectively. All data was recorded after the system reached a 20-min equilibrium.

2.5 Result Comparison

Figure 4 illustrated a result comparison between the simulation and experimental. An increase of the inlet flow rate has been observed to result in an elevation in both the heat transfer coefficient and pressure loss, as evidenced by the results of simulation and experimental. The maximum error between the heat transfer coefficient and

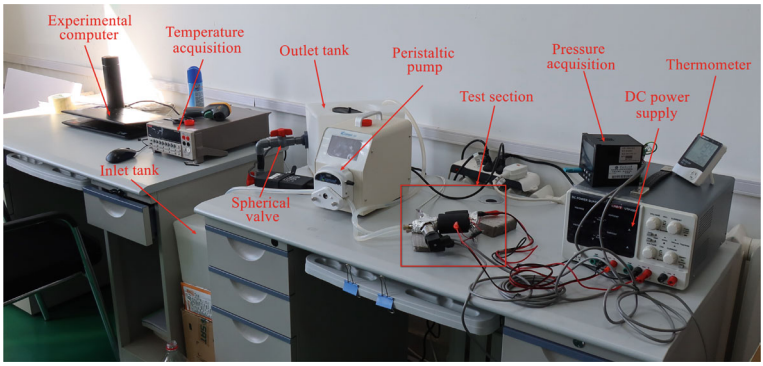
Table 2 Exhaustive data of numerical simulation

Working condition	Value (Unit)	Boundary	Working condition	Value (Unit)	Boundary
Inlet velocity of hot fluid	0.6 m/s	Velocity inlet	Left and right vertical surfaces along the X-axis	–	Wall
Inlet temperature of hot fluid	80 °C		Left and right vertical surfaces along the Z-axis	–	Periodic
Gauge pressure	0 Pa	Pressure outlet	Volume of board	–	Solid
Total hot fluid temperature reflux	80 °C		Volume of wrapping board	–	Fluid
Cold fluid inlet velocity	0.6 m/s	Velocity inlet	Turbulence model	–	Standard k–ε model
Cold fluid inlet temperature	20 °C				
Gauge pressure	0 Pa	Pressure outlet	Gravity Energy equation Turbulent kinetic energy	–	Applied
Total temperature of cold fluid reflux	80 °C				

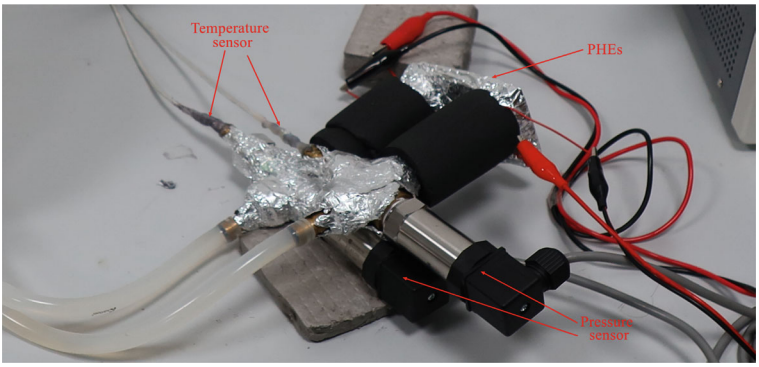
Table 3 Environment parameters of coolant

Environment parameters	Temperature	
	30 °C	70 °C
Specific heat capacity	3.319 kJ/(kg °C)	3.474 kJ/(kg °C)
Fluid density	1069 kg/m^3	1045 kg/m^3
Dynamic viscosity	2.94×10^{-3} Pa s	1.17×10^{-3} Pa s
Thermal conductivity	0.387 W/(m °C)	0.408 W/(m °C)

pressure loss in the results comparison between simulation and experimental in the tested flow range was 3.09% and 4.91%, respectively. The deviation between the experimental results and simulation results is within 5%, and we believe that the simulation results are valid.

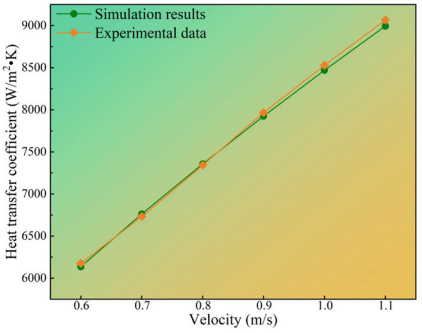


(a) Panorama of the test bench

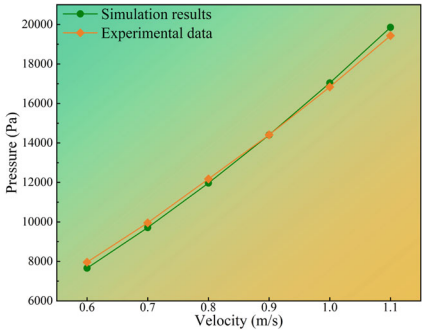


(b) Test section

Fig. 3 Experimental rig



(a) Heat transfer coefficient



(b) Pressure

Fig. 4 Result comparisons of heat transfer coefficient and pressure loss between numerical and experimental results

3 Energy Efficiency Index

The projected area A_p, and the effective area of a single board A_e of PHE are given by:

$$A_{\mathrm{p}} = L_{\mathrm{p}} \times B_{\mathrm{p}} \tag{1}$$

$$A_{\mathrm{e}} = A_{\mathrm{p}} \times \phi \tag{2}$$

The hydraulic diameter D_h from Reference [9] is expressed as:

$$D_h \cong \frac{2h_p + h_m}{\phi} \tag{3}$$

The amplification factor is defined in Reference [10] as:

$$\phi = \frac{A_e}{L_p \times B_p} \tag{4}$$

Prandtl number Pr and Reynolds number Re are calculated by:

$$\mathrm{Pr} = \frac{C_{\mathrm{p}}\mu}{\lambda} \tag{5}$$

$$\mathrm{Re} = \frac{\rho \times v \times D_{\mathrm{h}}}{\mu} \tag{6}$$

where ρ, v, C_p, μ, and λ are density, velocity, specific heat on constant pressure, fluid dynamic viscosity, and thermal conductivity, respectively.

Nu and f factor for the chevron PHE were introduced by Reference [11] and are written as:

$$Nu = 0.2267 \times \mathrm{Re}^{0.631} \times \mathrm{Pr}^{0.33} \tag{7}$$

$$f = 0.572 \times \mathrm{Re}^{-0.217} \tag{8}$$

$$f = 26.34 \times \mathrm{Re}^{-0.830} \tag{9}$$

Equation (7) is generally valid for $15 \leq Re \leq 1.5 \times 10^4$, Eqs. (8) and (9) are generally valid for Re>550 and $Re \leq 550$, respectively.

JF factor from Reference [12] is written as:

$$\alpha_{\mathrm{jf}} = \frac{j}{f^{1/3}} \tag{10}$$

Referring to References [13, 14], j and f are given as:

$$j = \frac{Nu}{\mathrm{Re} \times \mathrm{Pr}^{1/3}} = \frac{Nu \times \mu}{D_{\mathrm{h}} \times \rho \times v \times \mathrm{Pr}^{1/3}} \tag{11}$$

$$f = \frac{\Delta p}{L_{\mathrm{p}}} \frac{D_{\mathrm{h}}}{\rho v^2 / 2} \tag{12}$$

where Δp is pressure loss, ρ is fluid density.

Similar to Reference [15], ΔS_T, the entropy generation, induced by finite temperature differences is given as:

$$\Delta S_{\mathrm{T}} = m C_{\mathrm{p}} \ln \frac{T_o}{T_i} \tag{13}$$

where i and o indicate inlet and outlet, respectively.

For incompressible fluid flow, entropy generation ΔS_F caused by fluid friction is given by:

$$\Delta S_{\mathrm{F}} = \frac{\Delta p}{\rho} m \frac{\ln(T_{\mathrm{o}}/T_i)}{T_{\mathrm{i}} - T_{\mathrm{o}}} \tag{14}$$

The total entropy generation ΔS for a single side of PHE could be written as:

$$\Delta S = \Delta S_{\mathrm{T}} + \Delta S_{\mathrm{F}} \tag{15}$$

The heat transfer rate Q on the thermal side was written as:

$$Q = C_{\mathrm{p}} m (T_{\mathrm{i}} - T_{\mathrm{o}}) \tag{16}$$

According to Reference [16], Q for the thermal side could be also expressed as:

$$Q = h_{\mathrm{h}} \times A_{\mathrm{e}} (T_{\mathrm{b}} - T_{\mathrm{w}}) \tag{17}$$

where h_h is the thermal side heat transfer coefficient, and the mean temperatures of plate T_w and cold fluid T_b could be written as:

$$T_{\mathrm{w}} = \frac{T_{\mathrm{c,i}} + T_{\mathrm{c,o}} + T_{\mathrm{h.i}} + T_{\mathrm{h,o}}}{4} \approx 50\,^{\circ}\mathrm{C} \tag{18}$$

$$T_{\mathrm{b}} = \frac{T_{\mathrm{h,i}} + T_{\mathrm{h,o}}}{2} \tag{19}$$

where corner scripts h and c show the hot and cold fluids, respectively. By deriving Eqs. (16) and (17), the outlet temperature of cold fluid could be expressed as:

$$T_{\mathrm{o}} = \frac{2h_{\mathrm{h}}A_{\mathrm{e}}T_{\mathrm{w}} - h_{\mathrm{h}}A_{\mathrm{e}}T_{\mathrm{i}} + 2C_{\mathrm{p}}mT_{\mathrm{i}}}{h_{\mathrm{h}}A_{\mathrm{e}} + 2C_{\mathrm{p}}m} \tag{20}$$

The outlet temperature T_{o} in Eqs. (13) and (14) could be replaced by Eq. (20) and rewritten as Eq. (21):

$$\begin{aligned}\Delta S = {} & mC_{\mathrm{p}} \ln \frac{\frac{2h_{\mathrm{h}}A_{\mathrm{e}}T_{\mathrm{w}} - h_{\mathrm{h}}A_{\mathrm{e}}T_{\mathrm{i}} + 2C_{\mathrm{p}}mT_{\mathrm{i}}}{h_{\mathrm{h}}A_{\mathrm{e}} + 2C_{\mathrm{p}}m}}{T_{\mathrm{i}}} \\ & + \frac{\Delta p}{\rho} m \frac{\ln\left(\frac{2h_{\mathrm{h}}A_{\mathrm{e}}T_{\mathrm{w}} - h_{\mathrm{h}}A_{\mathrm{e}}T_{\mathrm{i}} + 2C_{\mathrm{p}}mT_{\mathrm{i}}}{h_{\mathrm{h}}A_{\mathrm{e}} + 2C_{\mathrm{p}}m} / T_{\mathrm{i}}\right)}{T_{\mathrm{i}} - \frac{2h_{\mathrm{h}}A_{\mathrm{e}}T_{\mathrm{w}} - h_{\mathrm{h}}A_{\mathrm{e}}T_{\mathrm{i}} + 2C_{\mathrm{p}}mT_{\mathrm{i}}}{h_{\mathrm{h}}A_{\mathrm{e}} + 2C_{\mathrm{p}}m}}\end{aligned} \tag{21}$$

The heat transfer coefficient, h_h, could be written as Eq. (22):

$$h_{\mathrm{h}} = \frac{Nu \times \lambda}{D_{\mathrm{h}}} \tag{22}$$

To comprehensively assess heat transfer and energy conservation, a quantitative EEI, the entropy generation per JF factor, was proposed as Eq. (23). In physics, it was expected to quantify how much the irreversible energy loss per thermal comprehensive performance is. Since the h_h and the Δp in Eq. (21) are respectively related to j and f, it could be finalized in terms of those two factors as Eq. (23):

$$\begin{aligned}\mathrm{EEI} = \frac{\Delta S}{\alpha_{\mathrm{jf}}} = {} & \frac{f^{\frac{1}{3}}}{j} mC_{\mathrm{p}} \ln\left[\frac{\lambda j A_{\mathrm{e}}(2T_{\mathrm{w}} - T_{\mathrm{i}}) + 2\mathrm{Pr}^{2/3}\mu T_{\mathrm{i}}}{\lambda j A_{\mathrm{e}} T_{\mathrm{i}} + 2\mathrm{Pr}^{2/3}\mu T_{\mathrm{i}}}\right] \\ & + \frac{f^{\frac{4}{3}}}{j^2} \frac{mL_{\mathrm{p}}v^2\left(\lambda j A_{\mathrm{e}} + 2\mathrm{Pr}^{2/3}\mu\right)}{4\lambda A_{\mathrm{e}}(T_{\mathrm{w}} - T_{\mathrm{i}})D_{\mathrm{h}}} \ln\left[\frac{\lambda j A_{\mathrm{e}}(2T_{\mathrm{w}} - T_{\mathrm{i}}) + 2\mathrm{Pr}^{2/3}\mu T_{\mathrm{i}}}{\lambda j A_{\mathrm{e}} T_{\mathrm{i}} + 2\mathrm{Pr}^{2/3}\mu T_{\mathrm{i}}}\right]\end{aligned} \tag{23}$$

4 Conclusion

A numerical analysis of the PHE was conducted to establish an EEI, considering heat transfer and energy conservation, and was validated by the corresponding experimental results. The EEI $= \Delta S/\alpha_{jf}$ was derived based on entropy generation, incorporating published correlations for the Nu and f factor, and expressed with the j and f factors. The major conclusions are as lists:

(1) The precision of the numerical analysis was validated by the discrepancy between the calculated and experimental values, which was found to be 3.09% for the heat transfer coefficient and 4.91% for pressure loss.
(2) EEI $= \Delta S/\alpha_{jf}$ could be derived with j factor and f factor for convenience and assist the result screening of the structural optimizations.
(3) The optimization at the same time considering both heat transfer and energy conservation would be a compromise instead of a coalition of them. The final optimal outcome would likely to strike a balance between thermal capability and energy conservation, depending on the specific goals and constraints of the system.

Acknowledgements This work is supported by National Science Foundation (No.52175063).

References

1. Tang LH, Zeng M, Wang QW (2009) Experimental and numerical investigation on air-side performance of fin-and-tube heat exchangers with various fin patterns. Exp Thermal Fluid Sci 33(5):818–827
2. Wen J, Li C, Hao H et al (2021) Numerical investigation on fin configuration improvement of 2 K sub-atmospheric plate-fin heat exchangers for the superfluid helium cryogenic systems. Appl Therm Eng 196. https://doi.org/10.1016/j.applthermaleng.2021.117290
3. Ma Q, Yang Y, Zuo Y (2021) Thermal performance optimization of heat exchanger with mixed cross-corrugated sinusoidal plate channels. Appl Therm Eng 195. https://doi.org/10.1016/j.applthermaleng.2021.117138
4. Tepe AÜ (2021) Heat transfer enhancement of fin-tube heat exchangers using punched triangular ramp vortex generator on the fin surface. Int J Heat Mass Transf 174. https://doi.org/10.1016/j.ijheatmasstransfer.2021.121326
5. Wang W, Shuai Y, Li B et al (2021) Enhanced heat transfer performance for multi-tube heat exchangers with various tube arrangements. Int J Heat Mass Transf 168. https://doi.org/10.1016/j.ijheatmasstransfer.2021.120905
6. Liu J, Cheng L, Pan D et al (2023) A general correlation for flat tube-and-fin radiator based on parametric CFD analysis. Numer Heat Transf Part A: Appl 1–26. https://doi.org/10.1080/10407782.2023.2190182
7. Ovando-Chacon GE, Ovando-Chacon SL, Rodriguez-Leon A et al (2020) Numerical study of nanofluid irreversibilities in a heat exchanger used with an aqueous medium. Entropy (Basel) 22. https://doi.org/10.3390/e22010086
8. Shamsdanesh R, Jafarmadar S, Javani N (2021) Exergy analysis of a plate heat exchanger with a new geometric design. J Therm Anal Calorim 145:1203–1213. https://doi.org/10.1007/s10973-020-10468-6
9. Raja BD, Jhala RL, Patel V (2018) Thermal-hydraulic optimization of plate heat exchanger: a multi-objective approach. Int J Therm Sci 124:522–535
10. Bi J, Liu J, Peng L, Jiang Y (2021) Signal-to-noise research on comprehensive performance of plate heat exchanger for commercial electric vehicle. Int J Therm Sci 165
11. Imran M, Pambudi NA, Farooq M (2017) Thermal and hydraulic optimization of plate heat exchanger using multi objective genetic algorithm. Case Stud Therm Eng 10:570–578
12. Webb RL (1994) Principles of enhanced heat transfer. Wiley, New York
13. Al Zahrani S, Islam MS, Saha SC (2020) Heat transfer augmentation in retrofitted corrugated plate heat exchanger (in English). Int J Heat Mass Transf 161:13, Art no 120226

14. Della Torre A, Montenegro G, Onorati A, Khadilkar S, Icarelli R (2019) Multi-scale CFD modeling of plate heat exchangers including offset-strip fins and dimple-type turbulators for automotive applications, (in English). Energies 12(15):20, Art no 2965
15. Shah RK, Sekulic DP (2003) Fundamentals of heat exchanger design. Wiley, Hoboken, New Jersey, p 927
16. Liu JX, Jiang YK, Wang BZ, He SM (2019) Assessment and optimization assistance of entropy generation to air-side comprehensive performance of fin-and-flat tube heat exchanger, (in English). Int J Therm Sci 138:61–74

Open Access This chapter is licensed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License (http://creativecommons.org/licenses/by-nc-nd/4.0/), which permits any noncommercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if you modified the licensed material. You do not have permission under this license to share adapted material derived from this chapter or parts of it.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.



The Simulation Researching for Some Vehicle Based on Dragging to Tray Carriage Load Case

He Li and Yongfu Xiao

Abstract In order to solve the issue of Some vehicle structure failure issue based on dragging to tray carriage load case, analysis the whole load case process. Infer the formula of the load case and obtain the key case. Taking advantage of explicit transient dynamic algorithm to simulation and resolve the typical load case. And the simulation result is completely consistent with the counterpart data from the test, which show that the mothed is effective. Based on the mothed, optimize the failure structure and realize them into the prototype, which is verified in the final test and it shows the design is effective. Realize the customization to generate simulation file automatically based on HYPERMESH secondary development TCL script, and improve modelling efficiency to at least 90%. The method in this paper can be an efficiency tool to develop vehicle body structure performance.

Keywords Tray carriage · Simulation · Failure · Cracking · Automatic

1 Introduction

During the vehicle development process, dragging performance is one of the important performance items. Ensuring the strength of the structure during vehicle towing is very crucial for product quality. At present, there are several government standards referring to the vehicle dragging device design, and some of them specifies different load angles and amplitudes putting on towing hook [1]. But in the reality usage scenario, it is more complicate loading case that varies of trailer types lead to difference of dimension parameters and kinematics parameters [2]. Up to now, vehicle dragging device design methods, are mostly based on above government

H. Li (✉) · Y. Xiao
FAW Global R&D Center, China FAW Co., Ltd., Changchun, China
e-mail: lihe1@faw.com.cn

Y. Xiao
e-mail: xiaoyongfu@faw.com.cn

© The Author(s) 2026

H. Saman et al. (eds.), *The 9th International Conference on Advances in Construction Machinery and Vehicle Engineering*, Lecture Notes in Mechanical Engineering,
https://doi.org/10.1007/978-981-96-5527-4_67

standard specification. Besides, Huang Li Ping also specified relative load case specification which refers to sea transportation case [3]. All above, is not able to cover the situation aiming at specific trailer type and parameters. It is urgently and necessary simulation method which refer to unique usage scenario matching specific trailer type and parameters [4–10].

2 Summary

In this paper, aiming at the rear bumper structure failure issue in the process of some SUV is dragged up to rescue trailer. At first, we analyzed the key load case based on above usage scenario, simulated and researched the key load case with FEM and recur the failure status. And then compared the simulation results with test counterpart response. Secondly, proposed the improvement structural concept and verified it with above simulation method. Finally, realized the improvement structural which is passed through the test verifying, and it shows the effectivity and practicability of the simulation method which simulates the dragging up to trailer load case in this paper. The whole process flow chart is shown in Fig. 1.

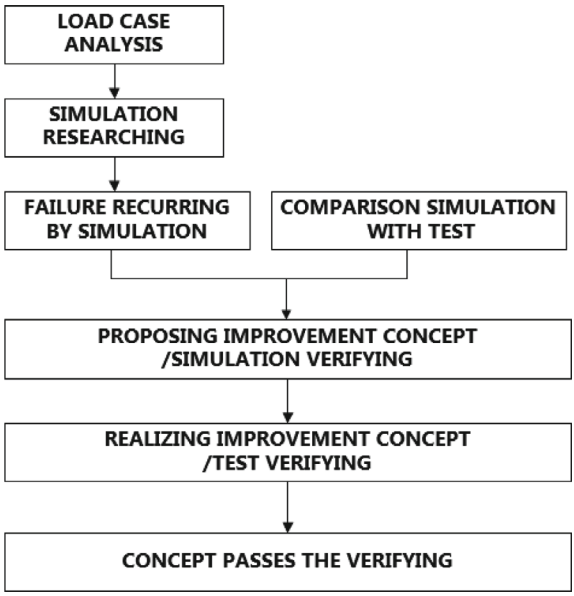


Fig. 1 Flow diagram

3 Load Case Analysis

The dragging up to trailer load case is decomposed as follows: (1) The plate of rescue trailer is descended to the ground, until its buffer plate contacting the ground to carry on the dragged vehicle. (2) Release the handbrake of the dragged vehicle, and driver sits down in the vehicle to modify the direction trace during the dragging process. (3) Fix the rear wheel with fix device in the main plate, and then take back the whole plate. During above process, angle between main plate and buffer plate is reducing to 0° as the buffer plate is detaching from ground. Relate to the whole plate device, vehicle revolve around the center of its rear wheel until the buffer plate is totally horizon to the main plate. All above, from the original status (as 1 describes), there are 2 key load cases: "dragging to plate load case (as 2 describes)" and "rising up plate load case (as 3 describes)", as shown in Fig. 2.

3.1 Dragging to Plate Load Case Analysis

Dragging to plate load case, which is defined as vehicle is dragged from original place to highest place in the whole plate, is shown in Fig. 3. The input parameters of this load case are as follows: (1) Verified vehicle parameter, include of wheel place, radius of tire (R) and wheel base (L), etc. (2) Dragging trailer parameter, include of the height between the fixed place of wire rope and surface of the main plate (L1), the distance between dragging limitation place and the fixed place of wire rope (L2),

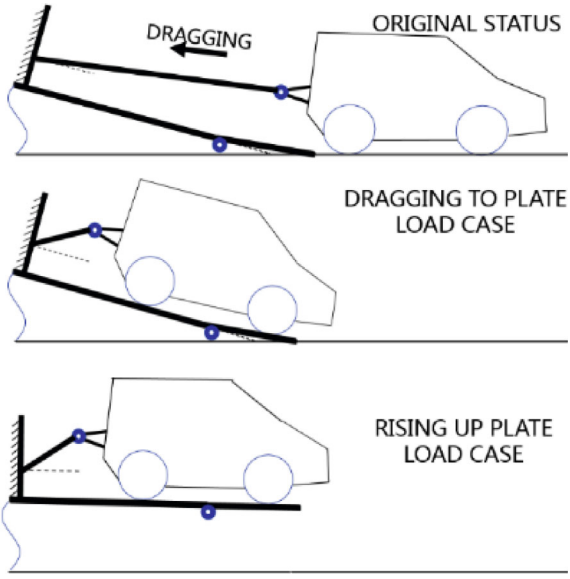


Fig. 2 Entire process of the load case

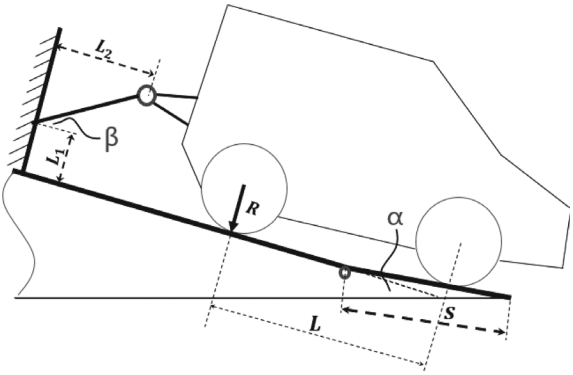


Fig. 3 Highest positon and key parameters

the length dimension of buffer plate (S), the degree between main plate and ground (α), the degree between buffer plate and ground (γ), and the angle between wire rope and surface of main plate according to the above parameters (β), etc.

According to the posture status and parameters shown in Fig. 3, obtain the load distribution and inferring conditions which is shown in Fig. 4, the tension force of wire rope is assigned as F_hook, the gravity of vehicle is assigned as F_mass, the coefficient of rolling friction is assigned as f. According to the X axis of local coordination, infer the formula based on force-balance theory:

$$F_{\text{hook}} \cos\beta = F_{\text{mass}} \sin\alpha + F_{\text{mass}} \cos\alpha * \text{f} \tag{1}$$

$$\text{F}_{\text{hook}} = \text{F}_{\text{mass}} * \left(\frac{\sin\alpha + \cos\alpha * \text{f}}{\cos\beta}\right) \text{F}_{\text{hook}} = \text{F}_{\text{mass}} \cdot \left(\frac{\sin\alpha + \cos\alpha \cdot f}{\cos\beta}\right) \tag{2}$$

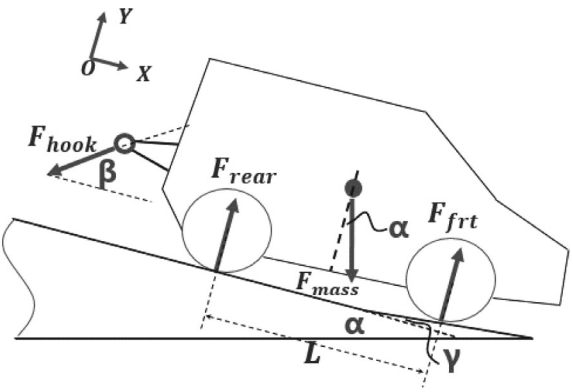


Fig. 4 Force condition of dragging to plate load case

3.2 Rising up Plate Load Case Analysis

After the vehicle is dragged to the highest place of the whole plate in the last load case, the rear wheel is fixed by lock device in the main plate. The load case simulates the following process that dragging plate rising to the statue of dragging plate is horizon. The limitation place of front wheel placing in the buffer plate is shown in Fig. 5. The description of the load case is as follows: relate to the whole plate device, vehicle revolve around the center of its rear wheel until the buffer plate is totally horizon to the main plate. And the forced degree of vehicle revolving is as the rising up plate case load, which is shown in Fig. 5. Infer the formula of the forced degree as follows:

$$\theta \approx \arcsin\left(\frac{s * \sin\gamma}{L}\right) \tag{3}$$

4 Simulation Researching

According to above load case analysis, to simulate the switching between 2 typical load cases, referring to many researching requirements include of monitoring the posture variation, loading course variation, and structure failure, etc. The explicit transient dynamic algorithm is more fitful to cater to requirements above.

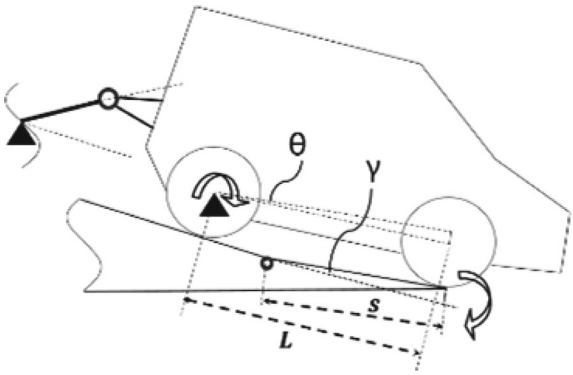


Fig. 5 Force condition of rising up plate load case

4.1 Explicit Transient Dynamic Algorithm

Transient dynamic algorithm is the numerical method, which takes the load of event function as input data, and takes the variable varying with time as output data. In this paper, apply the ABAQUS platform to researching the load case simulation.

The equilibrium equation of transient dynamic is as follows:

$$M\ddot{\delta}(t) + C\dot{\delta}(t) + K\delta(t) = F(t) \tag{4}$$

In above formula, the load vector, which is known function of time is assigned as F(t); the displacement vector, which is the function of time is assigned as $\delta(t)$; velocity vector, which is the derivative of first order of $\delta(t)$ to t, is assigned as $\dot{\delta}(t)$; acceleration vector, which is the derivative of second order of $\delta(t)$ to t, is assigned as $\ddot{\delta}(t)$.

The total kinetic energy of model is as follows:

$$T = \frac{1}{2}\dot{u}^T K > 0 \tag{5}$$

In above formula, iteration is assigned as n, time increment is assigned as, Δt β and γ are constant.

Based on above, infer the formula as follows:

$$m_{ij} = \frac{\partial^2 T}{\partial \dot{u}_i \partial u_j} = m_{ji} \tag{6}$$

It can be seen, stiffness metrics K shows that the positive semidefinite is symmetry and shows the system kinetic energy. Viscous damping is assigned as C, which is proportional to velocity. In ABAQUS platform, there are 3 different types of damp, and they respectively are direct modal damp, compound modal damp and Rayleigh damp.

The formula (4) is dynamic basic equation. In the solution procedure, it can be processed with semidiscretization method by ABAQUS. There are 2 main methods to solve the problem of dynamics, they respectively are immediate integration and mode superposition method. In this paper, the former method is applied to solve the formula (4), at the same time, central difference method is applied to calculate in ABAQUS/explicit solver. Apply the immediate integration method to solve the above differential equations, and obtain the displacement, velocity, acceleration vectors, and then solve the dynamics responses.



Fig. 6 Posture comparison of highest position between model and real

4.2 Simulation Researching

Simulate the whole process of dragging to plate load case and rising up plate load case, and solve the response course, such as wire rope tension force, structure strain and stress, displacement, etc.

First of all, model the whole mechanism within the ABAQUS/explicit platform: include of dragging beam and part of BIW structure, which is fractured to maintain the part adjacent to loading area. As shown in Fig. 6.

Secondary, model the simplify mechanism of wire rope system. 1D element, which type is CONN3D2, is applied to model wire rope model. The direction of vector from one endpoint of 1D element to the other, is as the X direction of the 1D element local coordination. One endpoint of 1D element is set as full constraint, and the other is bonded to towing hook.

Next, set property and material parameters to structure. And the material type of the whole beam and bracket is all the same. So it is necessary to set material failure parameters to above structure to monitor the failure behavior during the load case simulation process.

Finally, set the steps in order referring to load case analysis.

After solving step, check the harmony of transformation simulation result. Check internal force of 1D wire rope model. The first degree of freedom internal force is positive all the time during the time course, and the counterpart of the rest degree of freedom remain zero. Moreover, the transformation of structure in the loading process is consistent with reality situation. All above shows that the simulation mothed is correct.

4.3 Failure Issue Recurring Simulation

By simulating the whole load case process, the transformation cloud chart show that the location of failure element is totally same as the counterpart appeared in the test, as shown in Fig. 7.



Fig. 7 Failure comparison between simulation and real

4.4 Simulation Result Analysis

Wire rope tension force analysis: as shown in Fig. 8, force value increase to 1.3t when dragging vehicle with constant speed, and suddenly decrease to 1.1t when drag to highest position and stop. During the process, the failure appears and the force increase to peak value and suddenly fell down and shock on the eve of degree of BIW inclination leads to 0.88 degree (buffer plate is horizon to main plate), at the moment the wire rope tension force increase to 2.65t. All above shows that, the maximum wire rope tension force is overlay force between dragging to plate load case and rising up plate load case.

As shown in Fig. 8, compare static status plot to constant speed status plot, the peak values of force is almost the same, and that illustrates the current structure carrying capability is level about 2.65t. The characteristic of the load case is that, the forced displacement which is loaded in structure is definite value, based on this, if the stiffness of BIW structure is enough and without failure appears, the greater structure stiffness, the greater wire rope tension force, at the same time, the higher demanding standard of structure strength. Only in the condition that, the stiffness of the assembly of beam structure, towing hook bracket, BIW structure, is moderate

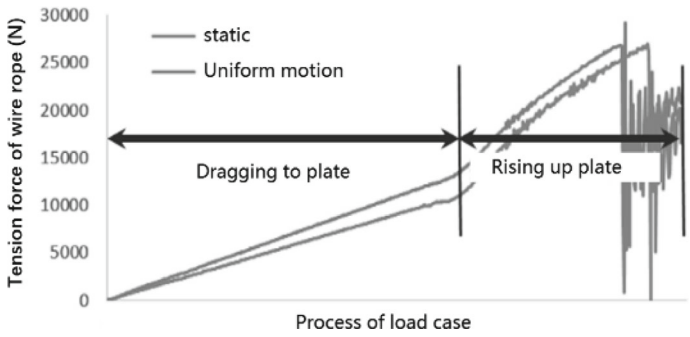
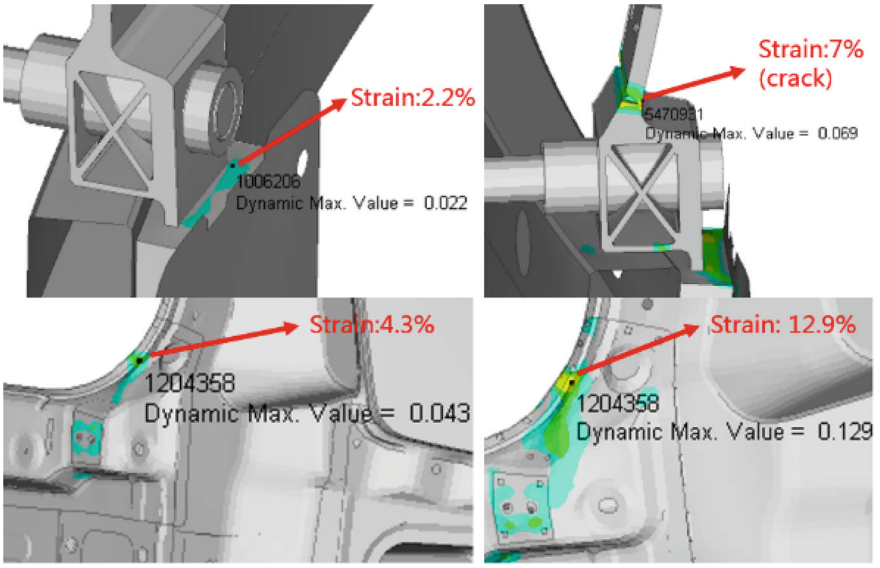


Fig. 8 Wire rope tension course

and rational matched. And all of them are able to absorb certain displacement and resistance strength damage due to above displacement they absorb, in this case the design concept is rational to maintain its strength performance.

Strength result analysis: the strain cloud chart of beam structure, towing hook bracket, BIW structure in 2 typical load cases is shown in Fig. 9, and strain result is shown in Table 1.

By the result shows we can learn that, the maximum strain of towing hook beam structure in dragging to plate load case is 2.2%; and the maximum strain of BIW structure is 4.3%. Based on result, the maximum strain of towing hook beam structure



(a)strength result of dragging to plate load case

(b) strength result of rising up plate load case

Fig. 9 Strength simulation cloud chart (strain)

Table 1 Simulation result

Structure\load case	Dragging to plate			Rising up plate				Strain limitation (%)
	Maximum strain (%)	Location	Wire rope tension force	Maximum strain (%)	Location	Remark	Wire rope tension force	
BIW	2.20	Back panel	1.3t	12.9	Back panel		2.65t	<15
Towing hook beam	4.30	Energy absorb box		7	Towing hook bracket	Crack		<7

in rising up plate load case increases to the limitation value 7% and failure appears in the towing hook bracket, at the moment the maximum strain of BIW structure increases to 12.9%.

4.5 Comparation Between Simulation and Test

Install the tension sensor between the wire rope and towing hook to monitor the test data of tension force loading in the hook, as shown in Fig. 10. And paste the strain gage in the risk location of strength based on simulation result, as shown in Fig. 11.

According to testing data, compare the response result of wire rope force and structure strain of BIW, as shown in Table 2.

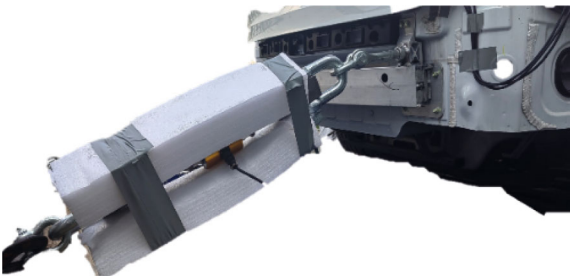


Fig. 10 Tension sensor arrangement



Fig. 11 The max principle strain direction of BIW structure risk location

Table 2 Validation between simulation and testing result

	Tension force of wire rope (t)	The max principle strain of BIW structure risk location (%)	Remark
Simulation	2.65	11.6	The max principle strain direction
Test	2.72	12.3	

5 Improvement Measure and Structure Concept Verifying

5.1 Proposals of Improvement Structure Concept

By above simulation research and result analysis, we can learn that the characteristic of the load case is the whole structure cater to constant forced displacement, and during the process of structure resistant above forced displacement, it is necessary to make sure rational matching stiffness among the towing hook beam, towing hook bracket and BIW structure to harmonize their displacement and strength. Based on this, propose the specific instrument measures as Fig. 12 shown.

The first instrument measure: enhance the location part of towing hook bracket structure which appeared failure in reality operation, to avoid reappear failure crack because of structure weakness;

The second improvement measure: enhance the energy absorber box structure to improve the capability that towing hook beam resistant global displacement, to avoid the part of structure where is link area between bracket and BIW suspend a lot due to the whole beam structure lack of stiffness;

The third improvement measure: embed aluminium blocks to beam cavity above where the top metal plate of beam is suspending the towing hook bracket, to improve



Fig. 12 Structure status after test verifying

Table 3 Simulation result of improvement structure concept

Strength result	Carrying capability		Strength result of structure					
			Towing hook beam			BIW		
	Tension force of wire rope (N)	Status	Maximum strain (%)	Limitation strain	Remark	Maximum strain (%)	Limitation strain (%)	Remark
Improvement concept	27,136	No crack	5.9	$\leq 7\%$	Beam base panel	14.3	≤ 15	Back panel

the capability that towing hook beam resistant global displacement, to avoid the part of structure where is link area between bracket and BIW suspend a lot due to the whole beam structure lack of stiffness.

Based on above three measures, enhance the weaknesses of towing hook bracket structure, at the same time decrease the possibility the assembly of beam displace globally to protect the BIW structure from excessive forced displacement.

5.2 Verifying of Improvement Structure Concept

Simulation verifying: according to above improvement measures, display the results separately in Table 3 shows. After structure improving, the maximum wire rope tension force increase from 2.65t simulation result before improving to 2.71t simulation result below improving, and no failure crack appears finally. And the maximum structure strain of beam and BIW is less then material limitation value. Due to structure enhancement and no failure crack appearing, the structure strain of BIW structure is greater than before status.

Test verifying: reality the above improvement measures to prototype, and verify it in final test. The test showed that there was no failure crack appearing and the improvement structure passed the test verifying. The structure status below test is shown in Fig. 12.

6 Automatic Flow Script Customization Based on Hypermesh Platform

TCL Language, which is the secondary development language of HYPERMESH interface, is a type of script language based on controlling and expanding application routine. There are a lot of characteristics such as simplified grammar, no compilation required, platform independent, etc.

Based on TCL script, write the automatic flow script including the entire process of dragging to tray carriage load case by taking advantage of HYPERMESH interface script language, and integrate the key steps such as system modelling of wire rope mechanism, property generating automatically, boundary conditions setup, application of load, creating load cases and analysis controlling, etc. All above are fully customized in the paper after mesh modelling step to decrease the error rate of manual modelling, improve simulation consistency. At the same time, the manual cost falls sharply and the modelling efficiency increase at least 90%.

The whole automatic processes are as follows. (1) Modelling wire rope mechanism; (2) Overall attitude rotating of the entire model; (3) Creation and definition of DRAGGING TO PLATE LOAD CASE; (4) Creation and definition of RISING UP PLATE LOAD CASE; (5) Outputting simulation files. The process below need to be coordinated necessary engineer interaction. The process design flow below based on TCL script language is as Fig. 13 shown.

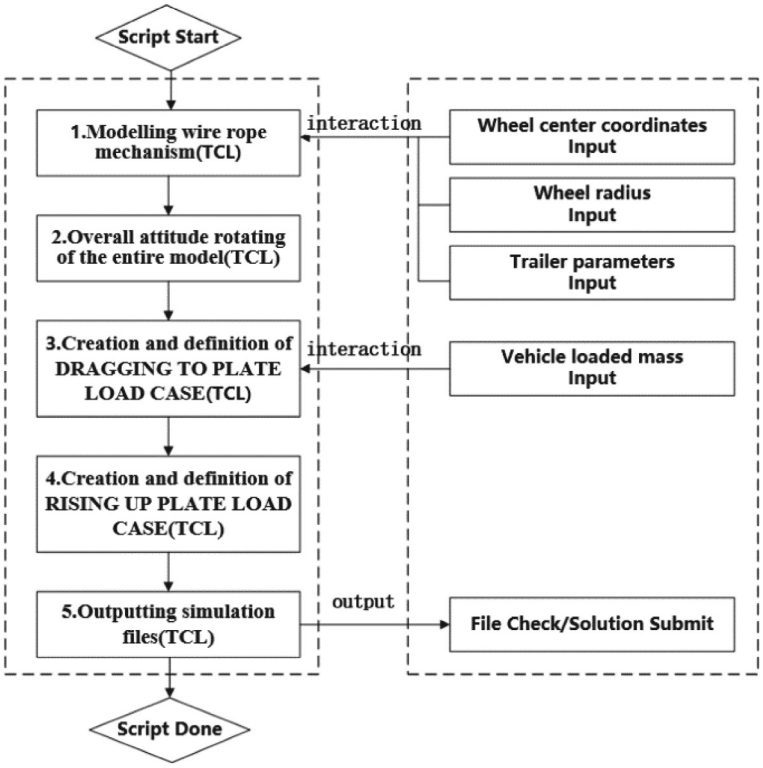


Fig. 13 Script routine design flow

7 Conclusion

There are different parameters to different types of dragging trailers, such as the mechanic type, dimension, the degree between plate and ground, etc. The parameters of trailer is almost the worst influence to vehicle structures, include of degree between main plate and ground, degree between buffer plate and ground and wire rope inclination degree when the vehicle is dragged to the highest position. In this paper, propose a simulation method, in which according different reality trailer parameters which are put into relative formula to solve typical load cases and simulation.

First of all, solve the load case that the vehicle is dragged up to trailer, and compose 2 key typical load cases, which are dragging to plate load case and rising up plate load case; next, propose a simulation method of above 2 typical load cases, include of modelling, loading and solving, and recur the failure crack which is high consistence with reality location by simulation measure, and compare the simulation result and test data and find that the comparison ratio between them is good consistence. All above illustrate the method is accuracy and effective. Then propose the improvement structure concept and verify it with simulation method below, reality the improvement structure concepts into prototype and verify it in the final test, and finally pass the test verifying and the reality problem is solved. In summary, the method in this paper propose a thought and method of BIW structure development as a necessary complement to sustain development of trailer function and structure.

References

1. GB/T 32087-2015, Lightweight car dragging device.
2. Huang LP (2020) Vehicle structure durability performance: theory and practice. China Machine Press, Beijing
3. Cui PX, Ran GQ (2023) Dragging device strength analysis of some lightweight vehicle. Vehicle Test Report 01:143–145
4. Zhang YJ, Li JY (2023) Technique improvement Research for towing hook strength toughness. Pract Technol Automobile 18:117–121
5. Li D, Zheng Z C, Gao Y, Li F, Ping S (2019) Lightweight design of some off-road vehicle towing hook base bracket based on topo optimization. Tractors Agricul Transp 46(05):48–55
6. Wu YB, Lin DQ (2017) Test method of lightweight vehicle dragging device. Pract Technol Automobile 02:176–181
7. Wang SJ, Bai XJ (2023) Structure design and analysis for dragging pool trailer. J Tianjin Univ Technol 01:9–15
8. Sorge F (2019) Motion and force analysis of slow-speed multi-trailer systems. Proc Inst Mech Eng C J Mech Eng Sci 233(8):2615–2635
9. Beregi S, Takács D, Stépán G (2016) Tyre induced vibrations of the car–trailer system. J Sound Vib 362:214–227
10. Widyotriatmo A, Nazaruddin YY, Putranto MRF, Ardhi R (2020) Forward and backward motions path following controls of a truck-trailer with references on the head-truck and on the trailer. ISA Trans 105:349–366

Open Access This chapter is licensed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License (http://creativecommons.org/licenses/by-nc-nd/4.0/), which permits any noncommercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if you modified the licensed material. You do not have permission under this license to share adapted material derived from this chapter or parts of it.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.



Research on Lightweight Design Methods of Cavity Structure



Ning Wang, Hongjun He, Baojun Yu, Yongfu Xiao, and Liwei Sun

Abstract In order to realize the lightweight of vehicle body structure, this paper proposes a design method based on regional structure lightweight optimization. Specifically, the D-pillar cavity structure of a vehicle is focused as the structure for optimization, and a parameterized model containing cross section dimensions is established. The work includes dimension optimization, structural topology optimization, connecting location and material thickness optimization, and it must meet all body performance requirements. Finally, the optimization results achieved 11% weight reduction for the D-pillar cavity structure. The torsional stiffness remains unchanged, and the modal, durability and safety performance meet the vehicle program design requirements.

Keywords Parameterization · Structure lightweight · Pillar cavity · Dimension optimization · Topology optimization

1 Introduction

Under the guidance of the "dual carbon" initiative, the automotive industry, as an important field for promoting energy conservation and emission reduction, is increasingly attracting attention to vehicle lightweighting. In "The Roadmap 2.0 for Energy Conservation and New Energy Vehicles" [1], a technology development roadmap has been developed for lightweight vehicles. The lightweight path of vehicles includes material lightweighting, structural lightweighting, and manufacturing process lightweighting. Considering cost factors, structural optimization is the most commonly used technical approach to achieve lightweighting of the vehicle body. The lightweight coefficient of the body-in-white is an important indicator for evaluating the level of lightweighting of the car body. The smaller the value, the higher the level of lightweighting. Its value is proportional to the weight of the body-in-white and

N. Wang (✉) · H. He · B. Yu · Y. Xiao · L. Sun
FAW Global R&D Center, Changchun, Jilin, China
e-mail: ounei@foxmail.com

© The Author(s) 2026

H. Saman et al. (eds.), *The 9th International Conference on Advances in Construction Machinery and Vehicle Engineering*, Lecture Notes in Mechanical Engineering,
https://doi.org/10.1007/978-981-96-5527-4_68

inversely proportional to the product of the torsional rigidity and the footprint area of the body-in-white. Therefore, improving torsional rigidity and reducing weight for the body-in-white is an important way to achieve lightweighting of vehicles. Based on past experience, the torsional rigidity of the body-in-white is more sensitive to changes in cross-sectional dimensions, while the weight of the body-in-white is more sensitive to changes in material thickness.

In recent years, scholars have used various methods to design lightweight structures for the body-in-white. For example, Zuo Wenjie et al. [2] set the minimum quality as the goal, with torsional rigidity, stress, and the first three natural frequencies as constraint conditions, the part thickness as the design variable, and established a body-in-white structure optimization model, achieving a reduction of 34.6 kg in body-in-white weight. Based on the precise concept model of the body-in-white, Wang Zhenhu [3] et al. optimized the dimensions of 19 key main sections, achieving a 2.7% reduction in vehicle weight. Ouyang Chenhui et al. [4] established a parameterization model of the car body, and successively carried out topological optimization, section optimization, and thickness optimization, achieving an actual decrease in the mass of the car body of about 1.7% in the early stage of development of a certain vehicle model. Wang Dengfeng et al. [5] conducted multi-objective optimization with the thickness of the side collision safety part as the design variable, resulting in a reduction in weight of 2.68 kg. Wang Chao et al. [6] identified the key force transmission path of the cab through topological optimization and strengthened the relevant structure. Based on the Morph method, they controlled the proportion of section nodes to achieve control of the cross-section dimension. Through multi-objective optimization design of the cab with comprehensive thickness parameters, they achieved a weight reduction of 12.8%.

All of the above studies have achieved good results in vehicle lightweighting, using techniques such as material thickness optimization, cross-section optimization, and topology optimization. However, in most cases, the existing research on material thickness optimization is based on the existing part structure. A few experts combine the optimization of cross-section dimensions and part thickness synchronously, but the cross-section dimensions are limited by the surrounding structure and connection relationship, with a small range of variable values, and the situation of overall position changes of the cross-section is not considered.

To further explore the lightweighting potential of the body-in-white structure, a lightweighting design method for the D-pillar cavity structure of a certain vehicle model has been proposed in the paper. This method is based on the implicit parameterization model of the body-in-white, and carries out cross-section optimization, topology optimization, part splitting, docking position and material thickness optimization. Combined with strength, modal and safety performance evaluation, a lightweighting body structure that meets the performance requirements is obtained.

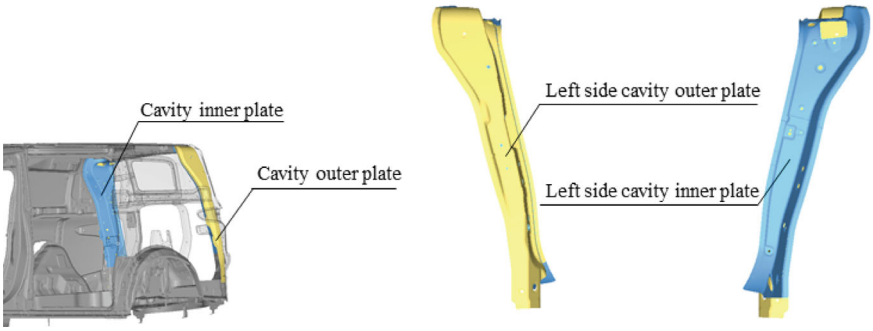


Fig. 1 D-Pillar cavity structure

2 Cross-Section Optimization Based on Implicit Parameterization Model

For regular sedans, only the ABC pillars are sufficient to meet the rigidity performance requirements. However, for longer vehicles, the torsional rigidity is relatively low and requires support from the D-pillar structure to improve rigidity. Therefore, the structure of the D-pillar and lightweighting design solutions have an important impact on the performance of the vehicle body. This section is based on the implicit parameterization model of the body-in-white, and focuses on the cross-sectional optimization design of the D-pillar structure of a certain vehicle model.

2.1 Optimization Area

This thesis focuses on the optimization design of the D-pillar pillar structure of a certain vehicle model. The D-pillar cavity is composed of inner and outer plates, with symmetrical structures on both sides. The thickness of the inner plate is 1.4 mm, and the thickness of the outer plate is 0.9 mm. The weight of the cavity is 12.2 kg, as shown in Fig. 1.

2.2 Parameterization Model and Variable Design

Based on the input conditions of the reference vehicle model and the preliminary development of the project, including the CAS surface and cross-sectional geometric data in the conceptual stage and the reference vehicle model finite element data, the SFE CONCEPT software is used to build a certain vehicle's body-in-white implicit

Fig. 2 Parameterized model

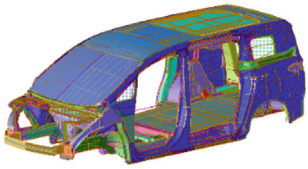
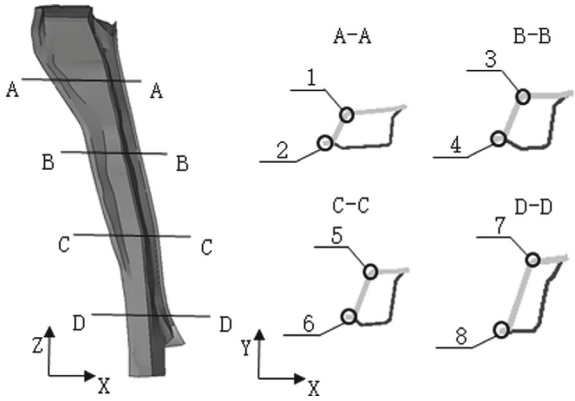


Fig. 3 Cross-sectional dimension variables



parameterization model, as shown in Fig. 2. Implicit parameterization models determine the position and shape of vehicle body parts through base points, baselines, and cross-sections. They connect complex vehicle body structures through topological relationships, allowing the original topological relationships and geometric continuity to be maintained even when various parameters change. By combining with finite element mesh automatic generation and updating techniques, rapid evaluation of model performance and optimization design can be achieved.

Four sections are taken from top to bottom in the D-pillar cavity, as shown in Fig. 3. Due to the influence of external factors such as the separation seam of the rear orifice, the structural characteristics of the inner side are selected as the dimension variables of the cavity. There are 8 design features and 12 cross-section dimension variables (symmetrical left and right D-pillar structures, with synchronized changes in cross-section variables). The variable range can be determined based on the structural spatial dimensions and layout conditions, as shown in Table 1.

2.3 Section Dimension Optimization

DOE experimental design is conducted using the optimized Latin hypercube algorithm to generate 150 sample points. The driving parameterization model generates the corresponding finite element model to calculate torsional rigidity and weight.

Table 1 Cross-sectional dimension variables and variable ranges

Number	Parameter name	Variable range (mm)
1	A-A cross-section No.1 position variable X-direction	(−20, 50)
2	A-A cross-section No.1 position variable Y-direction	(−-35, 50)
3	A-A cross-section No.2 position variable X-direction	(−50, 20)
4	B-B cross-section No.3 position variable X-direction	(−15, 40)
5	B-B cross-section No.3 position variable Y-direction	(−40, 50)
6	B-B cross-section No.4 position variable X-direction	(−50, 15)
7	C–C cross-section No.5 position variable X-direction	(−15, 30)
8	C–C cross-section No.5 position variable Y-direction	(−40, 50)
9	C–C cross-section No.6 position variable X-direction	(−50, 15)
10	D-D cross-section No.7 position variable X-direction	(−15, 15)
11	D-D cross-section No.7 position variable Y-direction	(−50, 50)
12	D-D cross-section No.8 position variable X-direction	(−50, 15)

The symmetrical structure changes synchronously when the variables change.

By fitting the response surface of cross-section dimension variables, torsional rigidity and weight, a second-order polynomial surrogate model as shown in Eq. (1) is constructed.

$$y = \beta_0 + \sum_{i=1}^{n} \beta_i x_i + \sum_{i=1}^{n} \beta_{ii} x_i^2 + \sum_{i=1}^{n-1} \sum_{j=i}^{n} \beta_{ij} x_i x_j + \varepsilon \tag{1}$$

In the equation β_0, β_i, β_{ii}, β_{ij} are unknown parameters, x_i is the i-th variable, and ε is the error.

In engineering, the fitting accuracy of the surrogate model is generally R^2 evaluated by the adjustment coefficient. The adjustment coefficient formula is shown as Eq. (2). When the adjustment factor is greater than 0.9, the surrogate model accuracy can replace the simulation model.

$$R^2 = 1 - \frac{\sum_{i=1}^{n} (y_i - \tilde{y}_i)}{\sum_{i=1}^{n} (y_i - \tilde{y})} \tag{2}$$

In the formula: n is the number of test sample points; y_i is the simulated value of the i-th response; $\widetilde{y}_i$ is the predicted value of the i-th response by the approximate model; and $\tilde{y}$ is the average value of the simulated values.

The R^2 of the optimized surrogate model for weight and torsional rigidity are both 0.999, meeting the accuracy requirements. Based on the surrogate model, the adaptive simulated annealing algorithm is used for structural optimization, with the maximum torsional rigidity as the optimization objective and no increase in weight as the constraint. The constructed optimization model is shown in Eq. (3).

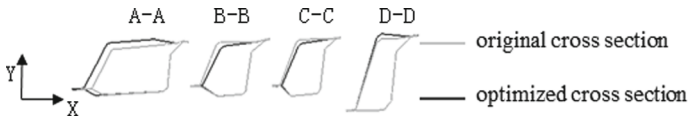


Fig. 4 Changes of cross-section

Table 2 Details of changes in cross-sectional dimension variables

Numbering of feature locations	X-direction change (mm)	Y-direction change (mm)
1	− 10	+ 15
2	− 20	0
3	+ 10	− 10
4	+ 5	0
5	+ 5	− 10
6	+ 5	0
7	0	+ 10
8	0	0

Table 3 Performance changes of cross-section optimization model

Performance	Initial model	Optimization model	Rate of change (%)
Cavity weight (kg)	12.20	12.18	− 0.16
Torsional stiffness of BIW (Nm/°)	–	–	+ 2.9

$$\begin{cases} findy(x_1, x_2, \ldots, x_i, \ldots, x_n) \\ \max C_t(y) \\ s.t.m(y) \le m_0 \end{cases} \tag{3}$$

In the equation: x_i is the i-th variable; n is the number of variables; C_t is the static torsional rigidity of the body-in-white, N·m/(°), m is the weight of the body-in-white structure, and m_0 is the weight of the original model, kg.

The comparison of the cross-section before and after optimization is shown in Fig. 4, and the variable changes are shown in Table 2. The changes in torsional rigidity and weight are shown in Table 3.

3 Topology Optimization and Part Splitting

Topology optimization refers to obtaining the optimal layout of materials within a given design area [7, 8], so as to achieve the best performance and weight ratio. This section focuses on topological optimization of the D-pillar cavity and obtaining

material layout. Based on the material retention, the original structure is broken and the inner and outer plates are split into parts. Compared to the overall change in material thickness of the structure, optimizing the material thickness after splitting the parts according to their importance is more targeted and yields better optimization results.

3.1 Topology Optimization

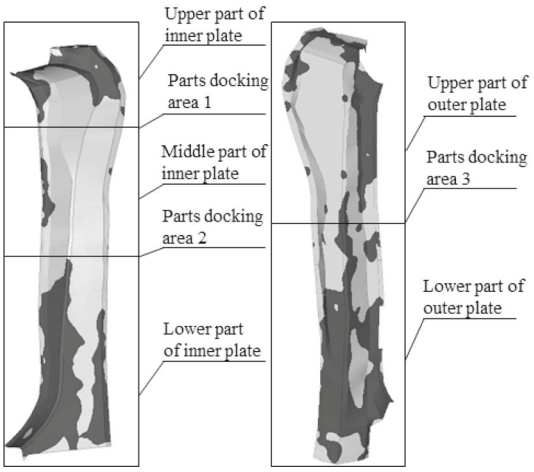
This thesis is based on the SIMP [9, 10] material density-rigidity interpolation model, and performs topology optimization on the optimized cross-section area structure to obtain the material distribution of the force-bearing structure within the design area. The topological optimization design area is the inner and outer plates of the D-pillar cavity after cross-section optimization. The relative density of the material is used as the design variable, and the lightest weight is the optimization goal. The mathematical model established with torsional rigidity as the constraint condition is shown in Eq. (4).

$$\begin{cases} findy(x_1, x_2, \ldots, x_i, \ldots, x_n) \\ \min m(y) \\ s.t.C_{t\min} \leq C_t(y) \leq C_{t\max} \\ 0 \leq x_{\min} \leq x_i \leq x_{\max} \leq 1 \end{cases} \tag{4}$$

In the equation: x_i is the i-th variable; n is the number of units within the design area.

3.2 Part Splitting

Based on the material retention in the topological optimization result cloud map, the parts are divided and designed. After the parts are divided, they are connected by laser welding and then formed by integral stamping. The cloud map of the internal plate topology optimization results is shown in Fig. 5a. The important materials are mainly concentrated at the upper and lower ends of the inner plate, so it is divided into 3 parts. The parts docking area 1 and 2 are respectively set at the locations where the material retention is relatively small on the upper and lower ends of the inner plate. Regarding the outer plate of the D-pillar cavity in this thesis, the topology optimization result cloud map is shown in Fig. 5b, and the important material retention positions are relatively scattered. Therefore, it is divided into two parts, and a docking area 3 is selected at a location with less material retention in the middle of the outer plate.



a) Inner plate of cavity b) Outer plate of cavity

Fig. 5 Topology optimization and part splitting results

3.3 Docking Position Optimization and Material Thickness Optimization

Implicit parameterization model variables include cross-section dimension variables, position variables, and material thickness variables, which can maintain the original topological relationship and geometric continuity even when various parameters change. It can achieve significant changes in cross-section dimension and position, and can simultaneously optimize cross-section position and material thickness. This section focuses on the comprehensive optimization of the docking position and material thickness of the split parts, and further explores the lightweighting scheme of the D-pillar cavity structure.

3.4 Variable Design

The docking area of the D-pillar cavity is designed as a variable in the Z direction of the vehicle coordinate system and the material thickness of the split parts.

The three docking areas shown in Fig. 5 are set with parameterization cross-sections that can vary along the baseline direction, ensuring that the distance between them and the adjacent cross-sections is greater than the variable range. The initial thickness of the parts is determined by the initial thickness values in each region. Among them, the initial thickness of 3 parts of the inner plate is 1.4 mm, and the initial thickness of 2 parts of the outer plate is 0.9 mm, as shown in Table 4, a total of 8 variables are designed.

Table 4 Range of values for material thickness and docking position variables

Number	Parameter name	Variable range(mm)
1	Thickness of the upper part of inner plate	(1.2, 1.4, 1.5, 1.6)
2	Thickness of the middle part of inner plate	(0.7, 0.9, 1.0, 1.2, 1.4, 1.5)
3	Thickness of the lower part of inner plate	(1.2, 1.4, 1.5, 1.6)
4	Thickness of the upper part of outer plate	(0.7, 0.8, 0.9, 1.0, 1.2)
5	Thickness of the lower part of outer plate	(0.7, 0.8, 0.9, 1.0, 1.2)
6	The Z-direction position change of No.1 docking area	(− 30, + 30)
7	The Z-direction position change of No.2 docking area	(− 30, + 30)
8	The Z-direction position change of No.3 docking area	(− 30, + 30)

3.5 Constructing Surrogate Model and Optimizing Results

The optimized Latin hypercube sampling algorithm is employed to generate 80 sample points, driving the parameterization model to generate a finite element model and calculate the torsional rigidity and weight. By fitting the response surface of variables and performance values, a second-order polynomial surrogate model is constructed. The weight and torsional rigidity R^2 values are 0.999 and 0.9852 respectively, both greater than 0.9, meeting the accuracy requirements.

Based on the surrogate model, the adaptive simulated annealing algorithm is used for structural optimization. The constraint condition is to achieve the required torsional rigidity, and the optimization objective is to minimize the structural weight. The constructed optimization model is shown in Eq. (5).

$$\begin{cases} findy(x_1, x_2, \ldots, x_i, \ldots, x_n) \\ \min m(y) \\ s.t.C_{t\min} \leq C_t(y) \end{cases} \tag{5}$$

In the equation: x_i is the i-th variable; n is the number of design variable.

The optimal solution from the optimized model is selected, and the changes of each variable are shown in Table 5. The optimal solution variable values are substituted into the parameterization model for scheme verification. The torsional rigidity and weight changes are shown in Table 6. This scheme can reduce the weight of the D-pillar cavity structure by 11.5% while meeting the torsional rigidity standard.

Table 5 Details of changes in material thickness and docking position variables

Number	Parameter name	Thickness or change value (mm)
1	Thickness of the upper part of inner plate	1.4
2	Thickness of the middle part of inner plate	0.9
3	Thickness of the lower part of inner plate	1.2
4	Thickness of the upper part of outer plate	0.9
5	Thickness of the lower part of outer plate	0.7
6	The Z-direction position change of No.1 docking area	+ 13
7	The Z-direction position change of No.2 docking area	− 6
8	The Z-direction position change of No.3 docking area	− 8

Table 6 Performance changes of optimization model

Performance	Initial model	Optimization model	Rate of change (%)
Cavity weight (kg)	12.2	10.8	− 11.5
Torsional stiffness of BIW (Nm/°)	–	–	+ 0.2

4 Performance Evaluation

The structural parts in the D-pillar area need to provide installation functions for parts such as the back door hinge, back door gas strut, and third-row seat belt retractor. The structural strength at the installation point is a key consideration for vehicle body durability simulation. The connection structure between the upper part of the D-pillar and the roof has an important impact on the first-order torsional mode of the vehicle body, and may also affect the safety roof crush condition. Therefore, it is necessary to evaluate whether the optimized solution meets the performance requirements for the first-order torsional mode, strength, and roof crush condition.

After load and simulation analysis verification of working conditions, the optimization scheme improves the first-order torsional mode of the body-in-white by 0.3%. The structure in the D-pillar area meets the requirements under 9 typical strength working conditions, and the displacement of the top pressure working condition loader meets the requirements. The relevant working conditions are shown in Table 7.

Table 7 Performance analysis results of optimization solutions

Working conditions	Evaluation criteria	Evaluation results
Turning	≤ Yield limit	Qualified
Reverse braking	≤ Yield limit	Qualified
Maximum braking	≤ Yield limit	Qualified
Maximum acceleration	≤ Yield limit	Qualified
Lateral impact	≤ Yield limit	Qualified
Forward handbrake	≤ Yield limit	Qualified
Backward handbrake	≤ Yield limit	Qualified
Road edge impact	≤ Yield limit	Qualified
Vertical 3G	≤ Yield limit	Qualified
First-order torsional mode	No decrease	Qualified
Top pressure condition loader displacement	≤ 127 mm	Qualified

5 Conclusions

This thesis focuses on the structural design of the D-pillar cavity of a certain car model, proposing a method for regional structure reconstruction and lightweighting design. Based on the implicit parameterization model of the body-in-white, this method combines cross-section dimension optimization, topology optimization, part splitting, docking position optimization, and material thickness optimization to design a lightweighting D-pillar cavity structure while ensuring the performance of the vehicle body. The optimized structure achieves a weight reduction ratio of 11%.

This thesis divides the inner and outer plates of the D-pillar cavity into parts based on the remaining amount of topology-optimized materials, providing more targeted optimization areas for subsequent thickness optimization. Comprehensive optimization is carried out for the docking position and material thickness of the split parts, further exploring the lightweighting design scheme of the D-pillar cavity structure. After modal, strength, and safety condition verification, the optimized scheme meets the performance requirements.

The lightweighting design method for the D-pillar cavity structure proposed in this thesis is applicable to other similar structures in the vehicle body and has a certain degree of scalability.

References

1. SAECCE (2021) Launch of technology roadmap for energy saving and new energy vehicles 2.0. China Machine Press, Beijing
2. WenJie Z, JiShun C, Yi Wen L et al (2017) Size optimization on plate thickness of BIW with constraints of stiffness, strength and frequency. Automot Eng 39(2):145–149

3. ZhenHu W, WanLin W, SongBo Z et al (2018) Size optimization on main cross-sections of body-in white based on conceptual model for car body. Automot Eng 40(8):904–911
4. YongKang T, YunXi G, Ping Z et al (2024) Sensitivity analysis and lightweight design of an automobile sliding door. J Mech Strength 1:122–128
5. DengFeng W, ShenHua L (2021) Lightweight optimization design of side collision safety parts for BIW based on Pareto mining. China Mech Eng 32(13):1584–1590
6. YanXing L, Tie W, ZhiYong J et al (2022) Multi-objective topology optimization design of hydrogen-fueled truck frame. J Mech Strength 44(5):1128–1133
7. Lee HA, Park GJ (2012) Topology optimization for structures with nonlinear behavior using the equivalent static loads method. J Mech Des 134(3):956–970
8. Patel NM, Kang BS, Renaud JE et al (2007) Crashworthiness design using topology optimization. J Mech Des 131(6):1–12
9. Guo F, Tang NJ, Brown IP (2022) Magneto-structural combined dimensional and topology optimization of interior permanent magnet synchronous machine rotors. IEEE Trans Ind Appl (6):7241–7250
10. Yao P, Lionel P, Yann LB et al (2023) SIMP-method topology optimization of ferrite structures in inductive power transfer systems. IEEE Trans Electromagn Compat 1:1744–1750

Open Access This chapter is licensed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License (http://creativecommons.org/licenses/by-nc-nd/4.0/), which permits any noncommercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if you modified the licensed material. You do not have permission under this license to share adapted material derived from this chapter or parts of it.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.



U-ATTENTION-NET: A U-shaped Structure for No-Reference Low-Light Image Enhancement Method



Zinong Zhang and Lei Han

Abstract Driver Monitoring System has become an integral part of modern vehicle safety systems. However, traditional monitoring systems face limitations in low-light conditions, especially at night or in weak light, making it challenging to accurately capture the driver's state. This paper proposes a U-shaped structure for no-reference low-light image enhancement, introducing channel and spatial attention mechanisms into the U-shaped structure. This lightweight neural network adjusts the brightness of individual pixels in images. Qualitative and quantitative experiments demonstrate that U-ATTENTION-NET achieves excellent low-light image enhancement results, with a significantly reduced number of parameters compared to traditional low-light image methods. U-ATTENTION-NET also boasts high inference speed, capable of meeting real-time frame-by-frame detection at 30 frames per second for camera applications.

Keywords U-shaped structure · Self-attention mechanism · No-reference learning · Unsupervised learning · Curve estimation

1 Introduction

During nighttime driving, drivers are more prone to fatigue due to circadian rhythms and the unique characteristics of nighttime environments. However, the poor lighting conditions during this period directly impact the performance of driver monitoring systems, significantly reducing their accuracy. In such low-light environments, monitoring systems struggle to effectively detect subtle signs of driver fatigue, leading to delayed detection and alerting of fatigue levels. This not only increases the risk to driving safety but also poses a potential threat to the health and safety of

Z. Zhang · L. Han (✉)
Beijing University of Posts and Telecommunications, Beijing, China
e-mail: hanl@bupt.edu.cn

Z. Zhang
e-mail: zzn1998@bupt.edu.cn

© The Author(s) 2026

H. Saman et al. (eds.), *The 9th International Conference on Advances in Construction Machinery and Vehicle Engineering*, Lecture Notes in Mechanical Engineering,
https://doi.org/10.1007/978-981-96-5527-4_69

the drivers. Low-light image enhancement techniques, by increasing image brightness, suppressing noise, and enhancing contrast, effectively improve the quality of low-light images. This enhancement results in clearer facial features of the driver, significantly boosting the accuracy and reliability of fatigue monitoring systems, thereby ensuring safer driving. Therefore, research on low-light image enhancement algorithms is of great importance.

Low-light image enhancement algorithms aim to improve visibility, perceptibility, and contrast for human vision, while for machine vision tasks, the goal is to enhance and enrich the brightness information contained in the image. Currently, low-light image enhancement algorithms are mainly divided into two categories. The first category includes traditional low-light image enhancement algorithms, such as Histogram Equalization [1], Gamma Correction [2], and the Retinex algorithm [3]. These methods use specific mathematical computations and theoretical assumptions to directly calculate the RGB pixel values of the image. However, these approaches have several limitations. Firstly, their fixed computational methods and theoretical assumptions perform poorly in different lighting conditions. Secondly, these methods only work well when the entire image is uniformly underexposed. They may struggle with images where local brightness and contrast vary significantly, leading to issues like overexposure and loss of image details. The second category comprises low-light image enhancement algorithms based on deep learning, which have emerged in recent years. Compared to traditional methods, these approaches offer higher accuracy, robustness, and speed [4]. Deep learning-based methods can be further divided into two subcategories based on the type of training data: reference-based methods and no-reference methods. In the reference-based methods, RetinexNet [5] is a typical example that applies the Retinex theory, using paired normal-light and low-light images to train its decomposition and adjustment networks. In the no-reference methods, EnlightenGAN [6] is a Generative Adversarial Network (GAN) trained with non-reference data. It consists of a generator that converts low-light images into enhanced high-light images, and a discriminator that evaluates the authenticity of the images, pushing the generator to produce more realistic outputs. Zero-DCE [7] is an unsupervised deep curve estimation model for low-light image enhancement, which first introduced the use of curve estimation to enhance low-light images. Zero-DCE++ [8] is an improved version of Zero-DCE that further optimizes the network architecture and training strategy, enhancing the image processing results. LUT-GCE [9] is another unsupervised deep curve estimation model for low-light image enhancement, proposing the use of a cubic global curve estimation for fast low-light enhancement. Additionally, LLE-NET and Gamma-NET [10] employ deep networks to estimate cubic curve correction parameters and Gamma correction parameters, combining curve-based enhancement strategies with traditional Gamma correction techniques to achieve rapid low-light image enhancement.

This paper considers that in tasks such as driver fatigue detection and object detection, low-light image enhancement algorithms serve as a preprocessing step for real-time visual tasks, requiring extremely high real-time performance. While using paired normal and low-light data for network training can achieve good low-light image enhancement results, it lacks flexibility in application, as constructing complex

paired datasets is necessary. Moreover, traditional GAN networks are difficult to train and often struggle to converge.

This paper proposes a low-light image enhancement algorithm called U-ATTENTION-NET. To avoid the training difficulties associated with traditional GANs and to improve the adaptability of the algorithm to different lighting environments, a zero-shot training method for low-light images is adopted. Secondly, to enhance the curve parameter estimation network's ability to extract brightness-related features from images, a U-shaped structure, consisting of an encoder and a decoder, is used. Spatial attention and channel attention mechanisms are also incorporated into this U-shaped structure. Finally, to reduce the computational load, parameter count, and forward inference time of the model, and to improve the efficiency of the curve parameter estimation network, depthwise separable convolutions are employed in both the encoder and decoder of the network.

2 U-ATTENTION-NET Low-Light Image Enhancement Method

2.1 Structure of the U-ATTENTION-NET

The structure of the U-ATTENTION-NET low-light image enhancement algorithm is shown in Fig. 1. The low-light image is input into the U-ATTENTION-BLOCK, which computes a set of curve estimation parameter maps. These computed curves are then used to iteratively enhance the input low-light image. This means that the input for each subsequent enhancement step is the result from the previous enhancement. Each enhancement step applies the curves to the image pixel by pixel.

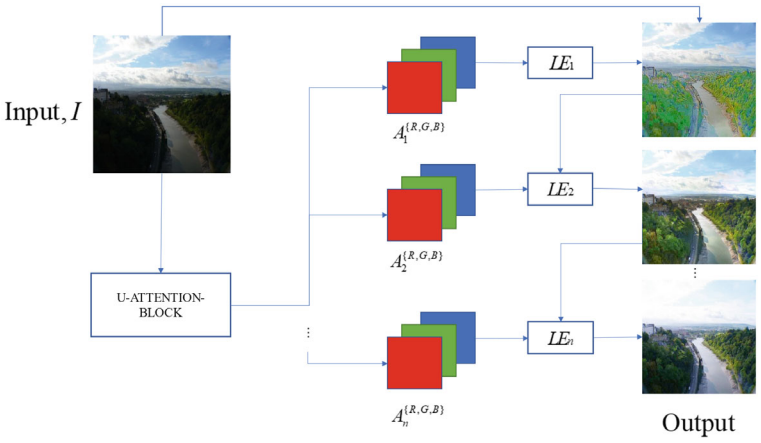


Fig. 1 U-ATTENTION-NET architecture diagram

2.2 LE Curve of the U-ATTENTION-NET

For each low-light image that needs enhancement, the curve is used to map each pixel in the image. The formula for the curve can be expressed as follows:

$$LE(I(x);\alpha) = I(x) + \alpha I(x)(1 - I(x)) \tag{1}$$

In this context, $I(x)$ represents the specific pixel value in each image, while α represents the value specified at the corresponding pixel position in the estimated parameter map. The formula above is for enhancing a single pixel in the image. For the entire low-light image, the overall *LE* curve formula can be expressed as follows:

$$LE(I(x);A\{R,G,B\}) = I(x) + A\{R,G,B\}I(x)(1 - I(x)) \tag{2}$$

In this context, $I(x)$ represents the RGB image, where for each pixel position in each channel, $A\{R, G, B\}$ contains the corresponding curve estimation value. For the iterative enhancement process of the entire image, the curve formula *LE* can be expressed as:

$$LEn(x) = LEn - 1(x) + AnLEn - 1(x)(1 - LEn - 1(x)) \tag{3}$$

$LEn(x)$ represents the result of the n th iteration of enhancement, $LEn - 1(x)$ represents the result of the $n - 1$ th iteration of enhancement, and *An* describes the curve parameter estimation map obtained by the network during the nth enhancement iteration, where the maximum value of n is 8.

2.3 Loss Function of the U-ATTENTION-NET

The loss function of the U-ATTENTION-NET low-light image enhancement algorithm is composed of four parts: *L*oss_*TV*, *L*oss_*SPA*, *L*oss_*COL* and *L*oss_*EXP*. The calculation formula is given by:

$$L\text{oss} = WTVL\text{oss_}TV + WSPAL\text{oss_}SPA + WCOLL\text{oss_}COL + WEXPL\text{oss_}EXP \tag{4}$$

where *WTV* is set to 1600, *WSPA* is set to 1, *WCOL* is set to 5, and *WEXP* is set to 10.

*L*oss_*TV* refers to the illumination smoothness loss, which aims to maintain the monotonic relationship between adjacent pixels. The calculation formula is:

$$L\text{oss_}TV = \frac{1}{N}\sum_{n=1}^{N}\sum_{c\in\xi}(\left|\nabla \text{x A}_n^c\right| + \left|\nabla y\ \text{A}_n^c\right|)^2, \xi = \{R,G,B\} \tag{5}$$

where N refers to the number of iterations, ∇x denotes the horizontal gradient, ∇y denotes the vertical gradient, and A_n^c represents the curve estimation parameter map obtained from the network. Here, n is the number of iterations, and c refers to the actual channel.

Loss_SPA refers to the spatial consistency loss, which aims to improve the spatial consistency between the generated image and the input image by maintaining the consistency of gradients in adjacent regions of the input and enhanced images. The calculation formula is:

$$Loss_SPA = \frac{1}{K}\sum_{n=1}^{K}\sum_{j\in\Omega(i)}(|(Yi - Yj| - |Ii - Ij|)^2 \tag{6}$$

where K denotes the number of local regions, $\Omega_{(i)}$ refers to the four adjacent regions centered around region i (including the top, bottom, left, and right), Y is the average intensity value of the local region in the low-light enhanced image, and I is the average intensity value of the local region in the original low-light input image.

Loss_COL refers to the color consistency loss, which is aimed at correcting potential color biases in the enhanced image and adjusting the relationships among the RGB channels. The calculation formula is:

$$Loss_COL = \sum\nolimits_{\forall(p,q)\in\varepsilon}(J^p - J^q)^2, \varepsilon = \{(R, G), (R, B), (G, B)\} \tag{7}$$

where J^p represents the average intensity value of the enhanced image in a specific channel, and ε refers to the set of pairwise combinations of the RGB channels.

Loss_EXP refers to the exposure control loss, which aims to prevent underexposed or overexposed areas and control the overall exposure level of the algorithm. The calculation formula is:

$$Loss_EXP = \frac{1}{M}\sum_{m=1}^{M}|Ym - E| \tag{8}$$

where M denotes the number of regions into which an image is segmented, Ym represents the average intensity value within each local region, and E is the average intensity value that an image should ideally have for proper exposure. U-ATTENTION-NET sets E to 0.6.

2.4 U-Attention-Block

U-ATTENTION-BLOCK is illustrated in Fig. 2. The overall network architecture employs a U-shaped structure. U-ATTENTION-BLOCK utilizes an encoder-decoder

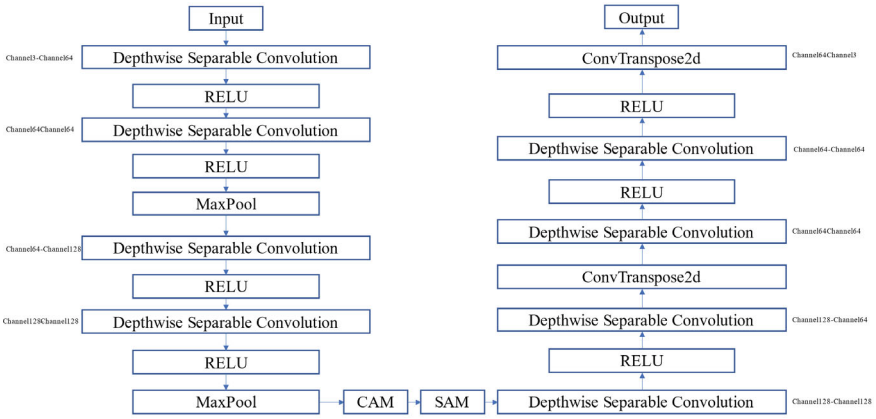


Fig. 2 U-ATTENTION-BLOCK

structure, enabling effective feature extraction and reconstruction. Depthwise separable convolutions are used in place of standard convolutions to reduce the number of network parameters. The network incorporates channel self-attention and spatial self-attention mechanisms in the middle part of the encoder-decoder. The channel self-attention mechanism enhances the network's feature representation capability, while the spatial self-attention mechanism improves the spatial detail information in the generated images, thereby enhancing the model's robustness.

3 Experiments

3.1 Experimental Details

The U-ATTENTION-NET system is configured with an Intel® Xeon® Silver 4116 CPU @ 2.10 GHz and a single NVIDIA GeForce RTX 2080 Ti GPU with 12 GB of memory. The CUDA version used is 11.6, PyTorch deep learning framework version is 1.13.1, and the Python version is 3.7.0.

The training dataset for U-ATTENTION-NET is selected from Zero-DCE [7], which consists of 2002 images. The batch size for training is set to 4, the learning rate is set to 0.0001, and the maximum number of epochs is set to 20.

In the experimental process, three classic low-light image enhancement algorithms—RetinexNet [5], EnlightenGAN [6], and Zero-DCE++ [8]—are compared qualitatively and quantitatively with the proposed U-ATTENTION-NET low-light image enhancement algorithm. The low-light datasets used for the experiments consist of four datasets: DICM [11] (44 images), LIME [12] (10 images), MEF [13] (79 images), and VV [14] (23 images). Since these datasets contain images of

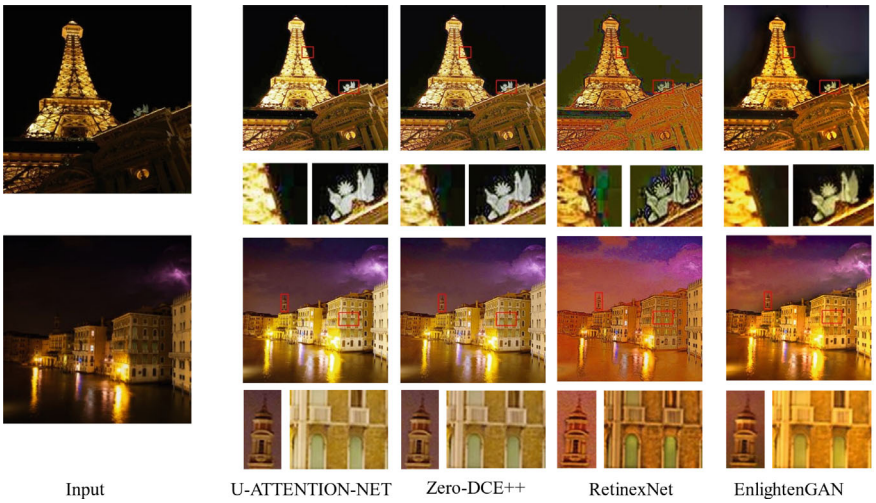


Fig. 3 Low-light image enhancement results of four methods

different sizes, to standardize the experiment, all low-light images are resized to a uniform size of 512 × 512 pixels using the LANCZOS [15] interpolation method prior to testing. LANCZOS interpolation significantly reduces the loss of image detail caused by resizing. The enhancement results of the four methods are shown in Fig. 3.

From the Fig. 3, it is evident that all four methods enhance the brightness of the original images. However, noticeable issues are observed: RetinexNet algorithm introduces artifacts and distortions at object edges, while EnlightenGAN algorithm results in less desirable color and darker outcomes.

3.2 Image Generation Evaluation Results

To quantitatively discuss the advantages of the U-ATTENTION-NET algorithm, four classic image generation quality evaluation metrics are used: Peak Signal-to-Noise Ratio (PSNR), Structural Similarity Index (SSIM) [16], Perceptual Image Quality Evaluator (PIQE) [17], and Natural Image Quality Evaluator (NIQE) [18]. Higher PSNR and SSIM values and lower PIQE and NIQE values indicate better image generation quality. The evaluation results of these four metrics are shown in Fig. 4.

From the Fig. 4, it can be seen that U-ATTENTION-NET performs exceptionally well in both PSNR and SSIM metrics, achieving the best results among the four algorithms. In the NIQE metric, U-ATTENTION-NET performs well, second only to Zero-DCE++. However, its performance on the PIQE metric is relatively average.

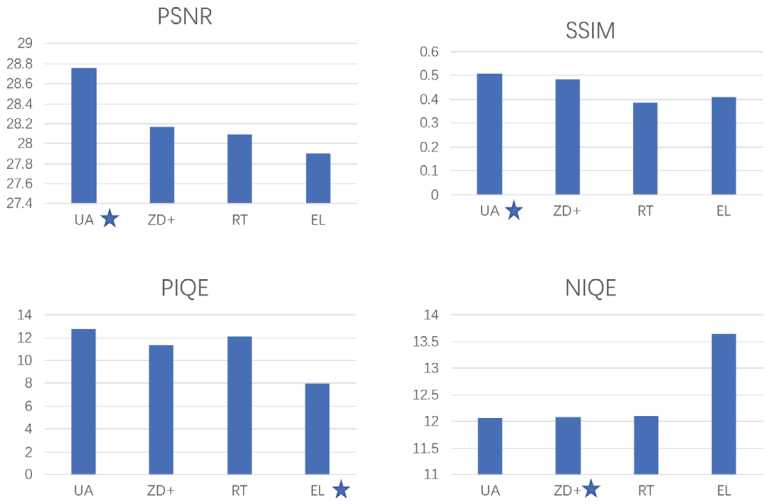


Fig. 4 Evaluation results of four metrics

3.3 Lightweight and Real-Time Performance Comparison

To verify the lightweight nature and real-time performance of the algorithm, we compare the number of parameters, floating-point operations (FLOPs), and forward inference time. Both FLOPs and forward inference time are calculated using (512, 512, 3) RGB images for forward inference. The comparison results are shown in Table 1:

From the table, it can be seen that U-ATTENTION-NET has a parameter count similar to that of Zero-DCE++, but due to the inclusion of attention mechanisms, its computational demand is higher. Compared to RetinexNet and EnlightenGAN, U-ATTENTION-NET has significantly fewer parameters and a noticeable reduction in forward computation time. The forward inference speed of U-ATTENTION-NET is 0.0056 s, which is a significant improvement over RetinexNet and EnlightenGAN. With typical camera frame rates of 30 frames per second, where each frame has a processing time of 0.033 s, U-ATTENTION-NET meets the image processing requirements for standard cameras.

Table 1 Comparison of lightweight characteristics for four low-light image enhancement algorithms

Method	Parameters (K)	GFLOPs	Time (s)
U-ATTENTION-NET	11.9	43.37	0.0056
RetinexNET	555	273	0.75
EnlightenGAN	8679	63.75	0.025
Zero-DCE++	10	2.64	0.0021

3.4 Discussion

To discuss the effectiveness of the attention mechanisms, we conducted three comparative experiments by omitting the channel self-attention mechanism, the spatial self-attention mechanism, and both mechanisms in the U-ATTENTION-BLOCK. The experimental results are shown in Table 2.

Adding self-attention mechanisms improves the PSNR and SSIM scores, but has a certain impact on the PIQE and NIQE results. Among the PSNR and SSIM metrics, U-ATTENTION-NET performs the best. For the PIQE metric, the best performance is achieved when no self-attention mechanisms are added. For the NIQE metric, the best performance is observed when the channel self-attention mechanism is not included (Fig. 5).

Table 2 Experimental results of self-attention mechanisms

Method	PSNR	SSIM	PIQE	NIQE
U-ATTENTION-NET	28.757	0.509	12.719	12.061
UA without Att	28.571	0.494	12.596	11.983
UA without Catt	28.671	0.497	12.623	11.926
UA without SAtt	28.625	0.495	12.643	11.985

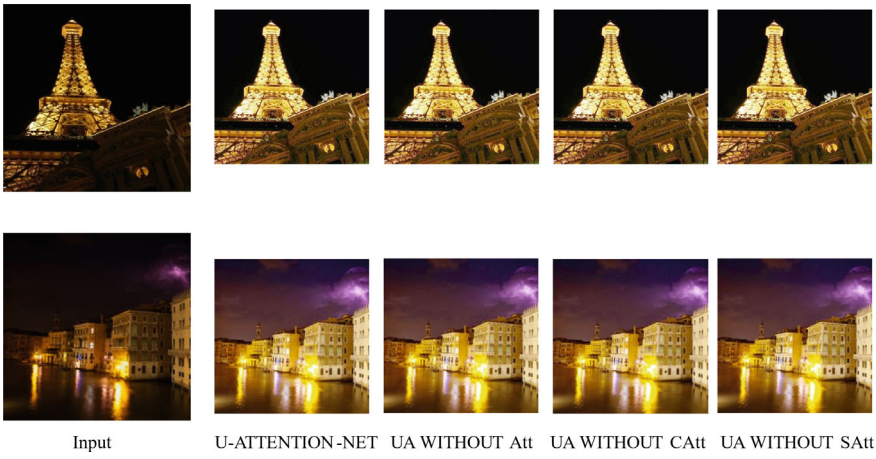


Fig. 5 . Enhancement results of four methods

3.5 Conclusion

In the field of low-light image enhancement, improving image quality while maintaining computational efficiency is one of the key challenges. This paper proposes a novel low-light image enhancement method, U-ATTENTION-NET, which constructs a U-shaped network structure and incorporates both channel self-attention and spatial self-attention mechanisms. U-ATTENTION-NET effectively captures the image's luminance information and spatial details by introducing a U-shaped network structure (including encoder and decoder) and self-attention mechanisms. The encoder and decoder use depthwise separable convolutions instead of standard convolutions, which reduces the model's parameters and computational load while maintaining efficient forward inference speed. Experimental results show that U-ATTENTION-NET performs exceptionally well on various quantitative evaluation metrics. As a lightweight neural network, U-ATTENTION-NET meets the requirements for real-time processing. U-ATTENTION-NET, with its lightweight design and efficient image enhancement performance, provides an effective solution for low-light image processing tasks. Its no-reference, real-time, and efficient characteristics make it highly promising for practical applications, particularly in scenarios requiring real-time processing, such as nighttime driver fatigue detection, autonomous driving, and nighttime surveillance.

References

1. Huang S, Ma K, Du J, Ma J (2021) Histogram equalization-based multi-scale contrast enhancement for image processing. IEEE Access 9:11572–11586
2. Jiang J, Huang S, Wang W, Zhao J (2021) Gamma correction-based deep learning framework for low-light image enhancement. IEEE Access 9:33445–33455
3. Chen Y, Chen C (2021) A novel adaptive multiscale retinex algorithm with color restoration for low-light image enhancement. IEEE Access 9:55232–55242. https://doi.org/10.1109/ACCESS.2021.3061002
4. Li C, Guo C, Han L, Jiang J, Cheng MM, Gu J, Loy CC (2021) Low-light image and video enhancement using deep learning: a survey. IEEE Trans Pattern Anal Mach Intell 2021(44):9396–9416
5. Chen W, Xu C, Koltun V, Wang Q (2018) Deep Retinex decomposition for low-light enhancement. In: Proceedings of the British machine vision conference (BMVC), pp 1–12
6. Dong J, Wang W, Pang Y, Wen S, Lu X (2019) EnlightenGAN: deep light enhancement without paired supervision. IEEE Trans Image Process 30:2340–2349
7. Li C, Guo C, Han L, Cong R (2020) Learning to enhance low-light image via zero-reference deep curve estimation. IEEE Trans Pattern Anal Mach Intell 44(5):2326–2340
8. Li C, Guo C, Han L, Cheng MM, Guo J (2021) Low-light image enhancement via breaking down the darkness. IEEE Trans Image Process 30:2636–2649
9. Wu C, Dong J, Tang J (2023) LUT-GCE: lookup table global curve estimation for fast low-light image enhancement. ArXiv preprint arXiv:2306.07588
10. Cao X, Yu J (2024) LLE-NET: a low-light image enhancement algorithm based on curve estimation. Mathematics 12(8):1228
11. Lee C, Lee C, Kim CS (2013) Contrast enhancement based on layered difference representation of 2D histograms. IEEE Trans Image Process 2013(22):5372–5384

12. Guo X, Li Y, Ling H (2017) LIME: low-light image enhancement via illumination map estimation. IEEE Trans Image Process 2017(26):982–993
13. Ma K, Zeng K, Wang Z (2018) Perceptual quality assessment for multi-exposure image fusion. IEEE Trans Image Process 24:3345–3356
14. Vonikakis V (2021) Busting image enhancement and tonemapping algorithms. Available online: https://sites.google.com/site/vonikakis/datasets (Accessed on 1 Jan 2021)
15. Liu S, Jiang K, Li K, Wang J (2023) Adaptive resolution enhancement for visual attention regions based on spatial interpolation. Sensors 23(14):6354
16. Wang Z, Bovik AC, Sheikh HR, Simoncelli EP (2004) Image quality assessment: from error visibility to structural similarity. IEEE Trans Image Process 13(4):600–612. https://doi.org/10.1109/TIP.2003.819861
17. Wang S, Wang Z, Bovik AC (2004) A new no-reference image quality metric based on perceptual information. IEEE Trans Image Process 13(4):546–558. https://doi.org/10.1109/TIP.2004.827165
18. Wang Z, Sheikh HR, Bovik AC (2012) No-reference quality assessment of natural images. IEEE Trans Image Process 21(8):3376–3387. https://doi.org/10.1109/TIP.2012.2188118

Open Access This chapter is licensed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License (http://creativecommons.org/licenses/by-nc-nd/4.0/), which permits any noncommercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if you modified the licensed material. You do not have permission under this license to share adapted material derived from this chapter or parts of it.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.



Research on Finite Element Modeling Method for Structural Strength Analysis of a C0-C0 Type Bogie Frame



Yuwen Huang, Xiaojuan Sun, Shijie Liu, Xinren Li, and Yu Wang

Abstract When performing strength analysis on a locomotive bogie frame, the frame structure is usually analyzed separately, and the constraints and loads are simplified or equivalent, which can lead to problems such as the inability to guarantee the accuracy of the results. This paper first establishes a complete bogie full finite element model (FEM) for frame strength analysis and compares two finite element modeling methods for frame strength calculation. Taking a C0-C0 wide-gauge locomotive as the research object, based on TB/T 3549.1-2019, JIS E 4207–2019 and ГОСТР 55,495-2013 standards, the working principles of the braking, traction and torsional loads are fully considered, and the supernormal and normal operation loading conditions are set. The static strength, fatigue strength and life of the frame are analysed. Results show that the maximum stress values of the full FEM under all calculation conditions are greater than those of the individual model of the frame, but the locations where they appear are different. In addition, from the fatigue life analysis, the service life obtained by the full FEM is two orders of magnitude smaller than that of the individual model of the frame. The research of this paper has reference significance for more accurate finite element modeling and provides a theoretical basis for frame strength design, further structural optimization and lightweight.

Keywords C0-C0 bogie · Frame · Finite element · Static strength · Fatigue strength · Fatigue life

Y. Huang · X. Sun (✉) · S. Liu
School of Mechanical Engineering, Taiyuan University of Science and Technology, Taiyuan, China
e-mail: sunxj@tyust.edu.cn

X. Li · Y. Wang
Technical Centre, CRRC Datong Electric Locomotive Co., LTD, Datong, China

© The Author(s) 2026

H. Saman et al. (eds.), *The 9th International Conference on Advances in Construction Machinery and Vehicle Engineering*, Lecture Notes in Mechanical Engineering,
https://doi.org/10.1007/978-981-96-5527-4_70

1 Introduction

As an important component of the locomotive, the bogie connects the car body at the top and the axle box and wheelset at the bottom. In order to ensure the stability and safety of the six-axle wide-gauge heavy-load locomotive, the structural fatigue strength analysis of the bogie frame must be carried out. When establishing a FEM, the choices of constraints and loading methods have great impact on the analysis results.

When analyzing the structural strength of the bogie frame, many studies choose to model the frame separately and perform reasonable constraints and loading according to the standard TB/T 3549.1-2019 [1]. The commonly used method is to simulate the primary suspension using spring element [2, 3] or displacement constraints. *Xiu* et al. [4] summarized the domestic and international strength tests and identification specifications for power bogie structures and found that braking and traction loads play an important role in predicting the fatigue life of the frame under heavy load conditions. In fact, torsional load also has a certain impact on frame fatigue [5]. Miao et al. [6] fully considered the effects of braking load, traction load and twisting load on the fatigue strength of the frame. Türe et al.[7] predicted the fatigue life of the weak position welding based on the main S–N curve method. Zhi [8] established a stress spectrum for fatigue reliability analysis and obtained the symmetrical cyclic equivalent stress based on Miner's law and fatigue damage equivalence principle. Dong [9] established a full FEM for a B0-B0 bogie, taking into account braking load, traction load, and twisting load, and performed strength analysis on the frame. However, when modeling, the wheelset was simulated using bar elements, and the twisting load was still loaded in the same way as the frame model alone.

This paper takes a certain type of six-axle wide-gauge heavy-duty locomotive as the research object, and establishes the structural strength finite element analysis model of the whole bogie and that of the individual frame, respectively. Based on Standards, TB/T 3549.1-2019 [1] and JIS E 4207-2019 [10], the study fully considers the effect of braking load, traction load and torsion load on fatigue strength, and sets up the abnormal loading and normal loading conditions. Moreover, two kinds of torsion loads are used for the two models, respectively. Fatigue strength analysis is carried out based on Goodman's assessment diagram. Fatigue life assessment is carried out using nCode Design-Life software to compare the effect of two FEMs on fatigue strength analysis. The research content of this paper provides guidance for more accurate finite element modeling and a theoretical basis for further structural optimization and lightweight of the frame.

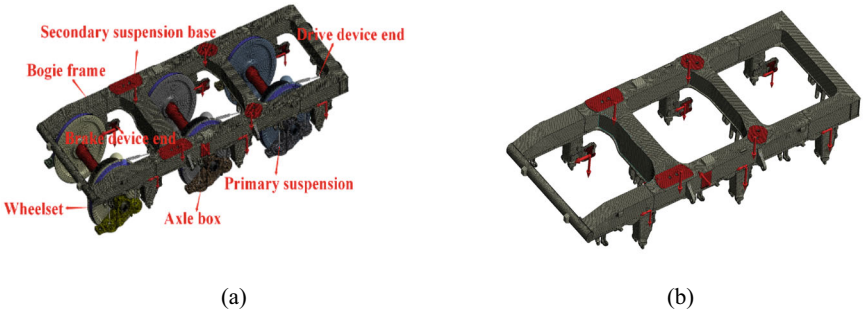
2 Finite Element Modelling

A single C0-C0 bogie of a locomotive is mainly composed of a frame, six axle boxes, three wheelsets, six primary suspension devices, several secondary suspension devices, braking devices and drive devices. The double-pull-rod axle box is connected to the frame by spherical bushing joints. The frame material is SM500MC. The structure and main parameters of the bogie studied in this paper can be found in reference [11].

2.1 Full FEM of the Bogie

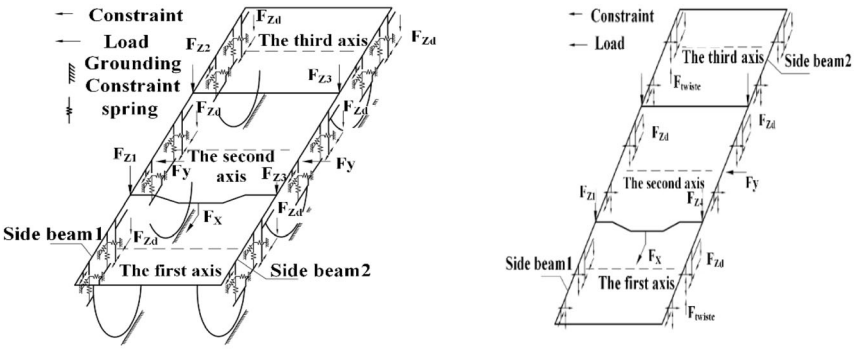
The model is meshed in Hypermesh software by a combination of automatic and manual meshing with a mesh size of 20 mm. Among them, the main part of the frame is divided by 4-node Solid187 elements, and the areas connected with the axle box, secondary suspension base, etc. are meshed by 5-node Solid185 elements. The secondary suspension base and the wheelset are divided by 8-node Solid185 elements. The axle box is divided by 4-node Solid187 element. The overall model includes a total of 1,434,241 elements and 609,053 nodes.

The bogie mesh model is imported into ANSYS workbench software to establish a full FEM, as shown in Fig. 1a. The steel spring of the primary suspension is simplified into a spring with three orthogonal stiffnesses, and is simulated using a vertical spring and two lateral ground springs. The axle-box pull rod joint is simulated using Bushing elements. The axial stiffness, radial stiffness, torsional stiffness, and bending stiffness are set according to the rubber joint parameters. The drive device is simplified as a rigid body, controlled by a mass point, with one end rigidly coupled to the wheel axle and the other end connected to the frame through a one-dimensional spring element; A revolute pair is established to connect the wheelset and the axle box to them. The constraints and load settings are shown in Fig. 2a.



(a) (b)

Fig. 1 Finite element modelling



(a) (b)

Fig. 2 Schematic diagram of constraint forces: F_Z—vertical load, F_y—lateral load, F_X—longitudinal load, F_{Zd}—vertical braking load, F_{Xd}—longitudinal braking load

2.2 FEM of the Frame

Based on the full FEM, the frame FEM is extracted separately, as shown in Fig. 1b. Using elastic constraints, a vertical elastic constraint is set under the primary suspension base, and the vertical stiffness is 0.13 Nmm^3. A lateral elastic constraint is set on one side of the primary suspension base, and the lateral stiffness is 0.80 Nmm^3. Longitudinal elastic constraints are applied to the bases of the first tie rod and the lower tie rod, respectively, and the longitudinal stiffness is 5.58 Nmm^3. The force loadings are set as the same as the full FEM. The constraint and load settings of the frame FEM are shown in Fig. 2b.

3 Static Strength Analysis

According to TB/T 3549.1-2019 [1], static strength analyses are mainly carried out under abnormal loading conditions and normal loading conditions.

3.1 Abnormal Loading Conditions

According to UIC 615-4-2003, five types of abnormal load conditions include vertical load condition (I), vertical load + lateral load condition (II), torsional load condition (III), no-load condition of one wheel (IV) and vertical load + braking load condition (V). According to the results of the static strength analysis, the locations where the maximum stresses occur are listed in Table 1.

Table 1 Comparison of maximum stress results for abnormal loading conditions

Working conditions	Full FEM of the bogie		FEM of the frame	
	Maximum stress/ Mpa	Locations	Maximum stress/ Mpa	Locations
I	174.5	Reinforcement of the tie rod seat on the left axle box of the second axle	131.8	Junction of the right primary suspension seat of the first axis and the upper plate of the tie rod seat
II	218	Bolt joint on the tie rod seat on the left axle box of the second axle	142.5	Junction of the right primary suspension seat of the first axis and the upper plate of the tie rod seat
III	160	Reinforcement of the tie rod seat on the left axle box of the first axle	133.06	Junction of the right primary suspension seat of the first axis and the upper plate of the tie rod seat
IV	192.6	Bolt joint on the tie rod seat on the left axle box of the first axle	144	Primary suspension seat of the second axle
V	208.8	Reinforcement of the tie rod seat on the right axle box of the second axle	180.5	Upper tie rod support seat on the axle box of the third axle

By comparing the maximum stress values obtained from the two models, firstly, it can be found that the maximum stress values of the full FEM are always greater than those of the frame FEM. Therefore, when calculating the static strength of the frame, using a FEM of the frame will result in a smaller maximum stress value. Secondly, the maximum stress areas of the two models under the five working conditions are not in the same positions. The maximum stresses of the full FEM are mainly concentrated on the reinforcements of the upper tie rod seat on the left side of the first and second axes, and the tie rod fixing positions. The maximum stresses of the FEM of the frame are mainly concentrated at the junction of the right primary suspension seat of the first axis and the upper plate of the tie rod seat, and at the primary suspension seat of the second axis.

4 Normal Loading Conditions

The results obtained under normal loading conditions are similar to those under abnormal loading conditions. Due to space limitation, only the maximum stress values and locations under 6 conditions are given, which are listed in Table 2.

4.1 Result Verification Between FEM and Test

A total of 100 measuring points were selected and arranged according to the actual loading conditions for static stress tests of the bogie frame. The test obtained the stress values on abnormal loading conditions and normal loading conditions, respectively. As shown in Fig. 3, the relative errors of the stress values of the two results under

Table 2 Comparison of maximum stress results for normal loading conditions

Working conditions	Full FEM of the bogie		FEM of the frame	
	Maximum stress/ Mpa	Locations	Maximum stress/ Mpa	Locations
I	122.1	Bolt joint on the tie rod seat on the right axle box of the first axle	94.1	Junction of the right primary suspension seat of the first axis and the upper plate of the tie rod seat
II	150.3	Bolt joint on the tie rod seat on the left axle box of the first axle	137.1	Reinforcement of the tie rod seat on the third axle box
III	176	Bolt joint on the tie rod seat on the left axle box of the first axle	143.1	Reinforcement of the tie rod seat on the third axle box
IV	187.1	Bolt joint on the tie rod seat on the left axle box of the first axle	172.2	Reinforcement of the tie rod seat on the third axle box
V	228.6	Bolt joint on the tie rod seat on the left axle box of the first axle	178.02	Reinforcement of the tie rod seat on the third axle box
VI	134.5	Bolt joint on the tie rod seat on the right axle box of the first axle	118.08	Reinforcement of the tie rod seat on the third axle box

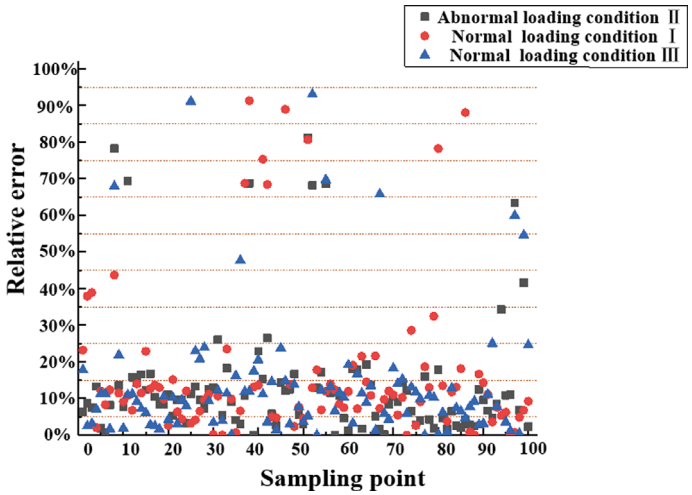


Fig. 3 Comparison of the FEM and experimental results

various working conditions are mainly below 25%. This is an acceptable result. It shows that the FEM of the frame established in this paper has sufficient precision.

5 Fatigue Analysis

5.1 Stress Fatigue Assessment

The S–N curve method is used to evaluate the fatigue strength of the frame stress obtained by the two models on normal loading conditions based on the Goodman diagram. As shown in Fig. 4, nodes in the stress concentration areas are selected at 15 locations, and the maximum and minimum stress values a_{max} and a_{min} of these nodes are obtained. The average stress a_m and dynamic stress amplitude a_s are calculated. The average stress and stress amplitude results are plotted in the Goodman diagram as shown in Fig. 5. Although the distribution of the two models in the evaluation diagram is different, they are all within the boundary line of the evaluation diagram, and the structure has good anti-fatigue characteristics.

5.2 Fatigue Life Assessment

According to engineering practices, the bogie frame belongs to the high-cycle fatigue analysis type based on the S–N curve. The fatigue life assessment is carried out based

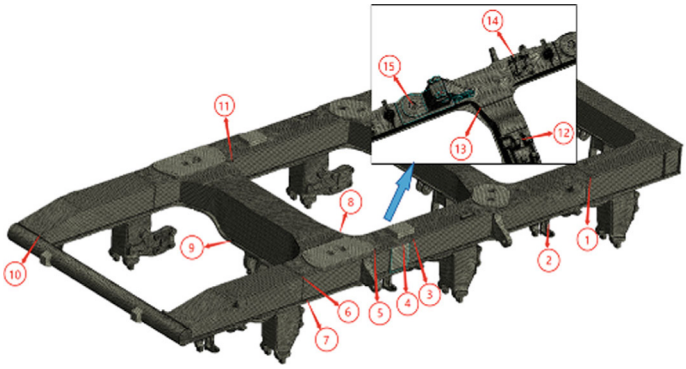


Fig. 4 Node selection location diagram

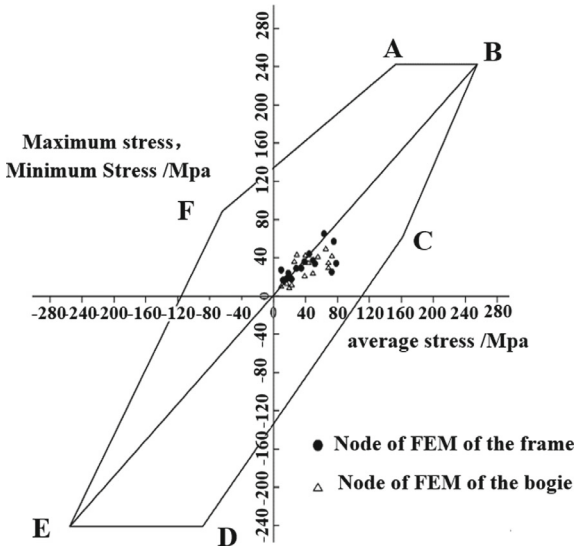


Fig. 5 Fatigue strength assessment diagram

on the nCode Design-life software. Results show that the minimum fatigue life of the frame using the full FEM is 1.79×10^7 times, and that using the frame FEM is 2.59×10^9 times. The weak areas of fatigue life of the two models are different. The weak areas in the full FEM are at the bends of the upper and lower side beams on both sides, as well as the primary suspension tie rod seat, etc. The weak areas in the frame model are mainly distributed at the bends of the upper and lower side beams on both sides, as well as the junctions of the secondary suspension base and the side beams. In practice, after the frame is subjected to the fatigue cycle test for 1×10^7 times no fatigue cracks appear. It can be indicated that the frame meets the fatigue strength requirements according to standards.

6 Conclusion

The paper developed two different modeling methods to study the static strength and fatigue strength of the bogie frame. The full FEM of the bogie and the FEM of the frame structure are established using finite element software, respectively. Based on the relevant railway standards, the two models are loaded and analyzed according to the abnormal loading conditions and normal loading conditions. The conclusions drawn are as follows: The accuracy of the frame model was verified by comparing it with the test results of the frame. The maximum stress value of the full FEM of the bogie is always greater than that of the frame model. The stress distribution of the two models under the same working conditions is quite different. The service life of the full FEM of the bogie is less than that of the frame FEM, but both are greater than that according to standards. The study in this paper aims to provide a more accurate finite element modeling approach and give a theoretical basis for further structural optimization and lightweight design of the frame. However, the nonlinearity of the primary and secondary suspensions is not fully considered in the FEMs, which is a possible extension of this work in the future.

Acknowledgements The authors would like to acknowledge the Key Science and Technology Project of Shanxi Province of China (No. 20181102002) and the National Natural Science Foundation of China (No. 52202478).

References

1. TB/T 3549.1-2019 Strength design and test accreditation specification for rolling stock. General
2. Sun H and Dai H (2023) Static strength analysis of the bogie frame of an EMU. Mach Build Automatio 52:96–99+103 (in Chinese). https://doi.org/10.19344/j.cnki.issn1671-5276.2023.05.022
3. Zhang J, MA J, Zhu Y (2021) Simulation analysis on strength and structure optimization of SKMB-200 bogie frame. Experimental Technol Manage 38:140–143+148 (in Chinese). https://doi.org/10.16791/j.cnki.sjg.2021.02.030
4. Xiu R, Spiryagin M, Wu Q, Yang S, Liu Y (2021) Fatigue life prediction for locomotive bogieframes using virtual prototype technique. Proc Inst Mech Eng J Rail Rapid Transit. F. 235:1122–1131. https://doi.org/10.1177/0954409720986
5. Xiu R, Spiryagin M, Wu Q, Yang S, Liu Y (2020) Fatigue life assessment methods for railway vehicle bogie frames. Eng Fail Anal. https://doi.org/10.1016/j.engfailanal.2020.104725
6. Miao C, Wang H, Song Y (2023) The correlation between static strength and fatigue strength test and simulation. J Phys Conf Ser (1). https://doi.org/10.1088/1742-6596/2660/1/012042
7. Türe S, Mugan A Master S-N (2024) curve fatigue life prediction of a railroad car bogie based on mesh insensitive structural stress. Method Eng Innovat 15–24. https://doi.org/10.4028/p-n8KYdv
8. Zhi P, Chen B, Li Y, Wang Y (2023) Fatigue reliability analysis of bogie frame based on track test. Mach Des Manuf 05:215–220+225 (in Chinese). https://doi.org/10.19356/j.cnki.1001-3997.20230228.004
9. Dong S (2021) Strength check and structure optimization of the bogie frame of permanent magnet direct drive electric. locomotive Railway Locomotive Car. 41:62–66+70 (in Chinese). https://doi.org/10.3969/j.issn.1008-7842.2021.01.12

10. JIS E 4207—2019. Rolling stock—Bogie— General rules for design of bogie frame strength.
11. Li H, Zhang W and Zhao M (2017) China-belarus freight II six-axle wide gauge electric locomotive bogie. Electr Drive Locomot 256:34–37 (in Chinese). https://doi.org/10.13890/j.issn.1000-128x.2017.03.102

Open Access This chapter is licensed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License (http://creativecommons.org/licenses/by-nc-nd/4.0/), which permits any noncommercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if you modified the licensed material. You do not have permission under this license to share adapted material derived from this chapter or parts of it.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.



Medium-Sized Truck Brake Design



Wang Tiantian and Qichen Zhong

Abstract In truck systems, braking is particularly critical for the safety performance of the vehicle. In order to improve the braking performance to ensure the safety of heavy vehicles, this study focuses on the optimization of front wheel brakes in medium trucks. Choose the disc brake, for its key parameters such as friction coefficient, braking force distribution coefficient, maximum torque, shaft load optimization calculation, completed the parts, including brake disc, brake clamp drawings and three-dimensional model, using ANS YS to part parts simulation analysis, to complete the medium truck brake brake performance improvement, to ensure the safety of the vehicle. It provides a solution for the current brake performance improvement, and also provides a reference for the future design of the truck brake system.

Keyword Medium-sized truck · Disc brake · Optimum design

1 Introduction

With the rapid development of transportation industry, the safety and reliability of medium truck brake system are getting more and more taken seriously. As the core component of the automobile safety system, the design of the brake directly affects the driving safety of the car. In this study, the design of the middle card brake is deeply discussed, and the disc brake is selected. By optimizing the brake structure, selecting high performance friction materials and reasonably distributing braking force and other key parameters, it aims to propose a brake design scheme with high braking efficiency and good stability and reliability. In addition, the present research results can also provide a reference for the brake design of other types of vehicles, and promote the continuous progress and development of automobile

W. Tiantian (✉) · Q. Zhong
School of Mechanical Engineering, Xijing University, Xi'an, China
e-mail: 1195114413@qq.com

© The Author(s) 2026

H. Saman et al. (eds.), *The 9th International Conference on Advances in Construction Machinery and Vehicle Engineering*, Lecture Notes in Mechanical Engineering,
https://doi.org/10.1007/978-981-96-5527-4_71

braking technology. Therefore, the study of medium card brake design has important practical significance and theoretical value.

2 Main Parameters of Brakes and Their Selection

2.1 Trucks Determination of the Brake Type

Compared to the drum brakes, the disc brakes show a significant advantage in the middle card design [1]. Its excellent heat dissipation performance effectively prevents the braking system from overheating, and ensures the stability and durability of the braking efficiency. At the same time, the disc brake responds quickly, can produce enough braking force in a short time, shorten the braking distance, improve the driving safety. In addition, its structure is simple, low maintenance cost, replacement and repair is more convenient. Therefore, for the frequent braking and high-speed driving card, the disc brake can better meet the high standards of braking performance and safety requirements.

2.2 Basic Parameters of Medium-Duty Trucks

The selected model is a medium-duty truck, which serves as a common reference for commercial vehicles on the market. Its specific data is shown in Table 1.

Effective Radius of Tires

$$D = 2B \times \frac{H}{B} + 25.4r = 1011.5\,\text{mm} \tag{1}$$

$$re = \frac{D}{2} + B \times \frac{H}{B} = 726\,\text{mm} \tag{2}$$

Table 1 Basic parameters of the vehicle

Dimensions	9000/2550/3600 mm	Vehicle weight	74550 kg
Axle distance	3300 mm/2000 mm	Gross weight	17550 kg
Max speed	89 km/h	Unloaded CG height	1500 mm
Torque	1010 N m	Loaded CG height	1200 mm
Power	195 kw		

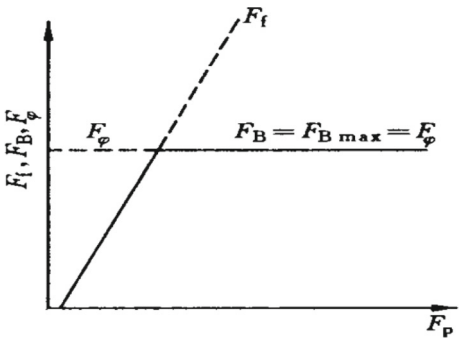


Fig. 1 Analysis of the braking process of vehicles

2.3 Determination of the Coefficient of Friction

Analyze the braking process of the vehicle on various surface conditions in Fig. 1, using a specific truck to determine the coefficient of friction.

Wheel lock-up during braking is critical for safety [2]. Front wheels usually lock up first, losing steering but maintaining stability, while rear wheel lock-up increases accident risk. This design sets the unloaded friction coefficient at 0.6 and loaded at 0.8 for effective braking. Analyzing these parameters improves performance and safety, ensuring safe truck operation and informing future enhancements.

2.4 Braking Force Distribution Coefficient Under Full Load

The normal reaction forces on the front and rear wheels are as follows: The braking force distribution coefficient [[3]] is the ratio of the front brake force to the total braking force:

$$\beta = \frac{\mathrm{F}\mu 1}{\mathrm{F}\mu} \tag{3}$$

$$\frac{F\mu 1}{F\mu 2} = \frac{\beta}{1-\beta} \tag{4}$$

The braking force distribution coefficient under full load can be determined as follows:

$$\beta = \frac{\phi 0 hg}{L} + \frac{b}{L} = 0.56 \tag{5}$$

2.5 Determination of Maximum Braking Torque Under Full Load on Different Road Surfaces

Referencing Table 2 for the coefficients of friction on different surfaces, a value of 0.85 is used for good road conditions, while 0.75 is applied for poorer surfaces. The following discussion will address these two surface conditions separately.

The maximum braking torque of the front and rear brakes is as follows:

$$T_{\mu 2\,max} = F_{xb2\,max} \times r \tag{6}$$

$$T\mu 2mzx = Fxb2\,max \times r \tag{7}$$

Substituting the vehicle parameters gives the maximum braking torque for the front and rear wheels:

$$T\mu 1\,max = \frac{G}{L}(b + Zhg)\phi r = 50817.59\,N \cdot m \tag{8}$$

$$T\mu 2\,max = \frac{G}{L}(a - Zhg)\phi r = 42830.96\,N \cdot m \tag{9}$$

Similarly, the maximum braking torque under poor road conditions can be obtained as:

$$T\mu 2\,max = 49262.67\ N \cdot m \tag{10}$$

$$T\mu 1\,max = \frac{\beta}{1-\beta} T\mu 2\,max = 62697.94\,N \cdot m \tag{11}$$

The axle load distribution when fully loaded is as follows.

Front axle load/Rear axle load:

$$G1 = \frac{b}{L}G = 6622\,kg \tag{12}$$

$$G2 = \frac{a}{L}G = 10927\,kg \tag{13}$$

Table 2 Coefficients of friction for different road surfaces

Surface	Peak coefficient	Sliding coefficient
Dry asphalt	0.8–0.9	0.75
Wet concrete	0.8	0.7
Dry dirt road	0.68	0.65
Ice	0.1	0.07

3 Design Calculations for Disc Brake Structure

3.1 Brake Disc Diameter

The brake disc diameter is typically 70–79% of the rim diameter [4]. For heavier vehicles, a 79% ratio is used to increase the effective radius, reducing caliper force and extending component life. The brake disc diameter can be calculated as:

$$D = 79\%Dr = 451.5\,\text{mm} \tag{14}$$

3.2 Brake Disc Thickness

Brake disc thickness is critical for performance and typically around 32 mm, with a wear limit of 29 mm. Thinner discs improve agility for city driving, while thicker ones are better for highway use [5]. For this design, a vented brake disc with a thickness of 38 mm is selected for the truck's environment.

3.3 The Inner and Outer Radii of the Friction Pad Are

The outer to inner radius ratio of the friction pad should not exceed 1.5. Consistent speeds at the edges during braking ensure uniform wear and maintain torque [6]. With a brake disc radius of 225.75 mm, the outer diameter is set at 220 mm, yielding a ratio of 1.4.Substituting the values gives an inner radius of 157.1 mm.

3.4 The Working Area of the Friction Pad is

The working area of the friction pad is essential for braking performance. A larger area increases friction force, reduces braking distance, and improves response time while enhancing thermal stability by dissipating heat [7]. It also lowers pressure and temperature, extending the system's lifespan. For disc brakes, the area is designed for weights between 1.6 and 3.5 kg/cm^2. For the selected truck at 17,550 kg, the friction pad is a sector with a 60°central angle.Therefore, the working area of the friction pad is:

$$\text{Re} = \frac{2\left(\text{R}_2^2 - \text{R}_1^2\right)}{3\left(\text{R}_2^2 - \text{R}_1^2\right)} = 191.5\,\text{mm} \tag{15}$$

After calculation, the average and effective radius values are similar, meeting design requirements. Thus, an average radius of 191.5 mm will be selected for the design.

3.5 Calculation of Friction Pad Wear Characteristics

Temperature, friction force, and brake disc materials affect friction pad wear, though theoretical calculations can't accurately determine it. Key influences are friction force and temperature. Intense braking generates heat, increasing pad wear. The specific energy dissipation rate measures energy dissipated per unit time per unit area to evaluate this load. The specific energy dissipation rate for the brake can be expressed as:

$$e2 = \frac{\delta m_a \left(V_1^2 - V_2^2\right)}{6tA_2}(1 - \beta) \tag{16}$$

During an emergency stop, the speed drops to 0 m/s. At this point, the specific energy dissipation rates for the front and rear brakes are as follows:

$$e1 = \frac{maV^2}{6tA1}\beta = 5.9\,w/mm^2 \tag{17}$$

The specific energy dissipation rate is 5.9 W/mm^2, below the maximum of 6.2 W/mm^2. This rate ensures proper friction pad wear, while higher rates may crack the brake disc.

4 Selection and Analysis of Main Components of Disc Brakes

4.1 Determination of the Brake Disc

The brake disc is vital for vehicle deceleration, using hydraulic pressure to create friction with the pads. Key design factors include strength, stiffness, and heat dissipation through internally ventilated discs. Gray cast iron (HT250 or higher) is selected for its performance and cost-effectiveness. Specifications include a flatness tolerance of 0.012 mm, surface roughness of 0.7–1.3, and parallelism tolerance under 0.05 mm. Its structure is illustrated in Fig. 2.

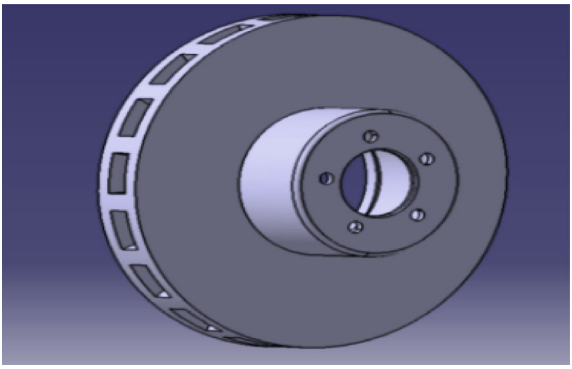


Fig. 2 Structure of the brake disc

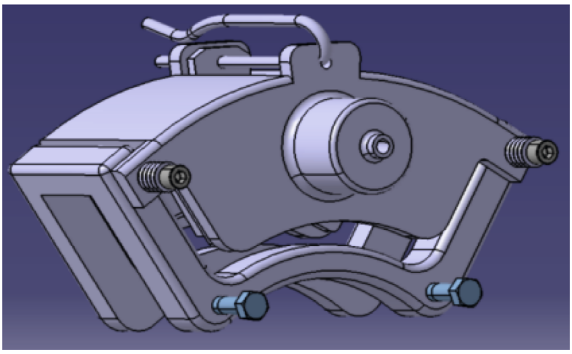


Fig. 3 Structure of the brake caliper

4.2 Determination of the Brake Caliper

The brake caliper is vital for vehicle braking performance and safety, pressing brake pads against the disc using high-pressure fluid. Materials should be strong, wear-resistant, and corrosion-resistant. This design uses ductile cast iron QT550-6 for the front axle, chosen for its low cost, durability, and ease of manufacture. Its structure is depicted in Fig. 3.

4.3 Determination of the Brake Pads

Brake pad includes a friction block and backing plate, slowing or stopping wheels by creating friction with brake disc. High braking pressure generates heat, requiring thicker friction block. Thickness is 15 mm, within typical ranges for cars and trucks. See Fig. 4 for structural dimensions. Friction materials are vital for vehicle safety, requiring high, stable friction, thermal stability, wear resistance, and low noise. Types include non-asbestos organic, semi-metallic, and ceramic. Friction coefficient ranges

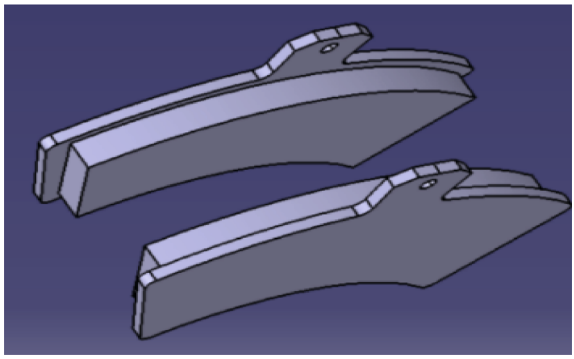


Fig. 4 Structure of the brake pad

0.3–0.5. Powder metallurgy materials are chosen, reducing failures and achieving a 0.4 friction coefficient.

4.4 3D Modeling of Disc Brakes

After conducting the aforementioned thorough and meticulous analysis process, we conducted a detailed examination and optimization of every critical aspect of the brake components. This step encompassed material selection and structural strength assessment. On this foundation, we embarked on the construction of the overall brake model using CATIA. By strictly adhering to the previous design concepts and parameter settings, we precisely assembled all the components according to the design requirements, ensuring the completeness and accuracy of the model. The model is illustrated in Fig. 5.

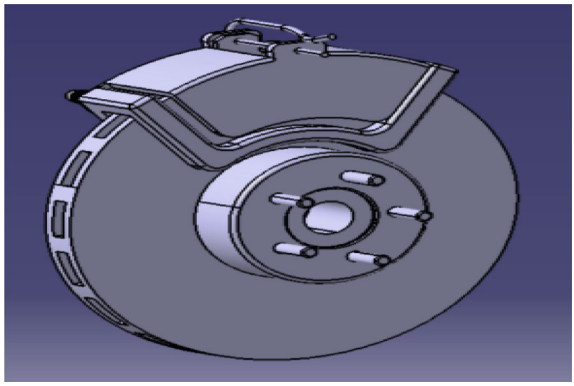


Fig. 5 Structure of the brake system

5 Finite Element Analysis of Main Components of the Brake System

The strength and stiffness of disc brake components are essential for reliability and safety. To evaluate the design, finite element analysis (FEA) will be performed on the brake disc and friction block using static analysis [8]. Models from CATIA will be imported into ABAQUS, where material parameters from Table 3 will be defined, and a mesh system created. After meshing, static analysis will assess loads and displacements, leading to the generation of stress and displacement contour plots for both components.

5.1 Static Analysis of Brake Discs

During the braking process, the brake disc not only has to endure the immense pressure from the brake caliper but also ensure that this pressure is evenly and effectively converted into braking force, enabling the vehicle to decelerate or stop smoothly and safely [9]. As clearly shown in the stress contour Fig. 6, the maximum braking torque endured by the brake disc during the braking process is 50817.59 N.m. Under this extreme operating condition, the peak stress value generated on the surface of the brake disc is 148 MPa. Through comparison, it is found that this stress level is significantly lower than the tensile strength of the material chosen for the brake disc (typically, high-quality brake disc materials possess tensile strengths far exceeding this value).

This result confirms the brake disc's rational design and suitable material choice, ensuring it meets braking strength needs. It maintains stability and safety under high braking forces, avoiding deformation or damage. Thus, the brake disc fully meets the braking system's stringent strength and durability requirements.

5.2 Static Analysis of Friction Linings

One surface of the friction lining directly receives the pushing force exerted by the piston, while its other surface is tightly in contact with and pressed against the brake disc. We use these conditions as the boundary conditions for mechanical analysis. By

Table 3 Main material parameters

Material	Elastic modulus	Poisson's ratio	Density (g/cm^3)	Min. Tensile strength (MPa)
Gray cast iron	1.2*1011 Pa	0.3	7.0	200
Powder metallurgy	1.0*1011 Pa	0.3	7.2	200–500

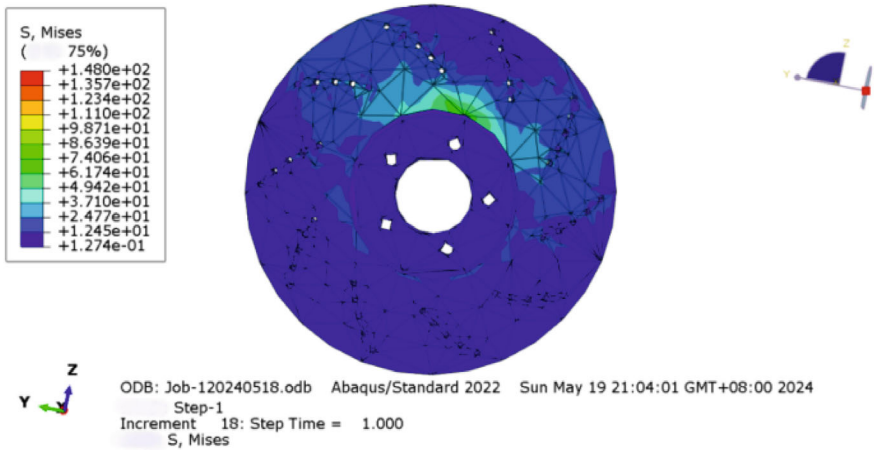


Fig. 6 Displacement contour map

carefully examining the stress contour plot of the friction lining shown in Fig. 7, we can clearly observe that the stress value generated on a single-sided friction lining is 2.13 MPa. This stress level is significantly lower than the tensile strength of the material used for the friction lining. When subjected to the stresses generated during braking, the friction lining fully meets the material's strength requirements, without undergoing deformation or damage.

In summary, the design of this brake friction lining is rational and safe, fully capable of meeting the stringent requirements for strength and durability imposed by the braking system.

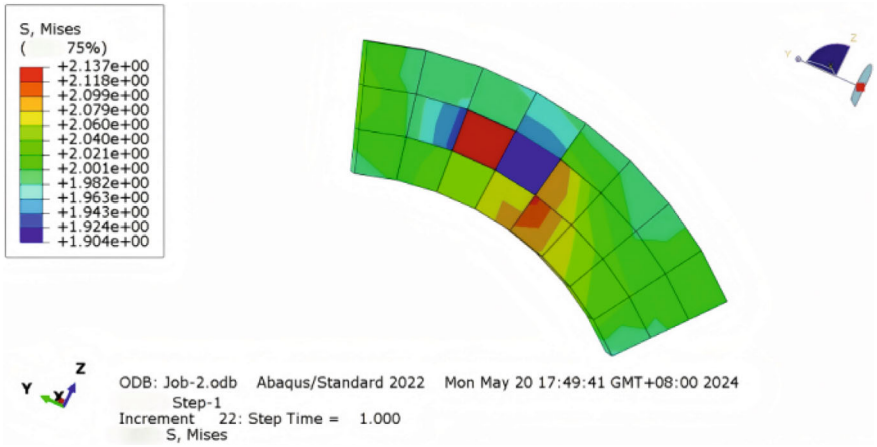


Fig. 7 Stress contour map

6 Conclusions

Based on the design of Chinese disc brake, the theoretical calculation, material selection and simulation optimization.

(1) Based on the braking principle and the use requirements of the card, the mathematical model is established. The key parameters of brake are calculated, including brake disc diameter, friction plate thickness and clamforce of brake clamp. These parameters provide the basis and guidance for subsequent material selection and simulation optimization.
(2) Consider the harsh conditions such as high temperature, high pressure and wear of the brake during work. Finally, HT 250 and other materials with excellent high temperature resistance and wear resistance were selected for the manufacture of brake disc and friction plate. The choice of these materials not only significantly improves the durability of the brakes, but also further improves their braking efficiency, ensuring the stability and reliability of the brakes in harsh working conditions.
(3) The simulation optimization uses the simulation software. By adjusting the structural parameters and material properties of the parts, the braking performance of the brake is optimized. The simulation results show that the optimized brake basically meets the high standard requirements of the middle card for the braking performance.

References

1. Yong X, Tao L (2018) Application of disc brake in Chongqing bus. Urban Public Transp 06:25–26
2. Lu D, Zhang W, Chen M (2020) Automobile brake system design and analysis. Mechanical Industry Press, Beijing
3. Bo D, Guangpo L, Guojun W et al (2023) Matching design of braking force distribution coefficient based on the statistical characteristics of braking strength. Sci and Technol Econom Guide 31(04):53–67
4. Lv X, Yin Q, Qu J et al (2024) Optimization design and research on processing technology of special structure brake discs. Light Ind Sci Technol 40(04):70–72
5. Yuan H (2024) Design of disc brake for small electric vehicles. Internal Combust Engine Accessor 13:67–69
6. Cai R (2019) Research on friction and wear performance and structural optimization design of disc brake friction linings. Jiangsu University.
7. Fieldhouse D, Ashraf N, Talbot C (2012) Measurement of dynamic pressure centers between disc and friction block contact surface under braking conditions. Trans Technol 26(03):34–42
8. Wei L (2014) NVH analysis and test study of disc brake based on complex eigenvalue. Chongqing University of Technology
9. Xuemin S (2017) Study on lightweight design of disc brake structure for heavy duty vehicles. Liaocheng University

Open Access This chapter is licensed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License (http://creativecommons.org/licenses/by-nc-nd/4.0/), which permits any noncommercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if you modified the licensed material. You do not have permission under this license to share adapted material derived from this chapter or parts of it.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.



Power System Design and Performance Analysis of an M100 Methanol Range Extended Electric Dump Truck



Xiqing Zhang, Yanbo Dai, Yafeng Ren, Yaojie Wang, and Shiwei Zhao

Abstract Dump trucks typically utilize diesel fuel, resulting in significant emissions of nitrogen oxides and particulate matter in their exhaust gas, which contributes to substantial environmental pollution. To support the strategy for achieving carbon peak and carbon neutrality, this study focuses on the range extended electric dump truck, using methanol as an alternative fuel to diesel. After completing the performance test of the M100 methanol range extender, the operating point of the range extender is determined. Considering the design requirements and performance indicators of the dump truck, the parameter design for both the power system and drive system is completed, and the 4-speed AMT is matched with the drive motor. The acceleration performance and fuel economy of the dump truck are simulated using AVL CRUISE. The results indicate that the acceleration time from 0 to 50 km/h is 7.73 s, and the maximum gradeability is 45.92%, the acceleration performance satisfies the design requirements. Under the CHTC-D cycle, the dump truck methanol consumption is 70.66 L/100 km, the electric energy consumption is 1.05 kWh/km, the equivalent diesel consumption is 29.79 L/100 km, and the driving range is 425 km, the fuel economy also satisfies the design requirements. The research on acceleration performance and fuel economy of the M100 methanol range extended electric dump truck lays a foundation for the future formulation and optimization of the control strategy for the range extender.

Keywords Range extended electric dump truck · Methanol range extender · Acceleration performance · Fuel economy

X. Zhang (✉) · Y. Dai · Y. Ren · Y. Wang · S. Zhao
School of Vehicle and Transportation Engineering, Taiyuan University of Science and Technology, Shanxi Taiyuan, China
e-mail: xiqingz@tyust.edu.cn

© The Author(s) 2026

H. Saman et al. (eds.), *The 9th International Conference on Advances in Construction Machinery and Vehicle Engineering*, Lecture Notes in Mechanical Engineering,
https://doi.org/10.1007/978-981-96-5527-4_72

1 Introduction

The dump truck is a kind of widely used commercial vehicle. With its compact body design and high cargo capacity, it has played a significant role in transporting urban waste soil and transferring building materials. The dump truck has a large load capacity and typically operates at low speed in low gear. However, this results in a low engine load rate and high fuel consumption rate.

Range Extended Electric Vehicle (REEV) is powered by the drive motor, while the range extender is not directly involved in vehicle drive. It only charges the power battery when its State of Charge (SOC) is low [1], ensuring that the range extender consistently operates in the high efficiency range. Additionally, due to the high torque characteristics of the motor at low speed, it can fully utilize its driving force to satisfy the operating conditions of the dump truck. Manfredi et al. [2] shown that the selection of cell chemistry and the constraints on battery weight can affect battery performance and the optimal energy management strategy of the REEV. Parkar et al. [3] developed a Modified Particle Swarm Optimization (Modified PSO) methodology aimed at minimizing fuel consumption and NOx emissions of the REEV. Through the coordination and energy transfer between the range extender, power battery, and drive motor, the range extended electric dump truck can significantly reduce fuel consumption and achieve better fuel economy compared to traditional fuel-powered dump trucks.

Methanol offers several advantages, including environmental protection, low exhaust pollution, low cost, and high safety. China's energy structure is characterized by an abundance of coal, a scarcity of oil, and limited natural gas resources. High-sulfur coal, which constitutes over 40% of the country's coal reserves, is not suitable for direct use as fuel in industrial applications such as power generation. However, it has been widely utilized in the production of industrial methanol. Currently, coal-derived methanol remains the predominant type of methanol produced [4]. The development of methanol engines can fully leverage China's coal resources and aligns with the country's basic national conditions.

From 1970s, China began researching methanol engines. Currently, it has achieved several technological breakthroughs and a small range of applications in passenger vehicles and commercial vehicles. Efforts are underway to carry methanol engines on construction machinery and large ships. However, there is relatively little research on methanol range extenders. Methanol engines face challenges such as high fuel consumption and low thermal efficiency [5]. Integrating a generator to create a vehicle-mounted methanol range extender can effectively address the existing technical issues associated with methanol engines while also compensating for the limited driving range of electric dump trucks.

Table 1 Basic parameters of the dump truck

Parameters	Numeric value
Drive form	6 × 4
Overall dimension/mm	7400 × 2550 × 3300
Wheelbase/mm	3800 + 1400
Curb weight/kg	14,000
Gross weight/kg	25,000
Frontal area/m^2	8.0
Drag coefficient	0.70
Final drive ratio	7.63
Methanol tank capacity/L	300
Tire size	12.00R20 16PR
Tire rolling radius/mm	545.4

2 Basic Structure of the Dump Truck

2.1 Basic Parameters of the Dump Truck

The dump truck is designed with a gross weight of 25 tons and adopts a 6 × 4 drive form. The basic parameters are shown in Table 1.

2.2 Performance Measures of the Dump Truck

The dump truck typically operates at low velocity in urban and rural areas. Due to its gross weight and operating conditions, the power requirement is significant. To accommodate various operating environments, the performance design measures of the dump truck are shown in Table 2.

Table 2 Performance measures of the dump truck

Measures	Numeric value
Maximum velocity/(km/h)	90
Acceleration time from 0 to 50 km/h/s	8.0
Maximum gradeability/%	45
Electric energy consumption/(kWh/km)	< 1.2
Methanol consumption/(L/100 km)	< 80

3 Design of Power and Drive Systems

3.1 Performance Test of the Range Extender

Based on an L6 natural gas engine, a spark ignition M100 methanol engine has been developed. The main technical parameters of the engine are shown in Table 3.

The M100 methanol engine serves as the power source, paired with a rated power of 150 kW Permanent Magnet Synchronous Motor (PMSM) to create the range extender for the dump truck. The main technical parameters of the range extender are shown in Table 4.

Table 3 Main technical parameters of the M100 methanol engine

Parameters	Numeric value
Type	4-valve/4-stroke/L6
Emission standard	China VI
Ignition mode	Spark plug ignition
Capacity/L	6.87
Cylinder bore/mm	108
Cylinder stroke/mm	125
Compression ratio	11.5:1
Intake form	TurboCharger + InterCooler
Spray form	Inlet multi-point electrojet
Rated power/kW	190
Rated torque/Nm	1000
Rated speed/(r/min)	2100
Technical route	Theoretical air–fuel ratio + EGR + TWC

Table 4 Main technical parameters of the range extender

Parameters	Numeric value
Rated power/kW	150
Rated voltage/V	540 V DC
Rated current/A	270
Rated frequency/Hz	50 / 60
Rated speed/(r/min)	1500 / 1600
Excitation mode	Brushless excitation
Voltage regulation range	$\geq 5\%$
Voltage fluctuating range	$\leq \pm 5\%$
Insulation class	H

The Zero-Carbon Power Comprehensive Test Platform (CJ560) is designed with a maximum test power of 560 kW, enabling a thorough and accurate evaluation of the acceleration performance and fuel economy of the 150 kW M100 methanol range extender. The arrangement of the range extender on the bench as shown in Fig. 1 and the test environment as shown in Fig. 2. The Zcanpro software communicates with the Electronic Control Unit (ECU) and the Motor Control Unit (MCU) via CAN communication to control the speed and torque of the range extender. Test data is measured through various instruments, including the YOKOGAWA WT1804E Power Analyzer and the CAMA Methanol Consumption Meter. The universal characteristics of the M100 methanol engine and range extender are shown in Fig. 3. The main performance parameters of the range extender are summarized in Table 5.

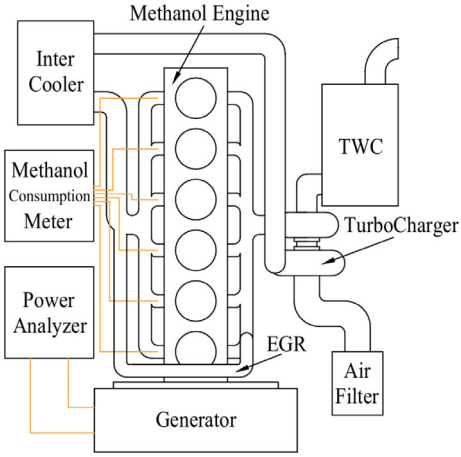


Fig. 1 Arrangement of the range extender on the bench



Fig. 2 Test environment

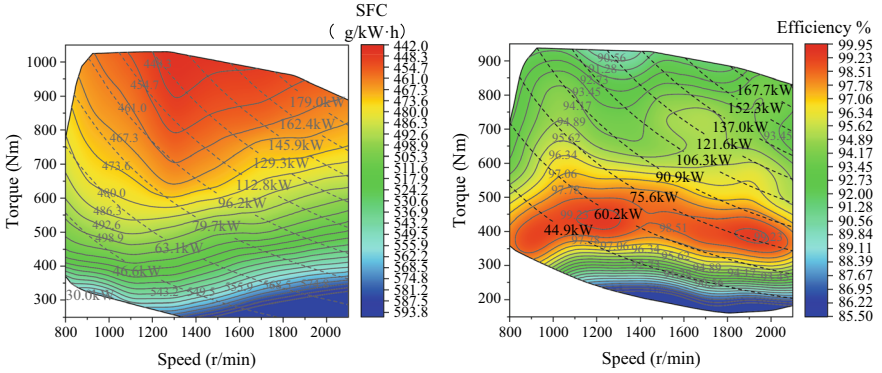


Fig. 3 Test result **a** Universal characteristic of the M100 methanol engine; **b** Universal characteristic of the range extender

Table 5 Main performance parameters of the range extender

Parameters	Numeric value
Maximum power/kW	182.57
Maximum torque/Nm	991.83
Minimum methanol consumption rate/(g/kWh)	436.72
Maximum thermal efficiency/%	35.85
Maximum power generation efficiency/%	99.93

3.2 Design of Drive Motor Parameters

The drive motor of the range extended electric dump truck must provide stable torque output at low speed and stable power output at high speed. The type of drive motor utilized is PMSM. P_{emax}, the peak power of the drive motor should satisfy the following conditions.

(1) P_{e1}, the demand power of the drive motor at maximum velocity.

When the range extended dump truck operates at maximum velocity in a uniform straight line on well-leveled pavement, both grade resistance and acceleration resistance can be neglected. The power of the drive motor should satisfy Eq. (1).

$$P_{e1} = \frac{v_{max}}{3600\eta_t}\left(m_f gf + \frac{C_D A v_{max}^2}{21.15}\right) \tag{1}$$

In the above equation, v_{max} is the maximum velocity 90 km/h; η_t is the transmission efficiency 0.95; m_f is the gross weight of the dump truck 25,000 kg; g is the gravitational acceleration 9.8 m/s^2; f is the tire rolling resistance coefficient 0.02; C_D is the drag coefficient 0.7; A is the frontal area 8 m^2. $P_{e1} = 185.39$ kW is obtained.

(2) P_{e2}, the demand power of the drive motor at maximum gradeability.

When the range extended dump truck operates at full load and uniform velocity on maximum grade road, the acceleration resistance can be neglected. The power of the drive motor should satisfy Eq. (2).

$$P_{e2} = \frac{v_\alpha}{3600\eta_t}\left(m_f g f \cos\alpha + \frac{C_D A v_\alpha^2}{21.15} + m_f g \sin\alpha\right) \tag{2}$$

In the above equation, v_α is the uniform climbing velocity 8 km/h; $\tan\alpha$ is the maximum gradeability 45%. $P_{e2} = 245.44$ kW is obtained.

The peak power of the drive motor should satisfy Eq. (3). P_{emax} is rounded, $P_{emax} = 250$ kW.

$$P_{emax} \geq \max[P_{e1}, P_{e2}] \tag{3}$$

P_{e0}, the rated power of the drive motor should satisfy Eq. (4).

$$P_{e0} = \frac{P_{emax}}{\lambda} \tag{4}$$

In the above equation, λ is the motor overload coefficient 1.6. P_{e0} is rounded, $P_{e0} = 160$ kW.

The relationship between the torque and power of the motor is described by Eq. (5).

$$T_e = \frac{9550 P_e}{n_0} \tag{5}$$

High power PMSM for vehicles is typically a low-speed motor. In the above equation, n_0 is the rated speed of drive motor 1500 r/min; both the peak torque T_{emax} and the rated torque T_{e0} of the drive motor are obtained and rounded, $T_{emax} = 1600$ Nm, $T_{e0} = 1000$ Nm.

3.3 Design of Transmission Parameters

(1) i_{gmax}, the slowest ratio of transmission should satisfy Eq. (6) and Eq. (7).

$$i_0 i_{g\,max} \geq \frac{r}{T_{e\,max}\eta_t}\left(m_f g f \cos\alpha + \frac{C_D A v_\alpha^2}{21.15} + m_f g \sin\alpha\right) \tag{6}$$

$$i_0 i_{gmax} \leq \frac{0.377 r n_{max}}{v_\alpha} \tag{7}$$

In the above equation, i_0 is the final drive ratio 7.63; r is the tire rolling radius 545.4 mm; n_{max} is the maximum speed of the drive motor 3400 r/min. $4.93 \leq i_{gmax} \leq 11.45$ is obtained.

(2) i_{gmin}, the top gear ratio of transmission should satisfy Eq. (8) and Eq. (9). $0.33 \leq i_{gmin} \leq 1.02$ is obtained.

$$i_0 i_{gmin} \geq \frac{r}{T_{emax}\eta_t}\left(m_f g f + \frac{C_D A v_{max}^2}{21.15}\right) \tag{8}$$

$$i_0 i_{gmax} \leq \frac{0.377 r n_{max}}{v_{max}} \tag{9}$$

According to engineering conditions, the transmission is designed to have four forward gears. Select 1 gear ratio $i_{g1} = i_{gmax} = 5.0$, 4 gear ratio $i_{g4} = i_{gmin} = 1.0$. To ensure that the drive motor operates in the high efficiency range, select $i_{g2} = 3.05$, $i_{g3} = 1.66$.

3.4 Design of Power Battery Parameters

The power battery directly supplies energy to the drive motor, which necessitates a high energy conversion rate. To minimize both the size of the power battery and weight of the dump truck, the onboard power battery pack must possess higher specific energy and specific power. The range extended electric dump truck experiences frequent charging and discharging cycles, requiring the power battery to support a greater number of charge and discharge cycles. Consequently, Lithium Iron Phosphate Battery has been chosen as the power battery for the range extended electric dump truck.

The dump truck is equipped with the range extender that starts when the power battery SOC is low. Consequently, when determining the capacity of the power battery, it is not necessary to consider the driving range of the dump truck, it should only satisfy the demand power of the drive motor in Eq. (10).

$$C_b \geq \frac{1000}{kU_m}\left(\frac{P_{emax}}{\eta_e} + P_A\right) \tag{10}$$

In the above equation, k is the maximum discharge rate of the battery 6.5 h^{-1}; U_m is the bus voltage 650 V; η_e is the efficiency of drive motor 0.9; P_A is the accessory power of the dump truck 20 kW. $C_b \geq 70.48$ kWh is obtained. Considering that reducing the frequency of battery charging and discharging can significantly enhance the service life of the power battery, select $C_b = 150$ kWh.

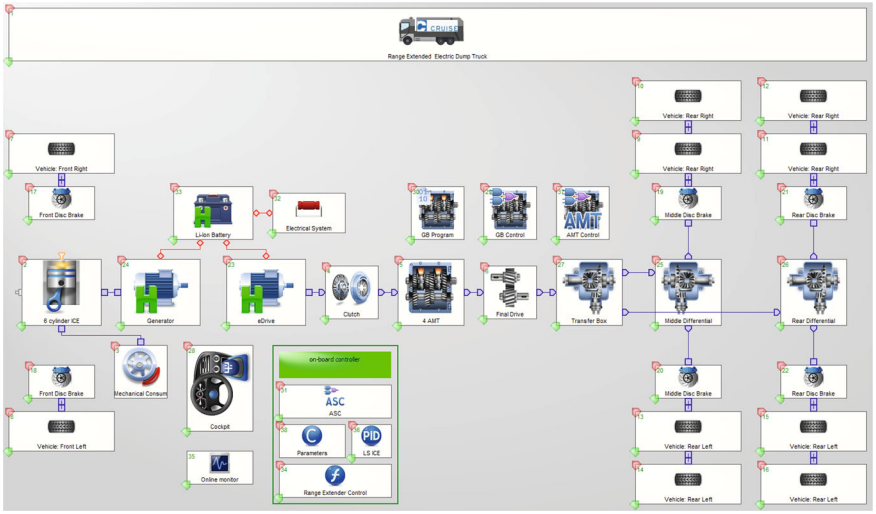


Fig. 4 Simulation model of the dump truck

4 Model Building and Simulation Analysis of the Dump Truck

4.1 Model Building of the Dump Truck

AVL CRUISE can simulate and analyze the acceleration performance and fuel economy of various types of vehicles, including traditional fuel vehicles and new energy vehicles. Utilizing the hybrid electric vehicle module in AVL CRUISE, the model of the range extended electric dump truck, as shown in Fig. 4, has been built. The parameters for the vehicle, engine, generator, power battery, drive motor, transmission, final drive, differential, brake, tire, and other components have been established and interconnected. The blue lines in Fig. 4 represent mechanical connections between components, while the red lines represent electrical connections, with components passing signals via the data bus.

4.2 Running Resistance of the Dump Truck

According to reference [6], the running resistance of the vehicle can be determined through the vehicle sliding test, as well as by Eq. (11).

$$v = \frac{m_f \times 2 \times \Delta v}{3.6 \times \left(A + Bv + Cv^2\right)} \tag{11}$$

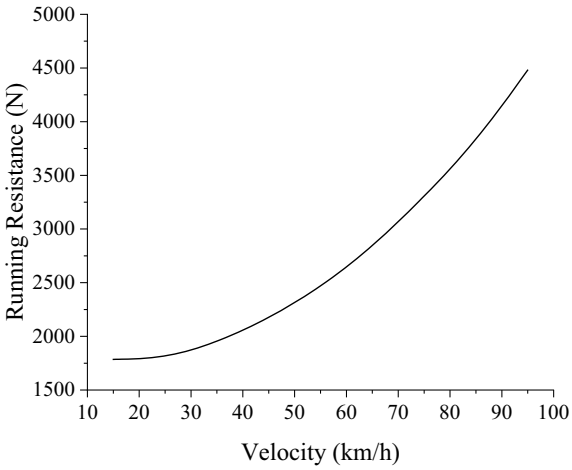


Fig. 5 Running resistance of the dump truck

In the above equation, v is the running velocity of the dump truck, which ranges from 15 km/h to 95 km/h; Δv is the deviation from velocity 5 km/h; A is the constant term of the running resistance 1677.2N; B is the monomial coefficient 13.74 N/(km/h); C is the quadratic coefficient 0.246 N/(km/h)^2. The running resistance of the dump truck is obtained by fitting as shown in Fig. 5.

4.3 Control Strategy of the Range Extender

(1) Operating point select of the range extender.

The rated power of the range extender is 150 kW. To achieve better fuel economy and emission, the operating point of the range extender is determined based on the performance test in Sect. 3.1. At this operating point, the speed is 1450 r/min, the torque is 988 Nm, the thermal efficiency is 35.6%, and the power generation efficiency is 93.84%.

The Function Module of AVL CRUISE is used to control the operating point of the range extender. To enhance the service life of the power battery, the range extender starts to charge the battery when the SOC falls below 30%. Conversely, when the SOC exceeds 80%, the range extender is turned off.

(2) Shift strategy of the transmission.

Among the various types of automatic transmissions, AMT offers several advantages, including high transmission torque, excellent efficiency, remarkable reliability, and low cost. Additionally, the shifting process is managed through electronic control, simplifying the operation for drivers and making it a popular choice for trucks. In AVL CRUISE, the AMT Control, Gearbox Control, and Gearbox Program are utilized to

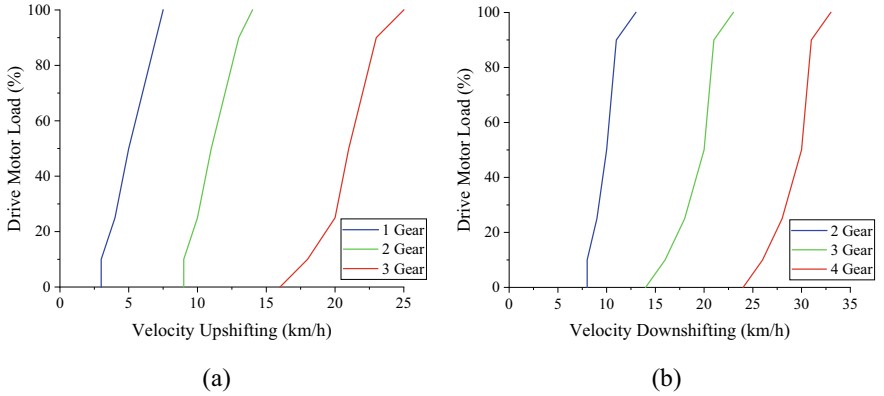


Fig. 6 Shift strategy of the AMT; **a** Upshift strategy; **b** Downshift strategy

control the operation of the 4-speed AMT. The transmission shift strategy is shown in Fig. 6.

4.4 Simulation Analysis

(1) Acceleration performance simulation of the dump truck.

The simulation results for the maximum velocity of the dump truck under no load are shown in Fig. 7a. The maximum velocity is 92.63 km/h and the acceleration time from 0 to 50 km/h is 7.73s. This performance satisfies the requirements both maximum velocity is 90 km/h and the acceleration time from 0 to 50 km/h is 8s. The simulation results for the maximum velocity of the dump truck under full load are shown in Fig. 7b. The maximum velocity is 92.59 km/h, with an acceleration time from 0 to 50 km/h of 13.17s.

The simulation results for the maximum gradeability of the dump truck under full load are shown in Fig. 8. The maximum gradeability in 1 gear is 45.92%, which satisfies the maximum gradeability requirement of 45% for the dump truck.

(2) Fuel economy of the dump truck.

CHTC-D is the latest dump truck cycle in China, featuring a total operating time of 1300 s and an operating range of 8.37 km. Within this cycle, the low-velocity condition lasts for 540 s, while the high-velocity condition lasts for 760 s. The maximum velocity is 71.4 km/h, and the average velocity is 23.19 km/h. The velocity curve for the CHTC-D cycle is shown in Fig. 9.

As shown in Fig. 10, a total of 33 CHTC-D cycles are conducted. The dump truck accumulates an operating range of 276.21 km. The initial SOC of the power battery is 80%. A total of three charge and discharge cycles are completed. The range extender

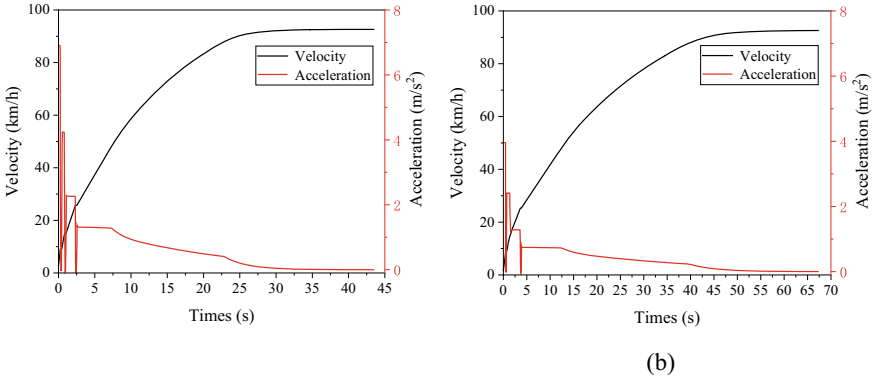
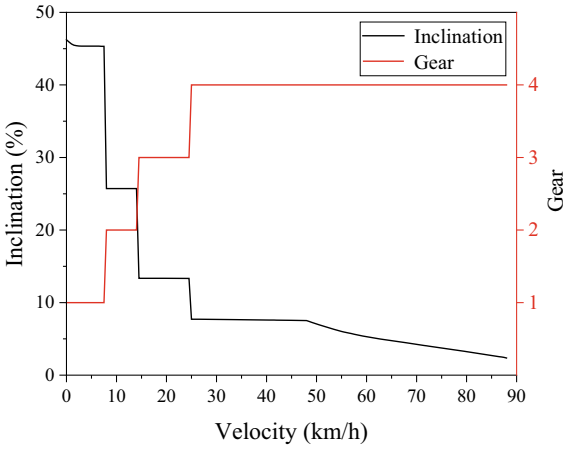


Fig. 7 Maximum velocity of the dump truck; **a** No load; **b** Full load

Fig. 8 Maximum gradeability of the dump truck



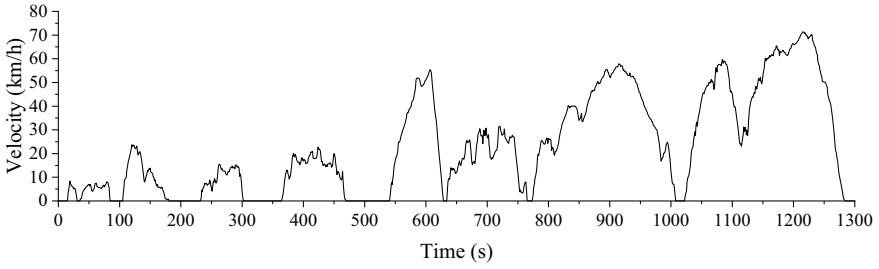


Fig. 9 CHTC-D cycle

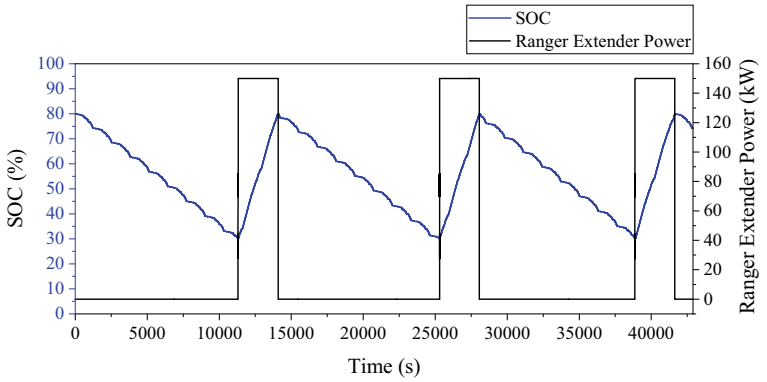


Fig. 10 CHTC-D SOC

is started three times, providing a stable power supply to the battery at 150 kW. At the end of the simulation, the battery SOC is 74.25%.

C-WTVC is developed by China based on the World Transient Vehicle Cycle (WTVC). It is applicable to various types of heavy commercial vehicles. The total operating time is 1800s, with an operating range of 20.52 km. This includes an urban condition of 900 s, a highway condition of 468 s, and a high-velocity condition of 432 s, with a maximum velocity of 87.8 km/h. The velocity curve for the C-WTVC cycle is shown in Fig. 11.

As shown in Fig. 12, a total of 16 C-WTVC cycles are conducted. The dump truck accumulates an operating range of 328.32 km. Similar to the CHTC-D cycle, the initial SOC of the power battery is 80%. A total of three charge and discharge cycles are completed. The range extender is started three times and the power is 150 kW. At the end of the simulation, the battery SOC is 70.29%.

As shown in Table 6, the dump truck satisfies the methanol consumption limit of 80 L/100 km under the CHTC-D and C-WTVC cycles. Additionally, according to reference [7], the diesel consumption limit for the dump truck with a gross weight of 25 tons under the CHTC-D cycle is 37.5 L/100 km. In this study, the equivalent

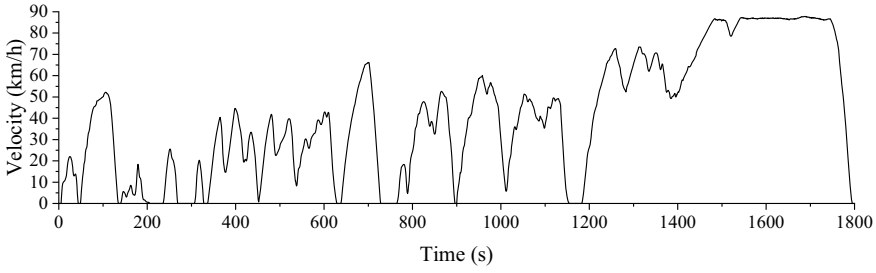


Fig. 11 C-WTVC cycle

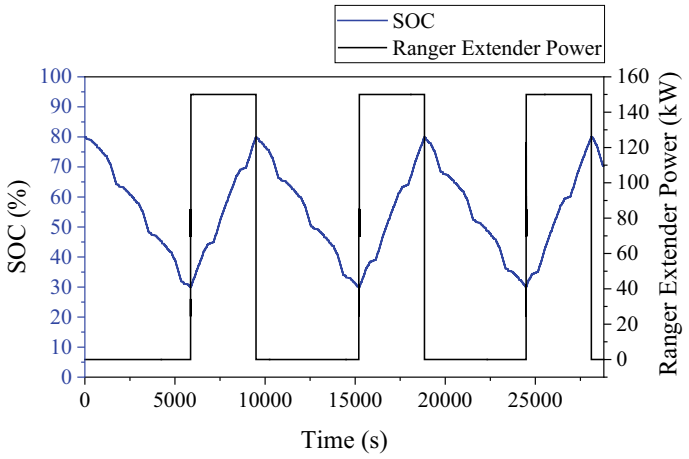


Fig. 12 C-WTVC SOC

Table 6 Simulation results of the fuel economy

Parameters	CHTC-D	C-WTVC
Methanol consumption/(L/100 km)	70.66	77.88
Electric energy consumption/(kWh/km)	1.05	1.17
Equivalent diesel consumption/(L/100 km)	29.79	33.89
Driving range/km	425	385

diesel consumption of the dump truck during the CHTC-D cycle is 29.79 L/100 km, which complies with national standard requirement.

5 Conclusions

(1) A performance test and evaluation of the M100 methanol range extender have been conducted using the bench test. The minimum methanol consumption rate of the range extender is 436.72 g/kWh, while the maximum thermal efficiency is 35.85%.

(2) The design of the parameters for the range extended electric dump truck accords to the basic specifications and performance indicators of the vehicle. By considering the operating environment and the lifespan of the power battery, establish the parameters for the power and drive systems of the range extended electric dump truck.

(3) The operating point of the range extender is determined through performance test. Simultaneously, to enhance fuel economy and satisfy the power demands of the dump truck, a 4-speed AMT with a maximum transmission ratio of 5.0 has been designed, along with its corresponding shift strategy.
(4) Through simulation, the performance of the dump truck satisfies the design requirements. The methanol consumption of the CHTC-D cycle is 70.66 L/100 km, and the equivalent diesel consumption is 79.44% of the national standard limit. Additionally, the price of methanol is approximately 40% that of diesel. Compared to similar diesel-powered dump trucks, the overall cost is lower.
(5) In order to reduce exhaust pollutant emissions from the range extender and comply with the China VI emission standard, the range extender utilizes an equivalent combustion mode and Exhaust Gas Recirculation (EGR), which correspondingly reduces thermal efficiency. Additionally, the range extender operates at a fixed point when generating electricity, all of them impacting the fuel economy of the dump truck. Consequently, the development of control strategies tailored to various driving conditions and power requirements will be the focus of future research.

Acknowledgements Supported by Shanxi Intelligent Transportation Production-Education Integration Postgraduate Joint Cultivation Demonstration Base (No. 2024JD10).

References

1. Zhang F (2017) Power system parameter matching and simulation optimization of range extended electric vehicle. Zhejiang University
2. Manfredi V, Ankur S, Matilde D et al (2022) Battery selection and optimal energy management for a range-extended electric delivery truck. SAE Int J Adv Current Pract Mob 5(3):1282–1291
3. Parkar O, Snyder B, Rahi A et al (2023) Modified particle swarm optimization based powertrain energy management for range extended electric vehicle. Energie 16(13)
4. Jinxia X, Wang L, Ye B (2017) Present situation, application, and industrialization trend of coal-derived methanol. Shandong Chem Ind 46(05):63–66
5. Yao C, Yao A (2023) Application status and prospect of methanol fuel. J Automot Safety Energy 14(05):521–535
6. GB/T 27840–2021: Fuel Consumption Test Methods for Heavy-Duty Commercial Vehicles.
7. GB 30510–2018: Fuel Consumption limits for Heavy-Duty Commercial Vehicles

Open Access This chapter is licensed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License (http://creativecommons.org/licenses/by-nc-nd/4.0/), which permits any noncommercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if you modified the licensed material. You do not have permission under this license to share adapted material derived from this chapter or parts of it.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.



Grayscale Image Classification Method for Auto Parts Based on MobileNet_V3_Small



Yujie Ye, Zhichao Hu, Long Diao, and Yunchao Wang

Abstract Classifying automotive parts during manufacturing is crucial for effective supply chain management and quality assurance. This paper proposes a grayscale image classification method for auto parts utilizing an enhanced MobileNetV3_Small model. Augment the original data set with data enhancement techniques. The spatial attention mechanism is added before the original SE attention mechanism to enhance the feature extraction capability of the MobileNetV3_Small network on grayscale images. Extensive experiments conducted on ten types of balanced datasets demonstrate that the model achieves an accuracy of 99.4%, with a loss rate of only 0.0054%. Furthermore, metrics including precision, recall, F1 score, and AUC all reach perfect scores of 1.00. These results comprehensively validate our proposed model's high accuracy and efficiency while showcasing its robust classification capabilities. Notably, the model's weight file is merely 6.2 MB in size, facilitating rapid inference and efficient deployment, meeting the industry demands for rapid classification of parts, and highlighting its advantages in adaptability for practical applications.

Keywords Grayscale image classification · MobileNetV3_small · Auto parts recognition

1 Introduction

In recent years, visual inspection algorithms have seen significant expansion in various fields. Automation and intelligent technologies are increasingly being integrated into the production and quality control of auto parts. Accurate classification of automotive parts is crucial to reduce costs and increase efficiency.

Y. Ye · Z. Hu · L. Diao · Y. Wang (✉)
Jimei University, Xiamen, China
e-mail: ychaowang@jmu.edu.cn

© The Author(s) 2026

H. Saman et al. (eds.), *The 9th International Conference on Advances in Construction Machinery and Vehicle Engineering*, Lecture Notes in Mechanical Engineering,
https://doi.org/10.1007/978-981-96-5527-4_73

Image classification techniques can be categorized into traditional methods and those based on deep learning. Currently, many factories continue to utilize conventional machine-learning approaches for image classification and defect detection. For instance, features are extracted by template matching, before employing classifiers like support vector machine [1], K-nearest neighbor [2], and naive Bayes [3] to complete the classification task. However, these traditional methodologies are increasingly being phased out by advanced manufacturers due to their limitations such as high false detection rate, poor adaptability, and dependence on manual design for feature extraction. Convolutional neural networks combined with deep learning, such as GoogleNet [4] and VGG [5] without the need to manually design feature extractors [6], show better adaptability and generalization ability.

However, the application of these techniques faces challenges such as small dataset size and sample imbalance. Meanwhile, as the network scales up, the increase in the number of parameters also tends to lead to overfitting [7]. To overcome these challenges, Chang et al. [8] enhanced a GAN and MobileNetV3-based method for strip defect classification, increasing prediction speed by 16.7% and accuracy by 2.62% through adjustments to convolutional layers and channel numbers. Based on the pre-trained ResNet50 model, Hu et al. [9] used transfer learning to classify copper alloy metallography, effectively addressing the challenge of limited dataset sizes by reconfiguring the fully connected layer. Liu et al. [10] used multi-scale separable convolution and dual-channel pooling to improve the ResNet50 network, which improved the learning speed by nearly 4 times compared with the original network model. Zhu et al. [11] enhanced MobileNetV3_Small with transfer learning to address low accuracy and slow convergence in classifying scarce PSPR capsule images. Therefore, lightweight networks that leverage transfer learning and channel attention mechanisms present an effective solution.

This study presents an image classification method that efficiently classifies auto parts by optimizing neural networks and adding spatial attention for enhanced feature extraction.

2 MobileNetV3_Small Model

MobileNetV3 is an efficient deep-learning model designed for vision applications in mobile and embedded devices. It comes in two versions: Large and Small to suit different performance needs. The model is optimized using Network Architecture Search (NAS) and Net Adapt algorithms to obtain smaller model sizes and lower computational costs. MobileNetV3 continues the design of deep separable convolution and adopts an improved inverted residual structure to further reduce the parameters and computational effort. In addition, it introduces the Squeeze-and-Excitation (SE) module to improve feature recognition and uses a new activation function, h-swish.

In some modules (as shown in Fig. 1), the Squeeze-and-Excitation (SE) module is added to enhance the feature representation, enabling channel-to-channel importance

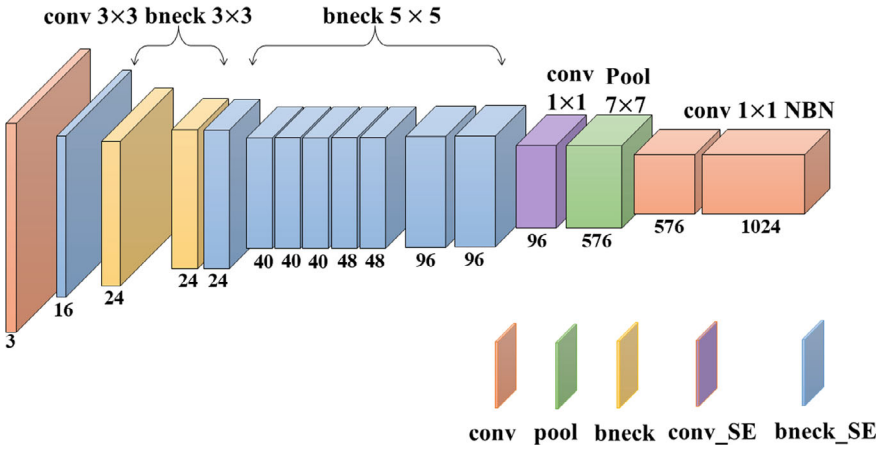


Fig. 1 MobileNetV3-small model structure

assessment through average pooling and compression and recovery of the two fully connected layers. On this basis, this paper adds a spatial attention mechanism in front of SE to enhance the learning of the current features of the image contour and increase the diversity of the extracted features.

By applying average pooling and maximum pooling in the last channel dimension, the channel average value and channel maximum value of the input feature map are calculated respectively, and then they are spliced in the last dimension to form a new tensor and perform a 1×1 convolution operation. Finally, the spatial attention weights are multiplied by the calculated attention weights with the original feature map inputs element by element to learn the spatial attention weights from the mean and maximum values.

3 Experiments and Result Analysis

3.1 Experimental Dataset and Evaluation Indicators

In this experiment, a public dataset of 50 auto parts was sourced from www.kaggle.com/datasets/gpiosenka/car-parts-40-classes?resource=download. 10 categories are selected for evaluation. Due to an imbalance in class distribution, and the number of pictures of each type does not exceed 200, the number of these 10 types of parts is equalized and the data augmentation is amplified by translation, rotation, zoom, and random brightness adjustments. Consequently, a total of 11,000 images from a dataset for classifying 10 types of auto parts were divided into validation, test, and training sets in a ratio of 1:1:8.

As for the evaluation metrics, accuracy, loss, precision, recall, F1, and AUC are selected to comprehensively assess the classification performance of the model from different perspectives. Accuracy measures the proportion of correct predictions made by the model, as shown in Eq. (1). where *TP* is the number of true positive examples, *TN* is the number of true negative examples, *FP* is the number of false positive examples, and *FN* is the number of false negative examples. Precision measures the proportion of samples that the model predicts to be positive in the positive category that are positive categories, as shown in Eq. (2), Recall measures the proportion of all samples that are positive to be correctly predicted by the model to be positive, as shown in Eq. (2), where the F1 score is the reconciled average of precision and recall to strike a balance between the two, as shown in Eq. (3).

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \tag{1}$$

$$Precision = \frac{TP}{TP + FP} \quad Recall = \frac{TP}{TP + FN} \tag{2}$$

$$F1 = 2 \times \frac{Precision \times Recall}{Precision + Recall} \tag{3}$$

3.2 Training Results and Visualization of Test Results

During the model training process, the parameters were configured as follows: the learning rate is 0.0003, the batch size is 16, and the epoch is 50. The results in Fig. 2a indicate that the verification accuracy is constantly approaching the curve of the training accuracy without significant deviation as training rounds increase. This suggests that features within the images can be effectively learned without overfitting. The validation loss curve in Fig. 2b is also approaching the training loss curve, indicating that the model's predicted values are increasingly aligning with true values.

Upon completion of model training, the corresponding checkpoint file is called for prediction by randomly selecting images from each type within the test set, and the result is illustrated in Fig. 3. As can be seen from the figure, the model can accurately identify the corresponding part types while highlighting feature extraction areas with high recognition accuracy and robust adaptability. Moreover, since processed images are single-channel grayscale images, which have less data and faster response speed than three-channel images, they can meet the actual needs of industrial production scenarios.

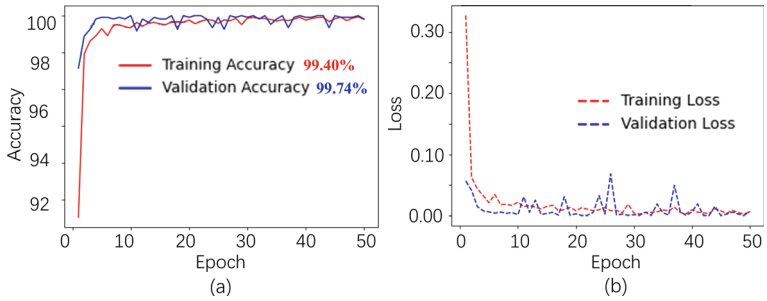


Fig. 2 Training curves of various indicators. **a** Accuracy curve. **b** Loss curve

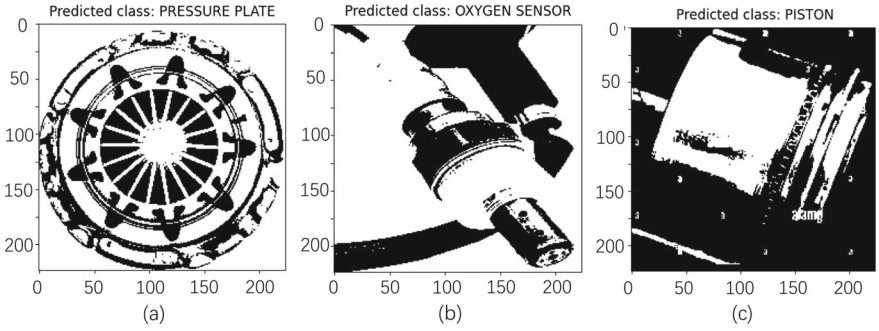


Fig. 3 Visualization of detection results. **a** Pressure Plate detection results. **b** Oxygen Sensors detection results. **c** Piston detection results

3.3 Comparison of Different Models

In the evaluation against ResNet18, ResNet50, ShuffleNet, and EfficientNet, our enhanced MobileNet_V3_Small model demonstrated optimal performance across accuracy, loss, AUC, model size, and FP + FN metrics. Figure 4 shows it reaching the highest accuracy, the quickest loss reduction, the highest AUC, and the best F1 score for consistent classification, striking an ideal balance between precision and recall.

Table 1 compares model efficiency and accuracy, highlighting that our model offers a compact size with the lowest FP + FN counts, outperforming ShuffleNet in classification tasks despite its smaller file size. The data confirm our model's superior performance for automotive part image classification in real-world production scenarios.

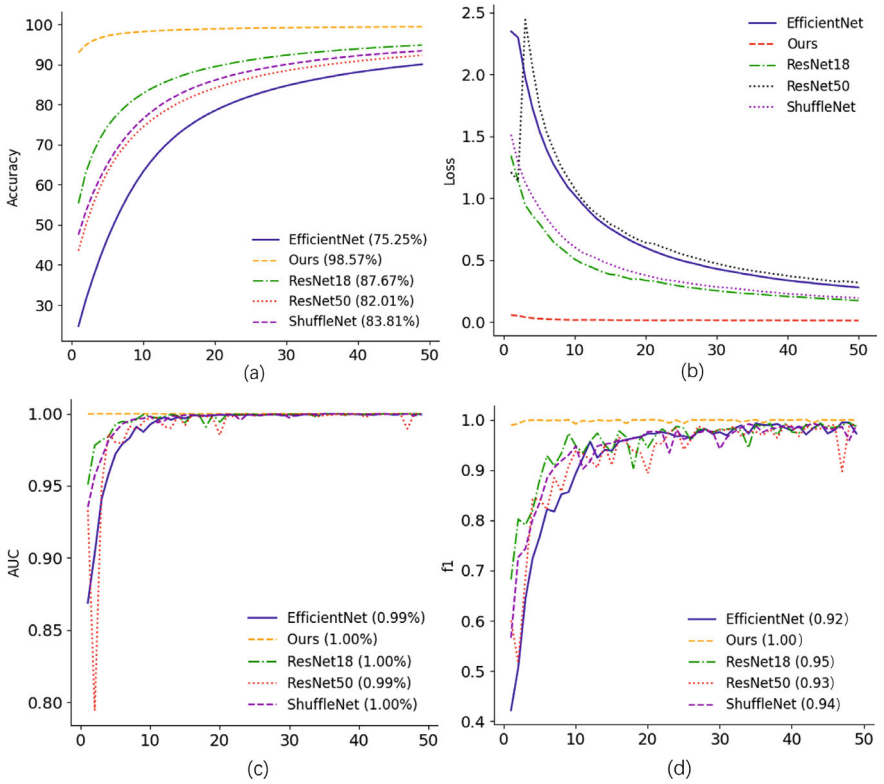


Fig. 4 Training results. **a** Accuracy curve in different models. **b** Loss curve in different models. **c** Auc curve in different models. **d** F1 curves in different models

Table 1 Comparative experiments with other models

Model	Precision	Recall	Weight size(MB)	FP + FN
EfficientNet	0.92	0.91	26.4	14 + 16
Ours	1.00	1.00	6.2	0 + 2
ResNet18	0.95	0.94	46.9	9 + 5
ResNet50	0.93	0.92	102.7	3 + 7
ShuffleNet	0.94	0.93	5.2	4 + 7

4 Conclusions

Aiming at the classification problem of auto parts in industrial production, an improved MobileNetV3_Small model was proposed to improve the detection accuracy of small sample gray images by optimizing the attention mechanism, without increasing the parameter and calculation burden. Image enhancement techniques

were used to improve the adaptability of the model. Experimental results achieved 99.4% classification accuracy.

Despite the lightweight design, the impact of the attention mechanism on the actual reasoning speed and energy consumption needs to be further studied, and the scalability of the model and potential integration with other industrial automation systems are also future research directions. In conclusion, while our work has made great strides in advancing the classification of automotive parts, there is still much to be explored and improved.

Acknowledgements This work was supported by the National Natural Science Foundation of China, China (No. 52375046).

References

1. Cortes C, Vapnik V (1995) Support-vector networks. Mach Learn 20(3):273–297
2. MacQueen J (1967) Some methods for classification and analysis of multivariate observations. In: Proceedings of the fifth Berkeley symposium on mathematical statistics and probability, volume 1: statistics
3. Rish I (2001) An empirical study of the naive Bayes classifier. In: IJCAI-01 workshop on empirical methods in AI
4. Sermanet P et al (2013) OverFeat: integrated recognition, localization and detection using convolutional networks. CoRR. abs/1312.6229
5. Simonyan K, Zisserman A (2014) Very deep convolutional networks for large-scale image recognition. CoRR. abs/1409.1556
6. Krizhevsky A, Sutskever I, Hinton GE (2017) ImageNet classification with deep convolutional neural networks. Commun ACM 60(6):84–90
7. Leiyu C et al (2021) Review of image classification algorithms based on convolutional neural networks. Remote Sens 13(22):4712–4712
8. Jiang C, Shengqi G, Hongyu S et al (2021) Strip defect classification based on improved generative adversarial networks and MobileNetV3. Laser Optoelectron Prog 58(04):221–226
9. Mingqi H, Qiuming L, Huiming C et al. Classification method of copper alloy metallographic diagram based on improved residual network and transfer learning. Nonferrous Met Sci Eng 1–13
10. Fan L, Rui C. Application of improved ResNet50 in image classification of kaolin ores. Mech Electr Eng Technol 1–7
11. Jin Z, Chuanhui Z, Changjiang Z (2023) *Papaver somniferum* and *Papaver rhoeas* classification based on visible capsule images using a modified MobileNetV3-Small network with transfer learning. Entropy 25(3):447–447

Open Access This chapter is licensed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License (http://creativecommons.org/licenses/by-nc-nd/4.0/), which permits any noncommercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if you modified the licensed material. You do not have permission under this license to share adapted material derived from this chapter or parts of it.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.



Response Characteristics of a Fourteen-Degree-Freedom Vehicle Under Four-Wheel Random Pavement Excitation



Lei Zou, Zhichao Hu, Haoyu Sun, and Yunchao Wang

Abstract In order to study the changes of vehicle dynamic response and tire dynamic load under different road levels, the time domain model of random road surface and the transfer function between four wheels and road surface were established by using filtered white noise method, and the dynamic response of vehicle on random uneven road surface was simulated. On this basis, considering the damping characteristics of the suspension system, the vehicle dynamics model of 14 degrees of freedom is established. The response of vehicle pitch Angle, roll Angle and tire dynamic load with road grade was studied. The results show that the dynamic response of vehicle state quantity fluctuates with the increase of road roughness, resulting in the gradual decrease of ride comfort and stability. Through the data analysis of vehicle state quantity and vertical force, it can be seen that the constructed vehicle model can reflect the vehicle state response under various road conditions, which provides a theoretical basis for the subsequent research of vehicle-road coupling.

Keywords Pavement random excitation · Vehicle-road coupling · Hardware-in-the-loop experiment

1 Introduction

Road surface irregularities, transmitted to the vehicle body through the tires, significantly impact the ride's smoothness, stability, and safety. Given that the vehicle-road interaction constitutes a quintessential nonlinear system, the development of an accurate road surface model is crucial for simulating the actual road-induced excitation experienced by vehicles [1, 2]. At present, the main methods of building random pavement unevenness models are the Fourier transform [3], the Harmonic superposition method [4], the AMRA model method [5], etc. Among them, the filtered white noise method [6] fits the white noise meeting certain conditions into the frequency response

L. Zou · Z. Hu · H. Sun · Y. Wang (✉)
Jimei University, Xiamen, China
e-mail: ychaowang@jmu.edu.cn

© The Author(s) 2026

H. Saman et al. (eds.), *The 9th International Conference on Advances in Construction Machinery and Vehicle Engineering*, Lecture Notes in Mechanical Engineering,
https://doi.org/10.1007/978-981-96-5527-4_74

function of the road roughness function after appropriate transformation, to obtain the state equation of the time–frequency power spectral density function, which is a more commonly used time-domain excitation modeling method for pavement roughness at present.

Liang et al. [7] conducted comparative experiments on various pavement excitation model modeling methods and showed that the pavement random excitation white noise time-domain model is relatively close to the real pavement unevenness. Huang et al. [8] conducted simulation experiments on various road surface grades based on the MATLAB/Simulink platform and analyzed the dynamic forces of tires under different road surfaces. In this paper, the time domain model of random pavement and the input model of four-wheel random pavement is established by using the filtered white noise method, and the accuracy of the model is verified. On this basis, a 14-DOF vehicle dynamics model including longitudinal, transverse, yaw, body pitch, roll, vertical, four-wheel rotation, and vertical motion was established. The vehicle-road interaction model was simulated in real-time using the hardship-in-the-loop (HIL) platform, and the dynamic response characteristics of vehicles on different road surfaces were studied. The effect of road excitation on vehicle vertical load is analyzed.

2 Fourteen-Degree-Freedom Vehicle

2.1 Pavement Time Domain Model

The measure of random pavement is pavement roughness, that is, roughness of the road surface. The smoothness of the road surface is the main cause of changes in vehicle posture and wheel load. In the process of establishing a time-domain model of a single-wheel road surface, according to the national standard "Vehicle Vibration Input—Representation of Road Surface Roughness" [9]. Combining Gaussian white noise to construct a road surface random excitation model, obtain the road surface excitation state equation:

$$\dot{q}(t) = -2\pi n_q u q(t) + 2\pi n_0 \sqrt{G q(n_0) u}\,\omega(t) \tag{1}$$

where: $\omega(t)$ is the random noise time-domain signal; $q(t)$ is the time-domain function of pavement unevenness, and n_q is the lower cutoff spatial frequency.

Due to the complexity of the model, the model can be built with simple vehicle driving conditions for experimentation first. Assuming that the vehicle travels in a straight line at a uniform speed and the front and rear wheels are on the same line, the rear wheel road excitation of the car is the lag of the front wheel road excitation over a period of time, and the lag time is according to, which can be expressed by the ratio of the sum of the front and rear axle distances $(a + b)$ the speed of the car. The relationship between the rear wheel pavement input $q_3(t)$ and the front wheel

pavement input $q_1(t)$ at moment t is shown in Eq. (2):

$$t_d = \frac{(a+b)}{u}, \quad q_3(t) = q_1(t - t_d) \tag{2}$$

Because of the retardation and left–right correlation between the four wheels, the expression between the four wheel road input can be obtained by using the second-order Pade approximation algorithm [10]. B stands for left and right wheel base, The road excitation of a four-wheel vehicle can be expressed as, Where q_1, q_2, q_3, and q_4 denote the road excitation inputs to the left front, right front, left rear, and right rear wheels of the vehicle.

$$\begin{cases} \dot{q}_2(t) = -\frac{2u}{B} q_2(t) + \frac{2u}{B} q_1(t) - \dot{q}_1(t) \\ \dot{q}_2(t) = -\frac{2u}{B} q_2(t) + \frac{2u}{B} q_1(t) - \dot{q}_1(t) \\ \dot{q}_2(t) = -\frac{2u}{B} q_2(t) + \frac{2u}{B} q_1(t) - \dot{q}_1(t) \end{cases} \tag{3}$$

2.2 Dynamic Model

During vehicle operation, lateral and longitudinal forces typically coexist. The tire's adhesion ellipse properties mean that the lateral deflection angle influences the longitudinal force magnitude, and vice versa, the longitudinal slip ratio affects the lateral force. This interaction creates a nonlinear coupling effect. To accurately simulate tire forces, the tire model in this study employs the Magic Formula [11].

In order to investigate the effect of pavement roughness on vehicle driving response and vehicle control stability, this subsection builds a vehicle-road coupling model based on the vehicle fourteen-degree-of-freedom dynamics model, as shown in Fig. 1.

The longitudinal, lateral and vertical movements of the car body obtained from the rigid body motion theorem are shown in Eq. (4):

$$\begin{cases} m_s a_x = (F_{x1} + F_{x2}) \cos\delta - \left(F_{y1} + F_{y2}\right) \sin\delta + F_{s3} + F_{s4} + m_s g \sin\theta \\ m_s a_y = (F_{x1} + F_{x2}) \sin\delta + \left(F_{y1} + F_{y2}\right) \cos\delta + F_{y3} + F_{y4} - m_s g \cos\varphi \sin\theta \\ m_s a_z = m_s \ddot{z}_b = F_{s1} + F_{s2} + F_{s3} + F_{s4} - m_s g \cos\varphi con\theta \end{cases} \tag{4}$$

The motion equations of pitch, roll and yaw are shown in Eq. (5):

$$\begin{cases} I_x \ddot{\theta} = (F_{s1} + F_{s3} - F_{s2} - F_{s4}) \frac{t}{2} + h_z \left(m_s a_y + m_s g \cos\varphi \sin\theta\right) \\ I_y \ddot{\varphi} = (F_{s1} + F_{s2}) a - (F_{s3} + F_{s4}) b + h_p (m_s a_x + m_s g \sin\varphi) \\ I_z \ddot{\psi} = (F_{x1} + F_{x3} - F_{x2} - F_{x4}) \frac{t}{2} + \left(F_{y1} + F_{y2}\right) a - \left(F_{y3} + F_{y4}\right) b \end{cases} \tag{5}$$

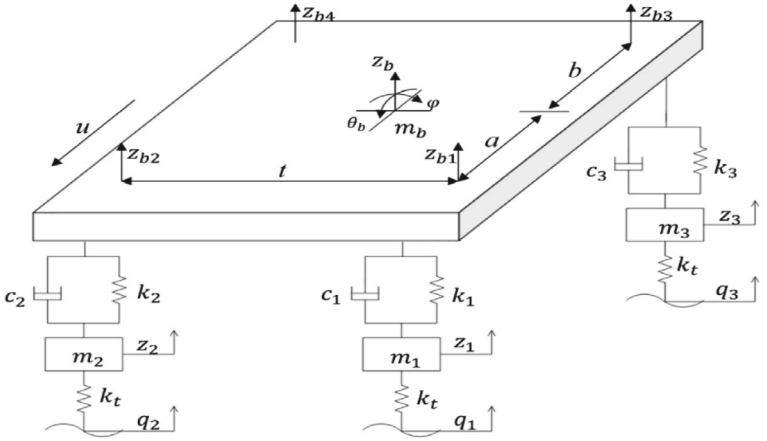


Fig. 1 Vehicle road coupling diagram

The differential equation of wheel rotation is:

$$J_i\dot{\omega}_i = T_i - F_{xi}R \tag{6}$$

The supporting force Fsi at the endpoints of the four wheels, where i = 1,2,3,4, that is, the vertical motion equation of the four wheels is:

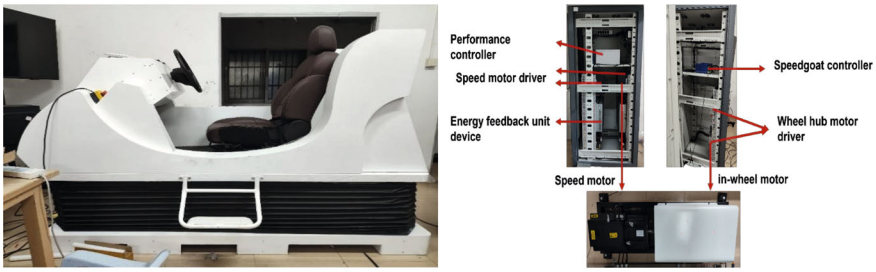
$$\begin{cases} F_{si} = c_i(\dot{z}_i - \dot{z}_{bi}) + k_i(z_i - z_{bi}) \\ m_i\ddot{z}_i = -F_{si} + k_{ti}(q_i - z_i) = c_i(\dot{z}_{bi} - \dot{z}_i) + k_i(z_{bi} - z_i) + k_{ti}(q_i - z_i) \end{cases} \tag{7}$$

The moment of inertia of the vehicle body about the x, y, and z axes are represented by I_x, I_y, and I_z, respectively. z_b denotes the displacement of the body's center of mass, m_s is the sprung mass, and mi is the tire mass. The body's roll, pitch, and yaw angles are denoted by θ, ϕ, and ψ, respectively. k_{ti} signifies tire vertical stiffness, k_i is the suspension spring stiffness, c_i represents the suspension damping coefficient, z_{bi} is the suspension corresponding body displacement, z_i is the vertical displacement of the unsprung mass, and q_i stands for the pavement excitation input.

The vertical tire load during vehicle operation comprises two components: the dynamic load and the static load, as depicted in Eq. (8). The static load's first part describes the distribution of the vehicle's gravitational force at the front and rear suspensions in a specific ratio when the vehicle is stationary.

$$G_f = \left[\frac{b}{2(a+b)}m_s + m_i\right]g \quad G_r = \left[\frac{a}{2(a+b)}m_s + m_i\right]g \tag{8}$$

The second component represents the dynamic load during vehicle motion, which accounts for the forces generated by body movements like pitch and roll due to steering, acceleration, and deceleration.



(a) (b)

Fig. 2 Hardware equipment. **a** Driving simulation platform. **b**.Overall hardware testing platform

$$F_{ni} = k_{ti}(z_i - q_i) \mp \frac{m_s a_y b h_1}{(a+b)t} \pm \frac{m_s a_x h_1}{2(a+b)} \tag{9}$$

3 Vehicle Coupling Model Real-Time Simulation Analysis

As shown in Fig. 2, The Speed-goat hardware-in-the-loop (HIL) experiment system was used to integrate the vehicle-on-road coupling model into a real-time HIL simulator. As shown in Fig. 3. Through real-time simulation under various working conditions, the dynamic response of vehicles to uneven road surfaces is analyzed. To further study the influence of road roughness on vehicle dynamic response, the speed was set at 90 km/h. On four different rough roads A, B, C, and D, the driving state response of the vehicle under the condition of double shift was observed.

It can be seen from Figs. 3 and 4 that the dynamic load response of vehicle tires is more intense and almost positively correlated with the increase in road roughness. The data from Table 1 show that when the vehicle speed is maintained at 90 km /h, the dynamic load response of the tire increases significantly with the increase of road roughness. It is worth noting that the maximum dynamic load response of Class D road surfaces is more than 5 times that of Class A road surfaces, and the maximum dynamic load response of a tire can account for 24.82% of its vertical load.

4 Conclusions

The time domain model of random pavement was established by using the white noise filtering method and verified on the MATLAB/Simulink platform. The accuracy and reliability of the model were verified by comparing it with the standard power spectral density function of B-class pavement. On this basis, based on the principles of vehicle dynamics and kinematics, a vehicle model with 14 degrees of freedom considering

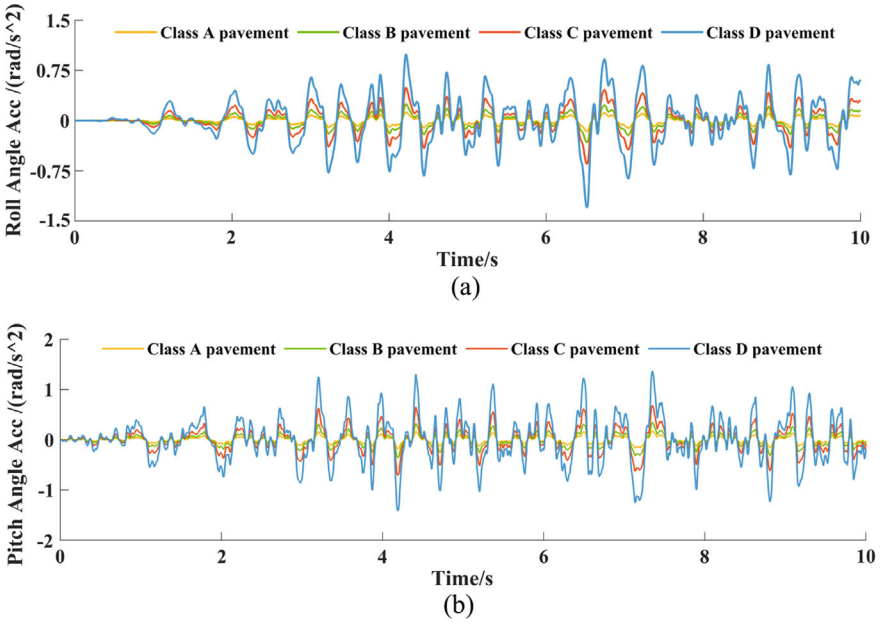


Fig. 3 State response **a** Pitch angle acceleration. **b** Roll angle acceleration

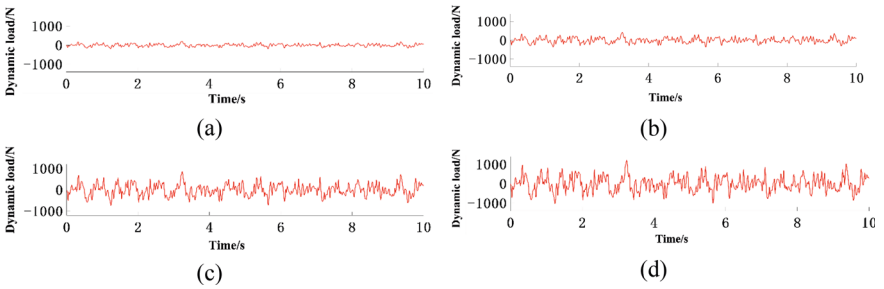


Fig. 4 Dynamic load response curve of tires under different classes of 90 km/h pavement. **a** Class A **b** Class B. **c** Class C **d** Class D

Table 1 Dynamic load response of tires under different classes of 90 km/h pavement

Pavement class/Dynamic load response	Response extremum	Response standard deviation	Maximum proportion of dynamic load (%)
Class A	216.6N	65.39N	5.52
Class B	433.2N	130.8N	10.64
Class C	866.3N	261.5N	18.93
Class D	1225.0N	369.9N	24.82

road excitation is constructed. The vehicle-road coupling model is used to evaluate the dynamic response of vehicles on different road surfaces. The research results show that the higher the road grade, the more obvious the response of vehicles, which has an adverse impact on the smoothness and stability of vehicles, resulting in a gradual decline in the smoothness and stability of vehicles. The dynamic load caused by road excitation can account for up to 24% of the tire's vertical load.

Acknowledgements This work was supported by the National Natural Science Foundation of China, China (No. 52375046),

References

1. Jun F, Xin Jie Z, Kong Hui G et al (2013) A frequency compensation algorithm of four-wheel coherence random road. Math Probl Eng 2013:1–12
2. Zhisheng Y (2009) Automobile theory, 5th edn. China Machine Press
3. Hongbin R (2016) Research on damping continuous adjustable hydro-pneumatic suspension and control technology. Beijing Institute of Technology
4. Paraforos DS, Griepentrog HW, Vougioukas SG (2016) Country road and field surface profiles acquisition, modeling and synthetic realization for evaluating fatigue life of agricultural machinery. J Terramech 63:1–12
5. Weidong X, Lei W, Yini S et al (2005) Application of random signal in the simulation of road roughness. J Vib Measur Diagn 155(02):126–130
6. Jin Y, Luo X et al (2013) Stochastic optimal active control of a half-car nonlinear suspension under random road excitation. Nonlinear Dyn 72(1–2):185–195.https://doi.org/10.1007/s11071-012-0702-x
7. Manning L, Wenfeng L, Xinfeng G (2024) Comparison of output response characteristics of automobile suspensions using time-domain modeling methods for road excitation. J Xuchang Univ 43(02):143–147
8. Dashan H, Bingqi W, Hai-liang L et al (2021) Mathematical model of pavement excitation for four-wheel vehicle. J Ordnance Equipment Eng 42(02):142–146
9. Xin J (2019) Research on the vehicle-road coupling simulation based on the automatic transmission object-in-the-loop system. Chongqing University of Technology
10. Guoren Z (2022) Research on robust control of electronically controlled air-suspension systems for vehicles. Jiangsu University of Science and Technology
11. Pacejka HB, Bakker E (1992) The magic formula tyre model. Veh Syst Dyn 21:1–18

Open Access This chapter is licensed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License (http://creativecommons.org/licenses/by-nc-nd/4.0/), which permits any noncommercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if you modified the licensed material. You do not have permission under this license to share adapted material derived from this chapter or parts of it.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.



Design, Modeling and Vibration Analysis of Piezoelectric Energy Harvesting Vehicle Suspension



Zewang Wu, Yunchao Wang, and Qi Hong

Abstract To address the challenges of energy recovery and vibration control performance optimization in vehicle suspension systems, this paper proposes a suspension system design based on piezoelectric energy recovery technology and conducts simulation and experimental studies on its characteristics. The system transfers the vibration energy generated during vehicle operation to a piezoelectric energy harvester through a hydraulic system, converting it into electrical energy for storage and utilization. Experimental results show that the maximum root mean square power of this piezoelectric energy harvesting suspension system can reach 0.33 mW under a 5 kΩ load resistance. Simulation analysis indicates that in step excitation vibration tests, the system demonstrates a faster vibration attenuation rate than traditional suspensions and provides greater damping force at low piston speeds. The results of this study offer new ideas and methods for optimizing vehicle suspension systems and provide significant reference value for the application of piezoelectric energy recovery technology in the automotive field.

Keywords Energy regenerative vehicle suspension · Vehicle vibration energy recovery · Co-simulation

1 Introduction

Global concern over the energy crisis has grown in recent decades, with solutions including the shift to renewable energy and energy recovery [1]. Transitioning from fossil fuels to renewable sources is complex, involving infrastructure development, technological research, and market adaptation [2].

Piezoelectric systems are widely used to capture energy from small rotational movements, commonly found in civilian and industrial applications [3, 3]. One method involves converting rotational movement into vibrations of piezoelectric

Z. Wu · Y. Wang (✉) · Q. Hong
Jimei University, Xiamen, China
e-mail: ychaowang@jmu.edu.cn

© The Author(s) 2026

H. Saman et al. (eds.), *The 9th International Conference on Advances in Construction Machinery and Vehicle Engineering*, Lecture Notes in Mechanical Engineering,
https://doi.org/10.1007/978-981-96-5527-4_75

materials, as wind drives blades to rotate and contact a piezoelectric beam to generate electricity [5]. Pozzi [6] designed a knee energy harvester, using gear teeth to move a piezoelectric cantilever, generating electrical energy [7].

The hydroelectric energy-harvesting damper was first proposed by Anderson and Avadhany from MIT in 2007 [8, 8], showing that heavy trucks could harvest 1.1 kW of energy. Lin Xu's research proposed a full-bridge hydroelectric damper with a damping coefficient of 32–91 kNs/m and energy-harvesting power of up to 220 W at 30% efficiency [10]. Zhang [11] developed a half-bridge damper with 73.2% hydraulic conversion efficiency, harvesting up to 33.4 W of power at 1.67 Hz and 50 mm excitation.

This study addresses the challenge of optimizing energy recovery and vibration damping in vehicle suspension systems. The proposed piezoelectric energy-harvesting device improves efficiency and damping performance, providing valuable insights for the design and optimization of such systems.

2 Principle of Operation of Piezoelectric Energy Recovery Suspension

The piezoelectric energy harvesting suspension system, shown in Fig. 1, consists of hydraulic cylinders, rectifier bridges, high- and low-pressure accumulators, hydraulic motors, and piezoelectric harvesters. During compression, the hydraulic cylinder forces fluid through the rectifier bridge, which powers the hydraulic motor, which in turn drives the piezoelectric harvester, which converts mechanical energy into electrical energy. The hydraulic fluid is then either stored in a low-pressure accumulator or returned to the hydraulic cylinder. During extension, the process is reversed. Piezoelectric harvesters use piezoelectric ceramics to generate electrical energy from mechanical stress.

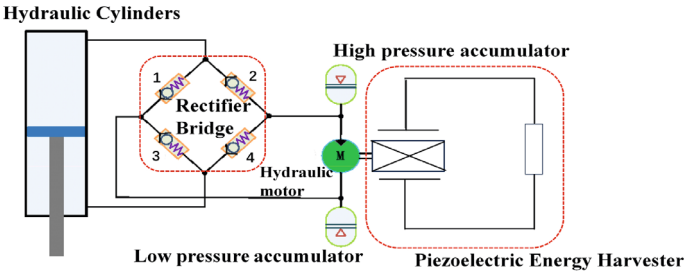


Fig. 1 Schematic diagram of piezoelectric energy-fed suspension

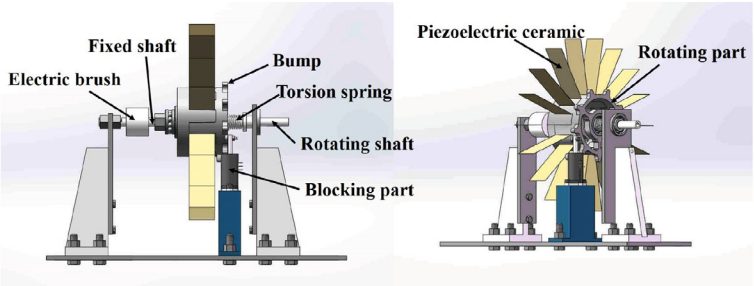


Fig. 2 3D model of the novel energy harvester

3 Design and Validation of Piezoelectric Energy Feeders

3.1 Structural Design of Piezoelectric Energy Feeders

This design employs a piezoelectric energy harvester based on collision-induced self-vibration, with its structure shown in Fig. 2. Piezoelectric energy harvesters convert mechanical energy into electricity through four main components: a rotating part with protrusions, a blocking part that interacts with the protrusions, piezoelectric ceramics that transform the mechanical energy from collisions into electrical energy, and a torsion spring that stores and releases energy. When the rotating part's protrusions hit the blocking part, the impact causes the spring to deform, storing energy. This energy is then released, continuing the rotation and the piezoelectric ceramics' vibration, which generates electricity. This electricity is then transmitted to an external circuit, allowing for continuous energy harvesting.

3.2 Mathematical Modelling of Piezoelectric Energy Feeders

The dynamics of the PEH were analyzed using a simplified model. The model is based on a spring-mass system interacting with a rotating part and a blocking component, as shown in Fig. 3. Assuming negligible frictional torque, the force acting on the system during collision can be described by the Eq. (1):

$$F_V = k_0 r \sin\theta + F_0 + m_0 \ddot{h} \tag{1}$$

When there is no collision between the rotating component and the blocking component, $F_v = 0$. Since the output impedance of the piezoelectric ceramic is relatively high, energy needs to be output through a load resistor R. Assuming the voltage across the load resistor is V_L, the power output of the electrical energy can be represented as:

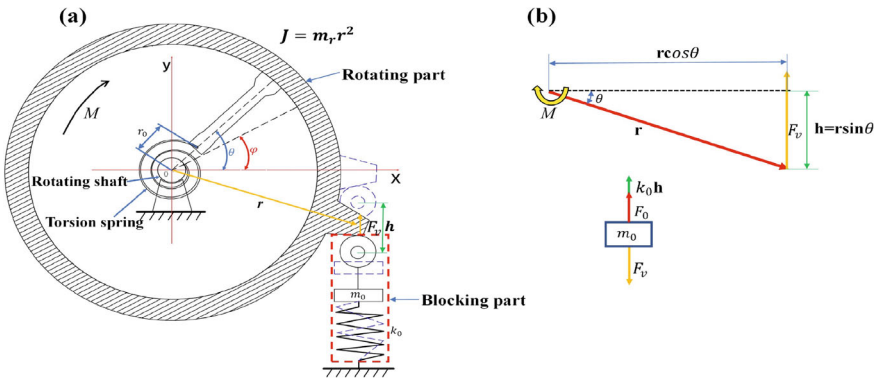


Fig. 3 **a** Schematic illustration and **b** model of the novel energy harvester

$$P = \frac{V_L^2}{R} \tag{2}$$

3.3 Piezoelectric Device Construction and Validation

The experimental setup for the non-contact collision-induced self-vibration piezoelectric energy harvester is shown in Fig. 4, which includes a motor, rotating part (radius: 350 mm), electric brush, blocking part, tachometer, PZT5 piezoelectric ceramic, and a voltage regulator. The harvester is mounted, the piezoelectric ceramic clamped, and rotational clearance adjusted for precise alignment. The ceramic's wires connect to an oscilloscope to capture real-time voltage. The blocking and rotating parts are aligned to maintain contact, and speed is measured using a tachometer and controlled by the voltage regulator. The setup utilizes the piezoelectric ceramic's self-vibration to generate electricity from the rotating part (mass: 0.5 kg, equivalent mass: 0.01 kg, 3 teeth), with the blocking part weighing 0.1 kg. Voltage data is analyzed via computer.

The effect of resistance on the power and voltage characteristics of the new harvester was investigated at a constant input shaft speed of 100 r/min as shown in Fig. 5. A range of resistances from 1 kΩ to 70 kΩ were tested to determine the optimum resistance for power output. The results show that the RMS power peaks at 5 kΩ with a maximum value of 0.33 mW.

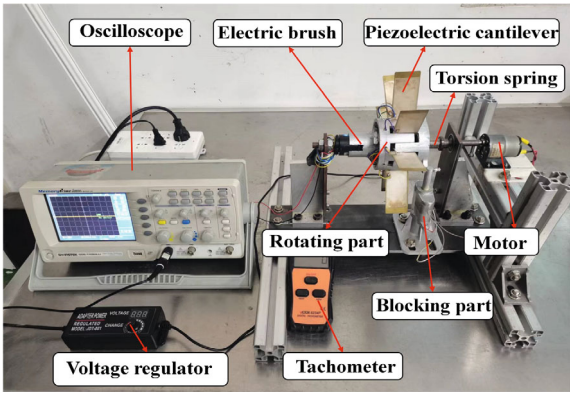


Fig. 4 Energy harvester

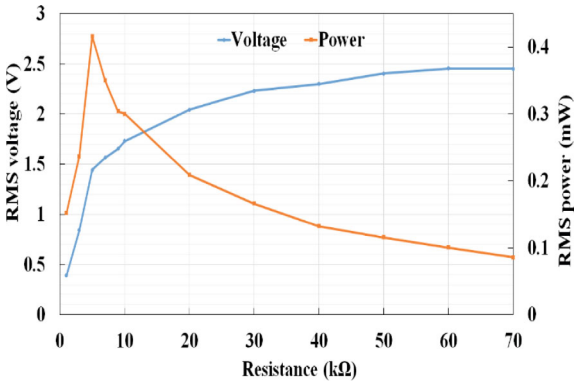


Fig. 5 RMS power and voltage

4 Simulation Analysis of Piezoelectric Energy Recovery Suspension

The co-simulation model and model parameters of the hydraulic damper and piezoelectric energy-harvesting device in AMESim and Simulink is shown in Fig. 6. Following the QC/T491—2018 standard, a sinusoidal excitation of 0.05 m amplitude and 1.67 Hz frequency was applied, and the resulting simulation data were analyzed.

As shown in Fig. 7, during the extension stroke, the damper's peak resistance is around 1700 N, meeting the QC/T491—2018 standard for a 40 mm diameter damper (1600 N to 10,000 N). In the compression stroke, the peak resistance is around 1000 N, also within the standard range (400 N–2200 N). The continuous variation in damping force results from the torque transmitted by the piezoelectric device to the hydraulic motor. Figure 8 shows that the piezoelectric suspension provides greater damping force at lower piston speeds, aided by the accumulator's filtering effect. With a resistance of 5 kΩ, the piezoelectric damper generates an average of 0.4 W of energy

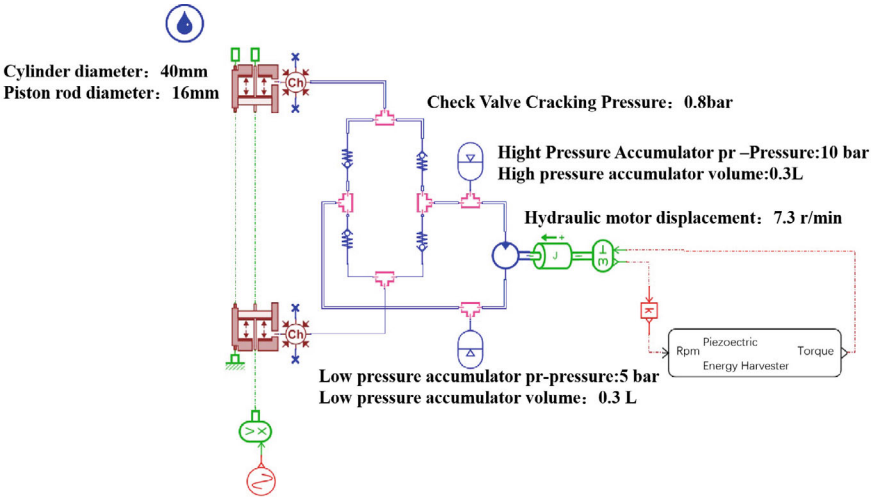


Fig. 6 Simulation model

while providing damping force (Fig. 9). A simulation model of the piezoelectric energy-harvesting damper was constructed for comparison with a traditional hydro-hydraulic suspension system. Both systems underwent a step excitation of 0.05 m. As shown in Fig. 10, the energy-harvesting suspension attenuates vibrations faster, stabilizing the vehicle body more quickly than the traditional system, although its peak displacement is slightly larger.

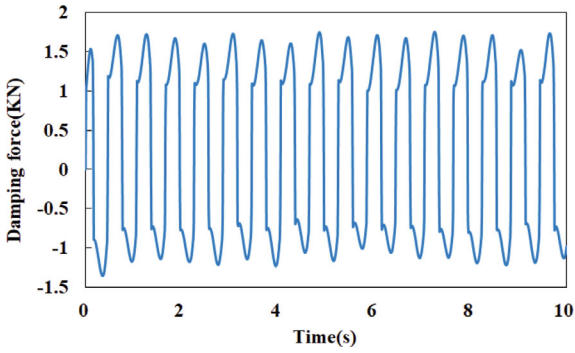


Fig. 7 Damping force acting on the piston

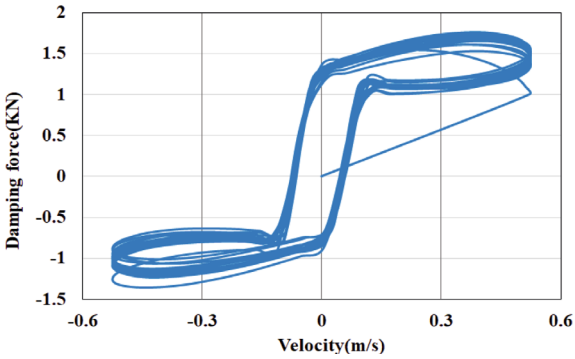


Fig. 8 Velocity diagram

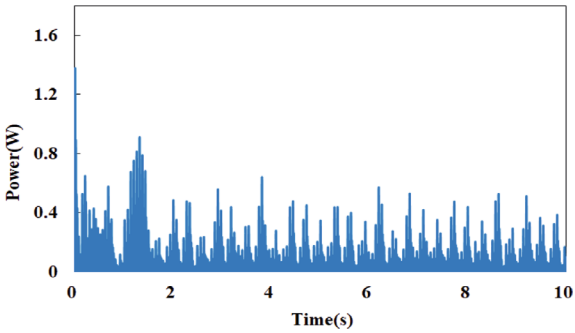


Fig. 9 Power diagram of piezoelectric energy feeder

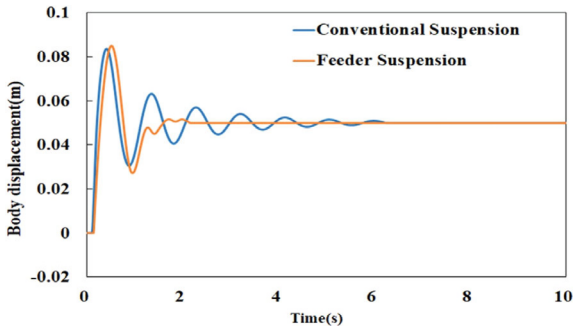


Fig. 10 Body displacement under step excitation

5 Conclusion

This study explores vibration energy recovery in vehicle suspension systems and proposes a piezoelectric energy recovery-based design, verified through simulation and experimentation. Key findings include:

(1) The piezoelectric suspension effectively converts vibration energy into electricity, enhancing fuel efficiency and outperforming traditional systems in vibration attenuation, as demonstrated by AMESim and Simulink co-simulation.
(2) Further optimization of the structural design is required to improve the energy conversion efficiency, and the experimental validation was carried out under specific conditions that may not be representative of all real-world situations.

Acknowledgements This work was supported by the National Natural Science Foundation of China, China (No. 52375046).

References

1. Hussain SA, Razi F, Hewage K et al (2023) The perspective of energy poverty and 1st energy crisis of green transition. Energy 275
2. Macedo DP, Marques AC (2023) Is the energy transition ready for declining budgets in RD&D for fossil fuels? Evidence from a panel of European countries. J Clean Prod 417:138102
3. Guo H, Wen Z, Zi Y et al (2016) A water-proof triboelectric–electromagnetic hybrid generator for energy harvesting in harsh environments. Adv Energy Mater 6(6):1501593
4. Cao R, Zhou T, Wang B et al (2017) Rotating-sleeve triboelectric-electromagnetic hybrid nanogenerator for high efficiency of harvesting mechanical energy. ACS Nano 11(8):8370–8378
5. Minh Tri Tien C, Seo Goo N (2010) Use of a piezo-composite generating element for harvesting wind energy in an urban region. Aircr Eng Aerosp Technol 82(6):376–381
6. Pozzi M, Zhu M (2011) Plucked piezoelectric bimorphs for knee-joint energy harvesting: modelling andexperimental validation. Smart Mater Struct 20(5):055007
7. Janphuang P, Lockhart RA, Isarakorn D et al (2014) Harvesting energy from a rotating gear using an AFM-like MEMS piezoelectric frequency up-converting energy harvester. J Microelectromech Syst 24(3):742–754
8. Anon. More power from bumps in the road. https://news.mit.edu/2009/shock-absorbers-0209
9. Avadhany SN (2009) Analysis of hydraulic power transduction in regenerative rotary shock absorbers as function of working fluid kinematic viscosity. Diss. Massachusetts Institute of Technology
10. Fang Z, Guo X, Xu L et al (2013) Experimental study of damping and energy regeneration characteristics of a hydraulic electromagnetic shock absorber. Adv Mech Eng 5:943528
11. Zhang Y, Zhang X, Zhan M et al (2015) Study on a novel hydraulic pumping regenerative suspension for vehicles. J Franklin Inst 352(2):485–499

Open Access This chapter is licensed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License (http://creativecommons.org/licenses/by-nc-nd/4.0/), which permits any noncommercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if you modified the licensed material. You do not have permission under this license to share adapted material derived from this chapter or parts of it.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.



Structural Optimization Design of Cutter Plate of Grapevine Burying Machine Based on Response Surface Methodology



Beibei Zhang, Julaiti Maitirouzi, Hongcheng Deng, Haileiqguli Yasen, and Wenjing Bai

Abstract In order to solve the problems of large volume, bulky structure and low efficiency of soil throwing of blade-type grapevine burying machine, parametric modeling of the key components of the vine burying machine cutter plate, on the basis of ANSYS Workbench static analysis, using the relevant parameter sensitivity method to screen out the design variables that have a greater impact on the performance of the cutter plate, and completing the design of experimental type by using the CCD, and fitting the design variables with the response surface of the output parameters through the nonparametric regression meta-model, and finally completing the optimization of the cutter plate based on the MOGA algorithm. The response surface of the design variables and the output parameters are fitted by the nonparametric regression meta-model, and finally the optimization of the cutter plate is completed based on the MOGA algorithm. The optimized cutter plate structure meets the working strength requirements, and the mass decreases by 19.2% and the equivalent stress decreases by 8.3% compared with the original structure. The simulation results show that the cutter plate is lightweight obviously, which verifies the feasibility of the response surface-based optimization design, and this research process has certain reference significance for the research in the field of structural optimization.

Keywords Grapevine burial · Response surface · Sensitivity · Experimental design · Optimization design

B. Zhang · J. Maitirouzi (✉)
College of Mechanical Engineering, Xinjiang University, Urumqi, China
e-mail: jurat@xju.edu.cn

H. Deng · H. Yasen · W. Bai
Tulufan Agricultural and Rural Mechanization Development Center, Tulufan Municipal Bureau of Agriculture and Rural Affairs, Tulufan, China

© The Author(s) 2026

H. Saman et al. (eds.), *The 9th International Conference on Advances in Construction Machinery and Vehicle Engineering*, Lecture Notes in Mechanical Engineering,
https://doi.org/10.1007/978-981-96-5527-4_76

1 Introduction

China's Xinjiang grape growing area is large [1], before the winter need to bury the vine to avoid damage to grape vines [2–5], manually burying the vine cost is high, low efficiency [6, 7], the existing burying the vine machinery model type variety, but no special adaptation to the climate and soil in Xinjiang [8, 9].

At present, a number of blade-type vine-burying machines have been designed for the northwestern region, such as the grapevine-burying machine 1PM-100 designed by Li Huanfang, which has an adjustable digging trench depth and mulching volume [10], and the grapevine-burying machine 3PMT-62 designed by Liu Wenying, which can achieve the effect of left and right sliding obstacle avoidance [11], but they only designed the machine, and they did not optimize and analyze the soil-throwing device of the grapevine-burying machine further, and the grapevine-burrowing machine 3PMT-62 designed by Liu Wenying was designed to avoid obstacles [12], but it did not analyze the strength and the influence of the soil-throwing device of the blade type on mulching. The blade-type soil-throwing device was studied and analyzed to determine the influence of the rotational speed of the soil-throwing wheel, the depth of soil entry and the angle of soil-throwing on the effect of mulching [12], but the strength of the soil-throwing device was not analyzed as well as further lightweighting was not carried out. The important structural component of the blade type soil throwing device is the cutter plate, so this paper takes the cutter plate as the focus of research. Cai Sheng pointed out that the structural lightweighting in agricultural machinery lightweighting has size optimization, shape optimization, topology optimization and multidisciplinary design optimization [13]. Hao Tengda, after parameterizing the hopper desanding vehicle, carried out finite element analysis of the hopper through the dem-fem method, and used the response surface method to optimize the structural design of the hopper [14].

Here we take the soil-throwing cutter plate of the grapevine burying machine as the research object, use ANSYS to carry out static analysis, and then through the sensitivity analysis of the relevant parameters, screen the factors that have a greater impact on the output parameters; based on the response surface method, in order to satisfy the strength requirements, the use of multi-objective algorithms on the structure of soil-throwing cutter plate is optimized to achieve a reduction in the mass.

2 Parametric Model

Blade type grapevine burying machine is a commonly used mechanical equipment for grapevine burying in Turpan area of Xinjiang, removing the tractor that provides the power, and its three-dimensional model is shown in Fig. 1, which is mainly composed of three-point suspension, telescopic obstacle avoidance device, transmission device, and soil throwing device. The tractor drives it forward, in which the cut and throw

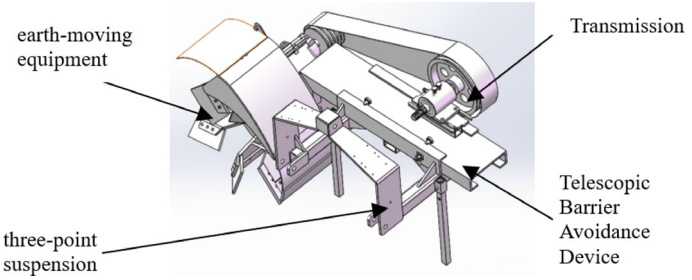
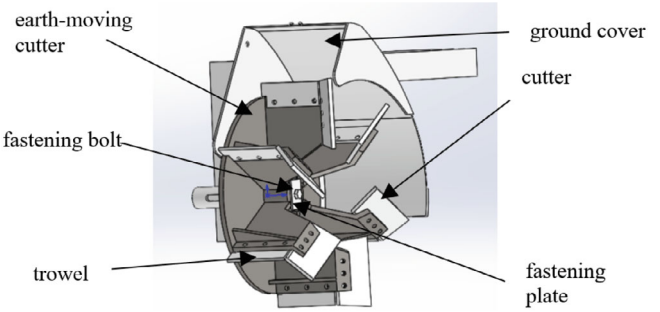


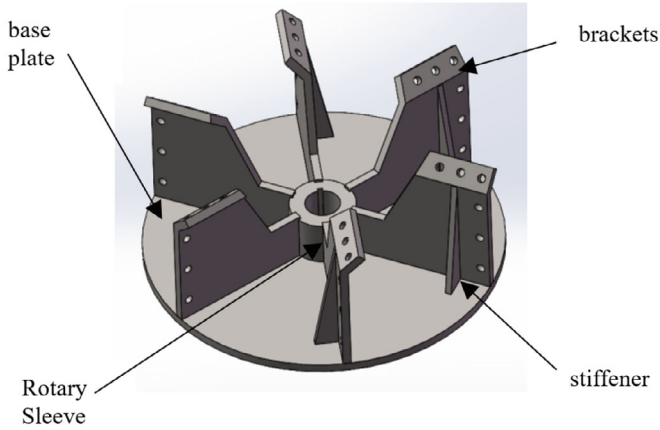
Fig. 1 Structure of the vine-burying machine

device is the core device of the grapevine burying machine, as shown in Fig. 2a. It mainly consists of soil throwing cutter plate, soil cutting knife, soil throwing knife, soil guide cover, fastening plate, fastening bolt and so on. When working, the cutter plate is driven by the output shaft of the reducer for rotary movement, the tractor drives the machine along the direction of the grapevine, cut and throw the soil device along the forward direction of the soil cut, through the soil throwing knife to give the soil a certain initial velocity, and in the soil guide cover device under the guidance of the soil thrown to the grapevine to realize the burying of the vine mulching soil. The soil throwing cutter plate is an important structural component of the soil throwing device, as shown in Fig. 2b. In the process of soil throwing, most of the soil and the knife and cutter plate on the soil throwing knife and cutter plate direct contact and load pressure on the cutter plate, making the cutter plate deformation and abrasion, so the structure of the cutter plate directly affects the whole grapevine burying machine operating quality and efficiency.

Parametric design of the cutter plate was carried out, firstly, based on the assembly requirements, pre-tests combined with agronomic requirements, the parameters available for optimization were screened. As shown in Fig. 3. Where P11 is the thickness of the bottom plate of the cutter plate, P6, P7 is the thickness of the blade holder, P1 is the thickness of the reinforcing bar, P10 is the outer diameter of the rotary sleeve, P2, P3 are the length and width of the reinforcing bar, P8 is the distance from the center of the rotation of the reinforcing bar, P5 is the bracket close to the side of the rotary shaft and the rotary axis of the angle of the rotary shaft, P4 is the bracket close to the height of the side of the rotary shaft, P9 is the height of the rotary sleeve. Where P9 is greater than p4 to optimize the thickness and cross-sectional dimensions of the cutter plate for lightweighting.

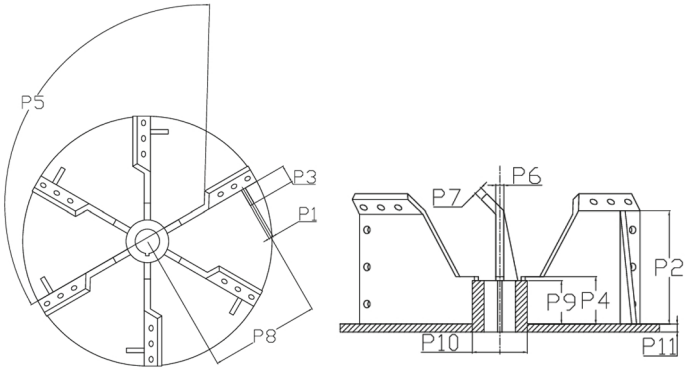


(a) Cutting and Throwing Device



(b) Cutter plate

Fig. 2 Structure of cut-and-throw device and cutter plate



(a) (b)

Fig. 3 Parameter dimensions of the cutterhead

3 Finite Element Analysis of Cutterheads

3.1 Discrete Element Modeling

In order to get the cutter plate under extreme working conditions closest to the actual external load, the 3D model was imported in edem software and the soil model was built with the discrete element parameters as shown in Table 1, and the discrete element parameters are referenced in literature [14–16].

According to the morphological characteristics of soil particles, in order to facilitate simulation and calculation, the soil particle model is simplified to a single ball, omitting bearings, bolts, nuts, etc. in the device. In the process of EDEM simulation, due to the presence of water molecules in the soil, so the soil particles have a bonding characteristic between them, and in order to simulate the interactions between the materials more accurately, the contact model between the soil is adopted as HertzMindlin with JKR Cohesion, and the Hertz-Mindlin (no slip) model was used between the soil particles and the overburden device. The contact parameters between materials and contact model parameters are shown in Table 2.

Table 1 Intrinsic parameters of materials

Sports event	Causality	Value
Soil particles	Poisson's ratio	0.38
	Shear modulus (Pa)	1×10^7
	Densities (kg/m^3)	1850
	Moisture content	10.46%
Earth-moving cutter plate	Poisson's ratio	0.3
	Shear modulus (Pa)	7.99×10^{10}
	Densities (kg/m^3)	7800

Table 2 Contact parameters and contact model parameters

Sports event	Causality	Value
Soil-soil	Coefficient of restitution	0.6
	Coefficient of static friction	0.6
	Coefficient of kinetic friction	0.4
	JKR(J/m^{-2})	8.7
Soil-throwing cutter plate	Coefficient of restitution	0.6
	Coefficient of static friction	0.6
	Coefficient of kinetic friction	0.05

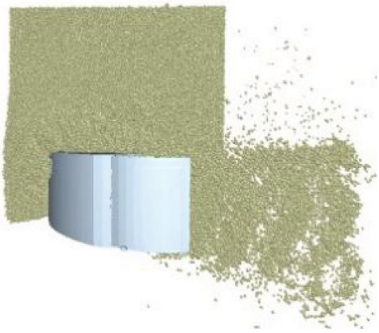


Fig. 4 Working model of the mulching device

Parameters of extreme working condition: forward speed is 1.2 km/h, tractor power output shaft is 540 rpm, transmission ratio is 2.2:1, the rotational speed of the cutter plate is set to 245.5 rpm in EDEM, and the forward speed is 0.333 m/s. Physical radius of the particles is 5 mm, and the radius of the contact is 6 mm, and the simulation step is set to 20%, and the grid size is 3 times of the particle radius. Considering the simulation time and its stability, the simulation step size is set to 20%, the grid size is 3 times of the radius of the particles, and the simulation time is set to 7 s. When the torque of the cutter plate reaches the maximum, i.e., the external load also reaches the maximum accordingly, the output of the load file of the mulching device at this moment. Figure 4 shows the working model of the mulching device.

Figure 4 shows the working model of the mulching device.

3.2 Finite Element Modeling

As the load on the cutter plate mainly comes from the throwing knife and the cutting knife, the parameterized model imported into the static structure module of ANSYS Workbench is equipped with the cutting knife and the throwing knife, and then the cutter plate is given material properties. Due to the strong comprehensive ability of Q345 steel, good plasticity and high yield strength, the cutter plate material selection of Q345 steel, Q345 steel material properties and some mechanical properties, as shown in Table 3.

Table 3 Material properties of Q345 steel

Yield strength	Modulus of elasticity $\times$ 10^5 MPa	Poisson's ratio	Densities $\times$ 10^3(kg·m^{-3})
345	2.1	0.3	7.8

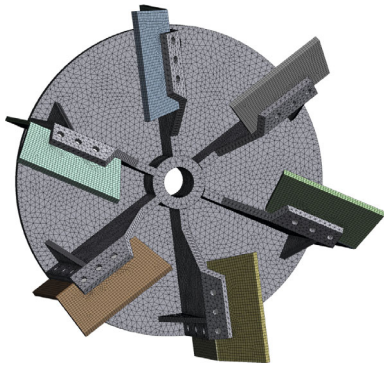


Fig. 5 Finite element mesh model

The automatic meshing is used to divide the mesh, the mesh size is 5 mm and the rest is kept as default, as shown in Fig. 5. The whole finite element model has 269,880 nodes and 106,918 cells.

3.3 Boundary Conditions and Loads

The cutter plate, as a mulching component, is subjected to loads mainly from the external pressure of the soil throwing knife and the soil cutting knife, and a small portion of the soil pressure on the bracket and the side wall is negligible, while the force of the soil throwing knife and the soil cutting knife comes from the reaction force of the soil on it, so it is added to the finite element simulation model of the soil throwing knife and soil cutting knife, in order to simplify the simulation model to improve the efficiency of the simulation, the cutter plate is connected with the soil throwing knife and the soil cutting knife. In order to simplify the simulation model and improve the simulation efficiency, the actual bolts connecting the cutter disk with the throwing knife and cutting knife are removed and connected in the form of a beam. The inner wall of the rotary sleeve of the cutter plate is used as a fixed constraint. The maximum pressure value of the soil throwing wheel is imported into ANSYS Workbench, and the load applied to the cutter plate is completed through the static coupling of Edem and Ansys.

The yield strength of Q345 profile is 345 MPa

$$[\sigma] = \sigma/n = 259 \text{ MPa} \tag{1}$$

The cutter plate meets the strength requirement when the maximum stress is less than 259 MPa, Take the safety factor n = 1.33[17].

Where: σ-material yield strength;

n-Safety coefficient.

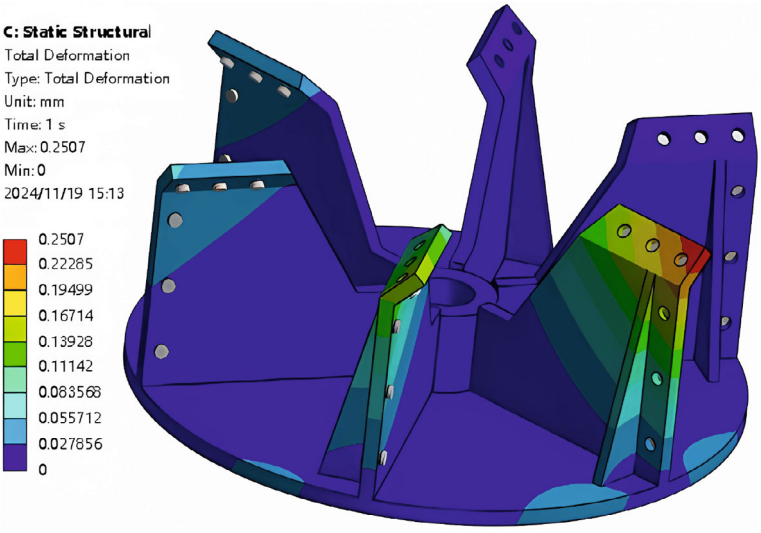


Fig. 6 Cloud view of cutter plate deformation

According to the above boundary conditions, the static characterization of the cutter plate is carried out, and the deformation cloud of the cutter plate and the equivalent force distribution cloud are obtained, as shown in Fig. 6.

From Fig. 6, the maximum deformation of the cutter plate occurs in the cutter plate bracket away from the outer side near the outer side of the position connected to the cutter knife, the maximum value of 0.2507 mm, the rest of the part of the deformation is not obvious, from Fig. 7, the maximum equivalent stress of the cutter plate is 52.564 MPa, the maximum position occurs in the bracket and the cutter plate connected to the reinforcing bar, which is located in the reinforcing bar close to the cutter plate and the rotary axis. The maximum equivalent stress is much less than the allowable stress of the material, the cutter plate to meet the strength and design requirements.

4 Optimization of Methodological Processes

Here the finite element-based response surface method is used to optimize the design of the cutter plate structure, and the optimization process is shown in Fig. 8. Firstly, the sensitivity analysis is carried out in the design size to screen out the variables to be optimized, then the experimental design is carried out to generate the sample points, and finally, the proxy model of the optimized variables and the response objective is fitted, and the optimal calculation of genetic algorithm is carried out on the proxy model to select the best optimization result.

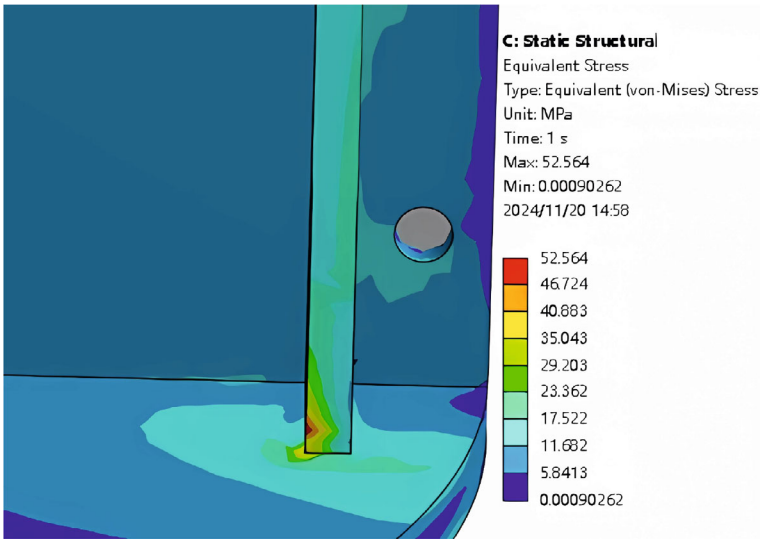


Fig. 7 Equivalent force on the cutter plate

5 Sensitivity Analysis of Parameters Related to Cutter Plate Structure

In optimal design, the design variables need to be optimized, but too many parameters will lead to an increase in the optimization time. Through the parameter sensitivity analysis, the influence of each parameter on the response can be distinguished, and the parameters that have less influence on the response surface target can be eliminated, so as to find the structural design parameters that have a significant influence on the response of structural characteristics among multiple parameters [18], and then carry out optimization on them. In this optimization, based on the finite element static analysis, the sensitivity analysis of the mass of the cutter plate and the maximum equivalent force is carried out to filter out the design variables with large influence. According to the three-dimensional modeling dimensions of the cutter plate, there are as many as 11 parameters in the optimization model, as shown in Fig. 9, where the value of P6 is equal to the value of P7, so the value of P7 is expressed by P6.

The Spearman method is used to analyze the sensitivity of 10 parameters. In order to ensure the accuracy, the number of samples should be about 10 times of the design variables, and here 100 groups of sample points are set to be solved; the automatic stop is started; the accuracy of the average value is 0.01 mm; the accuracy of the standard deviation is 0.02 mm, and the design variables of the cutter plate are the sensitivity to the mass and the maximum equivalent force, as shown in Fig. 10.

A positive vertical coordinate of the bar graph indicates a positive correlation between the design variables and the output parameters and vice versa. From Fig. 9, it can be seen that the parameters P5, P6, and P11 have the most significant effect on

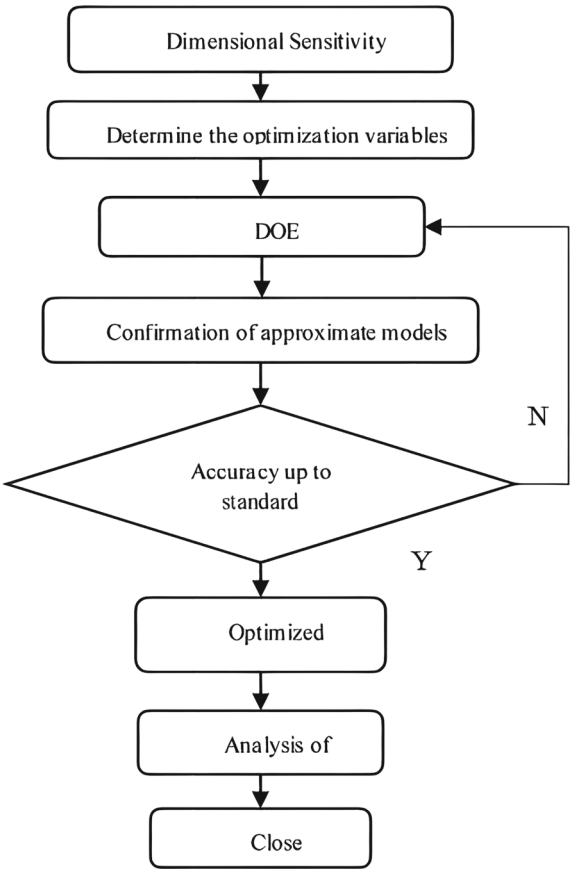


Fig. 8 Optimization flow chart

the mass of the cutter plate and are positively correlated with the mass of the cutter plate, which increases with the increase of the three dimensions. The design variables P1, P3, P6, P8 and P11 have a greater influence on the maximum equivalent force of the cutter plate. The design variables P1, P3, P6, P8 and P11 have a greater influence on the maximum equivalent force of the cutter plate, among which P8 is positively related to the maximum equivalent force of the cutter plate, and the larger the size of P8, the larger the maximum equivalent force of the cutter plate. According to the above results, the parameters with small sensitivity are ignored, and the thickness of the reinforcement P1, the width of the reinforcement P3, the angle of the bracket close to the side of the rotary axis P5, the thickness of the bracket P6, the distance of the reinforcement from the center of the rotary axis P8, and the thickness of the bottom plate of the cutter plate P11 are selected as the design variables in the optimization process. In order to ensure that no dimensional interference occurs in the original structure, the range of variables is determined based on the design experience of the cutter plate as shown in Table 4.

A	B	C	D
	A	B	C
1	ID	Parameter Name	Value
2	Input Parameters		
3	Geometry		
4	P1	DS_D10	10
5	P2	DS_D9	198
6	P3	DS_D8	44
7	P4	DS_D5	83
8	P5	DS_D6	121.5
9	P6	DS_D4	14
10	P7	DS_2D4	14
11	P8	DS_D7	245
12	P9	DS_D3	76
13	P10	DS_D2	96
14	P11	DS_D1	14

Fig. 9 Parameters related to the cutter

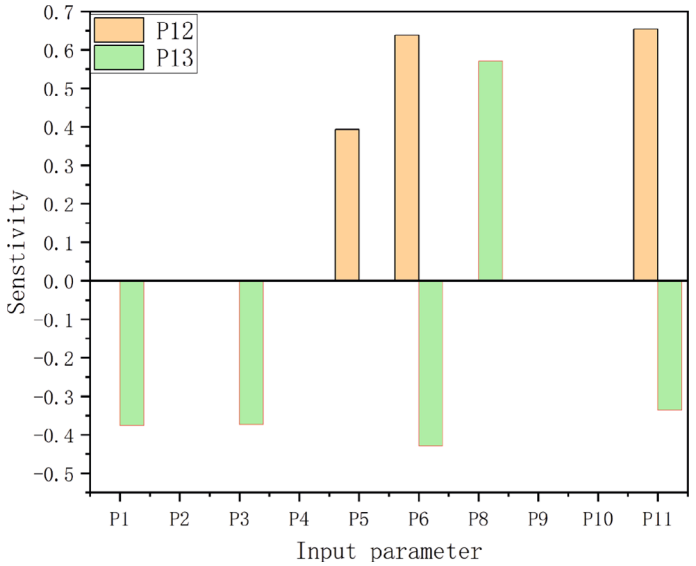


Fig. 10 Correlation sensitivity analysis

Table 4 Cutter plate optimization variables

Name	Optimization variables	Starting value	Scope of change
Thickness of reinforcement /mm	P1	10	7 ~ 18
Width of reinforcement /mm	P3	44	20–60
Angle of the bracket near the side of the rotary axis /(°)	P5	121.5	90–130
Bracket thickness /mm	P6	14	10–18
Distance of reinforcement from the center of rotation/mm	P8	245	175–245
Thickness of cutter base plate / mm	P11	14	10–20

6 Response Surface Construction

6.1 Design of Experiments

Response surface method is an optimization method based on experimental design, which improves the optimization efficiency and accuracy by fitting a nonlinear polynomial function to approximate the relationship between design variables and output parameters. The method can simplify the optimization process and shorten the computation time. The experimental design method is a method used to extract the most representative combination from the complete set of factor level combinations to be tested [6]. Here, the central composite design of experiments was chosen, which is a commonly used response surface design to quantify the relationship between the output response and the input response of one or more measurements [19], and the design type was set as self-defined. The design variables P1, P3, P5, P6, P8, and P11 with larger sensitivity after screening were used as experimental designs, and mass (P12) and maximum equivalent force (P13) were used as output parameters to obtain 45 sets of experimental sample points, and the experimental results, as shown in Table 5.

6.2 Response Surface Analysis

Based on the experimental design, a response surface model was fitted for the above six variables for mass and maximum equivalent force.

From Fig. 11, it can be seen that the quality of the blade is positively correlated with the design variables P5 and P6, in which the influence of the design variable P6 on the quality of the blade is more obvious, and from Fig. 12 the quality of the blade increases with the increase of the design variable P11, in which the influence of the design variable P11 on the quality of the blade is more obvious.

Table 5 Pilot point results

Serial number	Optimization variables						Response parameter	
	P1/mm	P3/mm	P5/ (°)	P6/mm	P8/mm	P11/mm	P12/kg	P13/MPa
1	12.5	40	110	14	210	15	59.314	44.81
2	7	40	110	14	210	15	58.288	55.327
3	12.5	40	110	14	210	15	60.34	40.509
4	12.5	20	110	14	210	15	58.148	52.229
5	12.5	60	110	14	210	15	60.48	44.207
6	12.5	40	90	14	210	15	56.754	46.293
7	12.5	40	130	14	210	15	62.656	42.804
8	12.5	40	110	10	210	15	52.718	59.147
9	12.5	40	110	18	210	15	65.85	37.246
10	12.5	40	110	14	175	15	59.317	38.967
11	12.5	40	110	14	245	15	59.314	49.859
12	12.5	40	110	14	210	10	49.74	61.933
13	12.5	40	110	14	210	20	68.888	28.903
……								
44	15.679	28.44	121.56	16.312	230.23	12.11	59.337	45.635
45	9.3209	51.56	121.56	16.312	230.23	12.11	59.499	46.71
46	15.679	51.56	121.56	16.312	230.23	17.89	72.095	27.317

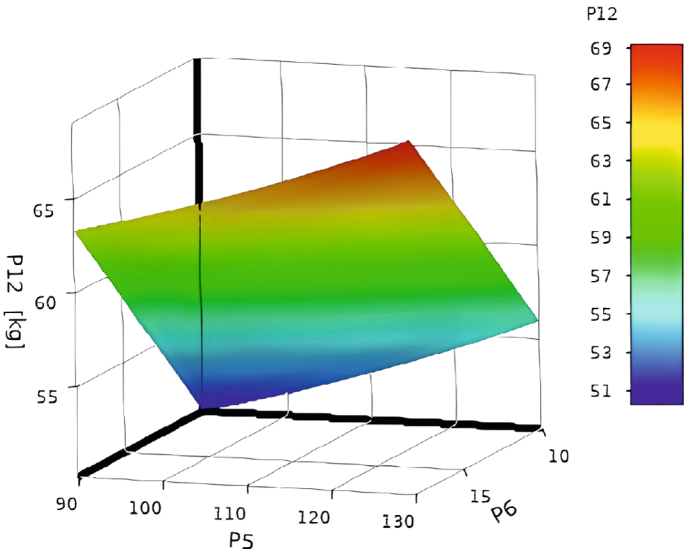


Fig. 11 P6, P5 and cutter plate mass response surfaces

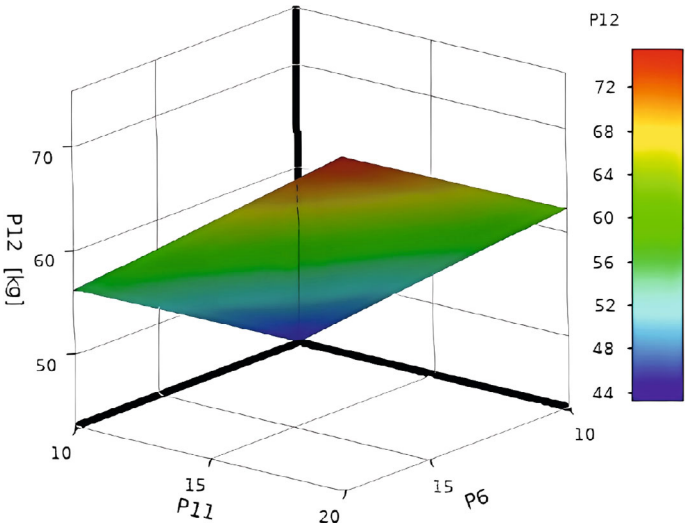


Fig. 12 P6, P11 and cutter plate mass response surfaces

From Fig. 13, it is known that the design variable P1 is negatively correlated with the maximum equivalent force and the relationship is significant; the design variable P3 decreases as the maximum equivalent force decreases first; from Fig. 14, it is known that the maximum equivalent force increases with the increase of the design variable P8 and decreases with the increase of P6 and the relationship of P6 is significant, and from Fig. 15, it is known that the maximum equivalent force decreases with the increase of the design variable P11.

The response surface accuracy was verified using the coefficient of determination (R2) and the goodness-of-fit curve. The coefficient of determination R2 is a commonly used metric to test the accuracy of the fit, and its expression is:

$$R^2 = 1 - \sum_{i=1}^{n} (y_i - \hat{y})^2 / \sum_{i=1}^{n} (y_i - \bar{y})^2 \tag{2}$$

The closer the value of R2 is to 1, the better the precision of the response surface fit and the more accurate the model. The coefficients of determination of the blade and the maximum equivalent stress are 1 and 0.99863, respectively, with high accuracy; the goodness-of-fit curve, as shown in Fig. 16. The predicted and observed values are close to the goodness-of-fit curve, which is a good fit and can be optimized with a response surface model.

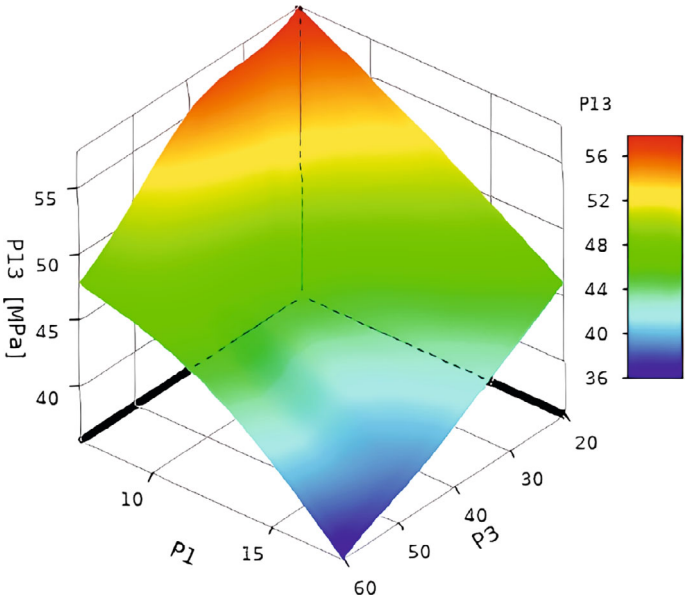


Fig. 13 Equivalent force response surface of P1 P3 and cutter plate

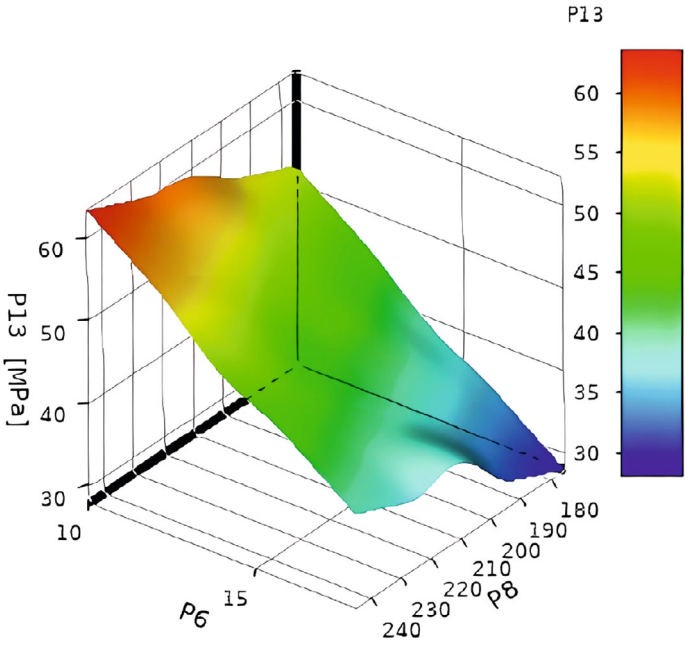


Fig. 14 Equivalent force response surface of P6 P8 and cutter plate

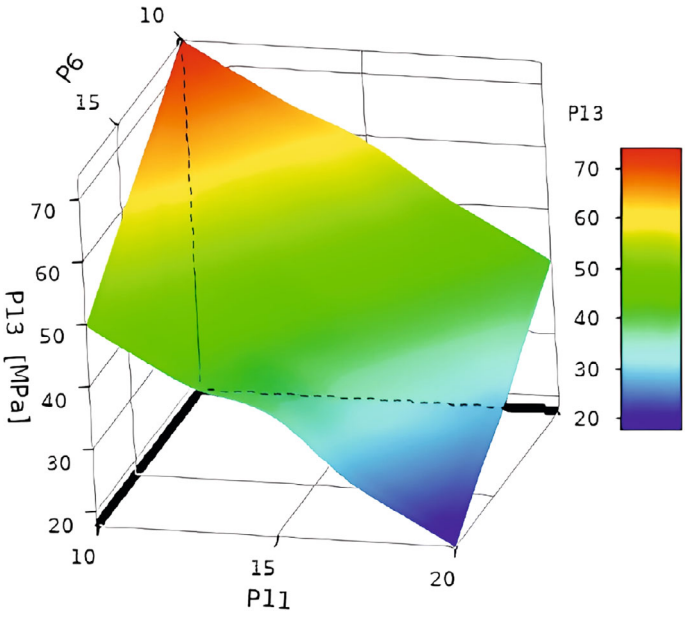


Fig. 15 Equivalent force response surface of P8 P11 and the cutter plate

Where: $\hat{y}$- the predicted value of the model;

y-Response Surface Optimization Variables Corresponding Values;
$\bar{y}$-Mean of experimental values.

7 Goal-Driven Optimization

7.1 Optimizing Mathematical Models

Objective-driven optimization is a multi-objective optimization technique that derives the "best" design point from a given set of samples. Different optimization objectives can be set for optimal design [20]. As the key motion mechanism of the grapevine burying machine, when optimizing the quality of the cutter plate, the optimal solution of the design variables should be achieved under the strength constraints. The optimization mathematical model of the cutter plate is [21]:

var x = [x1, x2, x3, x4, x5, x6]

min M = f (X)

s.t. $\sigma \leq 259$ MPa

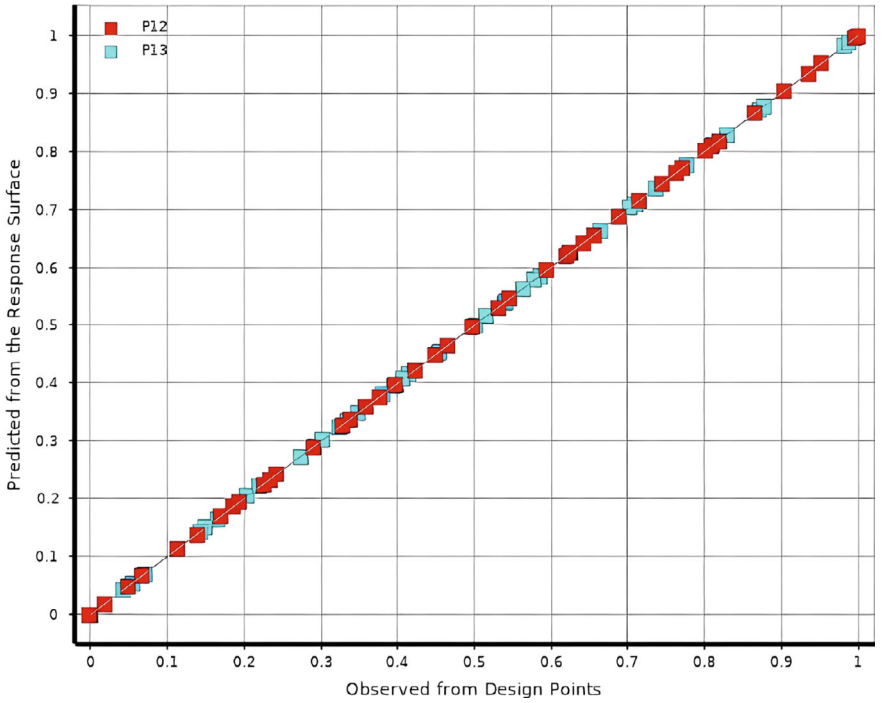


Fig. 16 Predicted versus observed values fitting curve

$7\ \text{mm} \leq \times 1 \leq 18\ \text{mm}$

$30\ \text{mm} \leq \times 2 \leq 60\ \text{mm}$

$90^\circ \leq \times 3 \leq 130^\circ$

$12\ \text{mm} \leq \times 4 \leq 16\ \text{mm}$

$175\ \text{mm} \leq \times 5 \leq 245\ \text{mm}$

$8\ \text{mm} \leq \times 6 \leq 20\ \text{mm}$

Where: x—Optimization of design variables;

M—Blade quality;

x1, x2, x3, x4, x5, x6—Optimization Variables P1, P3, P5, P6, P8, P11.

7.2 Genetic Algorithm Solving

Genetic algorithms simulate biological evolution, or "survival of the fittest", to optimize problems. It treats the possible solutions as a population, and through operations such as crossover and mutation, these solutions evolve to find a better solution.

The response surface model constructed above is optimally solved by using MOGA. To optimize the optimization, the number of initial samples is set to 600; the number of samples per sample iteration is set to 100; the maximum allowable Pareto percentage is set to 70%; the maximum number of iterations is 20; the percentage of convergence stability is set to 2%; and the maximum number of candidate points is set to 5.

7.3 Optimization Results

The optimization candidate points are shown in Table 6. All five candidate points satisfy the strength requirement and the model quality is reduced. Considering the degree of mass reduction, candidate point 1 is selected as the optimal design variable.

The dimensions of candidate point 1 are processed by taking P1 = 17.5 mm, P3 = 57 mm, P5 = 90.5, P6 = 11.5 mm, P8 = 184.5 mm, P11 = 11.5 mm, the above variables are used to regenerate the model of the cutter plate in 3D modeling software and carry out the static structural analysis, and the results are shown in Figs. 17 and 18. Under the same external load, the maximum deformation of the optimized cutter plate is 0.37 mm, the maximum equivalent force is 48.215 MPa, and the mass is 48.741 kg. Comparison of the mass and maximum equivalent force of the cutterhead before and after optimization is shown in Table 5.

According to Table 7, the maximum equivalent force of the optimized cutter plate is reduced by 8.3% compared with the pre-optimization; the deformation is increased to 0.37 mm, but the deformation is still within the permissible range of deformation, which is in line with the strength requirements; the mass of the cutter plate has been

Table 6 Optimization candidate points

Candidate point	Optimization variables						Response parameter	
	P1/mm	P3/mm	P5/ (°)	P6/mm	P8/mm	P11/mm	P12/kg	P18/MPa
1	17.585	57.136	90.067	11.707	187.97	11.493	48.267	47.866
2	17.636	56.923	90.623	11.703	184.45	11.294	47.938	48.746
3	16.955	52.12	90.003	11.773	187.79	11.714	48.394	47.87
4	16.97	51.895	91.197	11.745	187.63	11.712	48.355	47.982
5	17.586	57.854	91.628	11.737	186.44	11.459	48.468	47.736

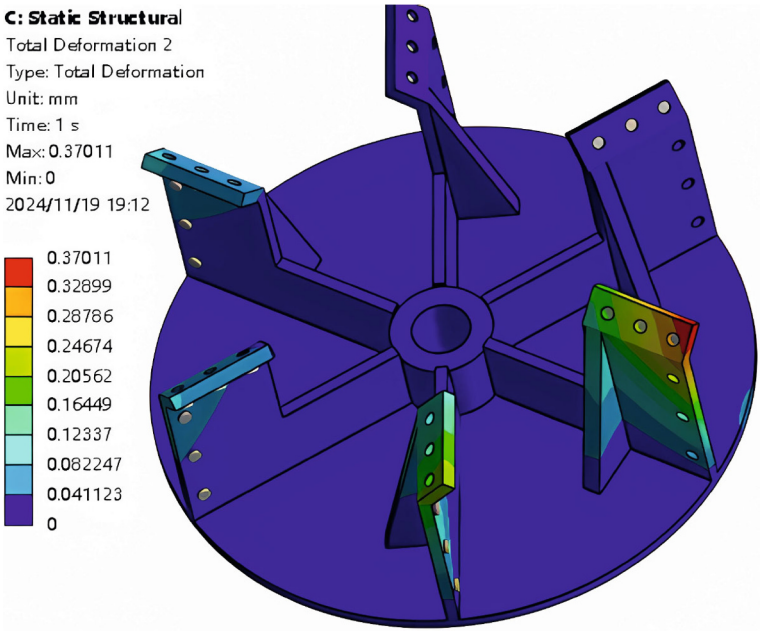


Fig. 17 Deformation after cutter plate optimization

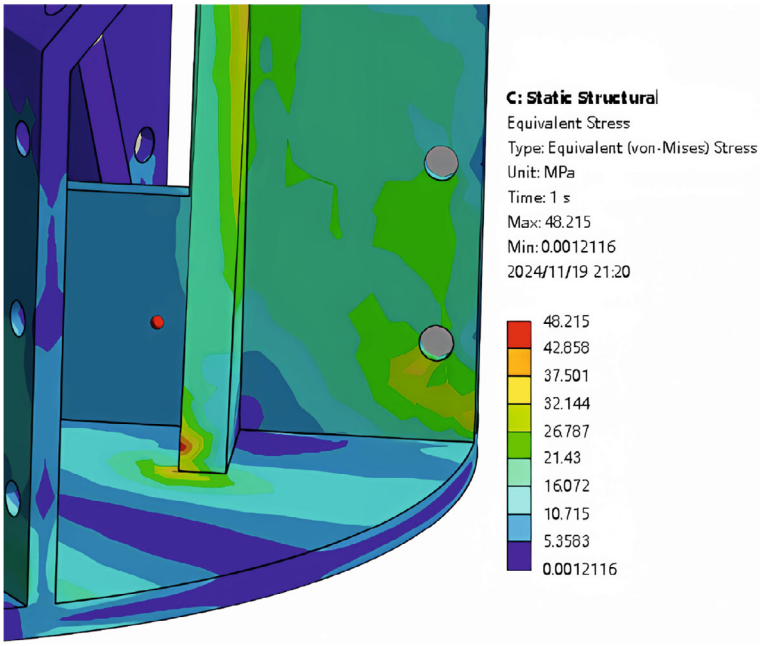


Fig. 18 Maximum equivalent force after cutter plate optimization

Table 7 Comparison of model performance before and after optimization

Goal	Pre-optimization	Post-optimization	Magnitude of change (%)
Mass /kg	60.297	48.741	19.2
Maximum equivalent force / MPa	52.564	48.215	8.3
Deformation /mm	0.25	0.37	48

reduced by 19.2%, and the weight of the cutter plate has been significantly reduced, which is obvious in the lightweighting.

8 Conclusion

(1) Based on the parameterisation of the cutter plate, a hydrostatic analysis was carried out and it was learnt that the strength of the cutter plate meets the working requirements.
(2) The sensitivity analysis of the design variables through the sensitivity of the relevant parameters concluded that the thickness of the reinforcing bar, the width of the reinforcing bar, the angle of the bracket near the side of the rotary axis, the thickness of the bracket, the distance of the reinforcing bar from the center of rotation, and the thickness of the side plate of the cutter plate have a greater impact on the performance of the cutter plate, and therefore the cutter plate optimization of the selection of these five parameters.
(3) The Central Composite Design methodology was used to experimentally analyze the design variables, establish a response surface model for the design variables and the objective function, and finally use the MOGA algorithm to structurally optimize the blade as feasible.
(4) The results show that the optimized mass is reduced by 19.2% compared with the pre-optimization, and the maximum equivalent stress is reduced by 8.3% compared with the previous one, which is a good optimization effect, and it can provide a relevant reference method in engineering optimization.
(5) This optimization uses the self-optimization module of Ansys Workbench, the optimization effect is limited, in the subsequent optimization, other experimental software such as Isight can be used to complete the optimization of high-quality response surface.

Acknowledgements This research was funded by: Digital modelling and optimal design solution for grapevine burying/opening machine, grant number (202407140014)

References

1. Bai WJ (2024) Current situation and problem analysis of mechanization of grapevine cultivation and production in Turpan City. Farm Mach 04:101–103
2. Bucur GM, Dejeu L (2016) Research on frost injury of new Romanian grapevine cultivars in the winter 2014–2015. Agric Agric Sci Procedia 10:233–237
3. Yang S, She H, Yang S (2018) Design and experimental validation of improved grapevine burying machine. Int J Agric Biol Eng 11:95–100
4. Frioni T, Green A, Emling JE, Zhuang S, Palliotti A, Sivilotti P, Sabbatini P (2017) Impact of spring freeze on yield, vine performance and fruit quality of Vitis interspecific hybrid Marquette. Sci Hortic 219:302–309
5. Li T, Kang J, Zhao WY, Sun W, Zhang FW, Zhang JL (2013) Design of a new grape vine overwintering burying machine. Acta Agric Boreali Occident Sin 22:151–156
6. Ning SC (2012) Selection and operation technology of grape vine burying machine. China Agric Mech Her 5:1–2
7. Wang TX, Yang JH, Su Y (2023) Application status of grapevine burying and clearing operation machinery in China. Sino-Overseas Grapevine Wine 6:117–123
8. Bi ZB (2016) Chinese and foreign grapes and wines. Xinjiang Agricultural University, Xinjiang
9. Ainur Y (2020) Xinjiang grape vine mechanization status quo and development direction. Agric Eng Technol 40:49–51
10. Li HF, He HYC, L Y (2019) Design and experiment of 1PM-100 grapevine burying machine. J Chin Agric Mech 40:13–18
11. Liu WY, Wu L (2015) 3PMT-62 type grapevine burying mechanization technology. Agric Mach Using Maint 5:100–101
12. Xu SJ (2019) Design and experiment of orchard straw mulching machine throwing mulching device. North West Agriculture and Forestry University, Shanxi
13. Cai S, Luo YH, Li QL (2021) State of the art of lightweight technology in agricultural machinery and its development trend. Engineering 57:35–52
14. Gao ZF, Hao TD, Zhen MJ (2024) Optimized design of hopper structure based on response surface methodology. Mach Des Manuf 8:11–16
15. Han SJ (2021) Design and testing of organic fertilizer deep applicator in Xinjiang vineyard. Shihezi University, Shihezi
16. Luo S, Yuan QX (2018) Parameter calibration of a discrete elemental method for vermicompost substrates based on the JKR bonding model. Trans Chin Soc Agric Mach 49:343–350
17. Liu RJ (2020) Design and test research of key components of railroad track desanding vehicle. Shijiazhuang Railway University, Shijiazhuang
18. Yao TB (2020) Research on structural lightweighting of stacker cranes based on response surface methodology. Xiangtan University, Xiangtan
19. Liu YH (2019) Optimized design of connecting rod structure based on response surface. North University of China, Taiyuan
20. Pu GY (2010) ANSYSWorkbench12 basic tutorial with example details. China Water Conservancy and Hydropower Press, Beijing
21. Li Y, Huang H, Guo RL (2019) Multi-level and multi-objective parameter optimization design of the machining center's vertical column structure based on the response surface method. J Mach Des 36:44–49

Open Access This chapter is licensed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License (http://creativecommons.org/licenses/by-nc-nd/4.0/), which permits any noncommercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if you modified the licensed material. You do not have permission under this license to share adapted material derived from this chapter or parts of it.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.



Multi-Element Coupling Model of Rail Grinding Vehicle



Zixuan Fan, Junjun Ding, Changquan Ding, Jianquan Zhang, Boxin Zhao, and Xiaoqian He

Abstract The operation of rail grinding vehicle is complex, involving theories of different disciplines such as dynamics, tribology, contact mechanics, hydraulics and control theory. To study the rail grinding process, a multi-software co-simulation model is established for the rail grinding vehicle is established in this paper. It includes four sub-models: dynamic sub-model of grinding trolley dynamics, rail-grinding wheel contact sub-model, hydraulic sub-model and control sub-model. The dynamic sub-model of grinding trolley can present the vertical displacement of grinding wheel, and the hydraulic system can present the contact pressure change according to the vertical displacement. According to the pressure change, the contact force and friction force between rail and grinding wheel are fed back to the dynamic sub-model to realize data circulation feedback. Finally, the dynamic parameters and hydraulic parameters are analyzed to verify that the model can reflect the real rail grinding process.

Keywords Rail grinding vehicle · Coupling model · Hydraulic · Control

1 Introduction

In recent years, with the rapid development of economy and society, the rail transit industry has developed rapidly. However, the accompanying problems of rail wear are becoming more serious. Rail vehicles will produce very complex loads and stresses during actual operation, resulting in different degrees of defects and damages on the rail surface and inside. This situation has a great impact on the comfort and safety of rail vehicles [1].

Z. Fan · J. Ding (✉) · B. Zhao · X. He
Southwest Jiaotong University, Chengdu, Sichuan, China
e-mail: dingjunjun@swjtu.edu.cn

J. Ding · C. Ding · J. Zhang
State Key Laboratory of Heavy Duty and Express High-Power Electric Locomotive, Zhuzhou, Hunan, China

© The Author(s) 2026

H. Saman et al. (eds.), *The 9th International Conference on Advances in Construction Machinery and Vehicle Engineering*, Lecture Notes in Mechanical Engineering,
https://doi.org/10.1007/978-981-96-5527-4_77

Rail grinding is an effective and active means to remove rail defects. This is to grind the defective layer to control the development of defects, and prolong the service life of rail through planned preventive grinding on the same time. This method has been widely used in many countries and regions [2].

The rail grinding vehicle appeared in the 1960s and soon became the workhorse for rail maintenance. The rail grinding vehicle grinds rails with its grinding wheels, so its dynamic performance in operation will interact with the whole grinding process. In addition, the grinding system consists of different modules, such as mechanical module, hydraulic module and control module. Therefore, the grinding process is complex and diversified. However, most existing research results are addressed to single module or joint simulation of a few modules. Such modeling methods could not well reflect the whole grinding process and did not present mutual influence caused by coupling of different parts, and there is some deviation in the analysis process. On the basis of existing research results, a multi-element coupling model of rail grinding vehicle is established in this paper [3].

2 Introduction to Operation Principle of Rail Grinding Vehicle

The GMC-96 × rail grinding vehicle is composed of 5 cars and it is 113.8 m long. Car 1 and car 5 are control cars with cabs, and there are three grinding trolleys under each car. Car 2 and car 4 are grinding cars, with three grinding trolleys under each car. Car 3 is a motor car to provide power for the grinding and travelling of the vehicle. The axle type is C0-C0. The maximum speed is 100 km/h, and accurate control can be realized in the low-speed range (3–24 km/h), with a control accuracy of ± 1.5 km/h.

The grinding trolleys is the main device for grinding. There are 12 grinding trolleys on GMC-96 × rail grinding vehicle, and each grinding trolleys is equipped with 8 grinding units, totaling 96 grinding units. Each grinding unit consists of five parts: deflection mechanism, lifting mechanism, grinding motor, grinding wheel and cradle.

During the grinding, the deflection mechanism fixes the grinding unit at a specified angle. The lifting mechanism controls the vertical lifting of the grinding motor and the grinding wheel through the hydraulic system to make the rail contact with the grinding wheel and provide a certain pressure. The grinding motor drives the grinding wheel to rotate rapidly to convert the vertical pressure into the grinding force. The rail grinding vehicle drives the whole system to move forward along the rail at a fixed speed. The above movements enable the constant power of grinding wheel for rail grinding. Each grinding wheel can be set at different angles and different power, and multiple grinding wheels can be arranged to form the target profile of rail.

3 Introduction to Overall Model

The model in this paper is established based on the above contents, and it consists of dynamic sub-model of grinding trolleys, rail-grinding wheel contact sub-model, hydraulic sub-model and control sub-model. The model principle is shown in Fig. 1.

Where, D_z, V_x, V_y and V_z are the vertical displacement and longitudinal, transverse and vertical speeds of the grinding trolleys respectively; Z_0 is the rail irregularity; Z_2 and V_Z are the vertical displacement and vertical displacement speeds of the grinding wheel; P_P is the pressure at the end of the rodless cavity of the hydraulic cylinder; I is the current of the grinding motor; F_N is the pressure provided to the rail surface vertically by the grinding wheel; f_x and f_y are the longitudinal and transverse grinding forces on the grinding wheel respectively; T is the grinding torque.

3.1 Dynamic Sub-Model of Grinding Trolleys

When the rail grinding vehicle is working, the lifting cylinder is at a low position to lower the grinder. At this time, the oil cylinder is free and unstressed. The vertical and lateral vibration of the carbody hardly affects the grinding trolleys, but only a traction force is provided in the longitudinal direction to make the grinding trolleys move at a constant speed. The rail grinding vehicle can therefore be neglected in the modeling.

In the grinding trolleys, the tilt cylinder is used to control the angle of the cradle. During operation, the cradle is set at a fixed angle and locked. There is no degree of freedom between the cradle and the grinding trolleys body. The cradle is linked to the grinding motor and grindstone through a feed cylinder and a guide post, with only vertical degree of freedom. The feed cylinder is modeled by AMEsim, and the force is fed back in real time according to the operating conditions. For other structural parameters and stress conditions, see reference [4].

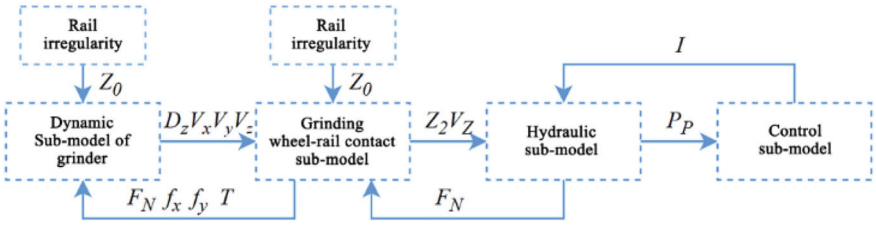


Fig. 1 Schematic diagram of rail grinding vehicle model

3.2 Rail-Grinding Wheel Contact Sub-Model

In the process of rail grinding, the grinding wheel advances along the rail to generate longitudinal grinding force; it laterally follows the displacement of the grinding trolleys to generate transverse grinding force; vertical displacement occurs along the irregularity of the vertical rail, driving the hydraulic rod to move and affecting the positive pressure in contact between the rail and the grinding wheel; the grinding motor drives the grinding wheel to rotate at high speed to form a rotary grinding force, generating a rotating torque, as shown in Fig. 2.

In the process of modeling, Simpack, AMEsim and Simulink are used for joint simulation. The rail contact force element is used to simulate the contact between rail and grinding wheel in Simpack, and the vertical displacement and velocity of hydraulic rod can be obtained by inputting rail irregularity. The change of positive pressure can be obtained by inputting the vertical displacement and speed of the hydraulic rod in AMEsim. Simulink is used to link AMEsim with Simpack, and input the positive pressure between rail and grinding wheel read from AMEsim, calculated transverse and longitudinal grinding forces and grinding wheel rotating torque into Simpack model. The specific logic diagram is shown in Fig. 3. D_z is the vertical position of the grinding wheel, and V_x and V_y are the longitudinal and lateral speeds of the grinding wheel respectively.

The grinding force of the grinding wheel consists of three parts: longitudinal grinding force, lateral grinding force and grinding torque. The grinding force consists of cutting deformation force and friction force. The grinding force generated by the contact between grinding wheel and rail can be obtained according to reference [5].

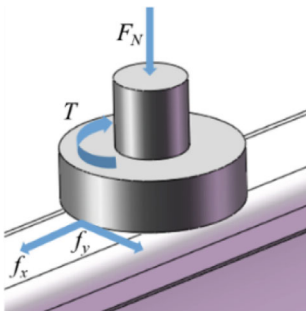


Fig. 2 Grinding wheel stress sub-model

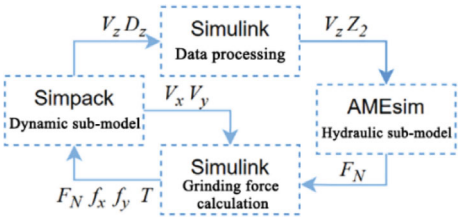


Fig. 3 Logic diagram of rail-grinding wheel contact sub-model

$$F = F_c + F_s = \frac{\upsilon P}{(\pi dn)^2} + \mu N \tag{1}$$

where, F_c and F_s represent the cutting deformation force and friction force respectively. The calculation results show that the friction force is much greater than the cutting deformation force. Therefore, the grinding force is calculated by amplifying the friction force by a certain factor.

The friction force generated by rail-grinding wheel contact is generally in the opposite direction of speed, so the lateral friction force fy and longitudinal friction force fx are as follows respectively [6]:

$$f_y = F_f \frac{V_y}{\sqrt{V_x^2 + V_y^2}} \qquad f_x = F_f \frac{V_x}{\sqrt{V_x^2 + V_y^2}} \tag{2}$$

The calculation formula of grinding torque is as follows:

$$dS = 4r \ \arcsin\left(\frac{D}{2r}\right) dr \quad R_1 < r < R_2 \tag{3}$$

$$T = \int ra\mu \frac{F_N}{S} dS \quad R1 < r < R2 \tag{4}$$

where, dS is the micro-annular area of rail-grinding wheel contact surface, D is the grinding width, r is the distance from the micro-annular to the center of grinding wheel, R_1 and R_2 are the inner diameter and outer diameter of grinding wheel, T is the grinding torque, a is the gain coefficient, μ, F_N and S are the friction factor, positive pressure and contact area between rail and grinding wheel [7].

3.3 Hydraulic System Sub-Model

The hydraulic system consists of feed cylinder, activation valve, pressure proportional valve, orifice, overflow device, accumulator and pump.

When the hydraulic system operates normally, the activation valve is fully open. In case of any abnormality, the activation valve is closed and the feed cylinder rises to prevent the pressure from being out of control and damaging the rail [8].

The feed cylinder is used to control the lifting and lowering of the grinding motor and the pressure F_N generated between the rail and the grinding wheel. The calculation formula of the force provided by the feed cylinder is as follows:

$$N = \frac{\pi P_P D_P^2}{4} - \frac{\pi P_r (D_P - D_R)^2}{4} \tag{5}$$

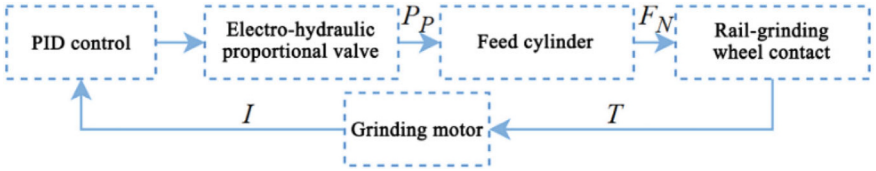


Fig. 4 Grinding pressure control system

where, N is the pressure that can be provided by the feed cylinder; P_P is the end pressure of rodless cavity, P_r is the pressure of rod cavity; D_P is the piston diameter; D_r is the piston rod diameter.

The vertical displacement and velocity of grinding wheel are input through joint simulation to simulate the displacement of grinding wheel caused by rail irregularity, driving the hydraulic rod to displace, affecting the pressure at the end of rodless cavity, and finally affecting the positive pressure generated by contact.

3.4 Control System

When the grinding system is running, the hydraulic cylinder pressure will fluctuate due to rail irregularity, affecting the rail grinding quality. Therefore, a pressure control system is necessary to keep the grinding pressure constant and realize constant power grinding. The power calculation formula is as follows [9]:

$$P = \sqrt{3} \times U \times I \times \cos\varphi \tag{6}$$

where P is the effective power of the motor, U and I are the voltage and current of the grinding motor, and $\cos\varphi$ is the power factor.

When a grinding command is received, the electrical signal activates the pressure proportional valve and activation valve. The hydraulic oil enters the rodless cavity end of the feed cylinder, and the piston rod extends until it contacts the rail. The pressure at the rodless cavity end rises until the power of the grinding motor reaches the set value. When the rail is irregular or voltage fluctuates, the control system controls the pressure at the rodless cavity end through the proportional valve and controls the positive pressure of the grinding wheel to stabilize the grinding power within a certain range and ensure the grinding quality. The control logic is shown in Fig. 4.

3.5 Overall Joint Simulation Model

Simulink is used to link the above sub-models to realize joint simulation.

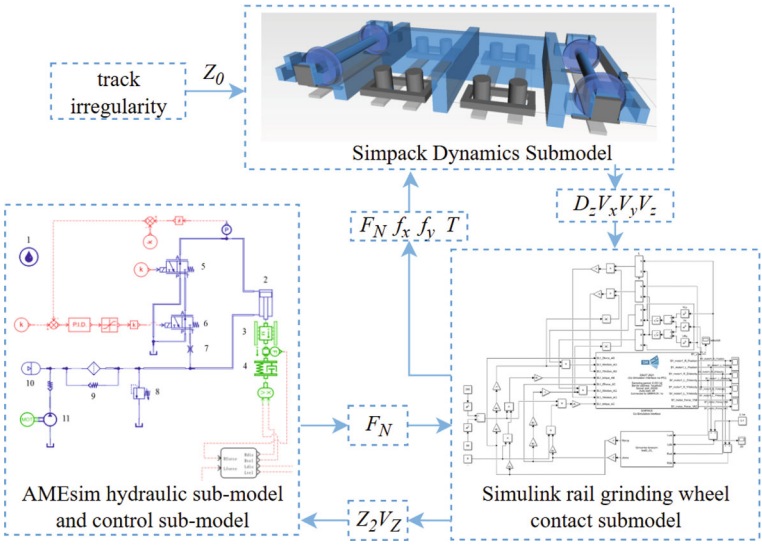


Fig. 5 Hydraulic-mechanical coupling model of co-simulation

Simpack is used to simulate the normal operation of the grinding trolley and output parameters of the grinding wheel. After data processing by Simulink, AMEsim reads the vertical displacement and speed of the grinding wheel, which reflect the pressure change in the rodless cavity and affect the output pressure change of the feed cylinder. The control system controls the pressure in the rodless cavity and finally outputs the controlled contact pressure between the rail and the grinding wheel. Simulink is used to calculate the grinding force and torque according to the contact pressure. It will input the contact parameters of rail and grinding wheels into the Simpack model to form a data closed loop and establish the real model of the grinding process. Figure 5 shows the joint simulation model.

4 Simulation Results

4.1 Dynamic Simulation Results

In order to verify the necessity of rail-grinding wheel contact in the dynamic model, a five-level rail spectrum of the United States is used to carry out dynamic tests on the models with and without contact force. The test results are shown in Figs. 6 and 7.

It is shown that, with the increase of speed, both derailment coefficient and wheel load reduction rate increase significantly. In addition, the indicators are high when

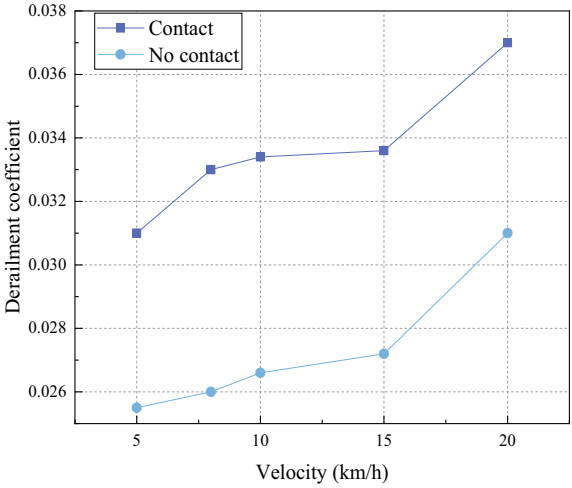


Fig. 6 Derailment coefficient

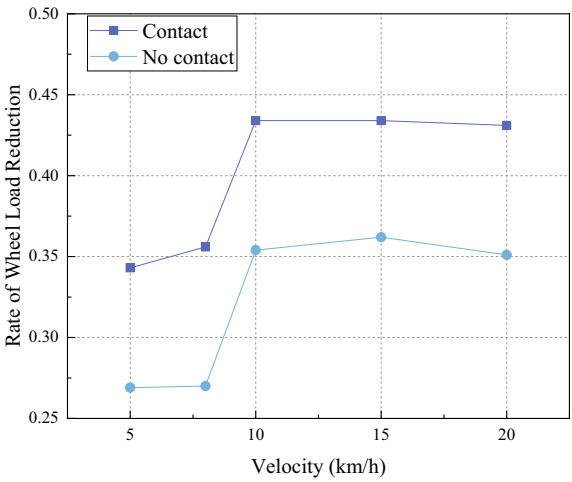


Fig. 7 Wheel load reduction rate

the contact force is taken into account, so the contact force and friction force should be taken into account when the grinding trolley is studied.

4.2 Simulation Results of Hydraulic System

The hydraulic system sub-model focuses on the end pressure of rodless cavity under the condition of rail irregularity and the reaction degree of hydraulic rod to rail irregularity. Figures 8 and 9 show the results for 1.7 Hz rail irregularity.

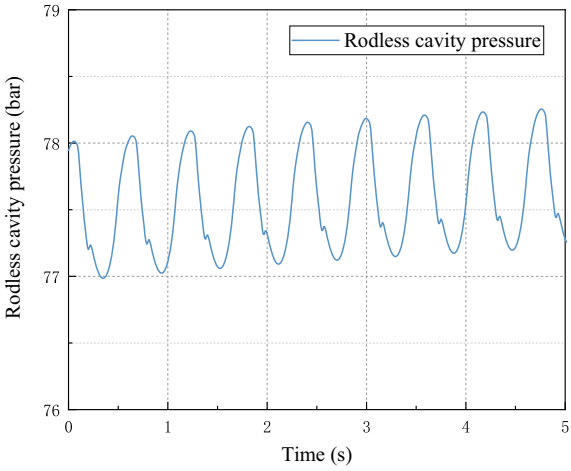


Fig. 8 Pressure at rodless cavity end at 1.7 Hz

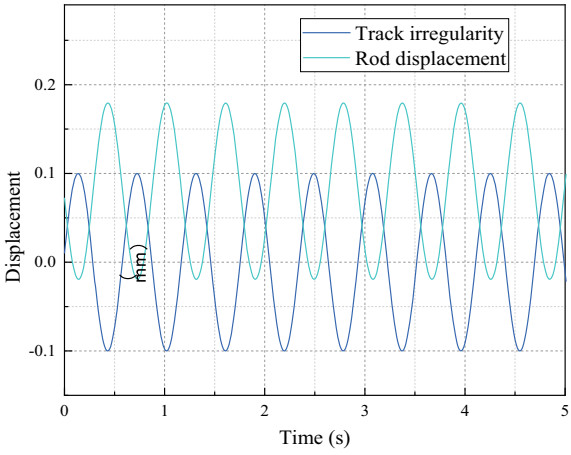


Fig. 9 Rail irregularity and hydraulic rod displacement

As shown in the figures, under the influence of 1.7 Hz irregularity, the pressure at the rodless cavity end fluctuates slightly, but it is stable in general. The fluctuation range is 1.05 bar, less than the pressure difference during actual grinding, thus realizing the control of pressure fluctuation. In addition, the displacement of hydraulic rod can well reflect the vertical position fluctuation caused by rail irregularity, verifying that the model is effective.

5 Conclusion and Prospect

In summary, the following conclusions are obtained:

(1) In this paper, a rail grinding vehicle model that takes different aspects into account is established. The model includes: dynamic sub-model of grinding trolley, rail-grinding wheel contact sub-model, hydraulic system sub-model and control system sub-model. It can well reflect the whole rail grinding process.
(2) The dynamic parameters of the coupled model and the uncoupled model are compared, and it can be concluded that the rail-grinding wheel contact force has an obvious influence on the dynamic performance of the grinding trolley. Therefore, the contact force and friction force should be taken into account in studies of rail grinding process.
(3) The maximum pressure change at the rodless cavity end of the hydraulic cylinder under 1.7 Hz rail irregularity is 1.05 bar, less than the actual grinding pressure difference on site. In addition, it is verified that the vertical position change of hydraulic rod can well reflect the fluctuation caused by rail irregularity. The model is validated accurate, and is suitable for the study of rail grinding process.
(4) In this paper, the data of other papers are used as verification, and field verification experiments will be conducted and a prediction model for the rail profile after grinding will be established in subsequent research.

Funding The open project of State Key Laboratory of Heavy-duty and Express High-power Electric Locomotive (GZKFKT2022-009); Sichuan Province International Science and Technology Innovation Cooperation Project (2022YFH0038); the Fundamental Research Funds for the Central Universities(2682024CG007).

References

1. Cannon DF (2003) Rail defects: an overview. Fatigue Fract Eng Mater Struct 26:865–886
2. Zhang W (2022) Development status and prospect of key rail grinding equipment and technology of grinding stone. China Mech Eng 33(19):2269–2287
3. Yuan Y (2023) Modeling and simulation of rail grinding wheel considering the random characteristics. J Mech Eng 59(8):213–223234
4. Zhai W (2007) Vehicle-track coupling dynamics, 3rd edn. Science Press, Beijing
5. Wang W, Guo J, Zhou K (2023) Rail grinding theory and technology in high-speed railway, 1st edn. China Railway Publishing House, Beijing
6. Popov (2019) Contact mechanics and friction physical principles and applications, 2nd edn. Tsinghua University Press, Beijing
7. Johnson KL (1985) Contact mechanics. Cambridge University Press, Cambridge
8. Wang P (2017) Experimental study on the press control system of Gmc-96x rail grinding train
9. Zeng L (2024) Mode of rail grinding based on mechanical-electric-hydraulic coupling. J Southwest Jiaotong Univ

Open Access This chapter is licensed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License (http://creativecommons.org/licenses/by-nc-nd/4.0/), which permits any noncommercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if you modified the licensed material. You do not have permission under this license to share adapted material derived from this chapter or parts of it.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.



Research on the Influence of Rack Railway Irregularity on the Rack Railway Engineering Vehicle



Han Lv, Junjun Ding, Changquan Ding, Jianquan Zhang, Haoxiang Weng, and Yang Qin

Abstract In order to study the influence of rack railway irregularity on the rack railway engineering vehicle, the dynamic model of the rack railway engineering vehicle is established, and the simulation data are compared with the experimental data to verify the accuracy of the dynamic model. With the split and integral models respectively established for the rack railway, the influence of the rack railway modeling method is studied, and the dynamic response law of the rack railway engineering vehicle under the rack railway irregularity and different running speeds is analyzed. The results show that compared with the integral model, the split model is more accurate in simulating the dynamic characteristics of the rack railway engineering vehicle. The step irregularity of the rack railway will increase the radial impact force of the rack and gear, leading to more serious unloading of the wheelset.

Keywords Rack railway engineering vehicle · Rack railway modeling method · Rack railway irregularity · Gear rack meshing · Dynamic response

1 Introduction

Mountainous areas in China, accounting for 33% of the country's total area, are often rich in tourism resources. Due to the limitation of topography and climate, those areas are difficult for several modes of transportation to adapt to. However, the rack railway is a form of rail transit suitable for steep mountainous areas [1]. In recent years, many scholars have carried out a series of studies on rack railway vehicles. Seo [2] designed a set of rack railway bogies with independent rotating wheels, which can run on curves with a minimum radius of 10 m. Zhao [3] studied the dynamic

H. Lv · J. Ding (✉) · H. Weng · Y. Qin
School of Mechanical Engineering, Southwest Jiaotong University, Chengdu, Sichuan, China
e-mail: dingjunjun@swjtu.edu.cn

J. Ding · C. Ding · J. Zhang
State Key Laboratory of Heavy Duty and Express High-Power Electric Locomotive, Zhuzhou, Hunan, China

© The Author(s) 2026

H. Saman et al. (eds.), *The 9th International Conference on Advances in Construction Machinery and Vehicle Engineering*, Lecture Notes in Mechanical Engineering,
https://doi.org/10.1007/978-981-96-5527-4_78

performance of rack railway vehicles under different center of gravity heights and moments of inertia. Samad [4] established a low-cost PM monitoring platform to monitor the concentration distribution of PM10, PM2.5 and PM1 along the Stuttgart Rack Railway. Chen et al. [5] studied the reliability and safety of fasteners in the rack railway system under the coupling effect of the vehicle-track-bridge system, and proposed improvement measures. Cheng [6] studied the influence of gear eccentricity on the dynamic performance of the rack railway vehicle. Lei [7] studied the contact stress response under different error levels and combined two types of errors, revealed the influence mechanism of the two types of errors on the contact stress distribution, and proposed countermeasures accordingly. Based on the finite element theory and the concrete plastic damage theory, Kong [8] established a refined analysis model of the rack railway-ballastless track-bridge to analyze the stress change and deformation law of the track structure under different slopes. Shang [9] built a test verification platform for the rack railway drive system.

In this paper, a model of the rack railway engineering vehicle is established, the influence of the rack railway modeling method on the rack railway engineering vehicle's dynamic response is studied, and the dynamic response law of the rack railway engineering vehicle at different running speeds is analyzed, providing a theoretical basis for the operation and maintenance of rack railway.

2 Rack Railway Engineering Vehicle-Track Dynamics Model

2.1 Vehicle System Dynamics Model

In this paper, based on the multi-body dynamics software SIMPACK, a dynamic model of the rack railway engineering vehicle is established. The gauge is 1000 mm, and the wheel-rail profiles are LM type and 50 kg/m type, as shown in Fig. 1. The dynamics model includes 1 car body, 2 frames, 2 motors, 2 gearboxes, 4 wheelsets, 4 running gears and 8 axle boxes.

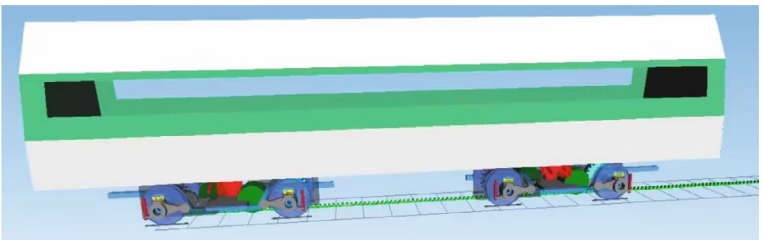


Fig. 1 Dnamics model of rack railway engineering vehicle

The gear rack meshing is modeled by the "Gear Pair" meshing force element, and the normal meshing force is expressed as:

$$F = \sum_{i=1}^{n} (k_i \delta_i + c_i \dot{\delta}_i) \tag{1}$$

where, n is the number of pairs of meshing teeth; k_i and c_i are respectively the meshing stiffness and damping of the pairs of meshing teeth i; δ_i and $\dot{\delta}_i$ are respectively the comprehensive deformation and relative speed of the pairs of meshing teeth i on the meshing line.

During calculation of the meshing stiffness of the gear and the rack railway, the typical "Weber-Bannaschck" formula is adopted [10]. According to the input gear parameters, the gear meshing stiffness is solved. The "Gear Pair" meshing force element also sliced the racks to divide the contact width into smaller parts. In addition, the meshing force element of "Gear Pair" also meshing force element also considers the friction force generated during meshing. The friction force satisfies the Coulomb's law.

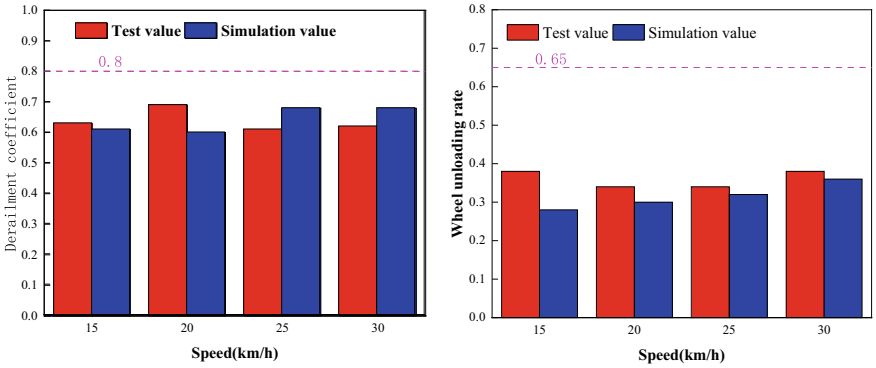
2.2 Model Validation

In order to study the operation performance of the rack railway engineering vehicle when it passes through the curve line, the operation test is carried out on the existing test line. The slope of the test line is 60‰, the curve length is 162 m, the transition curve length is 10 m, the curve radius is 400 m, and the superelevation is 0. The test line is shown in Fig. 2.

The comparison between test data and simulation data is shown in Fig. 3. It can be seen from the photo that the wheel unloading rate and derailment coefficient of the rack railway engineering vehicle on the test line meet the standard requirements,



Fig. 2 Test line photo



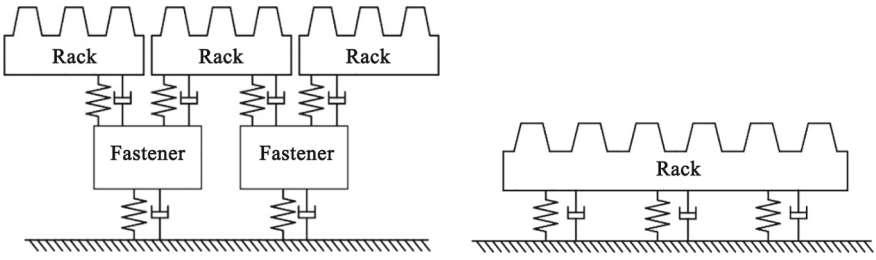
(a) Derailment coefficient (b) Wheel unloading ratio

Fig. 3 Comparison of safety indexes

and the simulation results are basically the same as the test results. In general, the dynamics model established can better predict the operation safety of the rack railway engineering vehicle on the actual line.

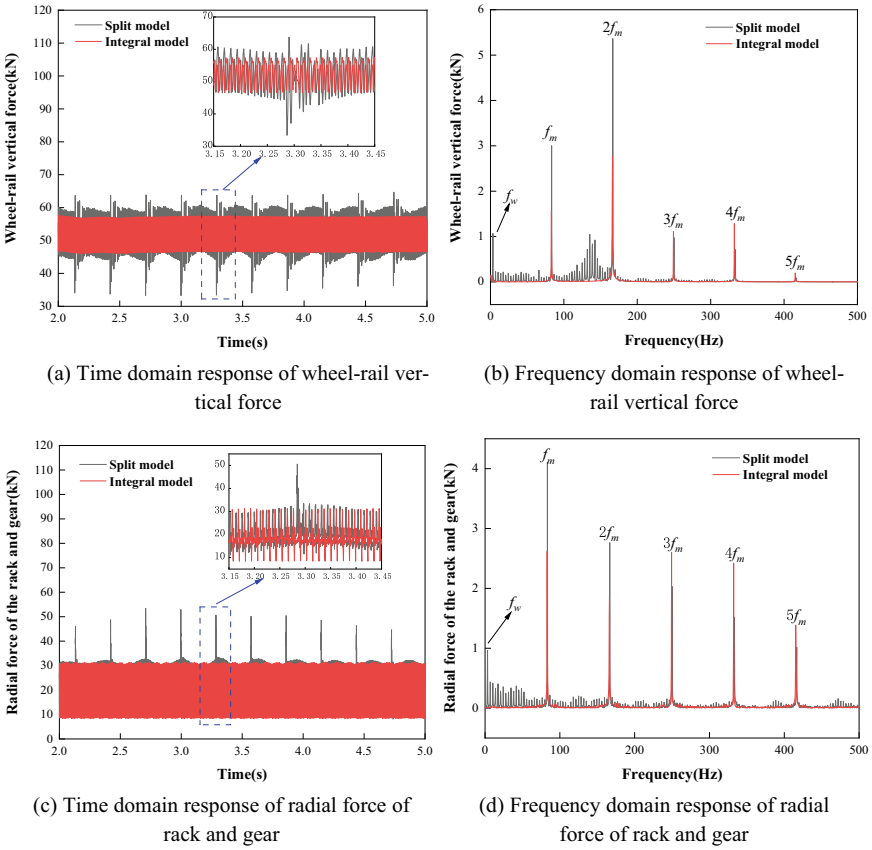
2.3 The Influence of Rack Railway Modeling Method on the Rack Railway Engineering Vehicle's Dynamic Response

There are two common rack railway modeling methods in dynamics model, namely split and integral models, as shown in Fig. 4. In this section, the dynamic responses of the wheel-rail vertical force and the radial force of the rack and gear are compared and analyzed for the two rack railway modeling methods. The results are shown in Fig. 5.



(a) Split model (b) Integral model

Fig. 4 Two calculation models



(a) Time domain response of wheel-rail vertical force

(b) Frequency domain response of wheel-rail vertical force

(c) Time domain response of radial force of rack and gear

(d) Frequency domain response of radial force of rack and gear

Fig. 5 Comparison of dynamic responses of two models

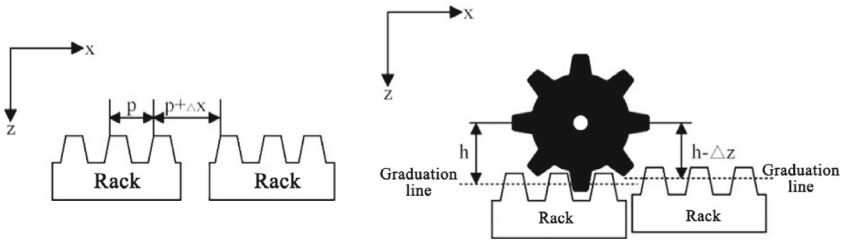
In terms of the time domain response, the split model shows a periodic peak, which is due to the fact that when passing through the fastener, the previous section of the rack is subjected to the meshing force, which leads to a relative displacement from the latter section of the rack, generating impact when passing through the fastener. However, the integral model does not reflect the periodic impact effect of the fastener. In terms of the frequency domain response, the meshing frequency *fm* and its frequency doubling appear in the dynamic responses of the wheel-rail vertical force and the radial force of the rack and gear of both models, which indicates that each system of the rack railway engineering vehicle is affected by the gear rack meshing. However, in the frequency domain response of the split model, the frequency *fw* of the wheelset passing through the rail fastener also appears, and the sidebands with *fw* as the interval appear on both sides of the meshing frequency and its frequency doubling, showing an obvious modulation phenomenon [11].

Compared with the integral model, the split model is more accurate in simulating the dynamic characteristics of the rack railway engineering vehicle. Therefore, this paper uses the split model for subsequent research.

3 The Influence of Rack Railway Irregularity

The step irregularity of the rack railway means that at the fastener connection, the front and rear racks are staggered in the longitudinal, transverse or vertical directions, resulting in a certain displacement difference. For the transverse step irregularity of the rack railway, in the design of the rack railway system, in order to avoid the reduction of the actual meshing width of the gear and the rack railway due to the axial misalignment, the width of the rack railway is designed to be slightly smaller than the width of the gear, thus eliminating the effect caused by axial misalignment [12]. Therefore, this section only studies the influence of the longitudinal step irregularity and the vertical step irregularity of the rack railway on the dynamic response of the rack railway engineering vehicle, as shown in Fig. 6.

The time domain response results of the rack railway engineering vehicle under the influence of the longitudinal step irregularity of the rack railway are shown in Fig. 7. When the wheelset does not pass through the fastener, the radial force of the rack and gear fluctuates in the range of 11 kN ~ 33 kN, and the wheel-rail vertical force fluctuates in the range of 38 kN ~ 60 kN. When the wheelset passes through the fastener, the running gear and the rack produce a radial impact force of 56.67 kN, which is transmitted from the hollow shaft to the axle through the hollow shaft bearing, resulting in the unloading of wheel-rail vertical force to 34.68 kN. However, under the influence of the longitudinal step irregularity of the rack railway, the wheelset unloading is more serious: The impact force of radial force of the rack and gear increases to 74.05 kN, and the unloading of wheel-rail vertical force further decreases to 27.25 kN. Therefore, it is suggested that the longitudinal step irregularity

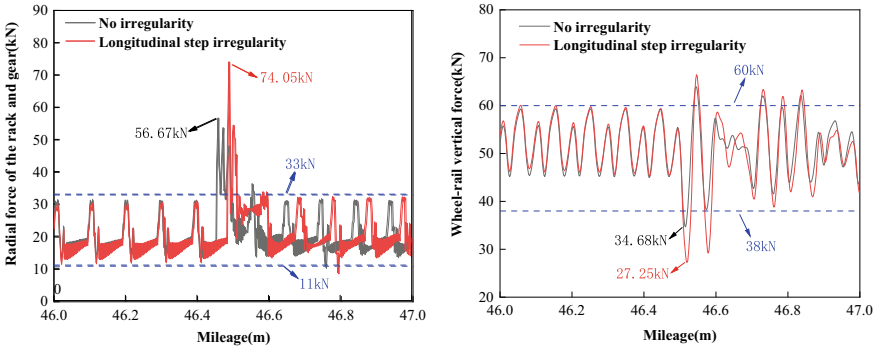


(a) Longitudinal step irregularity of rack railway (b) Vertical step irregularity of rack railway

Fig. 6 STEP irregularity of rack railway

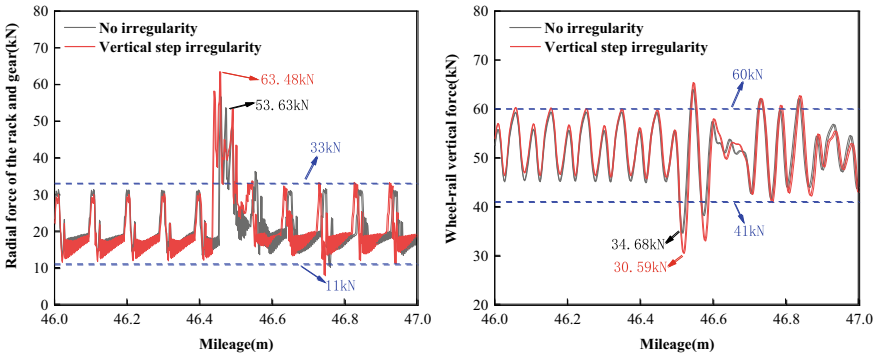
of the rack railway should be strictly controlled to ensure that the pitch of the rack railway at the fastener connection does not change significantly.

The time domain response results of the rack railway engineering vehicle under the influence of the vertical step irregularity of the rack railway are shown in Fig. 8. The vertical step irregularity of the rack railway will also cause a great impact on the gear and rack meshing, resulting in unloading of wheel-rail vertical force. Under the influence of the vertical step irregularity of the rack railway, the wheelset unloading is more serious: The impact force of the radial force of the rack and gear increases to 63.48 kN, and the unloading of wheel-rail vertical force further decreases to 30.59 kN. Therefore, it is recommended to strictly control the vertical step irregularity of the rack railway to ensure that the center distance of the fastener connection does not change significantly.



(a) Radial force of rack and gear (b) Wheel -rail vertical force

Fig. 7 Time domain response under the influence of longitudinal step irregularity of rack railway



(a) Radial force of rack and gear (b) Wheel-rail vertical force

Fig. 8 Time domain response under the influence of vertical step irregularity of rack railway

4 Conclusions

The main conclusions of this paper are as follows:

(1) Compared with the integral model, the split model is more accurate in simulating the dynamic characteristics of the rack railway engineering vehicle. Therefore, the rack railway should be established as a split model in the simulation.
(2) The longitudinal and vertical step irregularity of the rack railway will further increase the radial impact force of the rack and gear when the wheelset passes through the fastener, leading to more serious unloading of the wheelset.
(3) There are still some shortcomings in this paper: this paper only studies the influence of the step irregularity of the rack railway on the dynamic response of the rack railway engineering vehicle, but does not analyze the influence of other forms of irregularity of the rack railway. And the calculation condition is only straight line, without considering the curve line.

Funding The open project of State Key Laboratory of Heavy-duty and Express High-power Electric Locomotive (GZKFKT2022-009); Sichuan Province International Science and Technology Innovation Cooperation Project (2022YFH0038); the Fundamental Research Funds for the Central Universities(2682024CG007).

References

1. Xi L (2014) After a hundred years of rack railway. Traffic Transp 20(01):26
2. Seo S (2018) Development of a bogie for running on a track having a steep gradient and sharp curves. Proc Inst Mech Eng Part F-J Rail Rapid Transit 232(4):1121–1134
3. Zhao Guanchuang.: Influence of height of gravity center and moment of inertia on vehicle dynamic performance. Railway Standard Design 65(9):181–186,193(2021).
4. Samad A (2021) Particulate matter profiles along the rack railway route using low-cost sensor. Atmosphere 12(2):1–18
5. Chen Z (2024) Modified connection mode of rack system on super-slope bridge under extreme temperature load in rack railway. Eng Struct 301
6. Cheng X (2024) Research on the influence of gear eccentricity on the dynamic performance of rack railway vehicle. Machinery 51(5):43–49
7. Lei H (2024) Contact analysis and tooth profile optimization design of rack and pinion for rack-rail trains based on the finite-length line contact theory. J Mech Transm 48(6):124–132
8. Kong C (2024) Study on longitudinal mechanical properties of slab ballastless track on bridge with rack railway. J Cent South Univ Sci Technol 55(6):2404–2413
9. Shang Q (2024) Research and design of rack drive system test platform. Electric Drive for Locomotives
10. Japanese Mechanical Society Technical Data (1984) 'Gear Strength Design Data' Publication Association: Gear Strength Design Data. Mechanical Industry Press
11. Zhou J (2011) Gearbox mechanical vibration signal modulation analysis and its application. Inn Mongolia Sci Technol Econ (15):76–78
12. Hu R (2021) Structural characteristics and key structure parameter selections of mountain cog railway. J Railw Eng Soc 38(09):24–28+67

Open Access This chapter is licensed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License (http://creativecommons.org/licenses/by-nc-nd/4.0/), which permits any noncommercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if you modified the licensed material. You do not have permission under this license to share adapted material derived from this chapter or parts of it.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.



Path Planning System Design for Autonomous Haulage in Semi-Structured Environments



Hang Xu, Dawei Li, and Dawei Jiang

Abstract This paper provides a path planning system design for autonomous haulage applications through a Model-Based System Engineering (MBSE) approach. In the autonomous haulage system (AHS) of mining, the processes of loading, hauling, and unloading can be decomposed into multiple trajectory-following tasks with predefined trajectories and target positions. Supported by the perception, path planning, and vehicle control subsystems, the autonomous truck automatically follows the predefined trajectory to complete the production tasks scheduled by the central server. The path planning system provides collision-free, smooth, and efficient trajectories for the trucks to track. In cases with unknown obstacles and changing environments, the system offers re-planning features to ensure safety and efficiency. A four-level design scheme for a path planning system is obtained through a systems engineering process using the SysML and EAST-ADL architectures. This design allows engineers to build systems suitable for semi-structured environments, balancing both adaptability and performance.

Keywords Path planning · Autonomous haulage system · MBSE · SysML

1 Introduction

1.1 Autonomous Haulage System

The adoption of autonomous driving technology helps a haulage system in mining cut costs and improve productivity. Managed by a central control server, autonomous trucks, which are used for hauling and dumping on the site, can operate safely and continuously with a vehicle control system, GNSS, and equipped sensors. An essential autonomous haulage system (AHS) consists of autonomous trucks

H. Xu (✉) · D. Li · D. Jiang
Beijing Institute of Mechanical Equipment, Beijing, China
e-mail: intelligentxh@gmail.com

© The Author(s) 2026

H. Saman et al. (eds.), *The 9th International Conference on Advances in Construction Machinery and Vehicle Engineering*, Lecture Notes in Mechanical Engineering,
https://doi.org/10.1007/978-981-96-5527-4_79

for autonomous mission execution, a central control system performing operational management, and wireless communication networks ensuring the connection between the trucks and the central control system [1, 2]. In addition to these, auxiliary machines like excavators are also a necessary part of the system.

Either centralized or decentralized architecture could be applied to realize an AHS [3, 4]. A centralized AHS does the necessary calculations on the central control system in order to simplify the in-vehicle functionality and improve efficiency. With actuator control sent by the central control system, the vehicle end only needs to execute the operation, which helps reduce the requirements of in-vehicle equipment and keep any movement of a vehicle in control. However, such an operating model highly depends on communication performance. In circumstances of poor connection, the system scale and operation would be affected. On the contrary, a decentralized AHS enhances the smart functionality of each vehicle, and the central control system only focuses on the management task of telling each vehicle which mission to complete before the job starts. Such an operating model gives as much freedom to each vehicle, which leads to more potential problems if the vehicle end functionality is incomplete.

To balance the features of centralized and decentralized systems, the current AHS adopts an intermediate approach. The central control system on the server manages tasks and routes, while the in-vehicle control system handles obstacle avoidance and task completion. This setup aligns with the concepts of global and local planning in path planning. Most autonomous driving architectures do not fully utilize server computational resources [5–8], performing both global and local planning on the vehicle end. Therefore, developing a path planning system architecture that leverages the strengths of both server and vehicle systems would be valuable.

1.2 SysML

According to the International Council of Systems Engineering, System Engineering (SE) is an approach to system realization. It can help to document requirements and then gradually proceed with detailed design and system validation. Model-Based Systems Engineering (MBSE) is a methodology in SE that uses different model elements to illustrate system requirements, analysis, design, implementation, and verification. It brings a set of clear and transferable traceable documents for system construction. Systems Modeling Language (SysML) is a general-purpose modeling language for MBSE. SysML is a subset of the Unified Modeling Language (UML). Compared with UML, SysML is more flexible and expressive by removing many software-centric constructs. The SysML requirements diagram, which is not provided in UML, efficiently captures functional, performance, and interface needs. The types of SysML diagrams used in the project are introduced as follows.

- Requirement Diagram (Req): It specifies the requirements of the system as well as the relationships between the requirements. Also, other elements could show

up in the diagram to illustrate their relationship with particular requirements. It provides traceability to the system design and avoids redundancy.
- Block Definition Diagram (Bdd): It specifies the structural composition of the system by providing not only a structure block diagram but also relationships between blocks, which gives the designer a clear view of the system and subsystem construction. It benefits the hardware or software modular implementation of the system.
- Sequence Diagram (Sd): It shows the interactions between relevant system components and the messages delivered during the interactions. It gives a clear operating timeline of the system and subsystem.
- Activity Diagram (Act): It shows the actions performed by the system and subsystem, including the object flow and control flow. It gives a straightforward workflow of the system and subsystem.

1.3 East-Adl

EAST-ADL is an Architecture Description Language (ADL) for developing automotive embedded systems [9]. It is an abstraction of the vehicle system that breaks down the system model into four levels.

- Vehicle Level: Specify solution-independent vehicle content and functionalities.
- Analysis Level: Define entities, which could be subsystems and abstract interfaces, to satisfy the vehicle level requirements.
- Design Level: Provide detailed design of the subsystems mentioned in the analysis level, including corresponding software and hardware.
- Implementation Level: Provide implementation represented by AUTOSAR (Automotive Open System Architecture) components.

Since the system discussed is mostly based on a vehicle system, the four-level architectural breakdown could be an ideal reference. It provides a top-down approach beginning with user requirements to detailed implementation. To make it more suitable for the application here, the vehicle level is replaced by a system level, which consists of the vehicle and other main components of the system. Also, the Implementation Level would not use AUTOSAR components, since the generic system design instead of detailed implementation is what the paper focuses on.

2 Methodology

In order to provide a system engineering solution for path planning in the autonomous haulage scene, SysML is used to describe the modified four-level architecture. Referring to the EAST-ADL, the four-level architecture (Table 1) switches the design product from a single vehicle to an autonomous haulage system including vehicles.

Table 1 Four-level architecture

Level	Diagram content
System level	System composition Bdd
	Top-level Req
	Global path plan Sd
	Local path plan Sd
Analysis level	Detailed Req
	Planning Req
Design level	Planning Req
	Planning Bdd
	In-vehicle module Bdd
Implementation level	Planning Act
	Path package Bdd

The system level content specifies the system composition and the top-level requirements of the autonomous haulage system, as well as the actual path planning workflow between the components. According to the background of AHS [10], the system composition Bdd (Fig. 1) illustrates that the system consists of autonomous trucks, wireless communication networks, auxiliary machines, and a central control server.

A commander could interact with the whole system through the central control server to allocate haulage missions for trucks. The trucks, with the help of their controller system and sensors, shall finish their own missions automatically, safely, and efficiently. Thus, the overall requirements could be concluded as autonomous haulage, driving safety, and operation efficiency, leading to the top-level Req (Fig. 8) [11, 12]. The global path plan Sd (Fig. 4) illustrates the interaction process among the commander, central control server, wireless communication networks, and the autonomous truck. This process encompasses mission setting, mission generation, mission transferring, and mission execution. Similarly, the local path plan Sd (Fig. 5) depicts the interaction between the vehicle control system and sensors, covering positioning, sensing, and planning.

Processing to the analysis level, the system level requirements could be divided into sub-requirements including server requirements, perception requirements, planning requirements, and control requirements, which could be concluded in the detailed Req (Fig. 8). Since the path planning system is what we focused on, only the planning Req (Fig. 9) is shown here.

Continue to the design level, the planning requirements shown in the planning Req (Fig. 10) become more specified, leading to the conceptual design of the planning system, which could be concluded as the planning Bdd (Fig. 3). Also, the in-vehicle module Bdd (Fig. 2) shows the other important components of the vehicle control system, including but not limited to the perception module and control module, as a supplement to the planning Bdd.

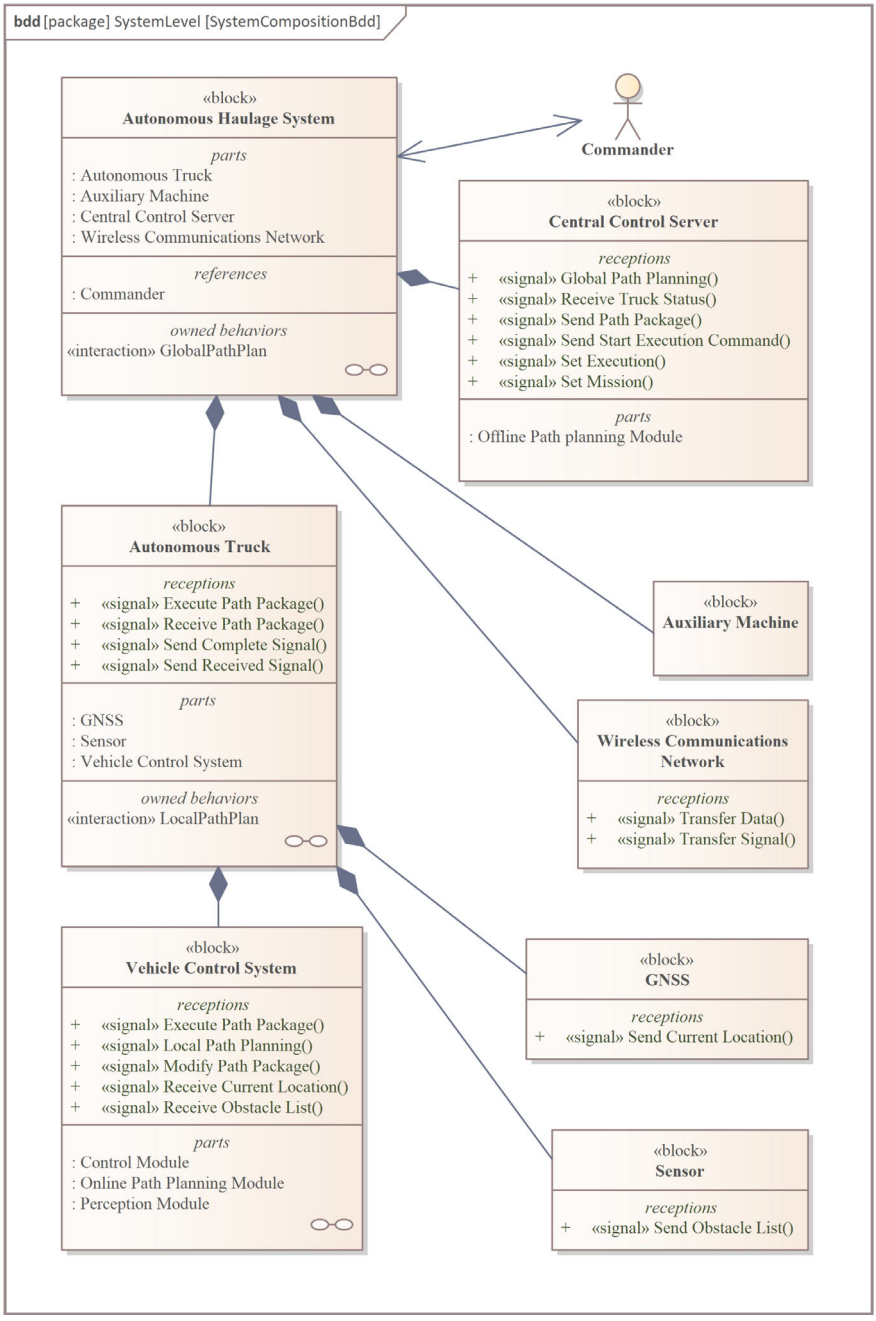


Fig. 1 System composition Bdd (system level)

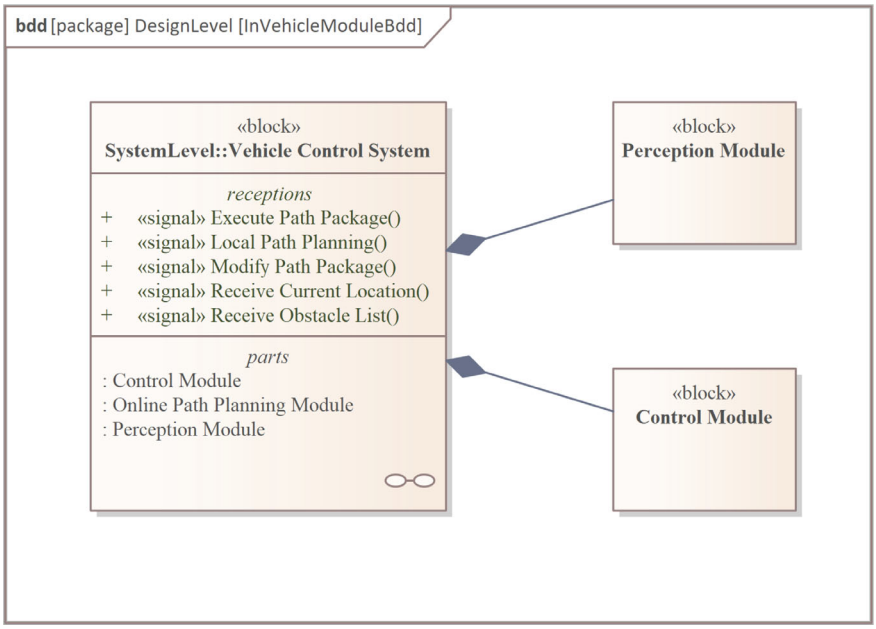


Fig. 2 In-vehicle module Bdd (design level)

Finally, the implementation level gives more design details of the path planning system by providing the path package Bdd (Fig. 7), which outlines what the path planning system should finally produce, and the planning Act (Fig. 6) that shows how the path package is generated.

The design process starts with the system level contents, which can be regarded as a modeling language representation of the AHS, describing the system components, requirements, and the execution process of the two path planning modes representing the centralized and decentralized system concepts. Based on the generalized information, from the analysis level to the design level, the requirements analyses become progressively more specific and focused on the path planning system. In response to the requirements, a path planning subsystem design is given as the design level contents. The conceptual design would be the key point to discuss, and how to code and deploy the system is beyond the scope of this paper. Nevertheless, for ease of comprehension, the implementation level contents are given as a supplement to the conceptual design, discussing the specific details of the designed subsystem, including the software logic flow and file system structure. Finally, the reasonableness of the whole path planning subsystem design would be tested by model completeness analysis. Thus, for the convenience of reading, all the requirement diagrams are shown in the model completeness analysis section, while the contents of the design level and implementation level will be shown in the system design section. Meanwhile, the sequence diagrams of the two path planning modes belonging to the

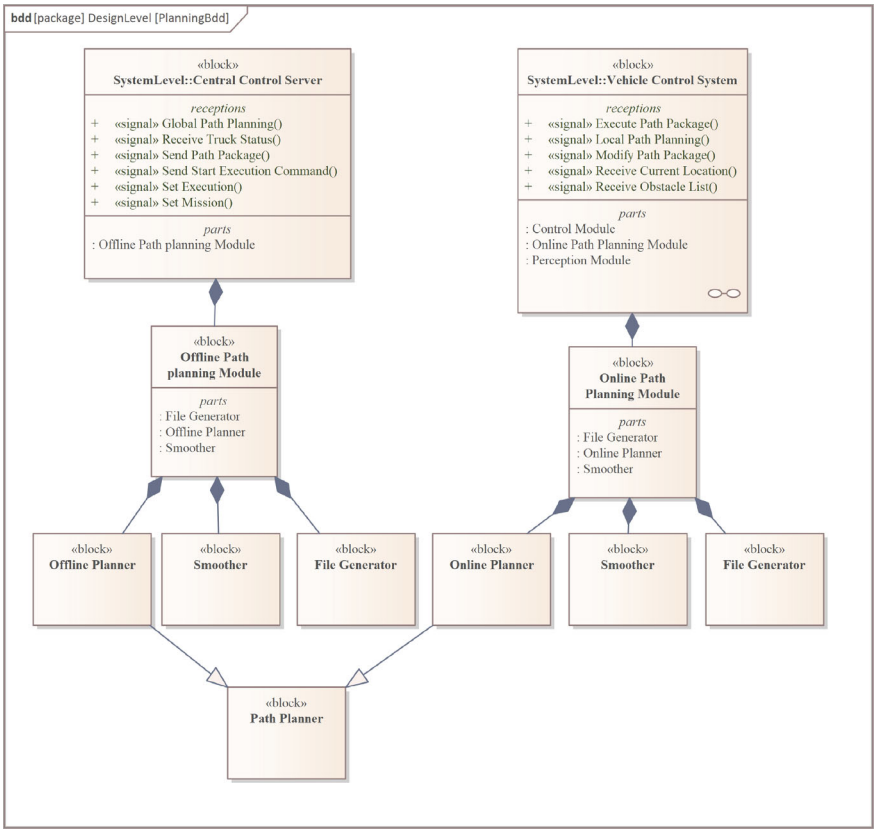


Fig. 3 Planning Bdd (design level)

system level are also shown in the system design section for a better understanding of the design.

3 System Design

To realize the on-site haulage, load and unload automation, the autonomous trucks shall travel autonomously between destinations or at the loading and unloading area. In this case, the self-driving scene could be described as a trajectory tracking process. A trajectory consists of sampling points, which is the path, of a curve that continuously connects the start and goal, and the time-relevant parameters, such as velocity and acceleration, of every sampling point. Thus, to ensure the autonomous haulage-load or haulage-unload mission could be executed safely, the path planning module shall generate a collision-free and vehicle constraint-satisfied (especially minimum

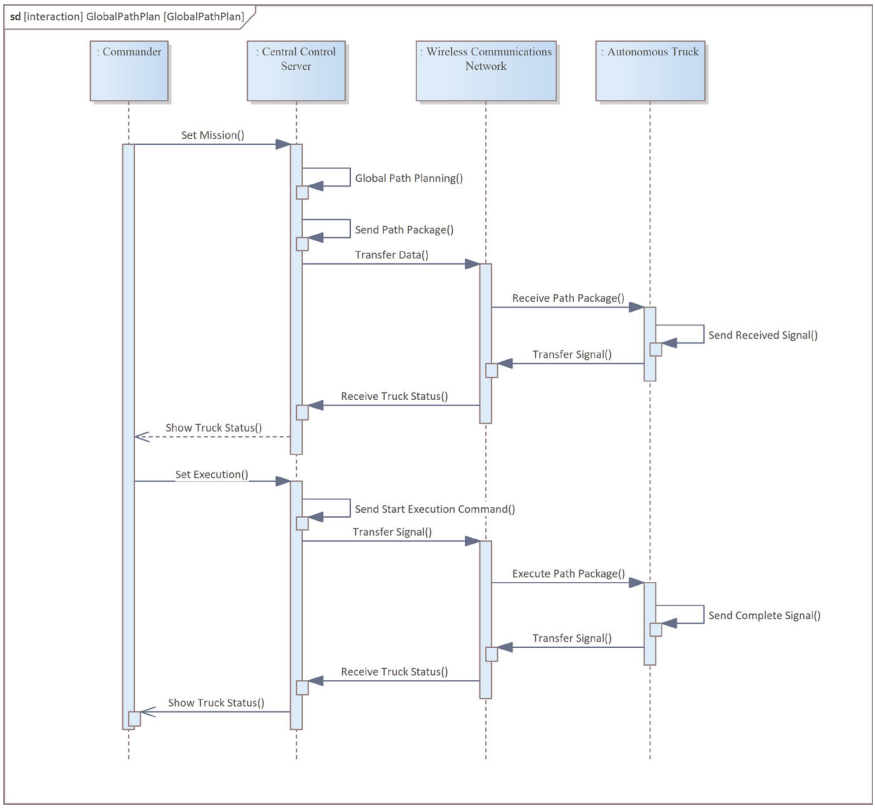


Fig. 4 Global path plan Sd (system level)

turning radius) route for autonomous trucks to follow [13–15]. Since a truck in the haulage operating scene always has to deal with massive loads, the path should be smooth and the footprints should be dense to prevent the truck from damaging itself. In addition, to compose a trajectory with the path footprints generated, the path planning module shall allocate reasonable turning velocity to avoid high-speed turning which may cause danger. Also, to make changing the gear state between Drive and Reverse easier, the trajectory should be split into multiple slices according to its forward or backward status. Additionally, for the obstacle bypass scene, when an unexpected obstacle appears in the current path, the path planning module shall provide a new trajectory replacing the old one to avoid collisions and ensure the haulage mission could be completed safely and efficiently.

Differing from the control and perception modules that are deployed on the vehicle control system, the path planning module has two parts separately deployed on the central control system and the vehicle end. Such a deployment strategy ensures the efficiency and safety features of the autonomous haulage system. The server and the vehicle part of the path planning module share the similar main feature of generating

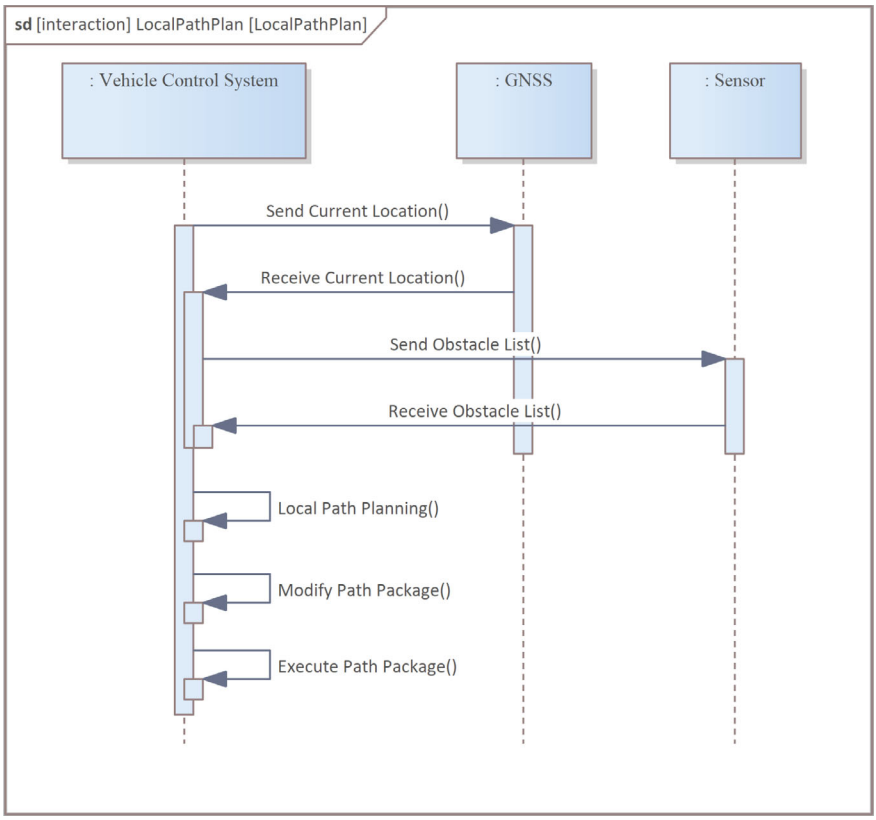


Fig. 5 Local path plan Sd (system level)

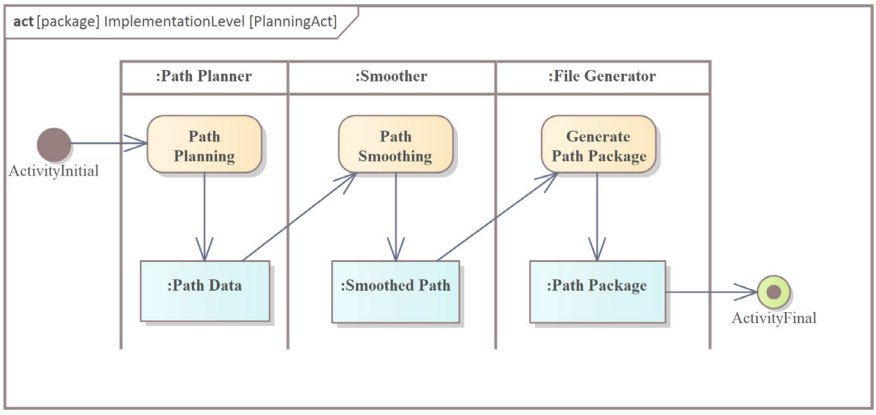


Fig. 6 Planning act (implementation level)

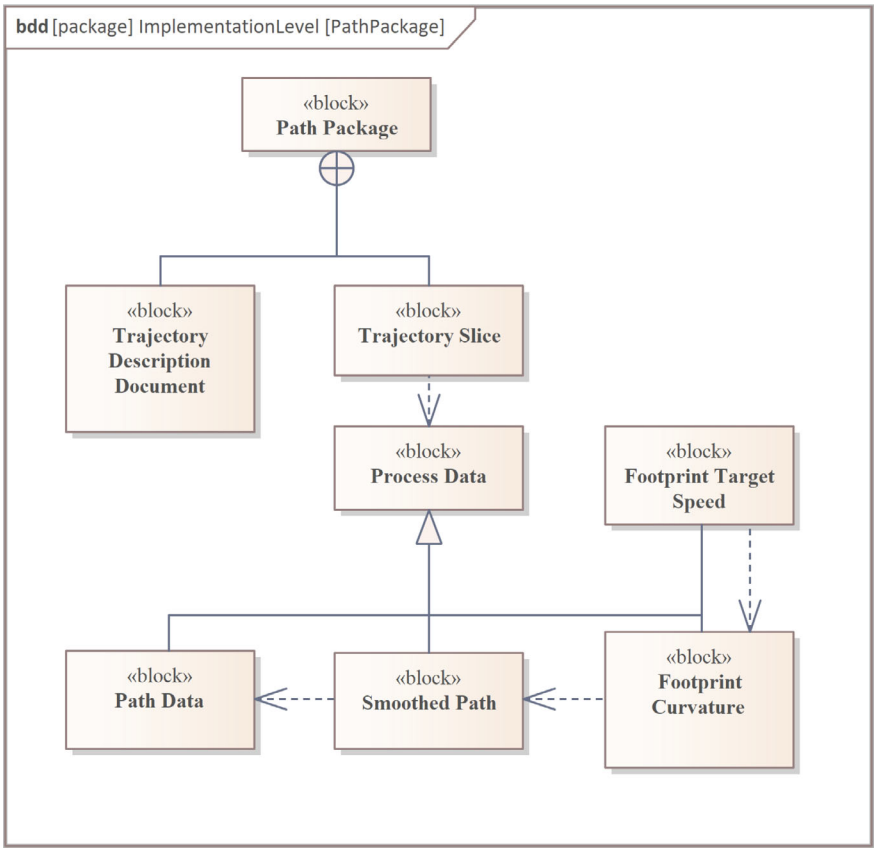


Fig. 7 Path package Bdd (implementation level)

path packages and both consist of three sub-modules: planner, smoother, and file generator.

Although the name of the planner is misleading, the off-line planner deployed on the server end could take advantage of the high computing performance of the server and calculate the long-distance global trajectory, while the on-line planner deployed on the vehicle end generates short-distance local trajectory by interacting with the control module and perception module [16].

The global path planning that the off-line planner is responsible for here specifically refers to generating the path starting from the vehicle's starting point through the haulage road between each loading and unloading area to the entry point of the loading and unloading area specified by the task. This section of the path has the characteristics of a long transportation distance, relatively obvious structural properties (still belonging to unstructured roads), and few unexpected obstacles. Therefore, after the path package is calculated on the server end, it is sent to the vehicle end together with the task description file through the wireless communications network.

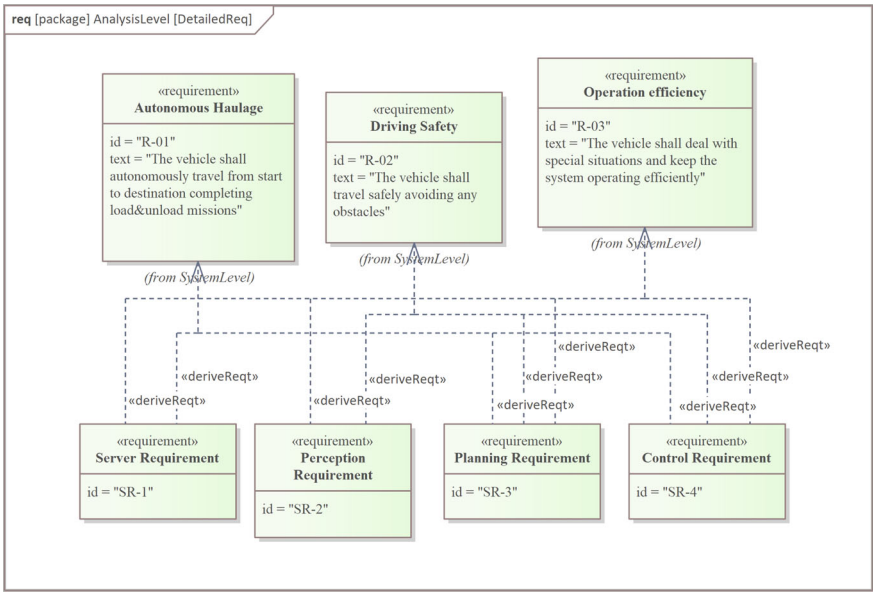


Fig. 8 Detailed Req (analysis level) & top-level Req (system level)

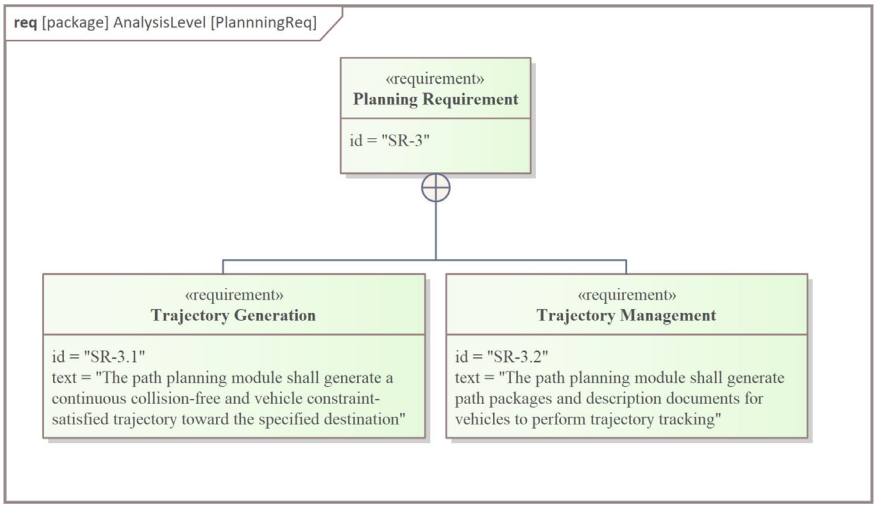


Fig. 9 Planning Req (Analysis Level)

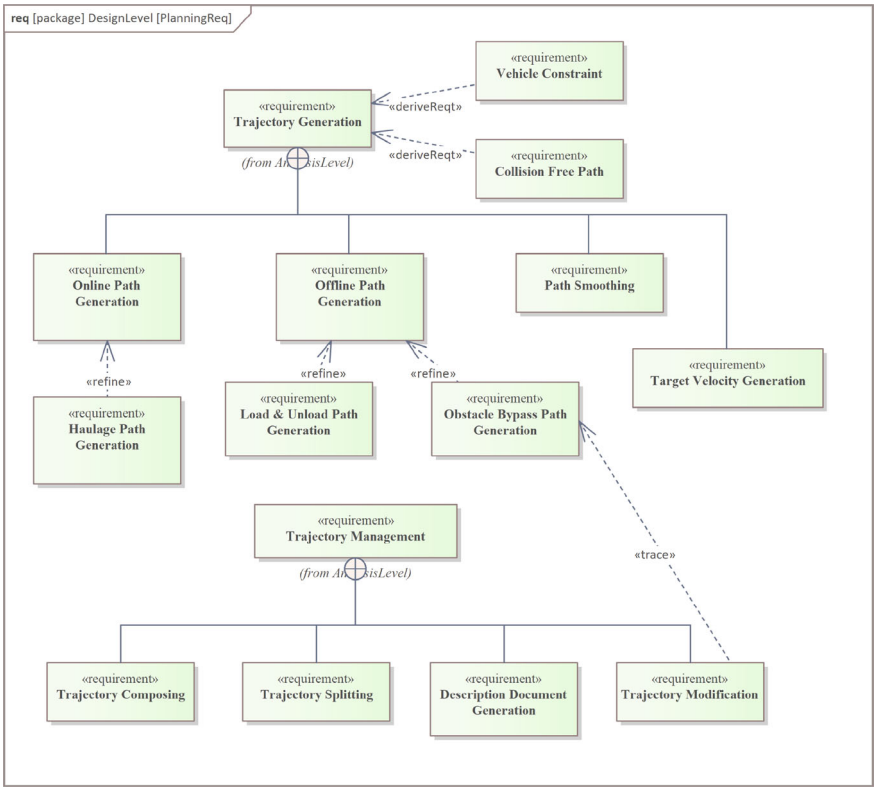


Fig. 10 Planning Req (design level)

Such a workflow can reduce the calculation amount on the vehicle end system during the mission process and save computing time. When a special scene occurs during the driving of the vehicle, the local planner on the vehicle end will be activated by the control module to deliver the on-line planning. It is also worth mentioning that, since the global trajectory does not consider any unexpected obstacles and the haulage routes are usually fixed, the calculated path package could be stored on the server for further reuse.

The local trajectory that the on-line planner is responsible for includes two scenes here. The first is the path planning in the loading and unloading area, which specifically refers to generating a two-segment path [17, 18]. One of the path segments is from the designated loading and unloading area entry point to the designated loading and unloading task spot, and the other path segment is from the task spot to the designated loading and unloading area exit point. The loading and unloading areas are usually typical unstructured areas with a short driving distance, where terrain always varies due to the execution of loading and unloading tasks and the changing positions of auxiliary machines. Also, the loading and unloading tasks are closely coupled with the features provided by the vehicle control module and perception

module. The second is the path planning for bypassing unexpected obstacles in the driving route [19, 20]. Here, the scene specifically refers to generating a path that could bypass the newly detected obstacles and return to the reusable footprints of the original path after the perception module detects the obstacles and informs the control module to slow down or stop the vehicle. In some special cases, the new path generated could be a brand new one that connects the current location with the original destination. But in most cases, the newly generated path section is short.

Thus, it is obvious that the above two scenes could be summarized as short-distance path planning in a non-static environment. Both situations need other modules to interact with the path planning module. Therefore, direct on-line planning on the vehicle end to generate a new path package or modify the original path package can reduce the redundant interaction between the vehicle and server ends, improving system efficiency and stability. Also, when the local trajectory is generated, various external factors are considered, so that the newly generated path's reusability is poor.

Although the planner sub-modules of the server and vehicle end path planning modules are different, the smoothers are mostly the same. The smoother sub-module aims to enhance the smoothness and resolution of the product paths of the planner sub-modules. The improvement of smoothness can reduce the curvature change rate of the path and make the steering of the vehicle more gentle, which could help reduce tire consumption [21]. The improvement of the resolution can make the footprints for tracking denser and reduce the deviation between the actual tracking path and the calculated path curve. In addition, the up-sampling of the path also enables the planner to generate paths with a larger interval between footprints, which could reduce the number of calculations and improve calculation efficiency.

The file generator sub-module is responsible for generating the path package according to the smoothed path calculated by the smoother sub-module. The generated path package will be read by the control module and used for the actual trajectory tracking. A path package consists of several path-slice files and a mission file. The mission file specifies the usage of each path-slice file, such as driving forward or backward. During the process of path package generation, the file generator sub-module can allocate the target velocity at each footprint according to its curvature in order to enhance the turning stability. Meanwhile, the file generator splits the smoothed path into path-slice files not only by their specified usage but also by a preset path length. By splitting a long trajectory into pieces, any excessively large file can be avoided, which helps reduce the running memory of the vehicle control unit and benefits the path traceback and modification. The file generator sub-modules of the server and vehicle end are slightly different, as the vehicle end sub-module requires extra features of path package reading and path data modification to satisfy the obstacle bypass functionality. In the obstacle bypass case, the vehicle end file generator shall locate and read the path slice in use and ensure that only the necessary path slices are modified to avoid redundant computation.

4 Model Completeness Analysis

The model completeness is assessed by the situation of how the requirements are satisfied. Since the high-level requirements are derived and broken down into simple and executable requirements, if all those sub-requirements are satisfied with the design, the model can be considered complete.

As shown in the detailed Req, the top-level system requirements from the system level could be derived into sub-requirements including server requirements, perception requirements, planning requirements, and control requirements. Among others, the planning requirement consists of the trajectory generation requirement and trajectory management requirement. Based on the design level planning Bdd and the system level global & local path plan Sd, all the requirements outlined in the design level planning requirements are met, as stated in the system design section.

5 Conclusion

In conclusion, with the help of SysML and EAST-ADL, a path planning system model specified for the autonomous haulage system is built. Following this model, software engineers could apply any algorithms in their programs to achieve a path planning system, which balances adaptability and performance, for semi-structured environment applications. However, this paper is limited to discussing the path planning system architecture and does not provide a comparison of the path planning algorithms that are suitable for this architecture. Thus, further research will be conducted in the future to explore the effectiveness of various path planning algorithms, such as those based on numerical optimization or graph sampling, when integrated with the proposed architecture to enhance system design and performance.

References

1. Hamada T, Saito S (2018) Autonomous haulage system for mining rationalization. Hitachi Rev 67
2. Gaber T, Jazouli Y, Eldesouky E, Ali A (2021) Autonomous haulage systems in the mining industry: cybersecurity, communication and safety issues and challenges. Electronics 10:1357. https://doi.org/10.3390/electronics10111357
3. Ishimoto H, Hamada T (2020) Safety concept and architecture for autonomous haulage system in mining. https://doi.org/10.22260/ISARC2020/0054
4. Matsui K, Adan F, Angeloudis P (2024) Decentralised motion planning for autonomous mining haulage trucks using prioritised multi-agent MPPI. IFAC-PapersOnLine. 58:154–161. https://doi.org/10.1016/j.ifacol.2024.07.333
5. Kurzer K (2016) Path planning in unstructured environments: a real-time hybrid a* implementation for fast and deterministic path generation for the kth research concept vehicle

6. Shan T, Wang W, Englot B, Ratti C, Rus D (2020) A receding horizon multi-objective planner for autonomous surface vehicles in urban waterways. In: 2020 59th IEEE conference on decision and control (CDC), pp 4085–4092
7. Zheng K (2017) ROS navigation tuning guide. ArXiv, abs/1706.09068
8. Kato S, Tokunaga S, Maruyama Y, Maeda S, Hirabayashi M, Kitsukawa Y, Monrroy A, Ando T, Fujii Y, Azumi T (2018) Autoware on board: enabling autonomous vehicles with embedded systems. In: 2018 ACM/IEEE 9th international conference on cyber-physical systems (ICCPS), pp 287–296
9. Lekamani V (2018) Model based systems engineering approach to autonomous driving: application of SysML for trajectory planning of autonomous vehicle
10. Gaber T. Autonomous haulage systems. https://encyclopedia.pub/entry/11186
11. Benlaajili S, Moutaouakkil F, Chebak A (2021) Infrastructural requirements for the implementation of autonomous trucks in open-pit mines. In: E3S web of conferences
12. Rada A, Kalacheva E, Savostian E, Razdorskaia V (2021) Digital technologies for safety and security risk reduction during open-pit mining. AIP Conf Proc 2402:070027. https://doi.org/10.1063/5.0071774
13. Pang J, Zhang S, Fu J, Liu J, Zheng N (2022) Curvature continuous path planning with reverse searching for efficient and precise autonomous parking. In: 2022 IEEE 25th international conference on intelligent transportation systems (ITSC), pp 2798–2805
14. Ke Y (2023) Comparative analysis of path planning algorithms and prospects for practical application. Highlights Sci Eng Technol.
15. Henning D, Schildbach G (2022) A continuous collision checking method for non-holonomic path planning in narrow environments. In: 2022 26th international conference on system theory, control and computing (ICSTCC), pp 594–600
16. Fedorenko R, Gurenko BV (2016) Local and global motion planning for unmanned surface vehicle. MATEC Web Conf 42:01005. https://doi.org/10.1051/matecconf/20164201005
17. Zhou Y (2022) Research on local trajectory planning of autonomous mobile robots. In: ITM web of conferences
18. Li B, Ouyang Y, Li X, Cao D, Zhang T, Wang Y (2022) Mixed-integer and conditional trajectory planning for an autonomous mining truck in loading/dumping scenarios: a global optimization approach. IEEE Trans Intell Veh 1–10. https://doi.org/10.1109/TIV.2022.3214777
19. Li B, Zhang Y (2019) Fast trajectory planning for off-road autonomous driving with a spatiotemporal tunnel and numerical optimal control approach. In: 2019 IEEE 4th international conference on advanced robotics and mechatronics (ICARM), pp 924–929
20. Zheng H, Dai M, Zhang Z, Xia Z, Zhang G, Jia F (2023) The navigation based on hybrid a star and TEB algorithm implemented in obstacles avoidance. In: 2023 29th International conference on mechatronics and machine vision in practice (M2VIP), pp 1–6
21. Schönnagel A, Dunkelberg N, Keppler F, Janschek K (2024) High-continuity path smoothing along obstacle boundaries applied to agricultural headland planning with kinematic constraints. In: 2024 32nd Mediterranean conference on control and automation (MED), pp 652–657

Open Access This chapter is licensed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License (http://creativecommons.org/licenses/by-nc-nd/4.0/), which permits any noncommercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if you modified the licensed material. You do not have permission under this license to share adapted material derived from this chapter or parts of it.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.



Digital Twin Based Hydraulic System Design



Yapeng Wen, Jiaqi Zhang, Yunzhong Deng, Rui Guo, and Yunfeng Gu

Abstract Hydraulic transmission has the advantages of simple structure, high transmission precision, bear high power, etc., widely used in engineering machinery, aerospace, automotive industry and other fields. However, there are problems such as difficult fault diagnosis, easy oil leakage, high maintenance cost, in order to monitor the real-time dynamics of the hydraulic system and the reliability of the operation, this paper proposes a hydraulic system design based on the digital twin system, and carries out a research around the real-time monitoring of the hydraulic digital twin system and the virtual reality synchronization. A hydraulic system digital twin framework is constructed, containing physical entity layer, virtual twin layer, twin data layer, information interaction layer and visualization application layer. The online operation of the hydraulic system twin is monitored jointly by sensing data and simulation data, and the real-time simulation of the digital twin realizes the visualization of the operation process of the hydraulic system, improves the reliability of the hydraulic system work, and provides a new research idea for the realization of the hydraulic system design based on the digital twin.

Keywords Hydraulic system · Digital twin · Visualization

1 Introduction

With the introduction of national-level manufacturing development strategies such as the U.S. Industrial Internet, Germany's Industry 4.0, and Made in China 2025, intelligent manufacturing has become a common trend and goal for the development of the global manufacturing industry [1, 2]. As the hydraulic system is a class of highly nonlinear system, the system interferes with each other between the circuits,

Y. Wen · Y. Deng · Y. Gu
AECC Sichuan Gas Turbine Establishment, Mianyang, Sichuan, China

J. Zhang · R. Guo (✉)
School of Mechanical Engineering, Yanshan University, Qinhuangdao, Hebei, China
e-mail: guorui@ysu.edu.cn

© The Author(s) 2026

H. Saman et al. (eds.), *The 9th International Conference on Advances in Construction Machinery and Vehicle Engineering*, Lecture Notes in Mechanical Engineering,
https://doi.org/10.1007/978-981-96-5527-4_80

so that the form of failure, failure mechanism is complex and varied; the system internal power transmission is closed, the parameter measurability is poor, the fault information is difficult to extract, resulting in the difficulty of diagnosing the fault of the hydraulic system. Therefore, the study of hydraulic system reliability and fault diagnosis is of great practical significance.

The concept of digital twin was first proposed by Prof. Grieves. M. W in 2003 in the product lifecycle management (PLM) course at the University of Michigan, USA, and was called "Mirror Space Model" at that time [3]. In 2006–2010, it was defined as "Information Mirror Model" and "Information Technology Model". In 2010, the National Aeronautics and Space Administration (NASA) introduced the digital twin concept for the first time in its space technology roadmap [4]. The U.S. Air Force Research Laboratory (AFRL) introduced the conceptual model of using digital twin technology for aircraft structural life prediction in 2011, and gradually extended it to airframe condition assessment studies by building a hyper-realistic, and combining it with historical flight monitoring data for virtual flights, to assess the maximum loads allowed and ensure airworthiness maximum loads to ensure airworthiness and safety, thereby reducing the burden of whole life cycle maintenance and increasing aircraft availability [5]. 2012 NASA released the "Modeling, Simulation, Information Technology, and Processing Roadmap," the digital twin into the public eye. 2014 digital twin theory and technology system was introduced, and accepted and promoted by the U.S. Department of Defense, NASA, Siemens, and other companies, due to GE, NASA, and Siemens [6]. 2015 General Electric Company based on digital twins to achieve real-time monitoring, inspection and maintenance of the engine. Gartneri, a world-renowned consulting firm, listed digital twins as one of the top 10 strategic technology trends for two consecutive years (2017 and 2018). The overall development of digital twins is illustrated in Fig. 1.

In this paper, the digital twin technology is combined with the hydraulic system design, based on AMEsim software to establish the digital twin body model of each hydraulic component and hydraulic system, and the fault simulation data of the model is processed and imported into the twin database Mysql as the original basis for fault diagnosis, and at the same time the real-time feature data of the actual

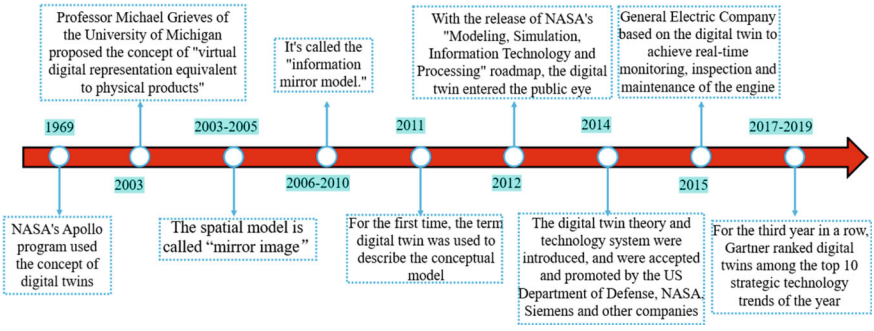


Fig. 1 History of digital twin development

hydraulic equipment is analyzed interactively with the twin data to update the real-time parameters of the virtual model. The twin data-driven, physical objects, virtual models and service systems interact and iterate, and can be a continuous process optimization of the functional system. Through Labview visualization interface, real-time data acquisition and fault diagnosis and other displays are completed. Based on the digital twin theory, it constructs all kinds of components and overall twin model of hydraulic system, which has certain research and academic value.

2 Digital Twin Hydraulic System Architecture Design

Based on the five-dimensional model of digital twin proposed by Prof. Tao Fei and combined with related design concepts [7], the digital twin technology and hydraulic system are combined to design the digital twin hydraulic system architecture. The digital twin system of hydraulic system mainly consists of: physical entity layer, virtual twin layer, twin data layer, information interaction layer and visualization application layer see Fig. 2.

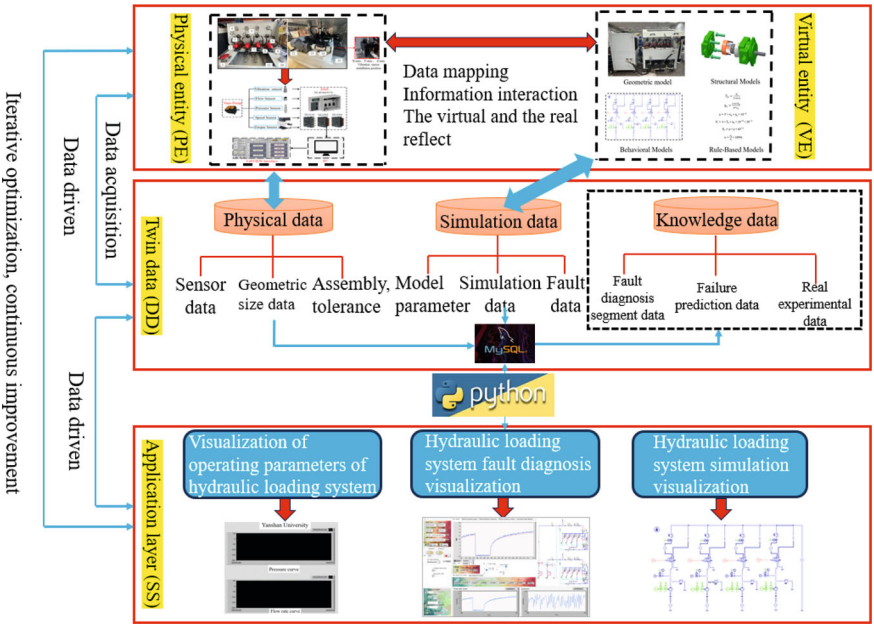


Fig. 2 Hydraulic system digital twin system framework

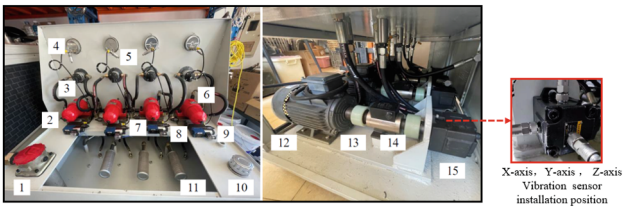
(1) Physical entity layer: including hydraulic system components, such as power components motor, hydraulic pumps, etc., actuating components motors, control components directional valves, flow valves, etc., and auxiliary components accumulators, etc.; through the entire system of pressure, flow, temperature, vibration, speed and torque and other types of sensors, monitoring the operation of the hydraulic system of the real data, to achieve the real-time dynamic perception of the hydraulic system operation process data.
(2) Virtual twin layer: contains geometric model, physical model, behavioral model, rule model; through the three-dimensional drawing software, such as UG, Solidworks, etc., etc., the proportion of drawing and rendering of the three-dimensional model of the hydraulic system, combined with the physical properties of the key hydraulic components, to achieve the simulation of the hydraulic system and the key components, the use of 3D MAX finally constructed a high-fidelity mirror of the operating process of the real hydraulic system.
(3) Twin data layer: it includes physical data, simulation data, knowledge data and other components. Twin data layer is the key to realize the physical entity layer and virtual twin layer and application layer data-driven.
(4) Information Interaction Layer: It includes the functions of data transmission, pre-processing, and storage. The information interaction layer is able to obtain the real-time state data of the hydraulic system, preprocess it into structured data, and transfer it to the visual interaction layer of the digital twin system for visual display. At the same time, this layer can also execute the real-time control commands issued by the operator to realize the control of the working state of the hydraulic system, so as to complete the task more efficiently.
(5) Visualization application layer: including the construction of a virtual reality form of the hydraulic system of the work scene, real-time dynamic display of the hydraulic system operation of the important performance data and fault diagnosis process, to achieve a multi-angle monitoring of the hydraulic system operation and digital management work.

3 Digital Twin Hydraulic System Design

3.1 Physical Entity Layer

Physical entity is to realize the digital twin system foundation, the research object of this paper is the hydraulic loading system of vane pump, hydraulic loading system physical entity shown in Fig. 3. The vibration sensor, rotational speed and torque meter, pressure sensor and flow sensor are installed in the test device to monitor the hydraulic system operation process in real time. The vibration sensors are installed in the direction of the X, Y, Z coordinate axes of the vane pumps, the axes are perpendicular to each other and intersect at a point in space.

In terms of hardware, Advantech industrial control computer combined with high-frequency data acquisition card is used to form a data acquisition system with multiple



1-oil return filter; 2-high pressure oil filter; 3-flow sensor; 4-pressure gauge; 5-pressure sensor; 6-check valve; 7-relief valve; 8-safety valve; 9-two-position, four-way solenoid reversing valve; 10-air filter; 11-suction filter; 12-motor; 13-coupling; 14-torque tachometer; 15-double-action vane pump

Fig. 3 Test bench hydraulic physical diagram

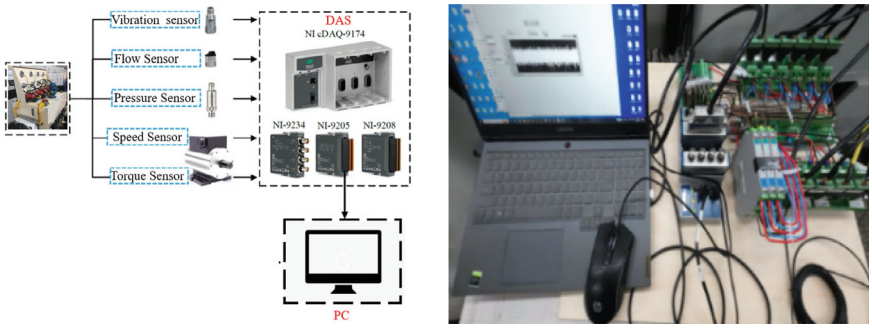
Table 1 Data acquisition elements of the test set

Number	Name	Models	Parameters
1	NI-9205 acquisition card	NI-9205	32 AI channels; 16-bit resolution
2	NI-9208 acquisition card	NI-9208	16 AI Channels; 24-bit Resolution
3	NI-9234 acquisition card	NI-9234	4 analog channels; 24-bit resolution
4	Pressure Sensors	PU5400	Range: 0-40 MPa; Accuracy: ± 0.5%
5	Flow Sensors	MG015	Range: -20 ~ 120°C; Accuracy: ± 0.5%
6	Vibration Sensors	YD-36D	Range: 0-2500 m/s2, frequency range: 1-12000 Hz
7	Temperature sensors	HSTL-103	Range: -25 ~ 85°C; Accuracy: ± 0.5%
8	Tachometer	CYT-302	Torque range: 0-50N-m, RPM range: 0-3000r/min

types of sensors. The parameters of the adopted sensors are shown in Table 1 and the data acquisition module is shown in Fig. 4. At the software level, this test is based on LabVIEW design data acquisition program, LabVIEW is responsible for obtaining various data generated during the test, which mainly contains the vibration waveforms in three directions of the vane pump, the flow and pressure waveforms in the hydraulic system, and the rotational speed and torque numerical display module of the motor. Through the panel, the status information of the tested vane pump can be visualized and problems can be found in time.

3.2 Virtual Twin Layer

Virtual twin layer is the physical world in the computer-aided technology (CAD), computer-aided engineering (CAE), computer-aided process engineering design (CAPP), etc. Virtual space exists with physical entities completely equivalent to the information model, can be based on the digital twin on the physical entity simulation

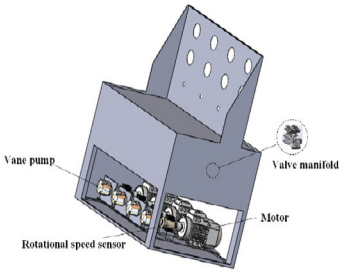


(a) Flowchart for data acquisition (b) Physical data acquisition diagram

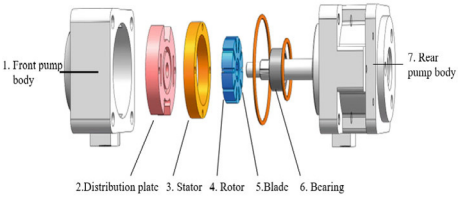
Fig. 4 Data acquisition module

analysis and optimization techniques. The hydraulic system consists of a geometric model, a physical model, a behavioral model and a rule model as shown in Fig. 5.

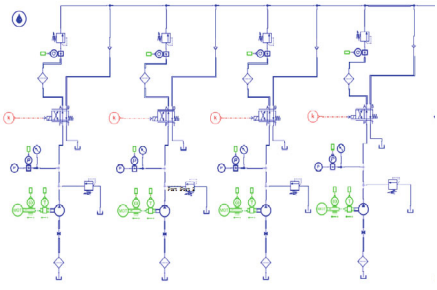
Geometric modeling typically involves the creation of a three-dimensional geometric model of a hydraulic system, including all of its components, such as pumps, valve manifolds, piping, and so on. Geometric modeling provides the physical structure and layout of the system in a digital twin and is the basis for other modeling. It ensures that the digital model agrees with the geometry of the actual physical



(a) Geometric modeling



(b) Physical/structural models



(c) Behavioral models

$$T_w = \frac{P_c}{2*\pi*n}$$

$$q_v = \frac{2*\pi*T_w}{p*\eta_m}$$

$$q = V * n_0 * \eta_0 * 10^{-3}$$

$$P_i = \pi * T_0 * n_0 * 10^{-3} * 30^{-1}$$

$$P_0 = p * q * 60^{-1}$$

$$\eta = \frac{P_0}{P_i} * 100\%$$

(d) Rule model

Fig. 5 Digital virtual twin of hydraulic system

system and provides a spatial reference for physical and behavioral modeling. Physical modeling cats the physical behavior of fluid flow in a hydraulic system and provides basic data for behavioral modeling. It helps to understand the dynamic performance and response characteristics of the system and enables the digital twin to simulate and predict the physical behavior of the real system. The geometric and physical models in this paper are based on Solidworks software for individual component level drawing and assembly drawing according to the dimensions of the physical entities, their contour shapes and their spatial locations in relation to each other. The geometric and physical models are shown in Fig. 5a and b

Behavioral modeling simulates the dynamic behavior of the hydraulic system in the digital twin to provide actual operational data for rule-based modeling. It helps to predict the performance of the system under various operating conditions and allows for troubleshooting and optimization. Behavioral modeling uses AMEsime software to simulate the hydraulic system digitally, selects the exact model according to the signal library, mechanical library, and HCD design library in the sketch model, selects the appropriate sub-model in the sub-model, and sets reasonable parameters for each component to set up reasonable simulation parameters. The behavioral model is shown in Fig. 5c

The rule model is based on theoretical knowledge, and the rule modeling defines the operating rules and logic of the hydraulic system, such as safety guidelines, maintenance intervals, and so on. Example of "Hydraulic Design Manual", "Mechanical Design Manual", etc., the basis of the hydraulic proprietary knowledge to follow the hydraulic design guidelines and mechanical fit, simulating the hydraulic system full life cycle operation status. Rule modeling provides sets of rules for operation and management, which are based on the results of physical modeling and behavioral modeling to ensure that the digital twin system can be operated in compliance with actual operating practices and safety guidelines. It provides support for automation and intelligence of the system. The rule model is shown in Fig. 5d.

3.3 Twin Data Layer

The hydraulic system twin data layer mainly consists of entity data, simulation data, knowledge data and other components. Entity data comes from the physical entity contains pressure, flow, temperature and other sensors collected by the system operation of the normal, fault and other real-time data; based on the Hydraulic Design Manual, Mechanical Design Manual, etc., the physical entity in the space of the geometric dimensions of the data, for the twin model modeling to provide a complete set of data; simulation data is based on the AMesim modeling, the design of the model parameters, and fault simulation data. The knowledge data is the data related to fault identification diagnosis and life prediction of the key components of the system. The above data are stored, called and updated by writing Mysql database scripts in Python to provide data support for visualization design. The overall twin database is shown in Fig. 6.

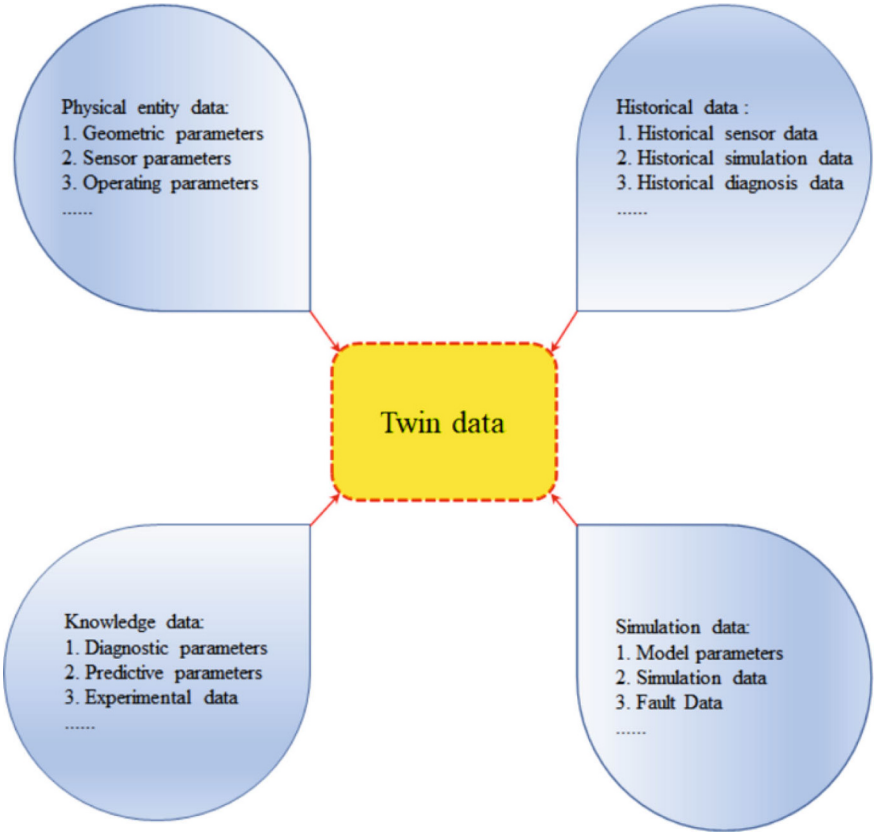


Fig. 6 Twin data layer collection

3.4 Information Interaction Layer

Information interaction is the key to connect the various parts of the data-driven architecture, data transmission, communication, preservation and storage. Hydraulic loading digital twin architecture which contains three parts of each other is synergistic interaction, the specific interaction mode is shown in Fig. 7. according to the hydraulic system operating conditions, the establishment of the system's twin model, the establishment of fault diagnosis simulation model, to complete the three aspects of interaction. See below for details:

(1) Interaction between the physical entity of the hydraulic system and the twin model: according to all kinds of sensors and data acquisition equipment to monitor the physical parameters of the hydraulic system in real time, correct the parameters of the simulation model and build a digital twin model to realize

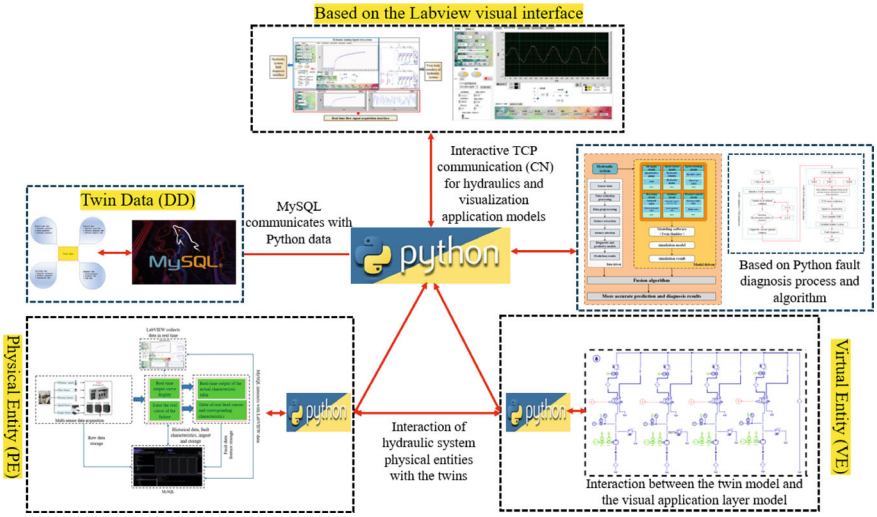


Fig. 7 Communication framework of information interaction layer

the accurate reflection and monitoring of the real-time status of the hydraulic system.

(2) Interaction between hydraulic system and visualization application model: Based on Python, the real-time data, operation status and performance parameters of the hydraulic system are transferred to the digital service model to realize remote monitoring and analysis of the equipment through visualization.

(3) Interaction between the twin model and the visualization application layer model: with reference to the normal and faulty conditions in the actual operation of the hydraulic system, simulate the operating state, performance characteristics and faults of the equipment, generate raw data and simulation data, and generate a fault diagnosis model; the service model models the parameters of the simulation model and carries out intelligent monitoring and analysis, and provides decision-making for fault diagnosis.

3.5 Visualization Application Layer

The realization of the visualization application layer of the hydraulic system is built on the basis of the physical entity layer, the virtual twin layer and the twin data layer. Through the fusion of the visualization application layer with the physical entity of the hydraulic system, the virtual twin, and the twin data, it realizes the visualization of the operation process of the hydraulic system, including the information of the collected data, the state of the equipment, and the type of faults. In the motion simulation of the digital twin system, the application layer fetches the parameter data, structure data, sensor data, etc. of the hydraulic system from the twin data for

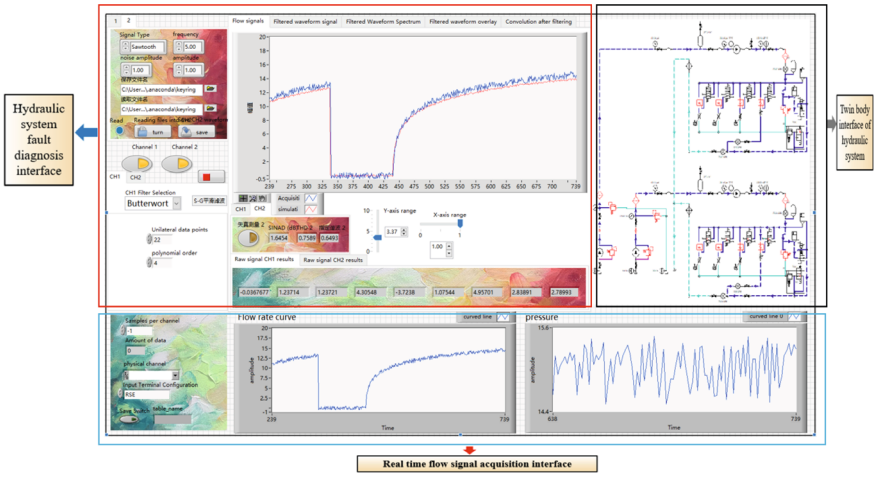


Fig. 8 Visualization interface

the purpose of dynamically refreshing the motion simulation visualization data in the human–computer interaction panel. The visualization interface is programmed through Labview software, and the visualization interface mainly includes: real-time comparison of pressure and flow curves based on real-time acquisition of various types of sensors and pressure and flow curves based on the simulated digital twin model to monitor the occurrence of faults see Fig. 8.

4 Conclusion

In this paper, the hydraulic system digital twin system framework is constructed, the hydraulic system physical entity layer is designed based on Solidworks 3D drawing software, the hydraulic system virtual twin model layer is established based on Amesim, Python is used as a bridge to store, call, and update the three types of data through Mysql database, and Labview software is used to program the hydraulic system visualization application layer, which realizes the visualization of the hydraulic system operation process real-time data and twin model data. The digital twin system can effectively support the digital and intelligent transformation of the hydraulic system, present real-time status, and provide a new way for real-time control, diagnosis and prediction. As this paper carries out the framework design of the hydraulic digital twin system, there may be problems in the detail design, and the follow-up will be based on this framework using real-time communication technology, in-depth study of real-time data acquisition and fault accuracy prediction in the operation process, so as to realize a more intelligent monitoring of the working process of the hydraulic system.

References

1. Xu B, Shen J, Liu S et al (2020) Research and development of electro-hydraulic control valves oriented to industry 4.0: a review. Chin J Mech Eng 33(1)
2. Zhang Q, Kong X, Yu B et al (2020) Review and development trend of digital hydraulic technology. Appl Sci-Basel 10(2)
3. Tao F, Zhang H, Zhang C (2024) Advancements and challenges of digital twins in industry. Nat Comput Sci 4(3):169–177
4. Junfeng Y, Yong Y, Xuecheng W et al (2023) Systematic review of digital twin technology and applications. J Vis Comput Ind Biomed Art 6(1):10
5. Mohd J, Abid H, Rajiv S (2023) Digital twin applications toward industry 4.0: a review. J Cogn Robot 3:71–92
6. Xin L, Du J, Bo T et al (2023) A systematic review of digital twin about physical entities, virtual models, twin data, and applications. J Adv Eng Inform 55
7. Wang JJ, Ye LK, Gao RX et al (2019) Digital Twin for rotating machinery fault diagnosis in smart manufacturing. Int J Prod Res 57(12):3920–3934

Open Access This chapter is licensed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License (http://creativecommons.org/licenses/by-nc-nd/4.0/), which permits any noncommercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if you modified the licensed material. You do not have permission under this license to share adapted material derived from this chapter or parts of it.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.



Research on Simulation and Optimization of Heat Sink for Motor Controller of Electric Forklift Truck



Jiyuan Mao, Zhenyuan Zha, Hao Chen, Xun Hao, Yuntong Wang, and Huan Ma

Abstract Taking the motor controller of an electric counterweight forklift truck as the research object, the heat dissipation path and thermal resistance of the controller were analyzed theoretically. The simulation experiment was designed from two aspects of heat sink design parameters and fin types. The finite element thermodynamic simulation experiment was conducted by establishing a simulation model to explore its influence on the heat dissipation performance of the motor controller. The results show that the fin thickness, number and height of the heat sink have a great influence on the heat dissipation performance of the heat sink. Among the three commonly used fin types of heat sinks, the saw-tooth heat sink has the best heat dissipation. Combined with the simulation experiment, the design of the heat sink is optimized, and the optimized MOSFET junction temperature is reduced by 4.9981 °C, and the optimization effect is obvious. Finally, the temperature field distribution of the heat sink was obtained by the infrared thermal imager in the real experiment, and compared with the simulation results which proved that the established simulation model is reliable.

Keywords Motor controller · Optimization · Thermal simulation · Heat sink · Fin type

1 Introduction

As one of the core components of new energy vehicles [1, 2], the performance of the motor controller will greatly affect the operating performance of the forklift, such as the output characteristics of the motor, the reliability of the control system, and the response characteristics of the forklift [3–5]. Appropriate heat dissipation structure

J. Mao (✉)
China Academy of Machinery Science and Technology, Beijing, China
e-mail: mjy200608@163.com

Z. Zha · H. Chen · X. Hao · Y. Wang · H. Ma
Machinery Technology Development Co., Ltd., Beijing, China

© The Author(s) 2026

H. Saman et al. (eds.), *The 9th International Conference on Advances in Construction Machinery and Vehicle Engineering*, Lecture Notes in Mechanical Engineering,
https://doi.org/10.1007/978-981-96-5527-4_81

and method can make the motor controller work stably at the appropriate temperature, which will enhance the stability of the motor controller and prolong its service life [6]. Moreover, compared with automobiles, the internal space of forklift trucks is narrow, and the heat generated by each component is difficult to be distributed to the external environment [7]. Therefore, the design of the heat dissipation structure and method of the motor controller is particularly important.

Ren et al. [8] carried out thermal simulation of a motor controller, explored the effect of printed circuit board (PCB) thickness, copper cladding rate, electronic component layout and other factors on the controller's heat, the results showed that the board thickness of 1.6 mm, copper cladding rate of more than 50% of the PCB has a better thermal performance. Cao et al. [9] carried out thermal simulation of the liquid-cooled heat sink structure of a permanent magnet motor controller, the results verified the reasonableness of its thermal structure design. Tripathi et al. [10] designed two heat sinks for motor controller and carried out thermal simulation on them respectively. The results showed that the new design of the heat sink is more effective. He et al. [11] used regression analysis to numerically simulate the heat sink of the range extender motor controller to investigate the effect of its fin thickness, length, and width on the junction temperature of the main component of the controller, Insulated-Gate Bipolar Transistor (IGBT). The results showed that the IGBT junction temperature is reduced from 72.29 °C to 71.76 °C after simulation optimization. At present, most of the research on the heat dissipation of the controller focuses on the layout of its own electronic components and the rationality of the design of its heat dissipation structure, and there are fewer studies on the optimization of its heat sink or the optimization effect is not obvious. Therefore, this paper will optimize the design of the controller heat sink to significantly improve its thermal performance as the focus of the research.

In this study, a thermal analysis is carried out for the studied motor controller and its heat sink. A simulation model is established for the design of the heat sink for simulation experiments. Using the one-factor control variable method, the effect of each design parameter on the heat dissipation performance of the heat sink is investigated, and the design parameters are optimized by the results of the simulation experiments. Combined with the optimization results, three commonly used fin types simulation experiment is carried out to explore the advantages and disadvantages of their thermal performances. The related research results have guiding significance for the design optimization of motor controller heat sink.

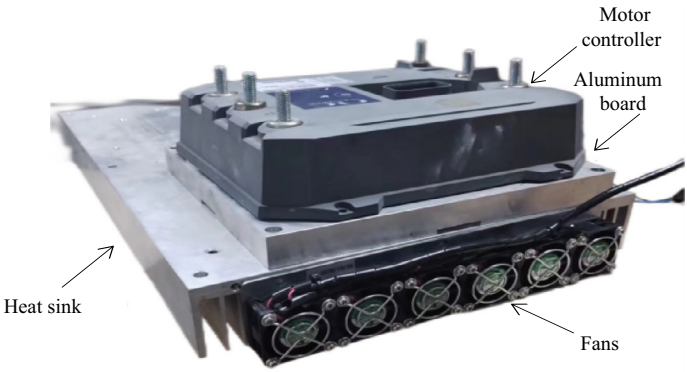


Fig. 1 Assemble structure of motor controller

2 Introduction and Analysis of the Motor Controller Structure

2.1 Introduction to the Controller Structure

The assembly structure of the motor controller studied in this paper is shown in Fig. 1. It consists of a PCB mounted on an aluminum plate and installed in a plastic enclosure. 96 metal–oxide–semiconductor field-effect transistors (MOSFET) encapsulated in the PCB have a high integration with large heat generation, which are the main heat source of the controller. The bottom surface of the PCB is aluminized and laminated with an aluminum board to improve heat transfer. The heat is transferred to the aluminum plate through the PCB and then to the heat sink through the thermal grease. Fans are assembled at end of the heat sink to enhance air convection to optimize heat dissipation. The heat sink selected for this physical experiment is a straight-fin heat sink. Two temperature sensors are installed on the PCB to measure the temperature of the PCB and MOSFETs.

2.2 Thermal and Thermal Resistance Analysis

The study of the heat transfer path is the basis of the thermal analysis of the motor controller. Thermal resistance and thermal path are considered in the analysis of the transfer path [12].

As shown in Fig. 2, the heat generated by the MOSFETs has two transmission paths. The path in turn through the PCB, aluminized layer, aluminum board, thermal grease, heat sink to the outside air. Another path is through the air inside the controller, the plastic shell to the outside air. The motor controller is installed in the closed forklift

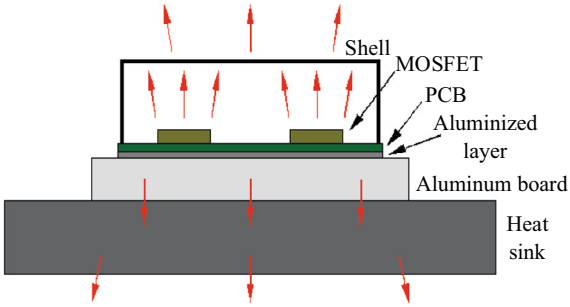


Fig. 2 Thermal path of the controller

cabin. Air in the cabin and the controller shell can be regarded as non-flowing, and the thermal conductivity of the air shell is much smaller than the heat sink, so this paper only considers the heat dissipation optimization of the heat sink thermal path.

3 Experiment Setups

3.1 Pretreatment

The simulation model of the motor controller is shown in Fig. 3. In order to reduce the simulation calculation, according to the principle of geometric model simplification, we make the following simplifications to reduce the model complexity:

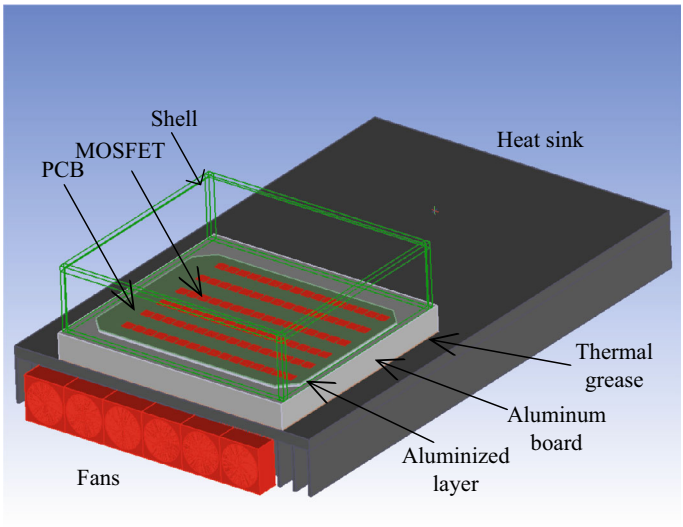


Fig. 3 Model of motor controller

Table 1 Table captions should be placed above the tables

Name	Thermal conductivity($W \cdot m^{-1} \cdot K^{-1}$)
Aluminum plate, heat sink	205
Thermal grease	1.2
Shell	0.6

(1) The components, grooves, and assembly holes on the plastic shell and aluminum board that have less effect on heat dissipation are simplified. The plastic shell is set to a uniform thickness.
(2) Small components inside the controller that generate less heat are ignored.
(3) Openings on the PCB that have a small effect on heat transfer are ignored.
(4) Air convection inside the controller is ignored.
(5) MOSFETs are simplified as heat source modules with uniform heat generation.

The rated power of this motor controller is 4.8 kW, and the heating efficiency of the MOSFET is about 2.4%, so it can be concluded that the heating power of the individual MOSFET is 1.2 W, and the overall heating rate of the controller is 115.2 W.

3.2 Boundary Condition Setups

Since the focus of this study is on the optimization of heat dissipation under the rated operating conditions of the controller, the steady state analysis at rated power is used. The radiative heat transfer is ignored with the heat path analysis in Sect. 2.2. The ambient temperature is set to 30 °C. The air intake of a single fan is 5.7 cfm. The thermal conductivity of other components are shown in Table 1.

4 Results and Analysis

4.1 Initial Simulation

The initial heat sink design parameters were 480 mm length, 300 mm width, 20 fins, 40 mm fin height, 2.5 mm fin thickness, 10 mm base height, and the fin type was straight.

Simulation was carried out according to the simulation model established in the previous section. The result is shown in Fig. 4. It can be seen that the distribution trend of MOSFET junction temperature is decreasing from the center part to the surroundings. The highest junction temperature is 61.4242 °C. The highest MOSFET

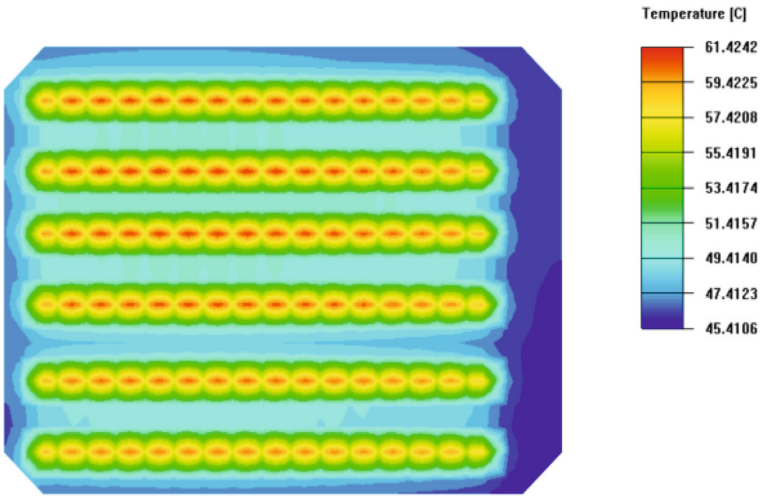


Fig. 4 MOSFET junction temperature under initial heat sink

junction temperature will be regarded as a index of the heat sink heat dissipation performance in the subsequent experiments.

4.2 Design Parameters Optimization

Optimization of heat sink was accomplished using the one-factor control variable method. The single-factor control variable method is a thought method specifically for multivariate problems, which breaks down a multivariate problem into multiple single-factor problems, changing the size of only one of the variables and keeping the other variables constant, so as to explore and discover the influence law for that single factor [12].

Usually, the space for forklift trucks to install controllers is relatively limited, and the heat sink area cannot be increased to optimize its characteristics. Without changing the length and width, the number of fins, the fin height, the fin thickness, and the base height were set as the variables. Their effect on the heat sink heat dissipation performance was investigated by controlling the changes of the variables of a single factor. The simulation experiment results are shown in Fig. 5.

In the case of changing the fin thickness alone, the MOSFET junction temperature decreases with the increase of fin thickness. However, if the fin thickness is too thick, the weight of the heat sink and the vehicle operating load will significantly increase. So when setting the fin thickness, the heat dissipation effect and the radiator quality should be taken into account to select the appropriate parameters.

In the case of changing the number of fins alone, the MOSFET junction temperature showed a significant decline with the increase in the number of fin at first and

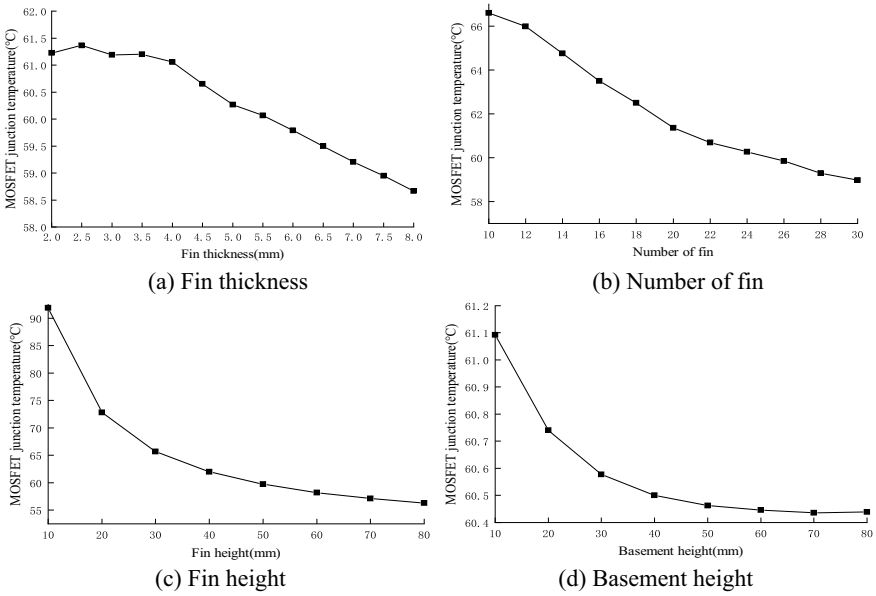


Fig. 5 Curve of MOSFET junction temperature under different design parameters of heat sink

then slowed down. The reason is that the heat conduction is much more difficult when the number of fins is less; then the number of fins increases to a certain extent, narrow air ducts will have a negative effect on the heat dissipation effect, which will slow down the trend of temperature drop. Therefore, when selecting the number of fins, the number of fins parameter of 14–20 should be selected.

In the case of changing the fin height alone, the MOSFET junction temperature change shows a significant decrease with the increase of fin height first, and then slow down the decreasing trend. Therefore, fin height parameter selection ensure the performance of the heat sink combined with the installation space and forklift load condition. The fin height parameter should be selected of 30–60 mm.

In the case of changing the base height alone, the MOSFET junction temperature decreases with the increase of the base height, and its decreasing trend shows a large decrease first and then a small decrease. However, the temperature difference is less than 1 °C during the change of substrate height. Therefore, it can be seen that the effect of the base height on the performance of the heat sink is relatively small, so when setting the base height, attention should be paid to reducing its height to reduce the overall mass of the radiator and the overall operating load of the forklift truck.

4.3 Fin Type Optimization

There are three fin types commonly used in real vehicles: straight, trapezoidal, and saw fin, as shown in Fig. 6. Their heat dissipation effects have not been specifically quantified and investigated. Therefore, under the same radiator design parameters, the three fin types with 2.5 mm end thickness were selected for simulation experiments to quantitatively compare and investigate their heat dissipation performance.

The results are shown in Fig. 7. The MOSFET junction temperature with the trapezoidal fin heat sink is 59.6100 °C, which is a decrease of 2.2544 °C compared to the straight fin heat sink, and with the saw fin is 58.0890 °C, which is a decrease of 3.7754 °C. It can be seen that the thermal performance of the saw fin heat sink is better than the other two. Therefore, it is recommended to use the saw fin heat sink in the controller heat sink design.

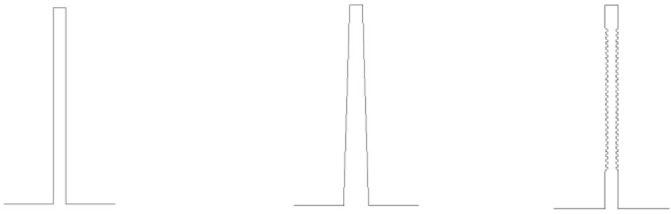


Fig. 6 Commonly used radiator fin types

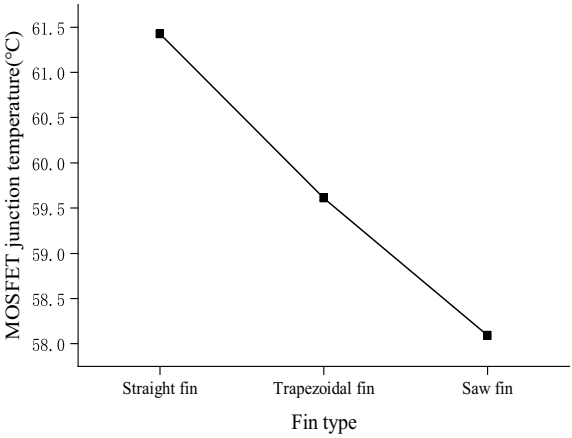


Fig. 7 Curve of MOSFET junction temperature under different fin types of heat sink

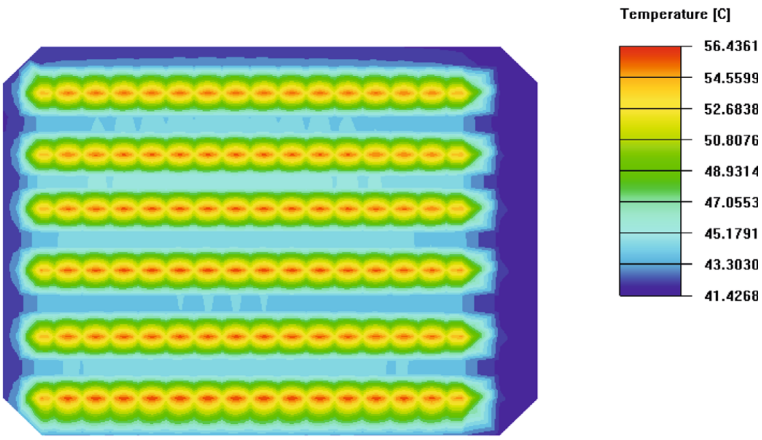


Fig. 8 MOSFET junction temperature under optimized heat sink

4.4 Optimization Verification

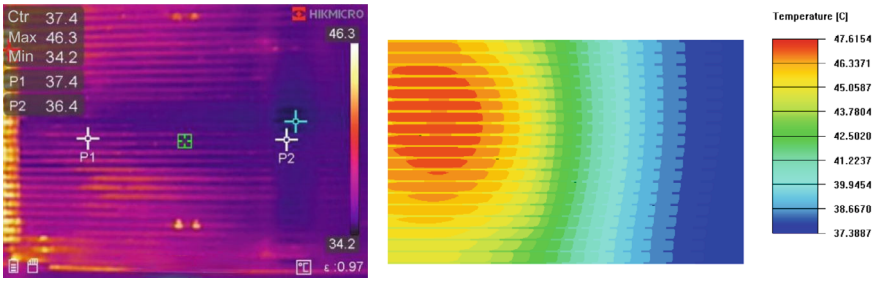
The initial heat sink was optimized by combining the results obtained from the previous optimization experiments for the design parameters and fin type. The length of 480 mm, width of 300 mm, number of fin 20, fin height of 50mm, fin thickness of 4 mm, base height of 10 mm, and fin type of saw teeth were selected. Simulation were carried out. The results are shown in Fig. 8. It can be seen that the MOSFET junction temperature is 56.4361 °C, which is 4.9981 °C lower compared to the initial heat sink, and the optimization effect is obvious.

5 Reality Experiment

The motor controller was connected to a motor load and the temperature data of the controller was monitored by a computer. After the controller reached thermal equilibrium, an infrared imager was used to obtain the temperature distribution on the surface of the controller and heat sink.

The temperature field distributions of the controller obtained from the simulation and the physical experiment are shown in Fig. 9. The obtained PCB temperature is 46.3 °C, and the MOSFET junction temperature is 59.0 °C. The simulation experiments measured the maximum temperature of the bottom plate under rated operating conditions is 47.6 °C, and the MOSFET junction temperature is 61.4 °C.

Meanwhile, we conducted simulation and physical comparison experiments for the controller position to verify the accuracy of the simulation model. Taking the long side of the heat sink as the axis, the direction away from the fan as the positive direction, and the position of the wide side of the heat sink near the fan group as



(a) Reality experiment result (b) Simulation result

Fig. 9 Result of confirmatory experiment

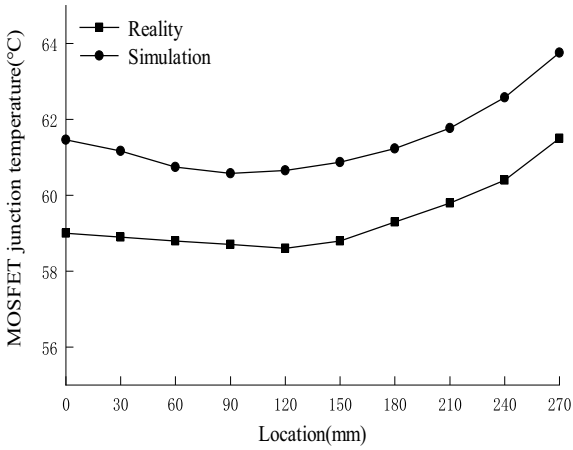


Fig. 10 Curve of MOSFET junction temperature under different location of controller

the zero point, changing the controller position and observing the MOSFET junction temperature. A comparison of the results of the physical and simulation experiments is shown in Fig. 10. The resultant deviation occurs at 0 mm, which is 4.1%, indicating that the simulation model is accurate. The reason for the error may be: the radiation was ignored; the structural details of the shell were ignored; the assembly holes of the aluminum plate and the heat sink were ignored, and so on.

6 Conclusion

This paper focused on the heat dissipation problem of a counterbalanced forklift motor controller. The heat dissipation path was theoretically analyzed. Thermal simulation experiments were conducted to optimize the design parameters and fin type of the heat sink, and the accuracy of the simulation model verified through physical experiments. The following conclusions were drawn:

(1) The main heat source in this motor controller is MOSFET, and its heating power is 115.2 W. The main heat dissipation heat path of the controller is MOSFET—PCB—aluminum plate—heat sink—air.
(2) The thickness, number and height of fin have a great effect on the heat dissipation performance of the heat sink. It is recommended to select the fin thickness parameter of 3.0–5.0 mm, the number of fins parameter of 14–20, and the height parameter of 30–60 mm in the design.
(3) Among the three commonly used radiator fin types, saw fin has the best heat dissipation, followed by trapezoidal fin. The MOSFET junction temperature of the saw fin heat sink has a temperature drop of 3.7754 °C compared to the straight fin. Therefore, the saw fin heat sink should be preferred in practical engineering applications.
(4) Combined with optimization experiment results, MOSFET junction temperature with optimized heat sink drops of 4.9981 °C. The optimization is effective.

This paper conducted a preliminary research on the design of motor controller heat sinks, but there are still deficiencies. Due to the mutual influence of design parameters, the single-factor control variable method may fall into the local better solution, so with the help of computerized methods such as machine learning may be a new research direction. In addition, the contact thermal resistance between the parts is also one of the factors affecting the heat sink, which is not covered in this paper. The optimization of it or the one-piece design of the controller and heat sink is also one of the important research directions in the future.

References

1. Tong L, Ding Y, Li X et al (2021) Cooling structure and thermal analysis of electric motor controller for vehicle. Small Spec Electr Mach 49(11):30–33+37
2. Huang F, Zhao J, Yuan Z et al (2020) Numerical analysis of mini-channel heat dissipation in power battery. J Therm Sci Technol 19(4):347–352
3. Wang S, Lu L, Lai J (2019) Calculation method of lGBT junction temperature for electric vehicle motor inverter and its verification. J Hefei Univ Technol Nat Sci 42(03):370–375
4. He Y, Chen X, Li R et al (2019) Numerical simulation on a semiconductor cooling system coupling with wolid-state fans. J Eng Thermophys 40(03):479–483
5. Yang M, Cheng M, Xiong Y et al (2024) CFD simulation analysis optimization and experimental verification of heat dissipation problem of electric vehicle motor controller. Int J Emerg Electr Power Syst 25(5)
6. Safari V, Kamkari B, Zandimagham M et al (2023) Transient thermal behavior of a passive heat sink integrated with phase change material: a numerical simulation. Int J Thermofluids 20:100454
7. Tang J, Wu X, Sun D et al (2024) Thermal analysis and optimization of radiator coupling in internal combustion forklift truck cooling system. J Hefei Univ Technol Nat Sci 47(03):317–323
8. Ren G, Tian F, Yang L (2013) The research of thermal design for vehicle controller based on simulation. Appl Therm Eng 58:420–429
9. Cao J, Shi C, Li D et al (2023) Thermal simulation analysis of vehicle permanent magnet motor vector control strategy and its controller. In: The proceedings of the 18th annual conference of china electrotechnical society. Springer Nature Singapore, Singapore, p 839–848

10. Tripathi K, Ghosh A, Mangal A et al (2023) Thermal study of MOSFET heat sinks for motor controller in neighbourhood electric vehicles. Int J Electr Hybrid Veh 15(3):256–271
11. He H, Hu W, Ruan X et al (2022) Optimization of motor controller's heat dissipation fins based on regression analysis. Sci Technol Eng 22(2):828–834
12. Zhang X (2021) Research on thermal analysis and heat dissipation optimization of motor driver. Hebei University of Technology

Open Access This chapter is licensed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License (http://creativecommons.org/licenses/by-nc-nd/4.0/), which permits any noncommercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if you modified the licensed material. You do not have permission under this license to share adapted material derived from this chapter or parts of it.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.



Li-Ion Battery Life Prediction Based on RF-SVM



Weijian Huang and Jixiang Zhou

Abstract Lithium-ion batteries have become a research hotspot in the field of battery life prediction due to their small size, light weight, high capacity and long life. In this study, a hybrid strategy combining Random Forest (RF) and Support Vector Machine (SVM) algorithms is proposed to predict the Remaining Useful Life (RUL) of lithium-ion batteries. The strategy exploits the integration properties and feature importance assessment capabilities of Random Forest, while combining the advantages of SVM in dealing with high-dimensional data and nonlinear problems. In addition, Random Forest features are used to reduce the risk of overfitting and improve the robustness of the model to outliers and noise. Through experimental validation using datasets provided by Toyota Motor Corporation and Stanford University, the results show that the hybrid strategy proposed in this study exhibits superior performance compared to the Random Forest SVM alone prediction method.

Keywords Lithium-ion battery · Life prediction · Random forest · Support vector mach

1 Introduction

Lithium-ion batteries exhibit notable advantages over alternative battery technologies, including lead-acid and nickel–cadmium batteries. These advantages include higher energy density, longer lifespan, lighter weight, and faster recharging speeds [1]. It is of great importance to accurately predict the lifetime of Li-ion batteries in order to enhance energy efficiency, contribute to the promotion of renewable energy sources, reduce the burden on the environment, accelerate the promotion of electric vehicles, improve the user experience of mobile devices, and advance the development of energy storage technologies.

W. Huang · J. Zhou (✉)
School of Automobile and Transportation Engineering, Guangzhou City University of Technology, Guangzhou, China
e-mail: 1354373750@qq.com

© The Author(s) 2026

H. Saman et al. (eds.), *The 9th International Conference on Advances in Construction Machinery and Vehicle Engineering*, Lecture Notes in Mechanical Engineering,
https://doi.org/10.1007/978-981-96-5527-4_82

Life prediction techniques for lithium-ion batteries are primarily classified into two main categories: model-driven and data-driven methods [2]. This paper proposes a fusion-type algorithm that combines the data-driven stacking strategy of Random Forest (RF) and SVM for enhancing the prediction performance of the remaining useful life (RUL) of lithium-ion batteries [3]. RF is used for feature extraction and pattern recognition, and its prediction results are used as inputs to SVM, which enhances the feature space and data comprehension of SVM. The integration property of RF helps to reduce the risk of overfitting and improve the model robustness and generalisation. Although integrated models are challenging to interpret, the prediction results provided by RF offer an additional level of interpretation for SVM. Through meticulous model design and parameter tuning, the method functions effectively in conjunction with one another to augment the model's feature representation, generalisation capacity and adaptability, particularly in the context of novel samples and evolving environments.

2 Theoretical Basis

2.1 Support Vector Machine (SVM)

Support vector regression (SVR) represents a variant of SVM that has been specifically designed to address regression problems [4]. The fundamental concept is to identify a hyperplane that encompasses as many data points as possible, with a minimum distance to the hyperplane and a maximum distance to the data points that do not lie on either side of the hyperplane. The objective of SVR is to maximise the width of the aforementioned tube, while allowing for some data points to violate this constraint, thereby enabling the prediction of battery life. The flowchart that illustrates the battery life process is provided in Fig. 1.

The optimisation problem for SVR can be expressed in the form of the following equations:

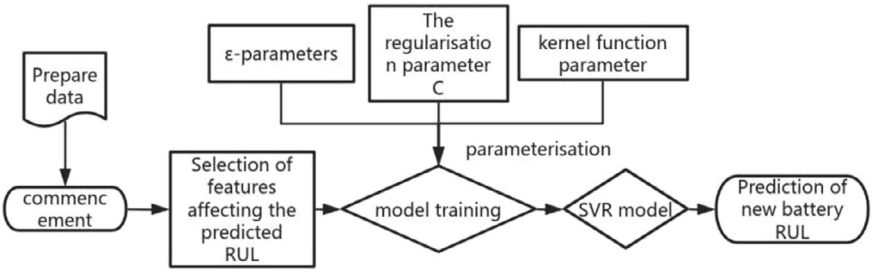


Fig. 1 Presents a flowchart of the SVM algorithm

$$\min_{\omega,b,\xi,\xi^*} \frac{1}{2}||\omega||^2 + C\sum_{i=1}^{N}(\xi_i + \xi_i^*) \tag{1}$$

$$s.t.y_i - (\omega x_i + b) \leq \varepsilon + \xi_i \tag{2}$$

$$(\omega x_i + b) - y_i \leq \varepsilon + \xi_i^* \tag{3}$$

$$\xi_i \xi_i^* \geq 0, \quad i = 1, ..., N \tag{4}$$

The following notation is employed: ω: the normal vector to the hyperplane. b: a bias term. ξ_i and ξ_i^* are slack variables that permit some data points to be situated outside the ϵ-tube, and C is a regularisation parameter that serves to balance the model complexity and the error. The parameter ϵ represents a predefined threshold that defines the width of the tube. By solving the optimisation problem, the regression model parameters ω and b can be obtained, which are then employed in the prediction of battery life for new data points.

2.2 *Random Forest (RF)*

Random Forest Regression represents the application of the Random Forest algorithm to regression problems. The method involves the construction of multiple decision trees, with the final prediction being obtained by averaging the predictions made by these trees [5]. The random forest process is illustrated in Fig. 2.

The construction process for each decision tree in a random forest is analogous to that of a classification problem. However, the final prediction is based on numerical outputs rather than category labels. The recursive selection of the best split features and split points, until the stopping condition is satisfied, is the mechanism by which each tree splits nodes.

The prediction formula for random forest regression is given by Eq. (5).

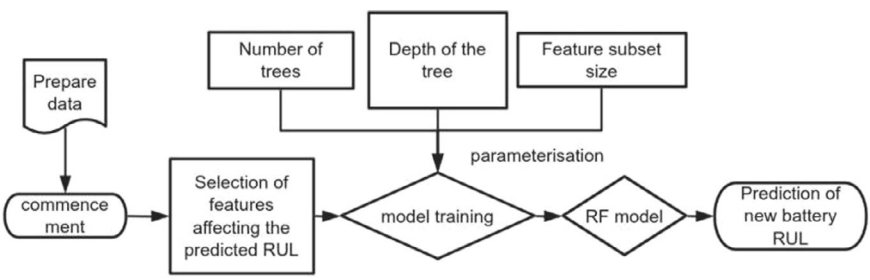


Fig. 2 Random forest flowchart

$$\hat{Y}(x) = \frac{1}{N}\sum_{i=1}^{N} h(x, \theta_k) \tag{5}$$

In Eq. (5), $\hat{Y}(x)$, this is a forecast of the input x. $h(x, \theta_k)$, this is the prediction of the kth decision tree for the input x. θ_k. In this context, the parameter in question pertains to the kth tree, with N representing the total number of trees. In practice, the random forest regression model arrives at the final prediction by aggregating the predictions of each tree.

2.3 Random Forest RF Fused with Support Vector Machines (SVM)

The capacity of Random Forest to manage intricate non-linear relationships and the capability of SVR to identify optimal hyperplanes in high-dimensional spaces are employed to enhance the overall model prediction performance. It is crucial to acknowledge that the efficacy of this integrated methodology hinges on the proficiency of the two models and the complementarity of the predictive outcomes between them. The construction of the model involved an initial prediction of battery life using a random forest regression model. The Random Forest method is an integrated learning approach that constructs multiple decision trees and aggregates their predictions to derive a final prediction.

The Random Forest model is trained using the training dataset, denoted by X, and the corresponding battery life, represented by Y. The model can be represented as Eq. (6):

$$\hat{Y}_{RF} = RFR(X) \tag{6}$$

Among them $\hat{Y}_{RF}$. This is the anticipated value of battery lifespan as forecast by the random forest model. The introduction of random forests as new features is a significant development in the field of machine learning.

Feature fusion entails the incorporation of the predictions of the Random Forest model as new features, which are then combined with the original features X to form a new feature set. The state of the model is illustrated in Eq. (7).

$$X_{\text{new}} = [X, \hat{Y}_{RF}] \tag{7}$$

X_{new}, the data set contains both the original features and the predicted values derived from the random forest model.

Building the SVR model: Using X_{new} as input, the final battery life prediction is made by running the Support Vector Regression model. SVR is a regression algorithm that searches for a hyperplane in a high-dimensional space to maximise the distances between different classes.

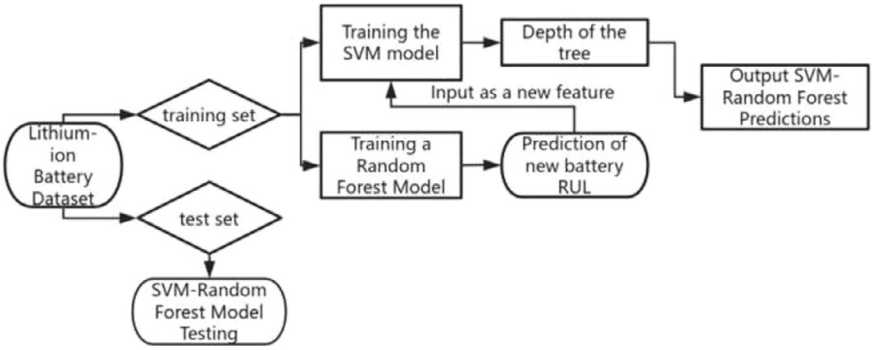


Fig. 3 Fusion flowchart for RF-SVM

Training the SVR model: The SVR model is trained using X_{new} and battery life Y. The model can be represented by Eq. 8:

$$\hat{Y}_{SVR} = SVR(X_{\text{new}}) \tag{8}$$

where $\hat{Y}_{SVR}$ is the final forecast.

The detailed steps of the RF-SVM based fusion method shown in the prediction phase diagram are shown in Fig. 3.

The internal resistance change, discharge current and discharge time data of the battery are extracted from the battery data set and pre-processed. The cycle capacity data set is divided into two parts, 20% of the total number of steps in the cycle period is used as the training set and the remaining 80% is used as the test set to test with the RF-SVM model, and the internal resistance change, charge voltage and discharge time data of the battery are used as the training data set for the RF model and the SVM model [6]. The training phase starts. The training set is used to train the random forest model, while the training set is also used to build the SVM model. The prediction phase starts. The RF model is used to predict the battery capacity at the time of the cycle, and the predicted value is used as a new feature input to train the SVM model to obtain the RF-SVM model for the final prediction of the battery cycle.

3 The Data Sets and Evaluation Criteria Are Presented Below

3.1 Data Set

The experimental data for the lithium-ion battery used in this paper were provided by Toyota Motor Corporation and Stanford University [7], The lithium-ion phosphate (LFP)/graphite battery was manufactured by A123 Systems (APR18650M1A) and cycled in a forced convection temperature chamber set at 48 °C, with a nominal capacity of 1.1 Ah and a nominal voltage of 3.3 V. The experimental charge/discharge data are shown in Fig. 4.

In order to assess the precision of the methodology employed in this study, the experimental data pertaining to lithium-ion batteries, which were collected on 30 June 2017 by Toyota Motor Corporation and Stanford University, were selected for analysis. From the aforementioned data, battery samples CH17, CH18, CH27 and CH28 were selected and subjected to preliminary processing in order to determine the number of charge/discharge cycles. Following processing, the charge/discharge cycle data for the CH17, CH18, CH27 and CH28 batteries yielded 525, 523, 504 and 490 sets of data, respectively. Figure 5 illustrates the relationship between the actual capacity of the batteries and the number of charge/discharge cycles.

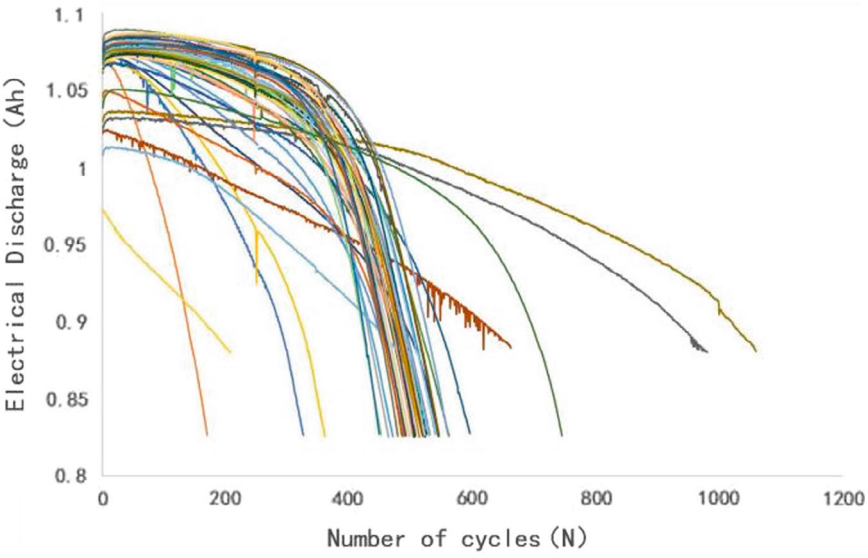


Fig. 4 Trend graph of battery capacity change with the increase of cycle charging times

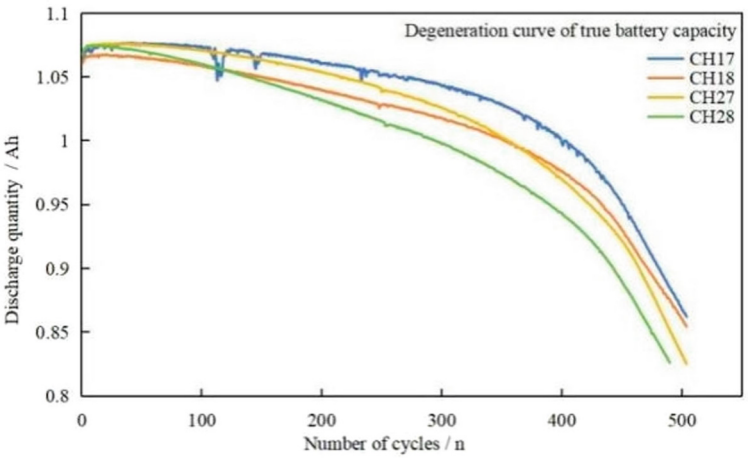


Fig. 5 Battery true capacity degradation curve

3.2 Evaluation Criteria

In order to more accurately demonstrate the efficacy of the RF-SVM fusion method proposed in this paper, the model accuracy is evaluated using the metric prediction error P and the root mean square error (RMSE). A battery is considered to be at the end of its useful life if its capacity is reduced to 80% of its original capacity. This is a global standard definition [8]. The value of P is defined as the absolute value of the difference between the predicted remaining useful life (RUL) and the actual remaining useful life (actual RUL), divided by the actual remaining useful life (actual RUL), in order to obtain a proportion of the relative error. This is demonstrated in Eq. (9).

$$P = \frac{\left|RUL_{pred} - RUL_{real}\right|}{RUL_{real}} \tag{9}$$

4 Simulation Results

4.1 Parameterisation

This paper presents the findings of four sets of simulations conducted on the CH17, CH18, CH27, and CH28 battery datasets, respectively. Each set of simulations employs a total of three methods, namely random forest (RF), support vector machine (SVM), and RF-SVM, to predict the same battery. The same battery is predicted

Table 1 RF-SVM initial parameters

Battery number	Number of trees	The regularisation parameter C
CH17	100	1
CH18	100	1
CH27	100	1
CH28	100	1

Table 2 RF-SVM debugging parameters

Battery number	Number of trees	The regularisation parameter C
CH17	50	2
CH18	50	2
CH27	50	2
CH28	50	2

with RF, SVM, and RF-SVM using the same initial training set, whereby the initial parameters of the models are kept synchronised. Firstly, the RF model is employed to ascertain the data capacity in the four batteries. The model is then constructed for prediction, and the predicted values obtained are utilised as feature inputs for SVM prediction. The initial parameters of RF-SVM are presented in Table 1.

Table 2 illustrates the value settings as determined through simulation debugging.

4.2 RUL Prediction Results for RF-SVM

The effective data from four batteries (CH17, CH18, CH27 and CH28) indicates that the prediction point for the batteries is set to 101 cycles. The capacity degradation predictions for lithium-ion batteries obtained using the three aforementioned methods are illustrated in Fig. 6.

The combined analysis of Fig. 6 and Table 3 indicates that the RF-SVM method yields RUL predictions that more accurately reflect actual values. This method also exhibits a lower error rate compared to RF and SVM alone. The RF-SVM approach demonstrates superior approximation of the lithium-ion battery degradation process as evidenced by RMSE data. Additionally, its predictions are not only closest to the true battery capacity but also show fluctuations that more closely mimic the actual capacity changes. Reference [9], RMSE of literature [10], RMSE is in a good range.

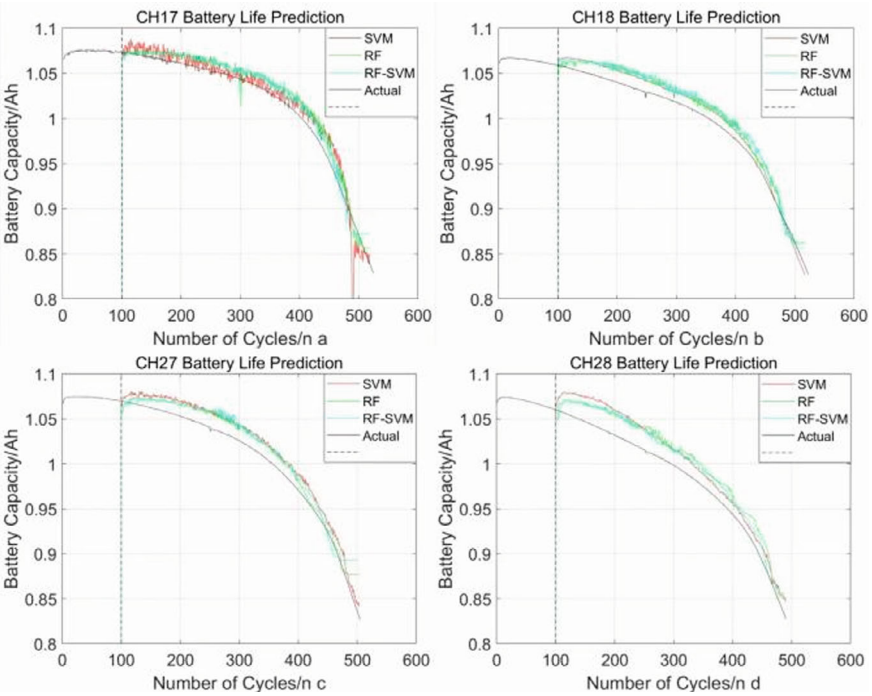


Fig. 6 Predicted capacity degradation of CH17, CH18, CH27, CH28 lithium-ion batteries

Table 3 Four cell lithium-ion battery RUL prediction results

Battery number	Forecasting method	RUL true value	RUL projections	Inaccuracies	P	RMSE
CH17	RF	494	497	3	0.006	0.0157
	SVM		490	4	0.008	0.0165
	RF-SVM		490	4	0.008	0.0143
CH18	RF	487	489	2	0.004	0.0178
	SVM		491	4	0.008	0.0162
	RF-SVM		489	2	0.004	0.0125
CH27	RF	477	480	3	0.006	0.0180
	SVM		483	6	0.012	0.0166
	RF-SVM		475	2	0.004	0.0147
CH28	RF	459	469	10	0.021	0.0273
	SVM		465	6	0.013	0.0324
	RF-SVM		460	1	0.002	0.0252

5 Conclusion

In this study, an innovative model incorporating RF-SVM technique is proposed for predicting the lifetime of lithium-ion batteries. The Random Forest algorithm is used to compute the predicted values and is integrated into the battery capacity degradation model as a feature input to the SVM. The method is suitable for analysing complex datasets, identifying patterns, improving prediction accuracy and adaptability, and providing an accurate and reliable solution for lithium-ion battery life prediction. Model validity is assessed by metrics such as error, relative error and root mean square error to compare the performance of the three algorithms.

The study demonstrates that the proposed fusion prediction method significantly reduces prediction error and enhances accuracy, particularly in experiments with four battery types. This method adeptly models the capacity degradation of lithium-ion batteries, markedly improving RUL prediction accuracy. Such progress aids in boosting battery efficiency, cutting costs and resource use, bolstering reliability and safety, and advancing electrification. These enhancements hold substantial practical significance and application value.

References

1. Wang M, Sun L, Guo P et al (2021) Thermal runaway characteristics of lithium iron phosphate energy storage battery module overcharge based on gas online monitoring. High Volt Technol 47(1):279–286. https://doi.org/10.13336/j.1003-6520.hve.20200227004
2. Suomenglei LJ, Ming Z et al (2020) Progress in life prediction of lithium-ion batteries. Battery Ind 24(05):255–263
3. Dai HF, Zhang YW, Wei XZ et al (2019) Research on residual life prediction of lithium-ion batteries. Power Supply Technol 43(12):2029–2035
4. Mehta C, Sant VA, Sharma P (2024) SVM-assisted ANN model with principal component analysis based dimensionality reduction for enhancing state-of-charge estimation in LiFePO4 batteries. e-Prime Adv Electr Eng Electron Energy 8100596
5. Sulaiman HM, Mustaffa Z (2024) State of charge estimation for electric vehicles using random forest. Green Energy Intell Transp 3(5):100177–100177
6. Li BJ, Han HX, Zhang WJ et al (2024) A review of remaining service life prediction methods for lithium-ion batteries. Energy Storage Sci Technol 13(04):1266–1276. https://doi.org/10.19799/j.cnki.2095-4239.2024.0098
7. Severson et al (2019) Data-driven prediction of battery cycle life before capacity degradation. Nat Energy 4:383–391
8. Cai Y, Chen W, Su Y et al (2021) A review of residual life prediction methods for lithium-ion batteries. Power Supply Technol 45(05):678–682
9. He N, Qian C, Li R (2022) Adaptive modelling and improved particle filtering for battery RUL prediction. J Harb Inst Technol 54(9):111–121
10. Chao XU, Liwei LI, Yuxin YANG (2020) Health state estimation of lithium-ion battery based on improved particle filtering. Autom Technol 12:19–24. https://doi.org/10.19620/j.cnki.1000-3703.20191421

Open Access This chapter is licensed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License (http://creativecommons.org/licenses/by-nc-nd/4.0/), which permits any noncommercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if you modified the licensed material. You do not have permission under this license to share adapted material derived from this chapter or parts of it.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.



Finite Element Analysis of Baja Racer Frame with Space Truss Structure



Qiang Ding, Jiatong Qin, and Biao Jin

Abstract This study investigates the three-dimensional modeling and failure analysis of the space truss structure in Baja racing frames. Using CATIA software, the cab and frame structures were designed based on Baja competition rules and automotive design principles, with materials specified. Strength assessments, modal analyses, and static stress analyses under various operational conditions were conducted using ANSYS software. The results indicate that the frame exhibits reasonable deformation under bending and twisting conditions, consistent with its structural and material properties. To prevent resonance at specific load frequencies, which could compromise structural integrity, we analyzed stress relief and deformation under different field conditions. The space truss structure maintains efficient and stable performance in extreme off-road environments, satisfying various field conditions and enhancing overall driving comfort. This study offers valuable insights for future design and research of Baja racing car body systems.

Keywords Space truss structure · Frame · ANSYS · Finite element analysis

1 Introduction

China Baja competition is a car design and production competition for undergraduates, which is divided into static and dynamic parts. With the development of the event, the performance requirements of Baja racing are becoming higher and higher, especially in terms of driving stability. At present, the development situation at home and abroad is as follows: Ye [1] used HyperWorks to carry out static analysis and optimize the frame structure size to realize lightweight, and then carried out simulation and experimental test. Lu [2] evaluated the performance of the frame in the competition through CATIA and ANSYS. NASA [3] was the first to study lightweight frame, and used NASTRAN software for finite element analysis and optimization design

Q. Ding · J. Qin · B. Jin (✉)
Guangdong University of Science and Technology, Dongguan, China
e-mail: 53340450@qq.com

© The Author(s) 2026

H. Saman et al. (eds.), *The 9th International Conference on Advances in Construction Machinery and Vehicle Engineering*, Lecture Notes in Mechanical Engineering,
https://doi.org/10.1007/978-981-96-5527-4_83

to reduce weight. Naresh et al. [4] analyzed and optimized the frame structure and used carbon fiber as chassis material. Araújo et al. [5] established a finite element model for the pre-processing of the frame model, and the static analysis showed that the equivalent stress was reasonable.

However, existing research often overlooks the driver's role. As race difficulty rises, increased vehicle resonance leads to decreased driving comfort, impacting driver performance, evident in struggles during 4-h endurance races. Therefore, thorough Bach frame design, material selection, strength verification, and modal analysis are crucial for enhancing performance and driving comfort.

2 Introduction to Geometric and Finite Element Models

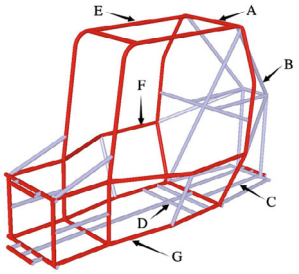
The frame of a car serves as its installation base and is crucial for enhancing performance. The space truss structure, widely used in competitions, offers simplicity in design, low-cost fabrication with basic tools, ease of crash repair, and robust rigidity and strength. For Baja cars, considering factors such as road conditions and maximum speed (60 km/h), carbon fiber offers limited benefits and high costs, making the space truss structure the optimal choice.

The frame (anti-roll frame) of the Baja racing car is designed according to competition rule 4, steel pipe (straight pipe or bent pipe) shall be used for the structural parts of the roller frame. The extension length of the straight pipe shall not exceed 1016 mm, and the support shall be added in accordance with rule B.4.2.4; The bending degree of the bend pipe does not exceed 30°, the extension length does not exceed 838 mm, and follows the requirements of the support structure, and the bending radius is less than 152 mm. The tangent point is the designated point, which is not a bending structural part. using CATIA software for 3D solid modeling. The basic geometry is sketched, and 3D features are created through stretching, rotating, and sweeping. Boolean operations are used for entity editing and decoration. Major structural parts are indicated in red (Ø25 * 3 mm), and minor parts in light gray (Ø25 * 2 mm). The frame's 3D model, shown in Fig. 1, is made of 4130 seamless steel pipe with superior physical properties. The designed frame is stored in STP format and imported into ANSYS for static structural analysis. The mesh size is set to 10 mm, using tetrahedral cells with 57,303 cells and 213,972 nodes.

3 Results and Analysis

3.1 Static Stress Analysis

Comprehensive analysis of bending and bending torsion conditions. During the analysis of racing car frame performance, two working conditions were considered:

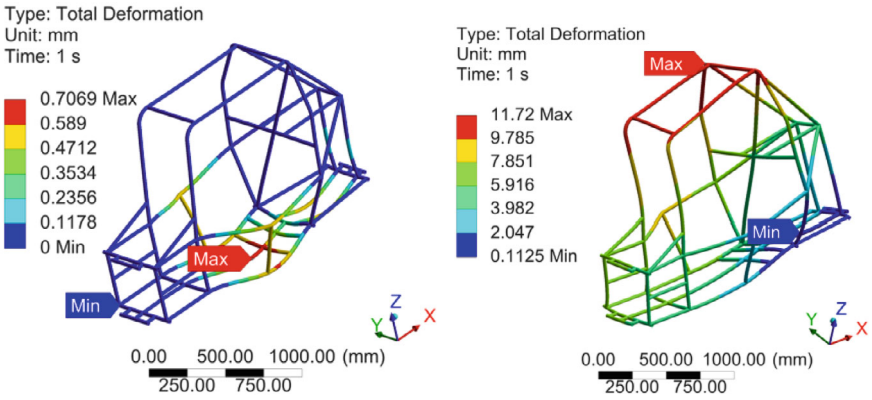


A—Anti-roll ring; B—Anti-roll ring support; C—Engine mounting bracket; D—Base support structure; E—Anti-roll hoop; F—Anti-side impact structural parts; G—Lower frame side members.

Fig. 1 3D model of frame with space truss structure

(1) Bending Condition: This simulates the frame's bending load under static or uniform-speed driving on a flat road. A gravity acceleration of 9.8 m/s^2 was applied along the negative z-axis, along with specific forces of 700 and 850 N on the seat and engine parts, respectively. All suspension connection points were constrained [6]. The resulting deformation, shown in Fig. 2a, exhibited a maximum displacement of 0.71 mm and a maximum stress of 161.7 MPa, indicating frame safety under this condition.

(2) Bending and Twisting Condition: This simulates the frame's bending and twisting deformation during high-speed turning due to centrifugal forces and wheel suspension misalignment. Similar loads were applied, but with a different constraint on the right rear suspension hard point [7]. The deformation, shown in Fig. 2b, showed a maximum displacement of 11.7 mm at the top of the right



(a) Bending condition (b) Bending and twisting condition

Fig. 2 Deformation contour of bending and bending torsion conditions

rear anti-roll ring, with a maximum stress of 609.5 MPa. Despite the severe conditions, the frame remained safe.

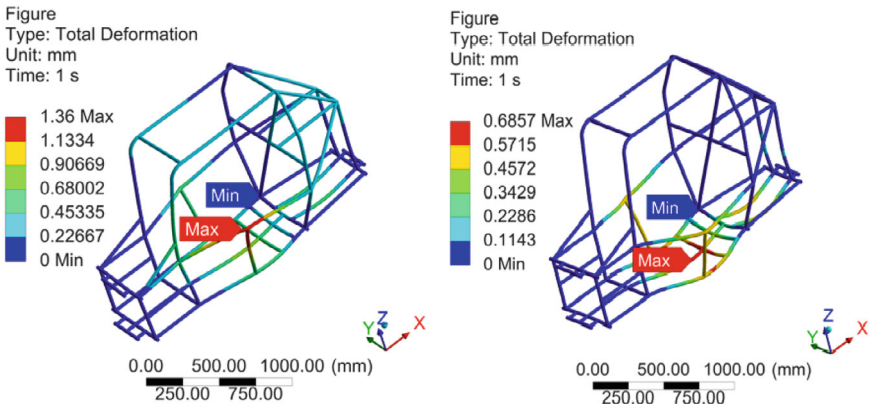
These two conditions are crucial for assessing the frame's performance and safety.

Comprehensive analysis of lateral and longitudinal acceleration conditions. In the analysis of a Baja racing car's working conditions, two key scenarios were evaluated: lateral acceleration during turning and longitudinal acceleration during emergency braking.

(1) Lateral Acceleration Condition: During left turns, a lateral force of − 1050 N on the y-axis (due to centrifugal effects) and gravity acceleration of − 9.8 m/s^2 on the z-axis were applied. Constraints were placed on the translational freedom of the suspension-frame connections. The resulting deformation, shown in Fig. 3a, included a slight twist in the anti-roll ring and a maximum displacement of 1.36 mm in the base support structure. The maximum stress, located at the USM anti-roll hoop, was 159.6 MPa, indicating frame safety under this condition.

(2) Longitudinal Acceleration Condition: During emergency braking, a force of 1275 N in the x-axis direction and gravity acceleration of − 9.8 m/s^2 on the z-axis were applied, with similar constraints on suspension connection freedom [8]. Figure 3b revealed a maximum deformation of 0.69 mm in the base support structure. The maximum stress, at the connection between the anti-roll ring and base, was 140.6 MPa, confirming frame safety under this condition as well.

This section examines the strength and stiffness of the frame's components across various working conditions, focusing on deformation under bending, bending-plus-twisting loads. Results indicate significant deformation at the joint between the right top of the RRH anti-roll ring and the tail of the RHO anti-roll hoop under bending and twisting. Additionally, components such as the seat-mounted and engine



(a) Lateral acceleration (b) Longitudinal acceleration

Fig. 3 Deformation contour of lateral and longitudinal acceleration conditions

Table 1 Natural excitation frequency table of Baja racing car

Item	Exciting frequency (Hz)
Uneven pavement	1–20
Engine running	14.58–31.66
Unbalanced transmission system	$\leq$ 10.46

bilge oblique rods exhibit notable deformation across multiple conditions, yet all deformation values remain within the material's acceptable range.

3.2 Modal Analysis

The modal analysis results were assessed using an analytical method that took into account external excitation frequency, the structure's natural frequency, and mode shape. Vehicle vibrations primarily arise from uneven road surfaces, engine operation, and power transmission imbalances. The Baja racing car's natural vibration frequencies were obtained through data review and calculation, as detailed in Table 1. Therefore, the first-order natural frequency of the designed frame should be outside the range of the above three excitation frequencies, so that it can meet the requirements of dynamic characteristics.

Figures 4, 5, 6, 7, 8 and 9 illustrate the frame's vibrational behavior at various frequencies. Specifically, at the first-order frequency of 37.1 Hz, the frame experiences lateral swinging at the rear. The second-order frequency of 48.9 Hz corresponds to back-and-forth swaying. The third-order frequency at 71.1 Hz results in lateral swinging of the front end. The fourth-order frequency of 75.1 Hz shows overall bending of the front end around the x-axis. The fifth frequency at 75.8 Hz exhibits forward–backward bending and torsional motion. Lastly, at the sixth-order frequency of 101.4 Hz, the frame undergoes bending around the y-axis. Detailed parameters are provided in Table 2.

It can be known from Table 2 that the natural frequency of the frame is mainly distributed between 37.1 and 101.4 Hz, which is much higher than the natural excitation frequency of Baja racing car. Therefore, the natural frequency of the frame is within a reasonable range, which improves the comfort and stability of the whole vehicle.

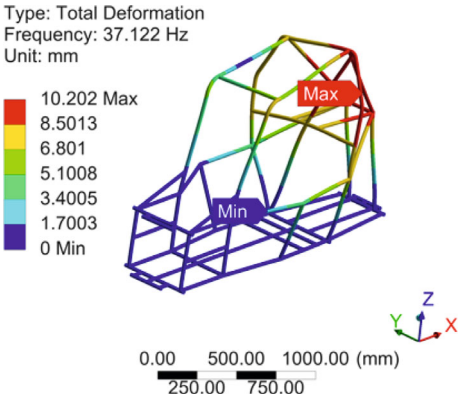


Fig. 4 First order amplitude

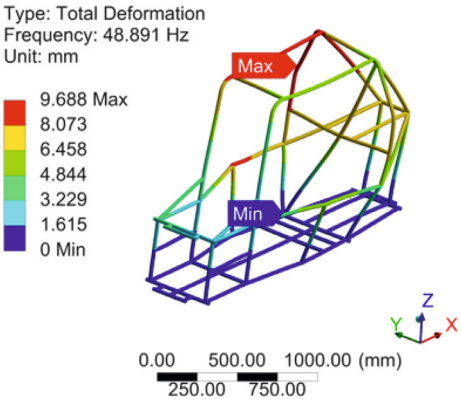


Fig. 5 Second order amplitude

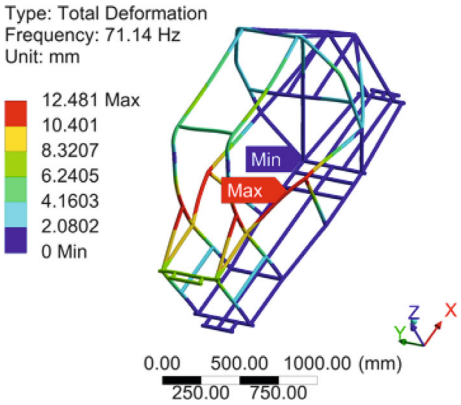


Fig. 6 Third order amplitude

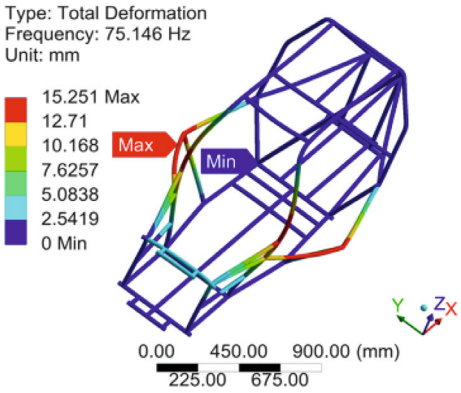


Fig. 7 Fourth order amplitude

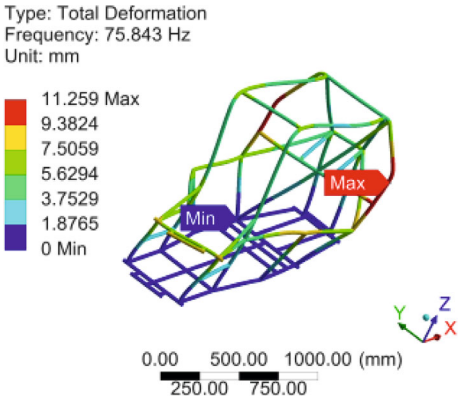


Fig. 8 Fifth order amplitude

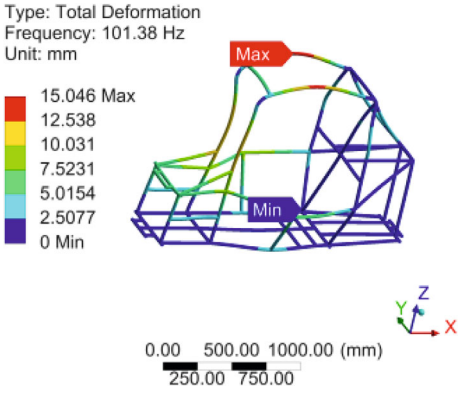


Fig. 9 Sixth order amplitude

Table 2 Natural excitation frequency of each order for the frame

Modal order	Natural frequency (Hz)
1	37.1
2	48.9
3	71.1
4	75.1
5	75.8
6	101.4

4 Conclusion

This study utilized CATIA and ANSYS software to design, verify strength, and analyze modal characteristics of a Baja racing frame with a space truss structure. FEA using ANSYS/Workbench confirmed that the frame's stress and deformation under various conditions are acceptable, and its natural frequencies exceed the vehicle's vibration frequencies. The results validate the frame's reliability, operational performance, and enhanced driving comfort in extreme off-road conditions.

The frame was constructed using CATIA 3D software, grounded in theoretical foundations yet requiring further enhancements. The racing car represents a sophisticated blend of rigidity and flexibility, necessitating future research to integrate flexible body dynamics for enhanced model accuracy and precise operational performance analysis. Additionally, field data collection during actual races will validate and refine the design, incorporating driver feedback and requirements to optimize real-race performance. In summary, through innovative exploration, Baja racing design aims to achieve multifaceted breakthroughs and significantly contribute to racing advancements.

Acknowledgements Our work is supported by the "Innovation and Strengthening Universities Project" (Research Category) funded by Guangdong University of Science and Technology, with project number GKY-2022CQPT-1.

References

1. Ye YH, Fu AJ, Chen ZY et al (2023) Simulation and optimization of Baja racer frame based on HyperWorks. J Guangxi Univ Sci Technol 3
2. Lu S (2019) Design and simulation of BSC racing car frame. Agric Equip Veh Eng 103–107:116
3. Guzel Z, Ekaterina G, Herman B (2020) Amperometric sensor based on MWNT and electropolymerized carminic acid for the simultaneous quantification of TBHQ and BHA. J Electroanal Chem 859
4. David Raja P, Naresh Ch, Phanisankar BSS (2020) Design and structural analysis of Baja frame with conventional and composite materials. Int J 23–31
5. Ward DC, Araújo P, Orlando D et al (2021) Structural analysis of the Baja SAE vehicle from UERJ technology faculty. In: SAE technical paper, pp 1–12

6. Wang ZX (2020) Frame optimization analysis of Bach off-road racing car. Sci Technol Innov 9:104–105, 107
7. Shi L, Zhu CD, Liu X et al (2020) Optimum design of the double roll rotary forging machine frame. Mechan Sci 11(1):101–1140
8. Vishwakarma A (2021) Design and finite elements analysis of ATV rollcage modeling. Int Res J Eng Technol (IRJET) 12(08):588–596

Open Access This chapter is licensed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License (http://creativecommons.org/licenses/by-nc-nd/4.0/), which permits any noncommercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if you modified the licensed material. You do not have permission under this license to share adapted material derived from this chapter or parts of it.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.



Influence of Hydro-pneumatic Suspension Structural Parameters on Off-Road Vehicle Ride Comfort

Ziqi Huang, Liqiang Zhao, Hao Xiong, Liangpeng Li, and Dingxuan Zhao

Abstract This research delves into how the design elements of hydro-pneumatic suspension systems (HPSS) impact the comfort of the ride, focusing on a particular brand of off-road vehicle. The suspension system was engineered, and its mathematical representation was formulated. The study explored the effects of varying the accumulator's initial pressure and the size of the throttle valve opening on the system's stiffness and damping properties. A dynamic simulation of the HPSS was developed using MATLAB-Simulink. By adjusting the damping forces to adapt to different terrains, significant enhancements in vehicle performance and comfort were realized. This methodical tuning not only boosts the stability of the drive but also reduces the disturbance caused by uneven road surfaces on the vehicle's dynamics. The results highlight the significance of refining damping force management to optimize both the vehicle's handling and the comfort of its passengers in practical driving conditions.

Keywords Hydro-pneumatic suspension systems (HPSS) · Damping characteristics · Stiffness characteristics · Off-road vehicle

1 Introduction

The suspension system is a fundamental component of vehicle design, fulfilling critical functions. It not only supports the vehicle's weight and maintains body stability but also plays a vital role in damping vibrations and enhancing passenger comfort [1–3]. The suspension system contributes to driving safety by ensuring optimal contact between the tires and the road, thereby improving grip and handling. Under complex

Z. Huang · L. Zhao · H. Xiong · L. Li · D. Zhao (✉)
School of Mechanical Engineering, Yanshan University, Qinhuangdao, Hebei, China
e-mail: zdx-yw@ysu.edu.cn

D. Zhao
State Key Laboratory of Crane Technology, School of Mechanical Engineering, Yanshan University, Qinhuangdao, China

© The Author(s) 2026

H. Saman et al. (eds.), *The 9th International Conference on Advances in Construction Machinery and Vehicle Engineering*, Lecture Notes in Mechanical Engineering,
https://doi.org/10.1007/978-981-96-5527-4_84

driving conditions, the performance of the suspension directly influences the vehicle's dynamic response and overall driving experience. However, the majority of vehicles are equipped with passive suspension systems, which rely on fixed spring and shock absorber settings that do not self-adjust based on road and driving conditions. This limitation can lead to compromised stability and comfort, particularly when navigating rough terrain or traveling at high speeds. Furthermore, the passive suspension's inability to effectively absorb shocks and control body attitude can result in significant vibrations and jolts, negatively impacting both driving safety and passenger comfort. In contrast, semi-active and active suspension technologies are garnering increasing attention as crucial advancements in modern automotive suspension systems. These systems are capable of real-time adjustments to suspension stiffness and height through intelligent control mechanisms, thereby significantly enhancing both vehicle handling and passenger comfort [4–10].

Semi-active suspension systems primarily utilize passive components, such as springs and shock absorbers, while adjusting the damping force in real time based on road conditions through an adjustable damping mechanism. In comparison to active suspension systems, semi-active suspensions feature a relatively simpler structure, lower cost, and ease of maintenance [11]. HPSS are engineered to endure extreme operating conditions, making them particularly suitable for off-road vehicles that often navigate rugged and uneven terrain. These systems effectively absorb ground shocks and vibrations, thereby providing a smooth driving experience [12, 13]. By adjusting structural parameters such as accumulator gas pressure, check valve overflow area, and throttle orifice diameter, the performance of off-road vehicles under rough road conditions can be significantly enhanced. Previous research has extensively investigated the impact of structural parameters of HPSS on vehicle performance. For instance, Mathieu et al. [14] proposed a method for selecting appropriate real-time suspension stiffness and damping, analyzing the dynamics of off-road vehicles in complex terrains to facilitate the traversal of high obstacles. Sang et al. [15] utilized theories of nonlinear vibration and fractional calculus to delve into the behavior of dual-chamber HPSS, examining how design parameters influence the system's stiffness and damping coefficients through computational modeling. Zheng et al. [3] concentrated their attention on the HPSS of a tractor's front axle, examining how variations in accumulator initial inflation pressure, initial inflation volume, and throttle flow diameter influence the tractor's pitch and roll stability, as well as ride comfort. Ghazaly et al. [16] studied an external damping valve-type HPSS using random road surfaces as input excitation, simulating the system in the AMESim environment to ascertain the influence of key structural parameter variations on vibration damping performance. Lee [17] utilized fluid dynamics theory to establish a parameterized model of a HPSS with a single air chamber, proposing a model that addresses high-frequency excitation, the study indicated that cavitation phenomena in HPSS, when subjected to high-frequency excitation, arise from excessive pressure differentials across the throttle orifice, leading to hydraulic chamber pressures falling below the vaporization pressure of the hydraulic fluid. The suspension damping force and displacement were compared across different temperatures through bench tests.

Furthermore, Tsuruga [18] Suruga proposed a theory of variable stiffness for hydro-pneumatic springs and canopy damping control for applications in wheeled hydraulic excavators, aimed at maintaining stability and attenuating vibration energy during operation. Lastly, Yin et al. [19] presented a HPSS that incorporates a rapid-switching valve, enabling the generation of two distinct spring forces and four varied damping forces.

In summary, many researchers employ intelligent algorithms to optimize the architectural parameters of HPSS, aiming to ensure smoothness and driving stability across various road surfaces. However, in practice, the majority of off-road vehicles are designed with a non-independent suspension structure for the rear wheels to enhance strength and durability in extreme off-road conditions. This design typically offers improved load capacity and stability, effectively adapting to uneven terrain while simplifying the overall structure and reducing maintenance costs. To ensure the accuracy of modeling and facilitate a better analysis of vehicle performance, this study adopts an independent suspension structure for the front wheels and a non-independent suspension for the rear wheels. A 7-DOF model of the vehicle's HPSS is constructed based on this configuration. Different pre-inflation pressures of the accumulators and throttle valve apertures are set to observe the effects of these structural parameters on the output force of the HPSS. A control group is implemented to compare the influence of various throttle valve apertures on the vehicle's performance evaluation metrics. The simulation data reveal that strategically enhancing the damping force when the hydraulic cylinder extends, and diminishing it during compression, can significantly lessen the vertical and pitch accelerations, as well as the dynamic disturbances to the suspension and the dynamic displacement of the tires on the vehicle body [20].

2 Structural Design of HPSS

The HPSS setup comprises a hydraulic cylinder, check valves, throttle valves, and an accumulator. The hydraulic cylinder is mounted such that its piston end is connected to the vehicle frame, and the opposite end, which is rodless, is linked to the wheel. Upon encountering irregularities in the road, the hydraulic cylinder experiences both extension and retraction movements. In the compression stage, some of the fluid from the rodless chamber is transferred into the accumulator, and another part moves through check valve 4 and throttle hole 3 into the rod chamber. During the extension stage, fluid from the accumulator returns to the rodless chamber, while fluid from the rod chamber flows back through check valve 1 and throttle hole 2 into the rodless chamber.

A diagrammatic representation of the HPSS's structure is shown in Fig. 1, where the blue dashed lines delineate the fluid flow path during the hydraulic cylinder's compression phase, and the red dashed lines show the path during the extension phase. The suspension's stiffness and damping properties can be modified by adjusting the accumulator's pre-charge pressure and the dimensions of the throttle openings.

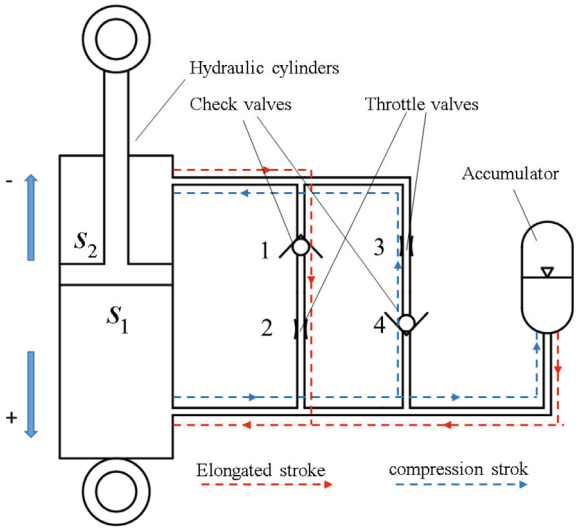


Fig. 1 Schematic diagram of HPSS structure

3 Mathematical Modeling and System Characteristics

The following assumptions are made to simplify the mathematical modeling and system characteristics analysis of the HPSS:

(1) The gas's volume and pressure inside the accumulator are considered unaffected by temperature variations.
(2) The hydraulic fluid is considered incompressible, and the oil–gas suspension system is assumed to be well-sealed and free from leakage.
(3) The effects of gas dissolution in the hydraulic oil are neglected in this analysis.

3.1 7-DOF Vehicle Dynamics Modeling

The subject of this research is a particular make of off-road vehicle. To maintain the integrity of the model and to enable a precise evaluation of the vehicle's dynamic behavior, a 7-DOF model has been developed. This model encompasses the vertical, pitch, and roll motions of the vehicle body, in addition to the vertical movement of each of the four wheels. A diagram of the 7-DOF vehicle configuration is displayed in Fig. 2. The off-road vehicle is characterized by a separate suspension setup for the front axle and a non-independent, solid rear axle with an integral bridge suspension system.

In the realm of vehicle dynamics modeling, it is common for researchers to assign fixed values to spring stiffness and damping coefficients, simplifying the suspension system as linear. However, the HPSS presents intricate non-linear dynamics. To

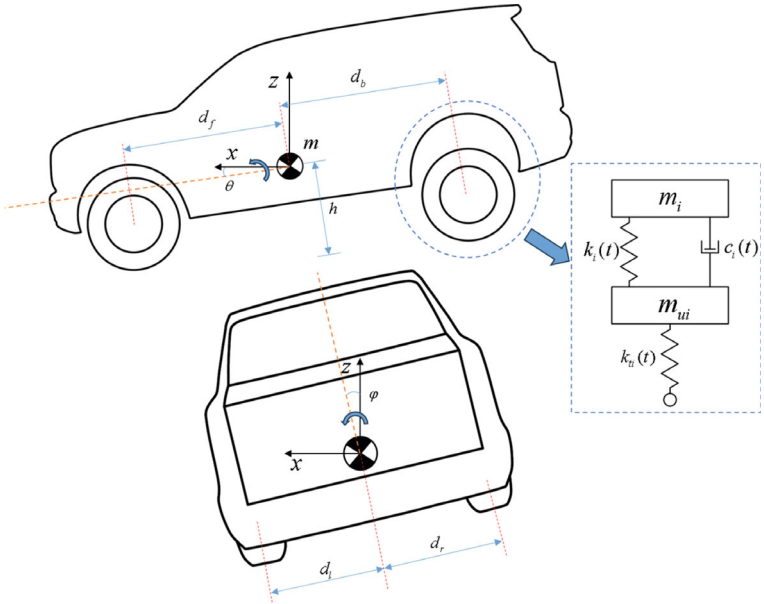


Fig. 2 7-DOF model of the vehicle

accurately capture these non-linear properties of stiffness and damping, this study adopts a model that incorporates adjustable elastic and damping forces.

According to Newton's second law, the dynamic equations for the vertical motion of the vehicle body can be expressed as:

$$\begin{aligned} mz'' = {} & k_{lf}(t)(z_{ulf} - z_{lf}) + k_{rf}(t)(z_{urf} - z_{rf}) \\ & + k_{rr}(t)(z_{urr} - z_{rr}) + k_{lr}(t)(z_{ulr} - z_{lr}) \\ & + c_{lf}(t)(z'_{ulf} - z'_{lf}) + c_{rf}(t)(z'_{urf} - z'_{rf}) \\ & + c_{rr}(t)(z'_{urr} - z'_{rr}) + c_{lr}(t)(z'_{ulr} - z'_{lr}) \end{aligned} \tag{1}$$

The dynamic equation for the lateral tilt motion of the vehicle body can be expressed as:

$$\begin{aligned} I_x\varphi'' = {} & [k_{lf}(t)(z_{ulf} - z_{lf}) + c_{lf}(t)(z'_{ulf} - z'_{lf}) \\ & + k_{lr}(t)(z_{ulr} - z_{lr}) + c_{lr}(t)(z'_{ulr} - z'_{lr})]d_l \\ & - [k_{rf}(t)(z_{urf} - z_{rf}) + c_{rf}(t)(z'_{urf} - z'_{rf}) \\ & + k_{rr}(t)(z_{urr} - z_{rr}) + c_{rr}(t)(z'_{urr} - z'_{rr})]d_r \end{aligned} \tag{2}$$

The dynamic equation for the pitch motion of the vehicle body can be expressed as:

$$\begin{aligned} I_y\theta'' = & [k_{rr}(t)(z_{urr} - z_{rr}) + c_{rr}(t)(z'_{urr} - z'_{rr}) \\ & + k_{lr}(t)(z_{ulr} - z_{lr}) + c_{lr}(t)(z'_{ulr} - z'_{lr})]d_b \\ & - [k_{lf}(t)(z_{ulf} - z_{rf}) + c_{lf}(t)(z'_{ulf} - z'_{lf}) \\ & + k_{rf}(t)(z_{urf} - z_{rf}) + c_{rf}(t)(z'_{urf} - z'_{rf})]d_f \end{aligned} \tag{3}$$

Due to the rear axle being a non-independent suspension, the geometric relationship of the relative height of the rear wheels can be observed from Fig. 2. The expression for the relative height of the rear wheels is:

$$d = z_{urr} - z_{ulr} = (d_0/2) \times \tan(\varphi) \tag{4}$$

Based on Eq. (4), the coupling motion relationship of the rear wheels can be derived. According to Newton's second law, the dynamic expression for the vertical motion of the unsprung mass is as follows:

$$\begin{cases} m_{ulf} z''_{ulf} = k_{tlf}(q_{lf} - z_{ulf}) - k_{lf}(t)(z_{ulf} - z_{lf}) - c_{lf}(z'_{ulf} - z'_{lf}) \\ m_{urf} z''_{urf} = k_{trf}(q_{rf} - z_{urf}) - k_{lf}(t)(z_{urf} - z_{rf}) - c_{lf}(z'_{urf} - z'_{rf}) \\ m_{urr} z''_{urr} = z''_{ulr} + d_0 \sec^2(\varphi) \tan(\varphi) \\ m_{ulr} z''_{ulr} = k_{tlr}(q_{lr} - z_{ulr}) - k_{lr}(t)(z_{ulr} - z_{lr}) - c_{lf}(z'_{urf} - z'_{rf}) \end{cases} \tag{5}$$

The relative displacement equation of the hydraulic cylinder pistons of HPSS is as follows:

$$\begin{cases} z_{lf} = z - d_f \sin(\theta) + d_l \sin(\varphi) \\ z_{rf} = z - d_f \sin(\theta) - d_r \sin(\varphi) \\ z_{rr} = z + d_b \sin(\theta) - d_r \sin(\varphi) \\ z_{lr} = z + d_b \sin(\theta) + d_l \sin(\varphi) \end{cases} \tag{6}$$

where m is the vehicle body mass, and $m_{uli}(i = lf, rf, lr, rr)$ represents the four unsprung masses, respectively; I_x, I_y are the lateral tilt moment of inertia and pitch moment of inertia, respectively; $k_i(t)(i = lf, rf, lr, rr)$ are the four suspension stiffness coefficients, respectively; $c_i(t)(i = lf, rf, lr, rr)$ are the four suspension damping coefficients, respectively; $z_i(i = lf, rf, lr, rr)$ are the vertical displacements of the four sprung masses, respectively; $z_{ui}(i = lf, rf, lr, rr)$ are the vertical displacements of the four unsprung masses, respectively; $q_i(i = lf, rf, lr, rr)$ are the four road excitation inputs, respectively; d_b, d_f, d_l, d_r are the center of mass-to-front-axis distance, center of mass-to-rear-axis distance, center of mass-to-left suspension distance and center of mass-to-right suspension distance, respectively; z, φ, θ are the vertical body displacements and rear wheel center relative height and rear wheel center clearance, respectively; d, d_0 are the rear wheel center relative height and rear wheel center clearance, respectively.

3.2 HPSS Modeling

HPSS, in contrast to conventional suspension systems, demonstrate non-linear properties in terms of stiffness and damping, which are affected by various elements including ambient conditions, the configuration of the hydraulic cylinder, and the operational status of the accumulator. This study delves into the effects of these diverse parameters on the HPSS's stiffness and damping traits by developing mathematical models and performing simulations using MATLAB. A compilation of the simulation parameters is provided in Table 1.

The following are the Mathematical models:

(1) Mathematical model of HPSS unit

Using the fully loaded suspension's equilibrium position as a reference point, or the origin, the direction of suspension compression is defined as positive. In this setup, the downward movement of the piston rod is considered a positive displacement, whereas the upward movement is considered negative.

Regarding the force analysis of the piston, the equation for the output force of the hydraulic cylinder within the HPSS can be formulated as follows:

$$F = p_1 s_1 - p_2 s_2 \tag{7}$$

where p_1 is rodless cavity pressure, p_2 is rodded cavity pressure, s_1 is rodless cavity area, s_2 is rodded cavity area.

(2) Mathematical model of accumulator

The HPSS's accumulator plays a pivotal role in regulating the flow of hydraulic fluid through the suspension's dynamic compression and extension. This mechanism involves the compression and decompression of the inert gas contained within the accumulator's gas chamber, which in turn enables the accumulation and release of energy. Nitrogen, known for its behavior closely resembling that of an ideal gas, is commonly utilized as the gas within the accumulator. Therefore, the mathematical

Table 1 Simulation parameters

Parameter name	Symbol	Unit	Parameter value
Flow coefficient	C_d	–	0.7
Oil density	ρ	kg/m^3	900
Check valve flow area	A_d	mm^2	20
Accumulator initial charging pressure	P_0	Mpa	3.5
Accumulator initial inflation volume	V_0	L	0.5
Gas Variable index	r	–	1.5
Piston rod diameter	R_1	mm	40
Hydraulic cylinder inner diameter	R_2	mm	22

representation of the accumulator can be derived from the thermodynamic principles governing the gas's behavior during its operation, as detailed below:

$$PV^r = P_0 V_0^r \tag{8}$$

where V_0 and P_0 represent the volume of the accumulator and the pre-inflation pressure, respectively; V and P are the volume of the gas and the gas pressure in the accumulator at a specific moment under working conditions, respectively; and r is the gas variability index.

(3) The change of gas volume is

$$V = V_0 - (s_1 - s_2)x \tag{9}$$

In the equation: x represents the piston displacement.

(4) The preload of the HPSS unit is

$$F_0 = P(s_1 - s_2) \tag{10}$$

HPSS stiffness characteristics. According to Eqs. (7)–(10), the gas elastic force of the HPSS unit can be expressed as:

$$F_k = \frac{P_0 V_0^r (s_1 - s_2)}{[V_0 - (s_1 - s_2)x]^r} \tag{11}$$

The derivative of the elastic force with respect to the displacement x represents the stiffness of the HPSS unit, which can be expressed as:

$$K = \frac{dF_k}{dx} = \frac{rP_0 V_0^r (s_1 - s_2)^2}{[V_0 - (s_1 - s_2)x]^{r+1}} \tag{12}$$

Maintaining all other factors at a constant level, the initial pressure within the accumulator was varied to 2.5, 3.0, and 3.5 Mpa. Figure 3 displays the resulting stiffness characteristic curves of the HPSS. As shown in the Fig. 3, there is a clear trend showing that higher initial pressures within the accumulator lead to increased stiffness in the HPSS.

HPSS damping characteristics. The flow equation through the throttle valve is:

$$Q = C_d \left(A_{j1} \left(\frac{1}{2} - \frac{1}{2} sign(\dot{x}) \right) + A_{j2} \left(\frac{1}{2} + \frac{1}{2} sign(\dot{x}) \right) + A_d \right) \sqrt{\frac{2(p_1 - p_2)}{\rho}} \tag{13}$$

where A_{j1} is the overflow area of throttle 1, A_{j2} is the overflow area of throttle 2.

The flow rate between the rodless chamber and the rod chamber of the hydraulic cylinder is:

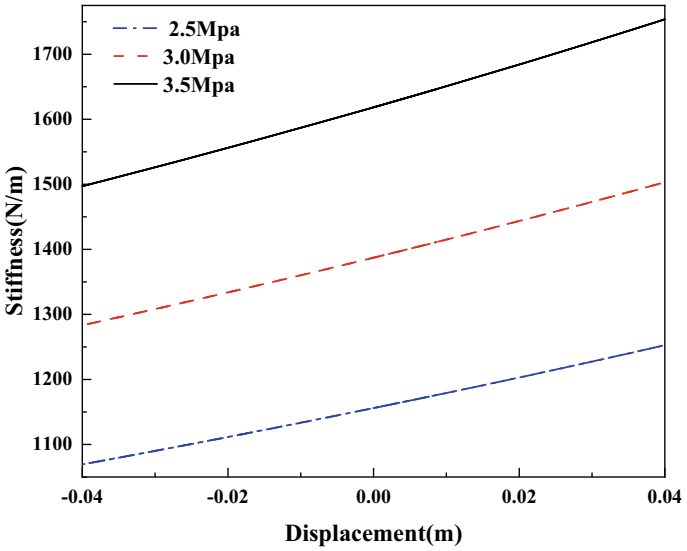


Fig. 3 HPSS stiffness characteristic curve

$$Q = s_2\dot{x} \tag{14}$$

The damping force of the HPSS is:

$$F_c = (p_1 - p_2)s_2 \tag{15}$$

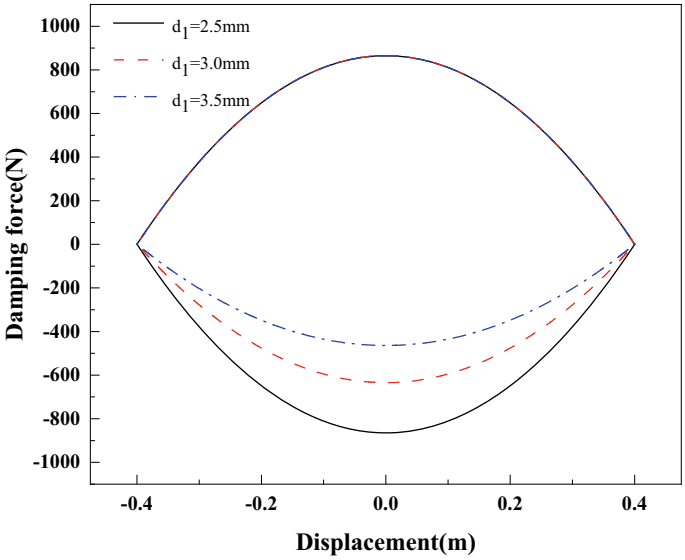
By simultaneously solving Eqs. (13)–(15), it can be obtained that:

$$F_c = \frac{1}{2}\rho \frac{s_2^3\dot{x}^2}{c_d^2[A_{j1}\left(\frac{1}{2} - \frac{1}{2}sign(\dot{x})\right) + A_{j2}\left(\frac{1}{2} + \frac{1}{2}sign(\dot{x})\right) + A_d]^2} sign(\dot{x}) \tag{16}$$

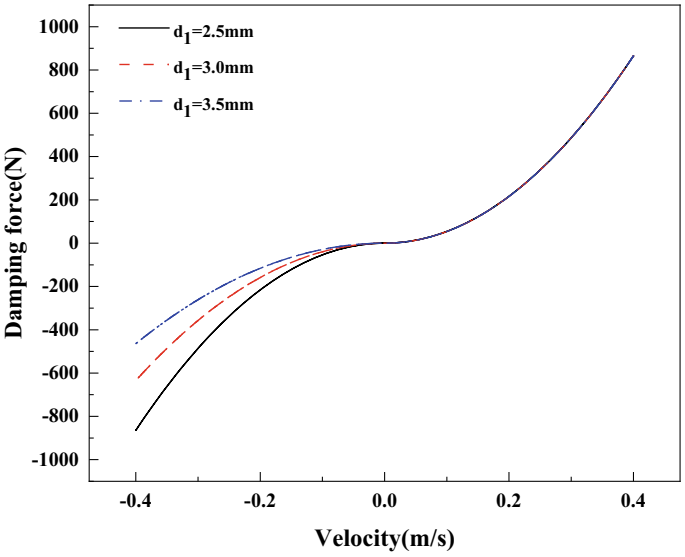
The damping coefficient is obtained by taking the derivative of the damping force with respect to velocity, and it can be expressed as follows:

$$C = \frac{\rho s_2^3\dot{x}}{c_d^2[A_{j1}\left(\frac{1}{2} - \frac{1}{2}sign(\dot{x})\right) + A_{j2}\left(\frac{1}{2} + \frac{1}{2}sign(\dot{x})\right) + A_d]^2} \tag{17}$$

While keeping all other variables unchanged, the throttle opening was adjusted to 2.5, 3.0, and 3.5 mm for the simulation experiments, leading to the damping characteristic curves of the HPSS as depicted in Fig. 4. The data trends revealed that an increase in the throttle valve's opening size corresponds to a reduction in the damping force of the HPSS, as well as a decrease in both the peak value and the rate of variation of the damping coefficient.



(a) Displacement characteristic curve



(b) Velocity characteristic curve

Fig. 4 Damping characteristic curve of HPSS

3.3 System Characterization

Drawing on the operational mechanics and mathematical framework of the HPSS, an extensive dynamic model of the vehicle's HPSS was constructed utilizing the MATLAB-Simulink platform. The stimulus for the vehicle's four wheels was set to a sinusoidal waveform, characterized by a peak-to-peak displacement of 5 mm and a frequency of 0.5 Hz. A compilation of the key parameters for the vehicle's model can be found in Table 2.

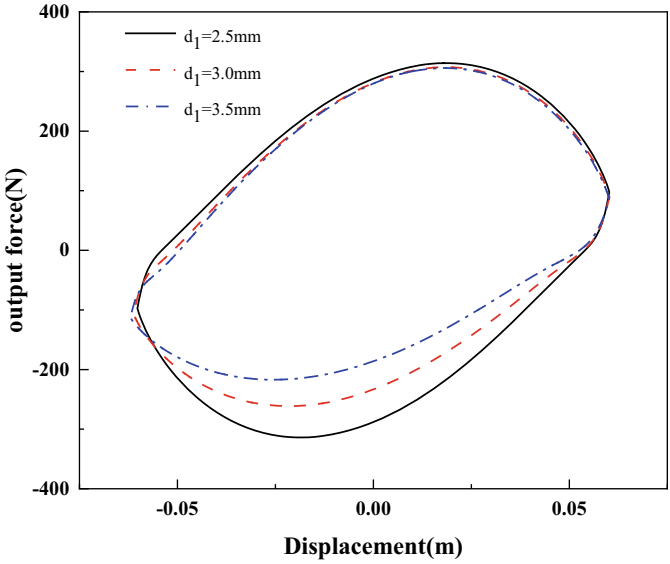
The influence of the HPSS's configuration on its force generation was investigated by adjusting the throttle valve's diameter and the accumulator's initial pressure. The tested diameters for the throttle valve were 2.5 mm, 3.0 mm, and 3.5 mm, corresponding to accumulator pressures of 2.5 MPa, 3.0 MPa, and 3.5 MPa.

Under a constant initial pressure in the accumulator, Throttle Valve 3 was fixed at a 2.5 mm aperture, while Throttle Valve 2's aperture was varied to 2.5, 3, and 3.5 mm. This setup was employed to study the impact of Throttle Valve 2's aperture on the suspension's force output. Figure 5 depicts the characteristic curves of displacement and velocity for the suspension's force output, revealing that the force varies with the piston rod's displacement and velocity in a nonlinear manner. With Throttle Valve 3's aperture held constant, the force output during the compression phase of the hydraulic cylinder remains stable; however, an increase in Throttle Valve 2's aperture results in a reduction of the suspension's force output.

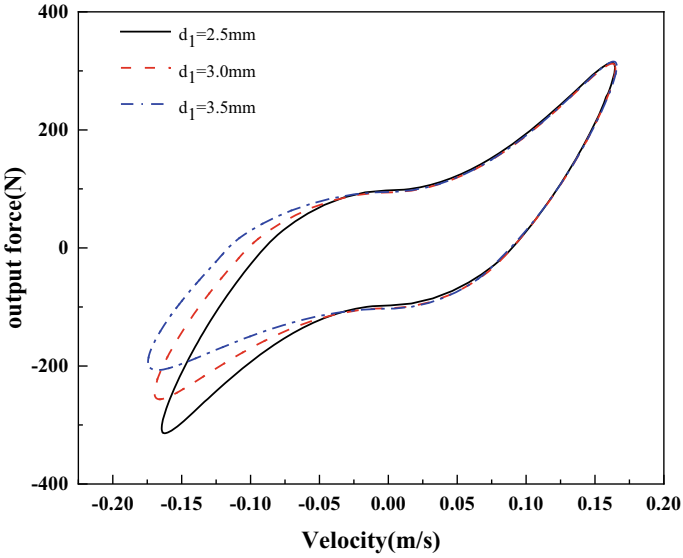
With Throttle 2's opening diameter fixed at 2.5 mm and Throttle 3's opening diameters tested at 2.5, 3, and 3.5 mm under a consistent initial inflation pressure of the accumulator, the experimental configuration aimed to assess the impact of Throttle 3's opening on the suspension's force output. Figure 6 presents the characteristic curves of the suspension's output force in terms of displacement and velocity, indicating a decrease in the suspension's force output as the aperture size of Throttle 3 increases.

Table 2 Vehicle model parameters

Vehicle parameters	Symbol	Numerical value
Body mass	m	2670 kg
Front axle unsprung mass	m_{ulf}, m_{urf}	36 kg
Rear axle unsprung mass	m_{ulr}, m_{urr}	36 kg
Lateral inertia	I_x	1639.68 kg/m^2
Pitch moment of inertia	I_y	6526.79 kg/m^2
Distance from center of mass to front axle	d_f	1 m
Center of mass to rear axle	d_b	1.5 m
Distance from center of mass to left side suspension	d_l	1.375 m
Distance from center of mass to right side suspension	d_r	1.375 m
Tire equivalent stiffness	k_{tlf}, k_{trf}, k_{tlr}, k_{trr}	10, 000 m/N
Rear axle distance	d	2.75 m

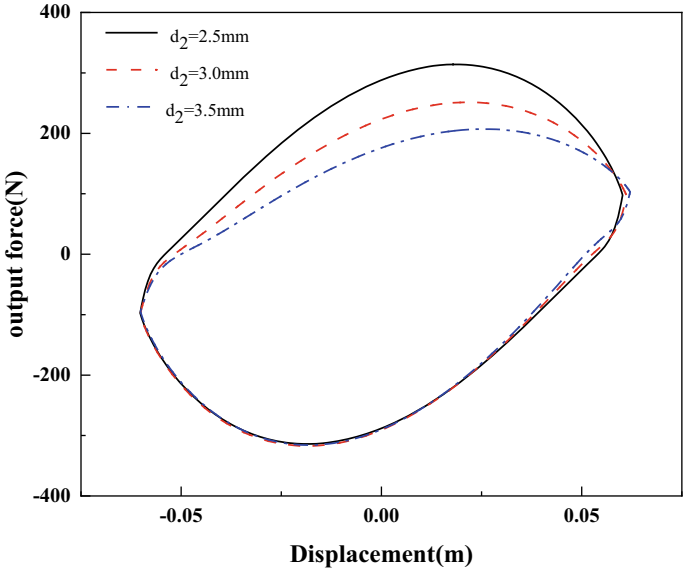


(a) Displacement characteristic curve

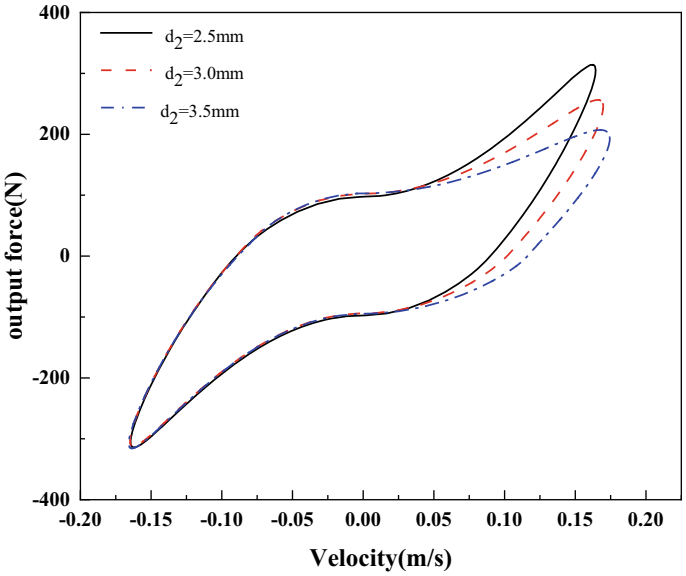


(b) Velocity characteristic curve

Fig. 5 Effect of throttle 2 orifice size on the performance of HPSS



(a) Displacement characteristic curve



(b) Velocity characteristic curve

Fig. 6 Effect of throttle orifice 2 diameter size on the performance of HPSS

To observe the impact of the accumulator initial inflation pressure on the suspension output force, the accumulator initial inflation pressure is adjusted to 2.5, 3, and 3.5 MPa when the throttle aperture size remains constant. Figure 7 displays the suspension output force's displacement and velocity characteristic curves. It is evident from Fig. 7 that as the accumulator starting inflation pressure rises, so does the suspension output force's magnitude.

4 Whole Vehicle Dynamics Simulation

The vehicle excitation input is the Class C random road surface; Fig. 8 displays the front and rear wheel excitation inputs, and Eq. (18) displays the Class C road surface unevenness time domain model.

$$z_g(t)' = -2\pi n_1 u z_g(t) + 2\pi n_0 \sqrt{G_q(n_0)u}\omega(t) \tag{18}$$

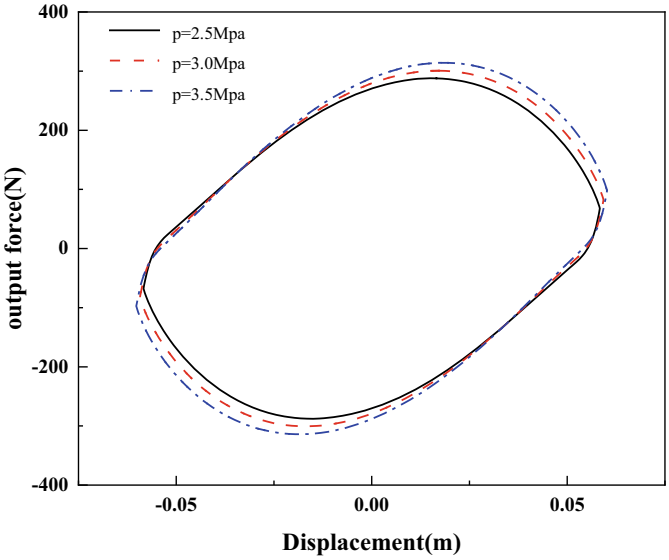
where z_g is the height of the road surface; n_1 is the lower cutoff spatial frequency, take the value of 0.01 m^{-1}; u is the vehicle speed, take the value of 20 m/s; n_0 is the reference spatial frequency, take the value of 0.1 m^{-1}; $\sqrt{G_q(n_0)}$ is the unevenness coefficient of the road surface, the value of the C-class pavement is 256×10^{-6} m^3; $\omega(t)$ is the white noise of the road surface, the noise power was set to 0.5, the sampling time was set to 0.005

Setting up the control group as shown in Table 3.

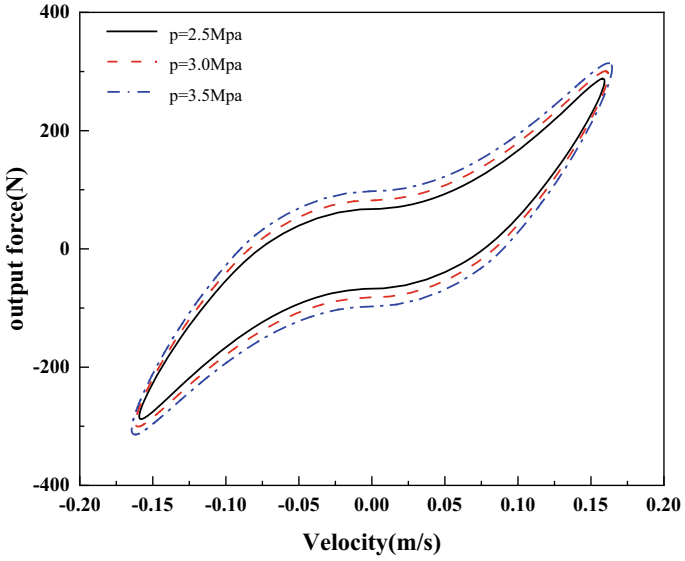
Figures 9, 10, 11, 12, 13 and 14 illustrate the simulation results for the vehicle's performance metrics, Table 4 offers a compilation of the root mean square values for these metrics. By strategically adjusting the throttle valve settings—with a reduction in the diameter of throttle valve 3's opening and an enlargement of throttle valve 2's opening—at a constant initial pressure in the accumulator, it is possible to effectively minimize the vehicle's body vertical acceleration, pitch acceleration, suspension dynamic disturbance, and tire displacement. In detail, this adjustment leads to a 7.1% reduction in the vertical acceleration of the vehicle body, a 2.9% decrease in pitch acceleration, a 10.2% reduction in the dynamic disturbance of the front suspension, and a 9% decrease for the rear suspension. Moreover, the dynamic displacement of the front tire is lessened by 3.5%, and that of the rear tire by 3.1%.

5 Conclusion

This research develops a comprehensive mathematical representation of the vehicle's HPSS and engineers its architecture. Utilizing MATLAB-Simulink, a dynamic model of the vehicle is constructed, which is grounded on this mathematical framework. The study investigates how the structural attributes of the HPSS influence the damping and elastic properties of the system, as well as various performance indicators of the



(a) Displacement characteristic curve



(b) Velocity characteristic curve

Fig. 7 Effect of accumulator pre-inflation pressure on the performance of HPSS

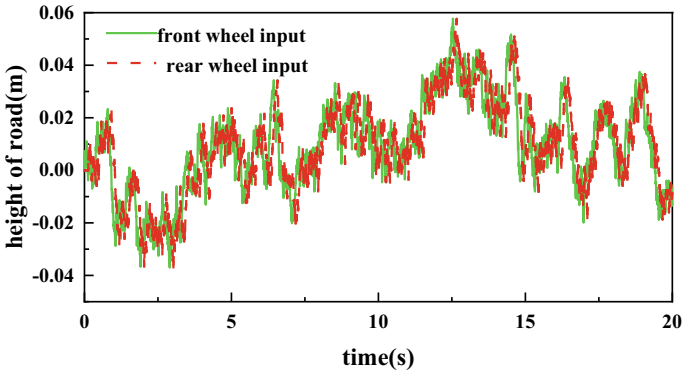


Fig. 8 Class C randomized pavement time domain map

Table 3 Control group

Parameter	d_2 (mm)	d_3 (mm)	p_0 (Mpa)
Mode1	3.5	2.5	3.5
Mode2	2.5	3.5	3.5

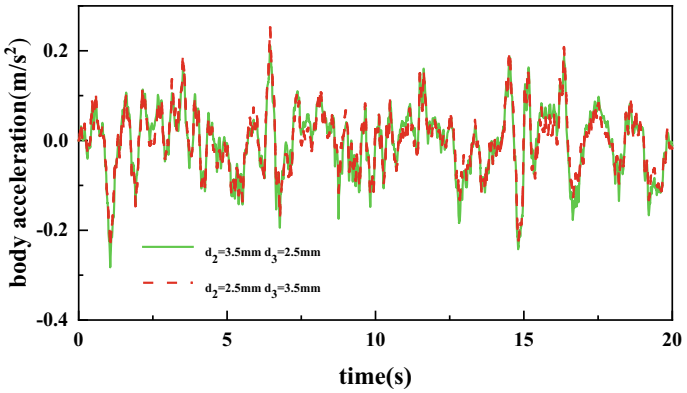


Fig. 9 Vertical body acceleration

vehicle. This is achieved by subjecting the suspension to both stochastic road surface inputs and sinusoidal excitation signals. The findings can be summarized as follows:

1. The HPSS exhibits non-linear stiffness and damping characteristics. Diminishing the diameter of the throttle opening enhances the suspension's damping capacity and output force, provided that the initial pressure in the accumulator is kept constant. Conversely, increasing the initial pressure in the accumulator raises the system's stiffness and output force, with the throttle opening held constant.

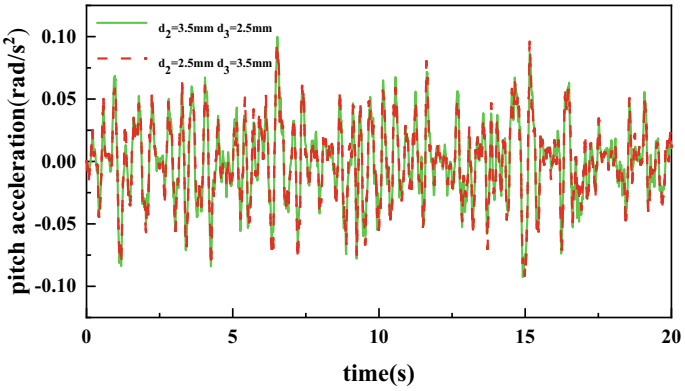


Fig. 10 Body pitch acceleration

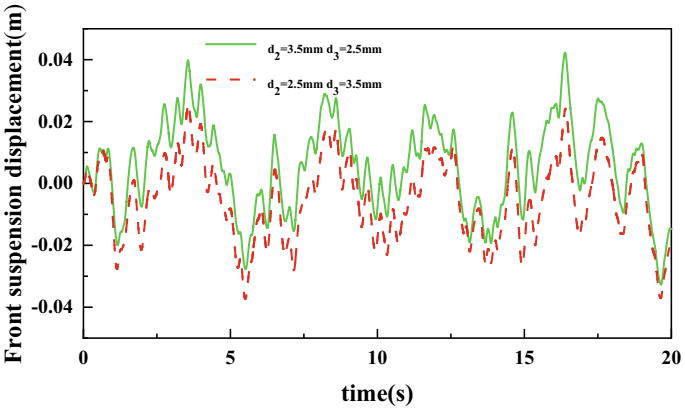


Fig. 11 Front suspension displacement

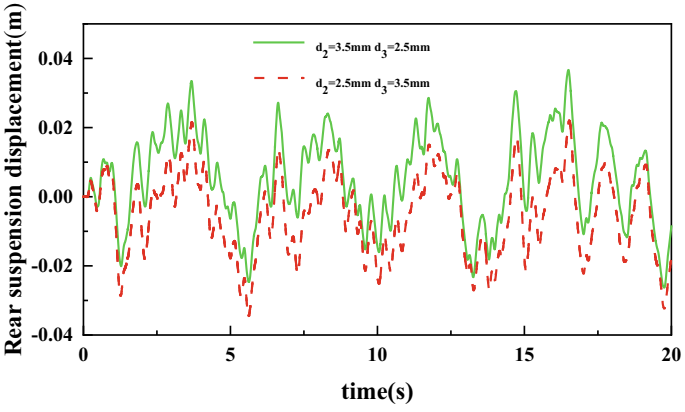


Fig. 12 Rear suspension displacement

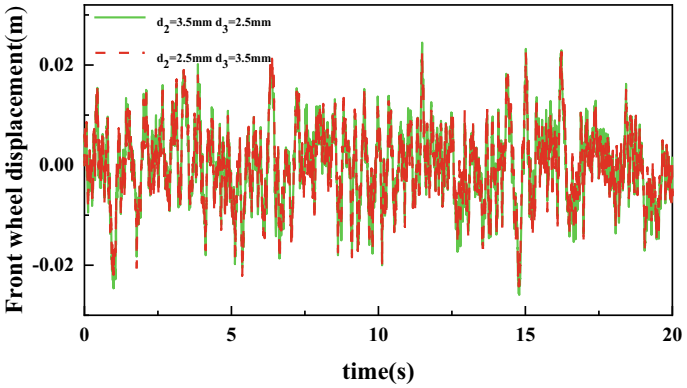


Fig. 13 Front wheel displacement

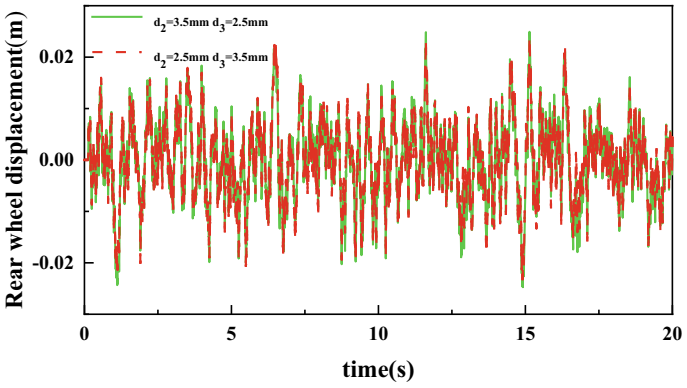


Fig. 14 Rear wheel displacement

Table 4 Performance indicators

Performance indicators	Vertical acceleration	Pitch acceleration	Front suspension dynamic disturbance	Rear suspension disturbance	Front wheel displacement	Rear wheel displacement
Mode1	0.0744	0.0304	0.0156	0.0143	0.00768	0.00767
Mode2	0.0691	0.0295	0.014	0.013	0.00741	0.00743

2. With the accumulator's initial pressure held constant, a significant reduction in vertical and pitch accelerations, suspension dynamic disturbance, and tire dynamic displacement is observed by reducing the aperture of throttle valve 2 and increasing that of throttle valve 3. This adjustment effectively amplifies

the damping force during the extension phase of the hydraulic cylinder and diminishes it during the compression phase.

References

1. Soliman AMA, Kaldas MMS (2021) Semi-active suspension systems from research to mass-market—a review. J Low Freq Noise Vib Active Control 40(2):1005–1023
2. Hurel Ezeta J, Mandow A, Garcia CA (2013) Active and semi-active suspension systems: a review. Revista Iberoamericana de Automática e Informática Industrial 10(2):121–132
3. Zheng E, Zhong X, Zhu R et al (2019) Investigation into the vibration characteristics of agricultural wheeled tractor-implement system with hydro-pneumatic suspension on the front axle. Biosyst Eng 186:14–33
4. Chunyu W, Litong H (2024) Study on modeling and control for a novel active suspension. Int J Autom Technol 1–8
5. Wang H, Lu Y, Tian Y et al (2020) Fuzzy sliding mode based active disturbance rejection control for active suspension system. Proc Inst Mech Eng Part D J Autom Eng 234(2–3):449–457
6. Ding Z, Zhao F, Qin Y et al (2017) Adaptive neural network control for semi-active vehicle suspensions. J Vibroeng 19(4):2654–2669
7. Prasad V, Pawaskar DN, Seshu P (2021) Controller design and multi-objective optimization of heavy goods vehicle suspension system by geometry-inspired GA. Struct Multidiscip Optim 64(1):89–111
8. Nguyen DN, Nguyen TA (2022) A novel hybrid control algorithm sliding mode-PID for the active suspension system with state multivariable. Complexity 2022(1):9527384
9. Pan H, Jing X, Sun W (2017) Robust finite-time tracking control for nonlinear suspension systems via disturbance compensation. Mech Syst Signal Process 88:49–61
10. Deshpande VS, Shendge PD, Phadke SB (2016) Nonlinear control for dual objective active suspension systems. IEEE Trans Intell Transp Syst 18(3):656–665
11. Han SY, Liang T (2022) Reinforcement-learning-based vibration control for a vehicle semi-active suspension system via the PPO approach. Appl Sci 12(6):3078
12. Kwon K, Seo M, Kim H et al (2020) Multi-objective optimisation of hydro-pneumatic suspension with gas–oil emulsion for heavy-duty vehicles. Veh Syst Dyn 58(7):1146–1165
13. Sim K, Lee H, Yoon JW et al (2017) Effectiveness evaluation of hydro-pneumatic and semi-active cab suspension for the improvement of ride comfort of agricultural tractors. J Terrramech 69:23–32
14. Deremetz M, Lenain R, Thuilot B (2017) Stiffness and damping real-time control algorithms for adjustable suspensions: a strategy to reduce dynamical effects on vehicles in off-road conditions. IFAC-Papers OnLine 50(1):1958–1964
15. Sang Z, Dong M, Gu L (2017) Numerical analysis of a dual-chamber hydro-pneumatic suspension using nonlinear vibration theory and fractional calculus. Adv Mech Eng 9(5):1687814017705797
16. Ghazaly NM, Moaaz AO (2020) Hydro-pneumatic passive suspension system performance analysis using amesim software. Int J Veh Struct Syst 12(1):9–12
17. Lee K (1997) Numerical modelling for the hydraulic performance prediction of automotive monotube dampers. Veh Syst Dyn 28(1):25–39
18. Tsuruga Y (1996) Study on hydraulic active suspension for wheeled hydraulic excavator. In: Proceedings of the JFPS international symposium on fluid power. The Japan Fluid Power System Society, vol 1996, no 3, pp 367–372
19. Yin C, Zhai X, Sun X et al (2022) Design and performance research of a hydro-pneumatic suspension with variable damping and stiffness characteristics. J Mech Sci Technol 36(10):4913–4923

20. Gao J, Li H (2023) Tuning parameters of the fractional order PID-LQR controller for semi-active suspension. Electronics 12(19):4115

Open Access This chapter is licensed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License (http://creativecommons.org/licenses/by-nc-nd/4.0/), which permits any noncommercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if you modified the licensed material. You do not have permission under this license to share adapted material derived from this chapter or parts of it.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.



Improved Particle Swarm Optimization for Rapid Tuning of Vehicle Suspension Damping



Hao Xiong, Zeguang Hu, Ziqi Huang, Liqiang Zhao, Haiwu Zheng, and Dingxuan Zhao

Abstract This study addresses the issues of low parameter setting efficiency and unclear evaluation criteria in suspension systems. A fast parameter optimization and tuning strategy based on the improved particle swarm optimization algorithm (IPSO) is proposed, and an objective function for evaluating suspension performance is constructed. The AMEsim and Simulink software are used to build a 1/4 vehicle semi-active suspension and dual-circuit passive suspension simulation model. Taking a Class C road as input, simulation calculations are performed on indicators such as vehicle body acceleration, suspension dynamic stroke, and tire dynamic load. The results show that the performance of the PID-controlled semi-active suspension system optimized by the IPSO algorithm is superior to that of the optimized dual-circuit passive suspension system, and relevant parameters can be quickly tuned. This study provides theoretical guidance and design basis for the optimization design of vehicle suspension systems, overcomes the subjectivity and inefficiency of traditional parameter settings, and has significant theoretical and practical value.

Keywords Improved particle swarm optimization algorithm (IPSO) · PID control · Semi-active suspension · Passive suspension · Compression damping · Rebound damping

1 Introduction

The development of automotive technology continuously promotes the improvement of vehicle performance, comfort, and safety. As one of the key systems, the suspension plays a crucial role in reducing the impact of road irregularities, improving vehicle handling, and ensuring smooth driving [1, 2]. With the continuous improvement of people's requirements for vehicle performance, ride comfort and safety closely related to the suspension system become increasingly important. How to

H. Xiong · Z. Hu · Z. Huang · L. Zhao · H. Zheng · D. Zhao (✉)
School of Mechanical Engineering, Yanshan University, Qinhuangdao, Hebei, China
e-mail: zdx@ysu.edu.cn

© The Author(s) 2026

H. Saman et al. (eds.), *The 9th International Conference on Advances in Construction Machinery and Vehicle Engineering*, Lecture Notes in Mechanical Engineering,
https://doi.org/10.1007/978-981-96-5527-4_85

improve the performance of vehicle suspension systems has become a key problem to be urgently solved in the field of vehicle engineering. In recent years, the suspension system has undergone significant development. From the traditional passive suspension system to more advanced active and semi-active suspension systems, the performance of the suspension system is continuously improved. Up to now, passive suspensions and semi-active suspensions have become the most widely used suspension systems in actual vehicle systems.

The development of ride comfort and handling of vehicle suspension systems is usually a process of repeated trials, which represents the optimization and tuning problem of suspension system parameters. In passive suspension systems, the damping coefficients required in the suspension compression process and rebound process are not the same. Different compression and rebound damping have a significant impact on the stability and handling of vehicle driving. This is also the design basis of compression and rebound oil circuit damping holes in real dual-circuit passive suspension systems. Therefore, how to quickly tune these damping parameters to achieve the best balance of various vehicle performance indicators is a key problem in the suspension design process.

The semi-active suspension system can adjust the damping coefficient in real time according to road conditions to meet the driving requirements of vehicles and thereby improve ride comfort and driving safety, so it has received great attention from the industry. The performance of the semi-active suspension largely depends on its control strategy [3]. To achieve an ideal suspension control effect, researchers have developed a variety of advanced control strategies based on the suspension dynamics model, such as adaptive control [4], sliding mode control [5], and model predictive control [6]. These strategies show excellent control performance in semi-active suspension systems. However, their control effects largely depend on the accuracy of the mathematical model of the suspension system. In practical engineering applications, the suspension system often exhibits nonlinear, uncertain, and complex dynamic characteristics, making it difficult to accurately describe the suspension dynamics model. Therefore, when the suspension dynamics model is difficult to accurately construct, model-based suspension control algorithms usually cannot achieve the expected control effect.

Proportional-integral-derivative (PID) controllers are widely used in the industrial field due to their simple structure, easy implementation, and no need to construct a controlled system model [7]. They are suitable for nonlinear controlled systems with limited underlying process knowledge and uncertain parameters. In engineering practice, coefficient setting often relies on the experience of engineers and a large number of tests. For controlled systems with uncertain parameters, different parameters will make the gain coefficients of PID controllers different under optimal performance. Setting coefficients based on experience is inefficient and lacks clear evaluation criteria. Therefore, for controlled systems with possible parameter changes, quickly and accurately determining the optimal gain coefficients is an important problem to be urgently solved in the field of industrial control.

Researchers have developed various intelligent optimization algorithms for gain coefficient tuning, such as the particle swarm optimization (PSO) algorithm [8], the

whale optimization algorithm (WOA) [9], and the beetle antennae search (BAS) algorithm [3], which are applied to different controlled systems. The particle swarm optimization algorithm is widely used due to its simple structure, easy implementation, and suitability for optimizing multi-dimensional nonlinear systems [10]. However, the inertia weight of the standard particle swarm optimization algorithm is a fixed value and cannot adjust the optimized value at the boundary. When facing complex optimization problems, it is easy to fall into local or boundary optima [11], and it needs to be improved to improve the optimization ability.

This study aims to solve the problems of low efficiency in setting gain coefficients and system parameters and unclear evaluation criteria in semi-active suspension systems and dual-circuit passive suspension systems. To this end, a fast parameter optimization and tuning strategy based on the improved particle swarm optimization algorithm (IPSO) is proposed, and an objective function for evaluating suspension performance is constructed based on suspension performance indicators. To verify the effectiveness of the proposed parameter tuning strategy, in this paper, the AMEsim and Simulink software are used to build a simulation model of a 1/4 vehicle semi-active suspension and a dual-circuit passive suspension. Taking a Class C road as the input of the model, simulation calculations are performed on the vehicle body acceleration (SMA), suspension dynamic stroke (SWS), and tire dynamic load (DTL) of different suspension systems. This study provides important theoretical guidance and design basis for the optimization design of vehicle suspension systems.

2 Model Building

2.1 Suspension Dynamics Model

In the semi-active suspension system, the vertical acting force is suppressed through the hydraulic damping shock absorber between the wheel and the vehicle body to reduce the vertical vibration of the vehicle body caused by uneven road surfaces during driving, thereby enhancing people's comfort when driving. As shown in Fig. 1, it is a two-degree-of-freedom dynamic model of a 1/4 vehicle semi-active suspension system. According to Newton's laws of motion, the dynamic equation of a 1/4 vehicle active suspension system is obtained as follows:

$$\begin{cases} M_s\ddot{Z}_s + C_s\left(\dot{Z}_s - \dot{Z}_t\right) + K_s(Z_s - Z_t) = 0 \\ M_t\ddot{Z}_t - C_s\left(\dot{Z}_s - \dot{Z}_t\right) - K_s(Z_s - Z_t) + K_t(Z_t - Z_r) = 0 \end{cases} \tag{1}$$

where M_s represents the sprung mass, M_t represents the unsprung mass, K_s is the elastic stiffness of the suspension spring, C_s is the suspension damping coefficient, K_t is the tire elastic stiffness; Z_s stand for the vertical displacement of the sprung mass, $\dot{Z}_s$ stand for the velocity of the sprung mass, $\ddot{Z}_s$ are the acceleration of the sprung mass;

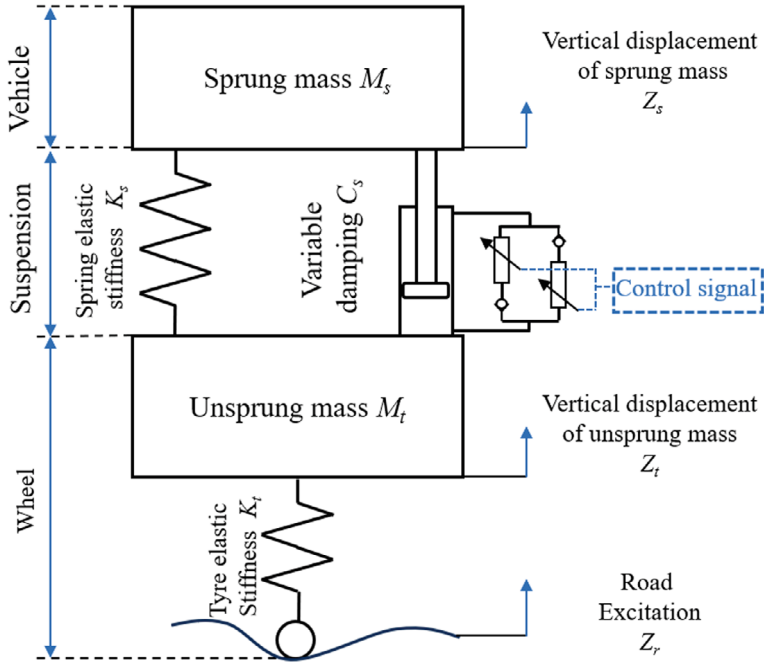
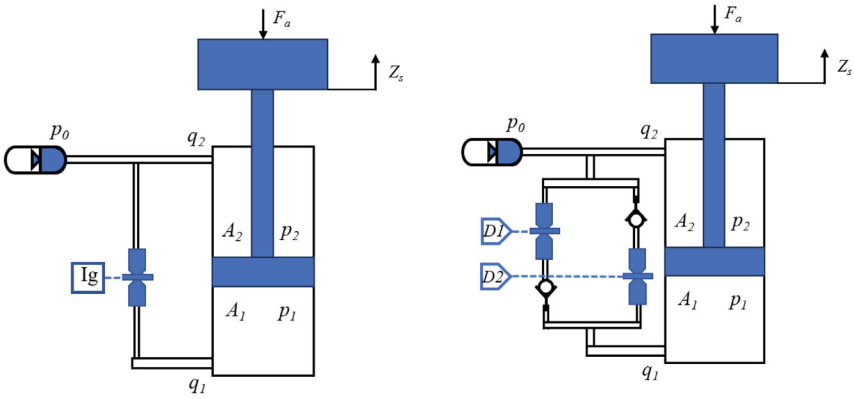


Fig. 1 1/4 suspension system model

Z_t, $\dot{Z}_t$ and $\ddot{Z}_t$ are indicative of the vertical displacement, velocity, and acceleration of the unsprung mass, respectively; Z_r is used to describe the road excitation.

There are two common types of hydraulic damping shock absorbers: electromagnetic and magnetorheological. Among them, the electromagnetic damping shock absorber realizes the change of damping during the compression or rebound process of the asymmetric hydraulic cylinder by controlling the size of the internal oil passage (equivalent to the damping hole). Compared with the magnetorheological damping shock absorber, the electromagnetic damping shock absorber has the advantages of low cost and compact structure, so it has been widely used in semi-active suspension systems. As shown in Fig. 2a, it is the schematic diagram of the electromagnetic damping shock absorber in the semi-active suspension. As shown in Fig. 2b, it is the schematic diagram of the hydraulic damping shock absorber in the dual-circuit passive suspension. D_1 and D_2 are the diameter of the damping hole of the compression or rebound oil circuit respectively; q_1 and q_2 are the flow into or out of the rodless or rod cavity of the hydraulic cylinder, respectively; p_1 and A_1 are the pressure and area of the rodless cavity of the hydraulic cylinder, respectively; p_2 and A_2 are the pressure and area of the rod cavity of the hydraulic cylinder, respectively; the actuating force F_a of the hydraulic cylinder is then given by:

$$p_1 A_1 - p_2 A_2 = F_a \tag{2}$$



(a) Electromagnetic damping shock absorber (b) Dual-circuit damping shock absorber

Fig. 2 Schematic diagram of damping shock absorber

The hydraulic cylinder load flow equation is:

$$\begin{cases} q_1 = A_1\left(\dot{Z}_s - \dot{Z}_t\right) + \beta_e^{-1} V_1 \dot{p}_1 \\ q_2 = A_2\left(\dot{Z}_s - \dot{Z}_t\right) - \beta_e^{-1} V_2 \dot{p}_2 \end{cases} \tag{3}$$

where β_e is the effective bulk modulus of the hydraulic oil; V_1 is the volume of the rodless chamber of the hydraulic cylinder, where $V_1 = V_{01} + A_1 Z_s$, and V_{01} is the initial volume of the rodless chamber; V_2 is the volume of the rod chamber of the hydraulic cylinder, where $V_2 = V_{02} - A_2 Z_s$, and V_{02} is the initial volume of the rod chamber.

The compression or expansion of the gas inside the accumulator, calculated by the ideal gas state change equation:

$$p_0 v_0^n = p_2 v_2^n \tag{4}$$

Among them, p_0 is the initial inflation pressure of the gas inside the accumulator, v_0 is the initial gas volume, n is the polytropic index (isothermal $n = 1$, adiabatic $n = 1.4$), and v_2 is the gas volume when the pressure is p_2. Then the oil flow rate Q of the accumulator can be calculated according to the change in its internal gas volume:

$$Q = v_0 - v_2 = \left[1 - \left(\frac{p_0}{p_2}\right)^{\frac{1}{n}}\right] \times v_0 \tag{5}$$

By changing the diameter of the variable damping hole, the hydraulic oil in the pipeline during the compression or rebound process of the hydraulic cylinder is throttled, and then the damping force during the compression or rebound process

is changed. The flow rate of the damping hole can be calculated by the following formula:

$$q = C_q \times \frac{\pi}{4} \times D^2 \times \sqrt{\frac{2|p_1 - p_2|}{\rho}} \tag{6}$$

where C_q is the flow coefficient, ρ is the hydraulic oil density, D is the diameter of the damping hole, then according to the flow relationship:

$$q = Q + q_2 = q_1 \tag{7}$$

The variable damping hole diameter and the input current are regarded as proportional links:

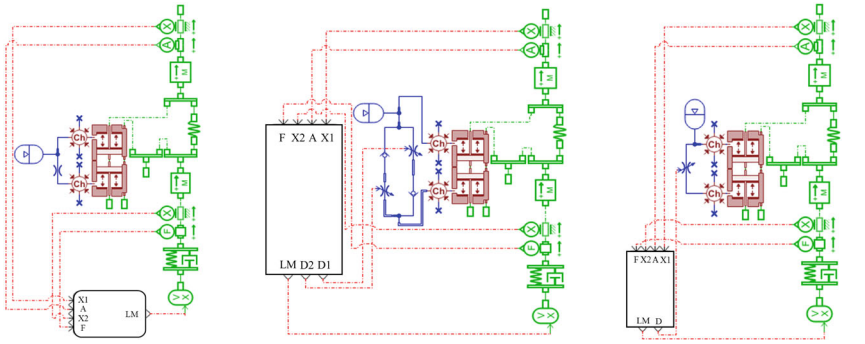
$$\frac{D}{I_g} = K_{xv} \tag{8}$$

where K_{xv} is the signal gain (mm/mA), I_g is the input current (mA).

2.2 Suspension Simulation Model Based on AMEsim

In this study, we use AMEsim and Simulink jointly to establish the simulation model of passive and semi-active suspensions, in which the controller part is established through Simulink. As shown in Fig. 3, using the 1D Mechanical, Hydraulic and Hydraulic Component Design (HCD) component libraries of AMEsim, a 1/4 vehicle suspension model is established. Figure 3a is a passive suspension model, in which the diameter of the oil circuit damping hole in the two chambers of the hydraulic cylinder is set to 10 mm; Fig. 3b is a passive suspension model with hydraulic cylinder compression oil circuit and rebound oil circuit; Fig. 3c is a semi-active suspension model. In addition, the controller and control current I_g and road excitation Z_r are established in Simulink and input into the AMEsim simulation model through the AMEsim-Simulink co-simulation module.

There are an acceleration sensor and a displacement sensor set at the sprung mass M_s, and a displacement sensor and a force sensor set at the unsprung mass M_t, , and the sensor data is transmitted to Simulink to calculate sprung mass acceleration (SMA), suspension dynamic stroke (SWS) and tire dynamic load (DTL). The parameters of the mechanical part in the active suspension system are shown in Table 1. The parameters of the electro-hydraulic servo system are shown in Table 2. When the vehicle is in a stationary state, the suspension system is in the middle position. At this time, the hydraulic damping shock absorber does not output damping force, and the vehicle is supported by the spring. Therefore, in the AMEsim model, the spring element is set with a preload to ensure that the suspension system is in an equilibrium state when the vehicle is stationary.



(a) Passive suspension (b) Dual-circuit suspension (c) Semi-active suspension

Fig. 3 1/4 suspension system simulation model

Table 1 Suspension system parameters

Name	Symbol	Numerical value
Sprung mass (kg)	M_s	550
Unsprung mass (kg)	M_t	50
Spring elastic stiffness (N/m)	K_s	20,000
Tire damping (N s/m)	C_s	1000
Tire elastic Stiffness (N/m)	K_t	100,000

Table 2 Parameters of electro-hydraulic servo system

Hydraulic element	Parameter	Numerical value
Hydraulic cylinder	Piston diameter (mm)	40
	Rod diameter (mm)	20
	Piston thickness (mm)	24
Variable damping hole	Input current range (mA)	0–10
	Maximum opening diameter (mm)	10
	Maximum flow coefficient	0.7
	Signal gain (mm/mA)	1
Hydraulic accumulator	Accumulator volume (L)	1
	Polytropic index	1.4
	Gas precharge pressure (*bar*)	9
	Gas pressure (*bar*)	10

3 Design of PID Controller Based on Improved Particle Swarm Optimization

3.1 PID Controller

The PID algorithm can realize the feedback control of the system only by relying on the input and output signals of the system. It has the characteristics of simple structure and easy implementation, so it has been widely used in various controlled systems. The functional expression of the output signal of the PID controller is:

$$u_0(t) = k_{p0}e(t) + k_{i0}\int_0^t e(t)dt + k_{d0}\frac{de(t)}{dt} \tag{9}$$

In the equation, $e(t)$ represents the deviation between the setpoint $r(t)$ and the output $c(t)$ of the controlled object, $e(t) = r(t) - c(t)$; t is the time variable, $u_0(t)$ is the output value of the PID controller; k_{p0}, k_{i0}, k_{d0} are the proportional, integral, and derivative gain coefficients of the PID controller, respectively.

In this study, the diameter of the variable damping hole is controlled by the PID controller to realize the change of the damping force of the semi-active suspension system. Taking the deviation $e(t)$ between the SMA $\ddot{Z}_s$ and the target value $r(t)$ as the input signal of the controller (abbreviated as e), and the variable damping hole control current I_g as the output signal. Considering that in actual devices, there is a certain delay in the change of the opening diameter of the variable damping hole, and it is not sensitive to high-frequency signals, so a "Filter" module is introduced into the controller for filtering. At the same time, considering the opening range of the variable damping hole and the inverse relationship that the smaller the damping hole, the more it can suppress the vertical acceleration, a "Saturation" module is introduced to limit the output signal, and the maximum control current $I_{g(\max)}$ is subtracted from the output signal to realize the inverse relationship. Then the PID controller of the semi-active suspension system as shown in Fig. 4 is obtained.

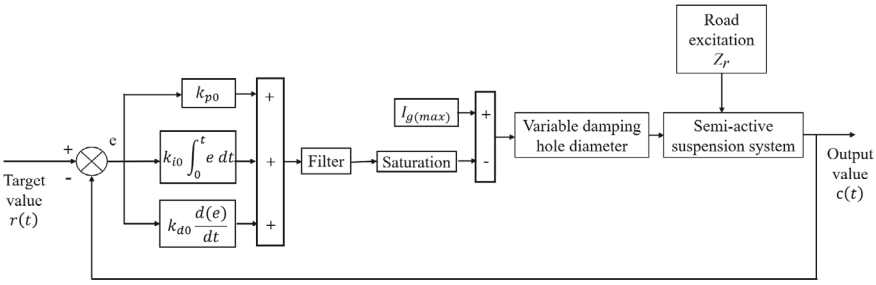


Fig. 4 Semi-active suspension PID controller

3.2 Improved Particle Swarm Optimization Algorithm (IPSO)

For a semi-active suspension system using PID control, its performance is directly related to the proportional, integral, and differential gain coefficients k_{p0}, k_{i0}, k_{d0} in the PID controller. In engineering practice, the determination of such coefficients usually depends on the experience of engineers and a large number of tests. However, the method of setting the control coefficients of the PID controller only relying on engineering experience has problems such as low efficiency and lack of clear evaluation criteria, and it is difficult to be widely applied in complex actual engineering systems. Therefore, how to quickly tune the coefficients of the PID controller has always been a key issue that relevant researchers focus on.

The particle swarm optimization (PSO) algorithm is an iterative optimization algorithm that can quickly search for the minimum value of the objective function and thus find the gain coefficients k_{p0}, k_{i0}, k_{d0} of the PID controller when approaching optimal performance. In the process of optimizing the PID control coefficients, the objective function is equivalent to the evaluation standard for the control effect of the controller. Therefore, designing a reasonable objective function is the key to optimization [12]. In addition, when dealing with complex optimization problems, the traditional particle swarm optimization algorithm cannot coordinate the global search ability and local search ability. The algorithm is not easy to converge and is easily trapped in local optima. Therefore, the particle swarm optimization algorithm needs to be improved to overcome these problems.

Objective function. Sprung mass acceleration (SMA) refers to the acceleration of the sprung mass in the vertical direction. It is an important indicator reflecting the ride comfort of a vehicle and also an important indicator for evaluating suspension performance. The calculation formula for SMA is:

$$Q_{SMA} = \ddot{z}_s(t) \tag{10}$$

For a suspension system, the smaller the absolute value of SMA, the better the performance of the suspension system. Therefore, the objective function can be constructed through sprung mass acceleration (SMA). In this study, with the passive suspension system as a reference, the absolute value integral of active suspension SMA is divided by the absolute value integral of passive suspension SMA, and then the objective function (fitness function) shown below is constructed:

$$L = \frac{\int_0^t \left[|Q_{SMA}|\right]_{PSO} dt}{\int_0^t \left[|Q_{SMA}|\right]_P dt} \tag{11}$$

where $\left[|Q_{SMA}|\right]_{PSO}$ is the absolute value of the SMA of the semi-active suspension; $\left[|Q_{SMA}|\right]_P$ is the absolute value of SMA of passive suspension.

Conventional Particle Swarm Optimization Algorithm. Assuming there is a D-dimensional target search space in which a population consists of N particles, the

position of the i th particle is represented as $x_i = (x_{i1}, x_{i2}, \ldots, x_{iD})$, and its velocity is represented as $v_i = (v_{i1}, v_{i2}, \ldots, v_{iD})$. Each particle records its own historical best position $P_{best} = (P_{i1}, P_{i2}, \ldots P_{iD})$ and the historical best position of the entire population $G_{best} = (G_1, G_2, \ldots G_D)$. The update formulas for particle velocity and position are as follows:

$$v_{id}^{k+1} = \omega v_{id}^{k} + c_1 r_1 \left[P_{best}(i) - x_i^k\right] + c_2 r_2 \left[G_{best} - x_i^k\right] \tag{12}$$

$$x_{id}^{k+1} = x_{id}^{k} + v_{id}^{k+1} \tag{13}$$

where v_{id}^k represents the velocity of the i th particle in the d th dimension at the k th iteration; x_{id}^k represents the position of the i th particle in the d th dimension at the k th iteration; ω is the inertia weight; c_1 and c_2 are non-negative learning factors, typically set to 2; r_1 and r_2 are uniformly distributed random numbers generated within the interval [0, 1].

The implementation of the Particle Swarm Optimization algorithm necessitates setting limits for particle velocity and position. Once the velocity and position of a particle exceed these defined limits, they are restricted by the boundary values of the limit range. Subsequently, the particle positions are input into the optimized function to compute the new individual optimal positions, the global optimal position, and the global optimal fitness value. The procedures described by Eqs. (12) and (13) are reiterated until the maximum number of iterations is attained or other specified conditions are fulfilled. At this point, the optimization search terminates, and the optimal solution along with the corresponding fitness value is output.

Improvement strategies for the particle swarm optimization algorithm. In the traditional Particle Swarm Optimization algorithm, the inertia weight ω is a fixed value. According to Shi's research, the inertia weight serves to balance the global and local search capabilities of particles. A larger inertia weight is beneficial for enhancing the global search ability of particles, allowing them to explore new regions, while a smaller inertia weight favors local exploration, enabling a refined search within local areas [13]. To balance the global and local search capabilities of particles, this study introduces the following dynamic adjustment strategy for the inertia weight:

$$\begin{cases} \omega = \omega_{max}, 0 < k \le 0.2M \\ \omega = \omega_{max} - (\omega_{max} - \omega_{min})\left[\frac{2(k-0.2M)}{0.6M} - \left(\frac{k-0.2M}{0.6M}\right)^2\right], 0.2M < k \le 0.6M \\ \omega = \omega_{min}, 0.6M < k \le M \end{cases} \tag{14}$$

where M represents the maximum number of iterations, k is the current iteration number, ω_{max} and ω_{min} are the maximum and minimum inertia weights, respectively. In the adjustment strategy of this study, the iteration process is divided into three stages. In the first stage ($0 < k \le 0.2M$), the inertia weight ω takes the maximum value, allowing particles to maintain a strong global search capability. In the second

stage ($0.2M < k \leq 0.6M$), the inertia weight ω decreases nonlinearly from the maximum to the minimum value, enabling particles to gradually acquire a strong local search capability. In the third stage ($0.6M < k \leq M$), the inertia weight ω takes the minimum value, allowing particles to maintain a strong local search capability and to conduct a refined search in the local area to seek the optimal solution.

In the iteration process of the algorithm, a particle's position in one or more dimensions might exceed the constraint boundaries. The traditional PSO method sets the out-of-bounds position of the particle to the boundary value. This can impact the population diversity in the early iteration stages, impede algorithm convergence, and potentially result in local or boundary optima [10]. To tackle this issue, this study uses a random reflection wall for out-of-bounds particles. The out-of-bounds particle x_{id} is randomly reflected back into the search space to maintain particle diversity and prevent convergence to boundary optima. The following strategy for the random reflection wall of out-of-bounds particles is presented.

If

$$x_{id} > U_d \&\& x_{id} - U_d \leq U_d - L_d$$

$$x_{id}\prime = U_d - rand(x_{id} - U_d), v_{id} = -v_{id}$$

else if

$$x_{id} > U_d \&\& x_{id} - U_d > U_d - L_d$$

$$x_{id}\prime = U_d - rand(U_d - L_d), v_{id} = -v_{id} \tag{15}$$

else if

$$x_{id} < L_d \&\& L_d - x_{id} \leq U_d - L_d$$

$$x_{id}\prime = L_d + rand(L_d - x_{id}), v_{id} = -v_{id}$$

else if

$$x_{id} < L_d \&\& L_d - x_{id} < U_d - L_d$$

$$x_{id}\prime = L_d + rand(U_d - L_d), v_{id} = -v_{id}$$

end

In the equation, U_d and L_d represent the upper and lower boundaries of the constraint for the d th dimension of all particles, respectively; *rand* is a random number between 0 and 1. The degree to which an out-of-bounds particle is reflected back into the search space is related to the distance it has exceeded the boundary, with

an element of randomness. Additionally, to prevent the particle from crossing the opposite boundary after reflection, the velocity of the particle in the d th dimension, v_{id} is set to its opposite value. The pseudocode for the IPSO algorithm is as follows:

begin
1. Initialize the number of particles N, the maximum number of iterations M, the maximum and minimum inertia weights ω_{max} and ω_{min}. Randomly initialize the velocities and positions of all particles within the specified limits
2. Calculate the fitness values (objective function values) for each particle, determine the best particle position for the entire swarm, and identify the historical best positions for each particle. Set k = 1
3. do
4. Calculate the inertia weight ω according to Eq. (14)
5. Update the velocities and positions of all particles based on Eqs. (12) and (13), as well as other constraint conditions
6. Adjust the positions and velocities of out-of-bounds particles according to Eq. (15)
7. Calculate the fitness values (objective function values) of the particles, and update the best particle position for the entire swarm and the historical best positions of each particle
8. k = k + 1
9. while (k ≤ M)
10. Output the optimal particle position and its corresponding fitness value
end

3.3 Optimization of PID Control Coefficient Based on IPSO

The principle of optimizing the PID controller of the semi-active suspension by the IPSO algorithm is shown in Fig. 5. Considering that the main function of the suspension system is to reduce the influence of uneven road surfaces on the vibration of the vehicle body, the vertical acceleration $\ddot{Z}_s$ of the sprung mass is taken as the output value and the target value $r(t)$ is set to zero. Then, iterative calculations are performed through the IPSO algorithm to find the minimum value of the objective function and then obtain a set of optimal parameters. The main steps of optimizing the PID controller by the IPSO algorithm are as follows:

Step 1: Set algorithm control parameters

A 3-dimensional target search space is constructed using the gain coefficients k_{p0},k_{i0}, k_{d0}. The position of the i th particle is represented as $x_i = (x_{i1}, x_{i2}, x_{i3})$ and the velocity as $v_i = (v_{i1}, v_{i2}, v_{i3})$. The 3 dimensions correspond sequentially to k_{p0}, k_{i0}, k_{d0}. The position constraint ranges are $0 \le k_{p0} \le 100$, $0 \le k_{i0} \le 50$, $0 \le k_{d0} \le 0.5$, and the velocity constraint range is uniformly set to $-2 \le v_i \le 2$. The particle swarm size is set to $N = 50$, the maximum number of iterations to

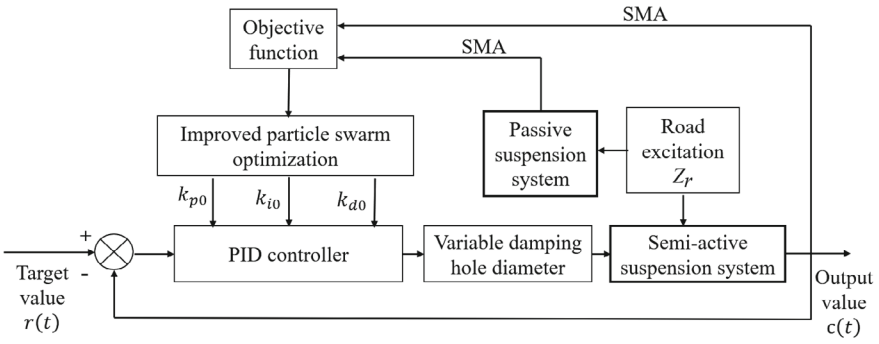


Fig. 5 Improved PSO PID control principle

$M = 50$, the maximum inertia weight to $\omega_{max} = 0.9$, and the minimum inertia weight to $\omega_{min} = 0.4$.

Step 2: Construct the fitness function

The quality of individual positions within the particle swarm is evaluated by the fitness function, and in this study, the fitness function is directly chosen as the objective function L.

Step 3: Initialize particles

Randomly initialize the velocities and positions of the particles, calculate the fitness values for each particle, and based on the particle position corresponding to the minimum fitness value, determine the best particle position for the entire swarm and the historical best positions for each particle.

Step 4: Iterative optimization

Calculate the inertia weight ω according to Eq. (14), and update the velocities and positions of all particles using Eqs. (12) and (13). For out-of-bounds particles, adjust their positions and velocities according to Eq. (15); calculate the fitness values for all particles, and update the best particle position for the entire swarm and the historical best positions for each particle.

Step 5: Output the optimal solution

Determine if the current iteration number m is greater than the maximum number of iterations M. If $m \leq M$, repeat Step 4 until $m > M$, then output the optimal particle position (x_{i1}, x_{i2}, x_{i3}), and the computation concludes.

3.4 Improved Particle Swarm Optimization of Dual-Circuit Passive Suspension Damping Aperture

A reasonable combination of compression damping and rebound damping can effectively improve the driving comfort and handling performance of vehicles, and the

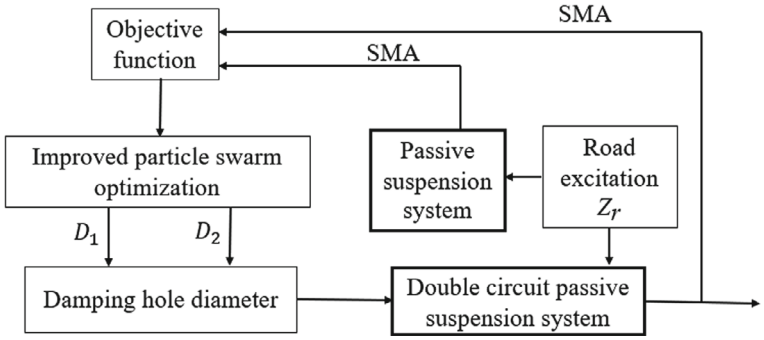


Fig. 6 Improved particle swarm optimization of dual-circuit passive suspension damping aperture principle

optimal damping parameters vary under different vehicle models and different usage scenarios. As shown in Fig. 6, it is the schematic diagram of optimizing the damping aperture D_1 of the compression oil circuit and the damping aperture D_2 of the rebound oil circuit in the dual-circuit passive suspension system by using the IPSO algorithm. In this optimization process, a two-dimensional target search space is constructed with the damping apertures D_1 and D_2, and the position limitation range is $0 \leq D_1 \leq 10$ and $0 \leq D_2 \leq 10$. The rest of the process is similar to the steps of optimizing the PID controller by the IPSO algorithm.

4 Simulation Analysis

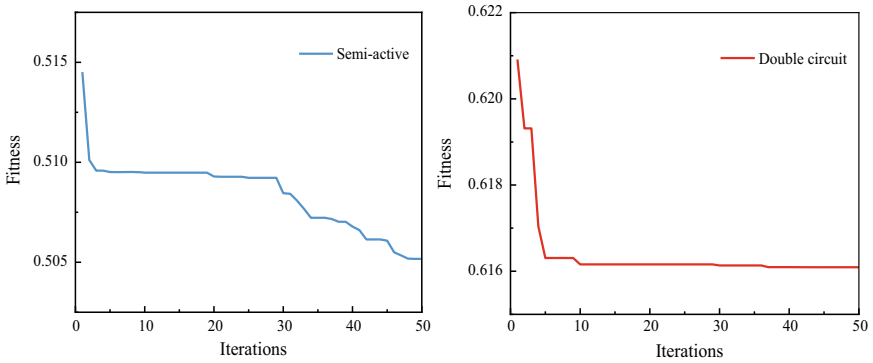
The time-domain model of a Class B road established based on the filtering white noise method is as follows:

$$\dot{Z}_r(t) + 2\pi \upsilon n_{min} Z_r(t) = 2\pi \sqrt{G_r(n_0)\upsilon} W(t) \tag{16}$$

where n_{min} is the minimum spatial frequency of the road surface, set to $0.01\ \mathrm{m}^{-1}$; υ is the vehicle speed, set to $20\ \mathrm{m/s}$; $G_r(n_0)$ is the road roughness coefficient, set to $64 \times 10^{-6}\mathrm{m}^{-3}$; and $W(t)$ is the standard Gaussian white noise. The IPSO algorithm is applied on this road surface to optimize the PID control coefficients of the semi-active suspension system and the compression and rebound damping apertures D_1 and D_2 of the dual-circuit passive suspension system respectively. The optimization results are shown in Table 3 and the fitness curves are shown in Fig. 7. It can be found that the fitness values of different suspension systems gradually decrease with the progress of iterations. In addition, the minimum fitness value of the semi-active suspension is significantly smaller than that of the dual-circuit passive suspension. This indicates that after optimization by the IPSO algorithm, the semi-active suspension system has more advantages in performance.

Table 3 Optimization results of parameters under C-class pavement

	Minimum fitness value	Parameter optimization result
Semi-active suspension	0.505	$k_{p0} = 72.95$, $k_{i0} = 49.98$,$k_{d0} = 0.27$
Double circuit passive suspension system	0.616	$D_1 = 1.43$, $D_2 = 1.89$



(a) Semi-active suspension (b) Dual-circuit passive suspension system

Fig. 7 Fitness curve optimized by IPSO algorithm under B-class pavement

Figures 8 and 9 respectively show the simulation curves of the performance index of the semi-active suspension system and the dual-circuit passive suspension system. The root mean square values of each performance index are shown in Table 4. Taking the performance index of the passive suspension as the standard, the semi-active suspension and dual-circuit passive suspension optimized by the IPSO algorithm have increased by 48.14% and 35.54% respectively in terms of the root mean square of sprung mass acceleration (SMA); in terms of the root mean square of suspension dynamic stroke SWS, they have increased by 51.94% and 44.96% respectively; in terms of the root mean square of tire dynamic load DTL, they have increased by 54.22% and 51.80% respectively. Observing the data, it can be seen that the performance indexes of the semi-active suspension system are all better than those of the dual-circuit passive suspension. In addition, the root mean square of SMA of the semi-active suspension is significantly better than that of the dual-circuit passive suspension, while the root mean square values of SWS and DTL are very close. This is mainly because the fitness function L is built with SMA as the reference index. Thus, it can be seen that establishing an appropriate fitness function is the key to optimizing the performance of the suspension system by the IPSO algorithm. In conclusion, the IPSO algorithm can quickly tune the control coefficients of the PID controller of the semi-active suspension and the damping aperture of the dual-circuit

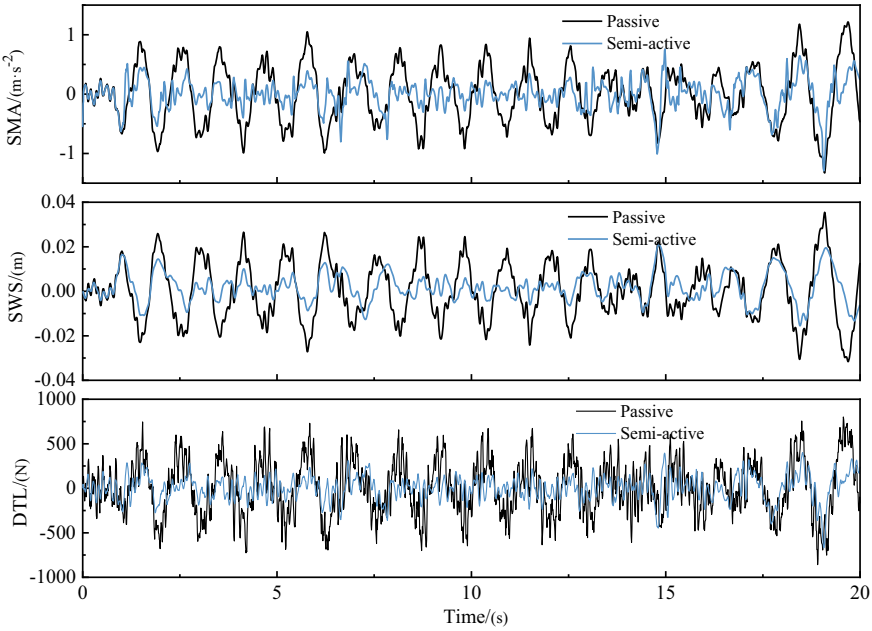


Fig. 8 Semi-active suspension performance simulation curve

passive suspension system, and can effectively improve the performance indexes of the suspension system.

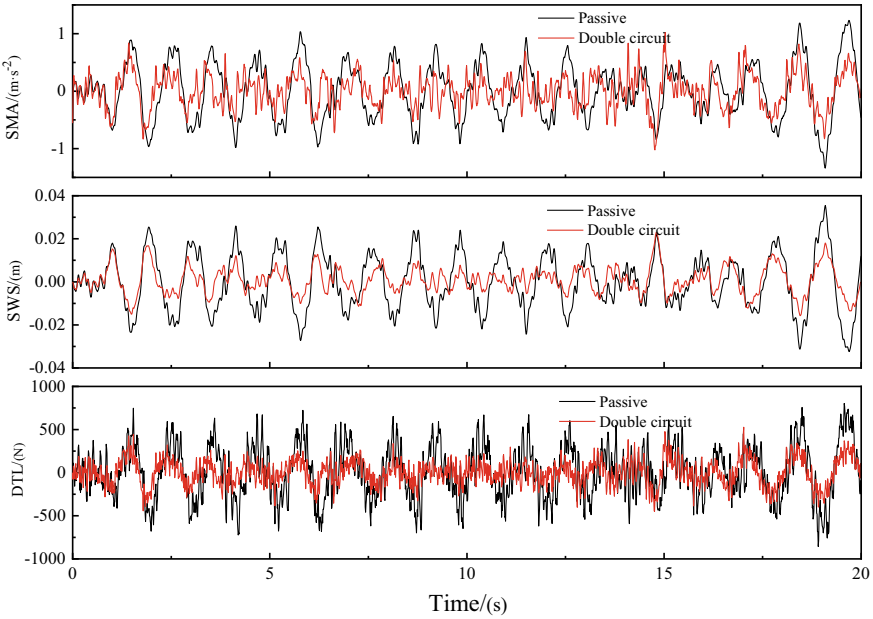


Fig. 9 Performance simulation curve of Dual-circuit passive suspension

Table 4 Root mean square value of the performance index of C-class suspension under road surface

	Passive	Semi-active	Dual-circuit
$SMA/(\mathrm{m\,s^{-2}})$	0.484	0.251	0.312
$SWS/(\mathrm{m})$	0.0129	0.0062	0.0071
$DTL/(\mathrm{N})$	303.76	139.05	146.40

5 Conclusion

This study proposes a fast parameter tuning strategy based on the improved particle swarm optimization algorithm (IPSO) for semi-active suspension systems and dual-circuit passive suspension systems. Firstly, a simulation model of semi-active suspension and dual-circuit passive suspension is constructed, and the sprung mass acceleration (SMA) of the suspension is quantified, and then an objective function for evaluating suspension performance is established. Taking a Class C road as the model input, simulation calculations are carried out on the performance indicators of different suspension systems. The simulation results show that the PID-controlled semi-active suspension system optimized by the IPSO algorithm is superior in performance to the dual-circuit passive suspension system optimized by the IPSO algorithm. At the same time, the study also found that the IPSO algorithm can quickly and effectively tune the control coefficients of the semi-active suspension PID controller and the damping aperture of the dual-circuit passive suspension system. This research

provides novel ideas and methods for the optimization design of vehicle suspension systems, overcomes the subjectivity and inefficiency inherent in traditional parameter setting based on engineering experience, and has significant theoretical value and practical application value.

References

1. Abbas SB, Youn I (2024) Optimal control of a semi-active suspension system collaborated by an active aerodynamic surface based on a quarter-car model. Electronics 13(19):3884. https://doi.org/10.3390/electronics13193884
2. Duan Y, Wu J, Zhang Y (2025) Kinematics characteristics of unsprung mass in a double wishbone suspension based on velocity transformation. Acta Mech Sin 41(2):523661. https://doi.org/10.1007/s10409-024-23661-x
3. Wang L et al (2024) Research on gate opening control based on improved beetle antennae search. Sensors 24(13):4425. https://doi.org/10.3390/s24134425
4. Yoon D-S, Choi S-B (2024) Adaptive control for suspension system of in-wheel motor vehicle with magnetorheological damper. Machines 12(7):433. https://doi.org/10.3390/machines12070433
5. Chen H, Liu Y-J, Liu L, Tong S, Gao Z (2022) Anti-saturation-based adaptive sliding-mode control for active suspension systems with time-varying vertical displacement and speed constraints. IEEE Trans Cybern 52(7):6244–6254. https://doi.org/10.1109/TCYB.2020.3042613
6. Liu Z, Si Y, Sun W (2024) Ride comfort oriented integrated design of preview active suspension control and longitudinal velocity planning. Mech Syst Sig Process 208:110992. https://doi.org/10.1016/j.ymssp.2023.110992
7. Joseph SB, Dada EG, Abidemi A, Oyewola DO, Khammas BM (2022) Metaheuristic algorithms for PID controller parameters tuning: review, approaches and open problems. Heliyon 8(5):e09399. https://doi.org/10.1016/j.heliyon.2022.e09399
8. Ambroziak A, Chojecki A (2023) The PID controller optimisation module using fuzzy self-tuning PSO for air handling unit in continuous operation. Eng Appl Artif Intell 117:105485. https://doi.org/10.1016/j.engappai.2022.105485
9. Swethamarai P, Lakshmi P, Gokul Prassad S (2022) Whale-optimized fuzzy-fractional order controller-based automobile suspension model. Eng Optim 54(7):1110–1130. https://doi.org/10.1080/0305215X.2021.1913735
10. Zhang Y, Wang S, Ji G (2015) A comprehensive survey on particle swarm optimization algorithm and its applications. Math Probl Eng 2015:1–38. https://doi.org/10.1155/2015/931256
11. Chiou J-S, Tsai S-H, Liu M-T (2012) A PSO-based adaptive fuzzy PID-controllers. Simul Model Pract Theory 26:49–59. https://doi.org/10.1016/j.simpat.2012.04.001
12. Sahib MA, Ahmed BS (2016) A new multiobjective performance criterion used in PID tuning optimization algorithms. J Adv Res 7(1):125–134. https://doi.org/10.1016/j.jare.2015.03.004
13. Shi Y, Eberhart RC (1999) Empirical study of particle swarm optimization. In: Proceedings of the 1999 congress on evolutionary computation-CEC99 (Cat. No. 99TH8406). IEEE, Washington, DC, pp 1945–1950. https://doi.org/10.1109/CEC.1999.785511

Open Access This chapter is licensed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License (http://creativecommons.org/licenses/by-nc-nd/4.0/), which permits any noncommercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if you modified the licensed material. You do not have permission under this license to share adapted material derived from this chapter or parts of it.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.



Mobile Carrier Platform Design



Hao Wang, Youqing Li, Rui Guo, Zebing Fan, Dianmin Chen, Pengfei Chen, Chenglong Yu, and Jingyi Zhao

Abstract Aiming at the problem that the traditional ground test bench cannot carry out dynamic test, a single motor-driven duplex pump-multi-way valve-hydraulic motor mobile transportation platform is designed. According to the design steps, firstly, the mechanical structure of the transportation platform was designed and a three-dimensional model was built. Then on the basis of the model, the hydraulic components were selected and calculated and the hydraulic principle circuit was designed. With the electrical control system and touch screen to control the load-sensitive proportional multi-way valve to realize the whole vehicle operation of the transport platform, and finally carried out the physical assembly and debugging of the transport platform. Through debugging, it has accomplished the functions of forward, backward, left turn, right turn, in-situ turn and 30°hill climbing under flat ground condition. The research results can promote development of test technology capability and have important military and engineering application value.

Keyword Mobile transportation platform · Load sensing proportional multi-control valves · Stability studies

H. Wang · P. Chen
AECC Sichuan Gas Turbine Establishment, Mianyang, Sichuan, China

Y. Li
Nanjing University of Aeronautics and Astronautics, Nanjing, Jiangsu, China

R. Guo (✉) · D. Chen · C. Yu · J. Zhao
School of Mechanical Engineering, Yanshan University, Qinhuangdao, Hebei, China
e-mail: guorui@ysu.edu.cn

Z. Fan
School of Power and Energy, Northwestern Polytechnical University, Xi'an, Shanxi, China

© The Author(s) 2026

H. Saman et al. (eds.), *The 9th International Conference on Advances in Construction Machinery and Vehicle Engineering*, Lecture Notes in Mechanical Engineering,
https://doi.org/10.1007/978-981-96-5527-4_86

1 Introduction

At present, all countries in the world are facing the energy crisis and environmental pollution problems, countries have to carry out research and development of renewable energy and environmental protection [1]. Wheeled chassis, as one of the more widely used construction machinery, is used in a large number of agricultural, transportation and other fields, at present, most of the traditional wheeled chassis are powered by diesel engines [2]. Diesel engines have the disadvantages of low energy conversion rate, pollution, noise, etc. which not only cause certain pollution to the atmosphere, but also aggravate the energy crisis in China with the continuous fuel consumption [3]. Therefore, electrification of wheeled chassis has become one of the technological paths of choice for the development of the industry.

In agriculture, there is a huge potential for development in small area environments such as hills and mountains in the south. The electric chassis is compact, small, flexible and easy to maneuver, which is more suitable for such rugged terrain operations [4]. In order to reduce the workload of fruit farmers and protect the orchard environment, electric chassis are also commonly used in orchard operations to promote the large-scale development of orchards and enhance the level of integrated management of orchards [5]. Wu of Huazhong Agricultural University has designed a hub-motor-driven mountain forest fruit and tea plantation transporter, which provides a solution for complex terrain operations [6].

In the field of transportation, Zhang of Shenyang University of Technology designed an electric greenhouse wheeled transporter to meet the intelligent needs of agricultural vehicles in greenhouses [7]. Li of the Highway Research Institute of the Ministry of Transportation and Communications, in order to ensure the safe transportation of aero-engines, the whole vehicle and local structure of the aero-engine transporter design [8].

According to the relevant research on wheeled chassis, the wheeled chassis commonly used in the form of direct drive motor and the use of hydraulic transmission system drive and so on. Yang et al. from Taiyuan University of Technology designed a walk-behind all-terrain hexapod robot utilizing DC motors for direct control, together with a servo for turn control [9]. Reiser et al. at the University of Hohenheim, Germany, investigated a rotary electric weeding machine for orchards, using an electric motor as a power source instead of hydraulic power. To meet the requirements of energy saving and emission reduction, and suitability for complex terrain [10]. The wheeled chassis of this design is controlled by motor drive, and its electric system is compact, flexible and suitable for rugged roads.

Aiming at the problem that it is difficult to carry out dynamic tests on the traditional ground test bench, this paper develops a single motor-driven duplex pump-multi-way valve-hydraulic motor mobile transportation platform.

2 Mechanical Structure Design and Principle

The mobile transportation platform mainly includes frame chassis, traveling mechanism, hydraulic system, and electrical control system. The overall three-dimensional model is shown in Fig. 1.

In this design, the front axle is the steering axle and the rear axle is the drive axle. The front axle adopts hydraulic trapezoidal steering system to realize its function of steering. The rear axle is driven by double hydraulic motors, when the transport platform needs to move forward or backward, the motor-driven pump supplies oil to the hydraulic motors through the valve, thus realizing the forward and backward functions of the transport platform.

The specific working principle is to send command signals to Delta PLC through Delta touch screen, and the PLC outputs signals to control the opening sizes of the electromagnetic relief valves and load-sensitive proportional multi-way valves on both sides, which in turn adjusts the size of the flow rate in the hydraulic circuit and controls the rotational speed of the hydraulic motors on both sides. In this process, the 60 V lead battery is converted to 220 V through the modified string inverter to supply power for the transportation platform. The specific parameters of the mobile transportation platform are shown in Table 1.

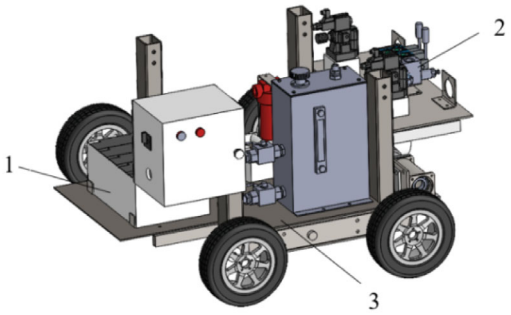


Fig. 1 Overall design 3D model diagram. 1—Electrical control system module; 2—hydraulic system module; 3—frame chassis module

Table 1 Design requirements for mobile transport platforms

Norm	Vehicle weight	Load	Systems stresses	Peacefully tempo	Ramps tempo
Parameters	330 kg	20 kg	10 MPa	3 km/h	1 km/h

3 Hydraulic System Design

3.1 Overall Hydraulic System Design

The mobile transportation platform system consists of hydraulic components such as power unit, duplex pumps, electromagnetic relief valves, globe valves, check valves, high-pressure filters, load-sensitive proportional multi-way valves, gear motors and other hydraulic components. The hydraulic system schematic diagram is shown in Fig. 2, and the relevant parameters are shown in Table 2.

(1) Before starting the duplex pump, first close the 1, 2 two high-pressure ball valves, check the tank to ensure that the tank level is normal and there is no leakage.
(2) Open shut-off valves 1 and 2, close shut-off valve 9, start the single-phase asynchronous motor, and use electromagnetic relief valves 5 and 6 to regulate the pressure of the two modes respectively. The low pressure is adjusted to 8 MPa and the high pressure is adjusted to 10 MPa.
(3) When the car needs to stop at a ramp, shut-off valves 13, 14, 15, and 16 can be closed, thus sealing off both motors for temporary parking.
(4) 11 is a load-sensitive proportional multi-way valve with inlets and outlets P and T. Motor inlets and outlets are A and B, respectively.
(5) The two solenoid relief valves, 5 and 6, correspond to two operating conditions, and only one can operate and the other unload at the same time.

Hydraulic oil from the tank, through the duplex pump to provide power, through the high-pressure filter, into the load-sensitive proportional multi-way valve, and finally through the two motors back to the tank, forming a closed hydraulic circuit, thus driving the movement of the carrier platform.

Specific working scenarios are as follows: before the system starts, open the shut-off valves 1 and 2 to make the oil and hydraulic duplex pumps, close the high-pressure ball valve 9 to electromagnetic relief valves 5 and 6 to regulate the pressure, and the preparatory work is completed. When the transport platform moves forward in a straight line, the oil passes through the load-sensitive proportional multi-way valve 11, making the two hydraulic motors 17 and 18 rotate synchronously. When acceleration is required, adjust the throttle port of the load-sensitive proportional multi-way valve larger to realize acceleration. Conversely, the throttle port of the load-sensitive proportional multivalve is adjusted downward to realize deceleration.

3.2 Principle of Load Sensitive Proportional Multiway Valve

This load-sensitive proportional multidirectional valve consists of three connections, a head connection and two working connections. The hydraulic system schematic diagram is shown in Fig. 3.

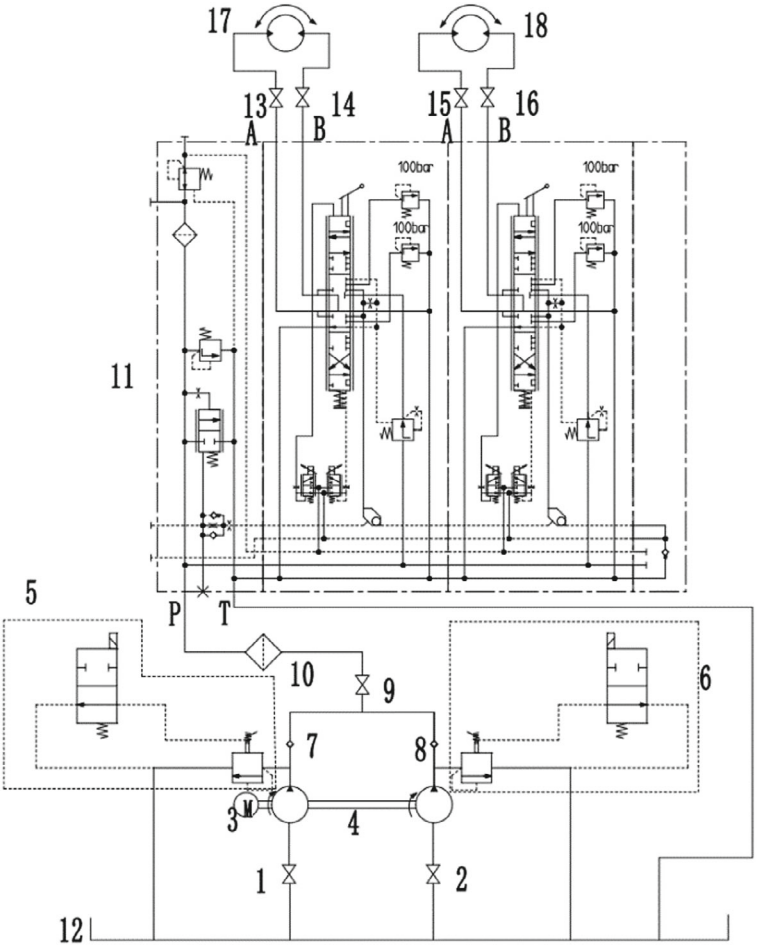


Fig. 2 Hydraulic schematic diagram of mobile transportation platform

Table 2 Names of hydraulic components

Number	1, 2	3	4	5,6	7,8
Element name	Shutoff valve	Single-phase asynchronous motor	Duplex pump	Electromagnetic relief valve	Single-phase valve

Number	10	9, 13, 14, 15, 16	11	12	17, 18
Element name	High-pressure oil filter	Shutoff valve	Load-sensitive proportional multi-way valve	Tank	Swing motor

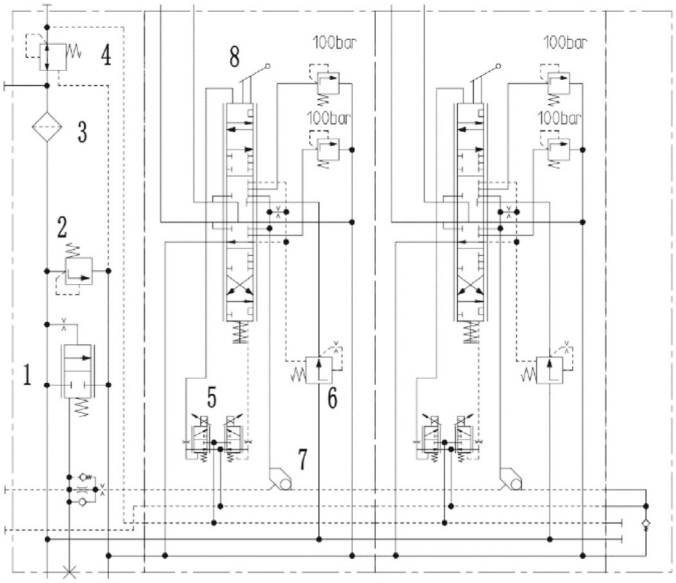


Fig. 3 Hydraulic schematic diagram of load sensitive proportional multi way valve. 1—three-way flow valve; 2—relief valve; 3—oil filter; 4—pressure reducing valve; 5—electromagnetic proportional reversing valve; 6—pressure compensator; 7—shuttle valve; 8—proportional directional valve

Element 1 is a three-way flow valve that compares the inlet pressure to the operating pressure. When the operating pressure through the shuttle valve is greater than the input pressure, the three-way flow valve tends to close and more flow is used to drive the load to work. When the input pressure is greater than the pressure output from the shuttle valve, the three-way flow valve tends to open and more flow is unloaded through the return port.

Element 2 is a relief valve, which provides protection.

Element 3 is an oil filter, which acts as a filter.

Element 4 is a pressure-reducing valve, through which oil passes to control the reversing and proportional control of the two directional valves.

Element 5 is a solenoid proportional directional valve that controls the direction change and proportional control of 8.

Element 6 is a pressure compensator that acts as a load-sensitive device to maintain a consistent differential pressure through the reversing valve.

Element 7 is a shuttle valve, through which the maximum pressure of the two working couplings is screened for comparison in 1.

Element 8 is a proportional directional valve, which is commutated through 8 to regulate the flow.

4 Electrical System Design

The electric drive of the mobile transport platform can control the platform movement by adjusting the output of the motors, realizing the stability control of the mobile transport platform maneuvering and preventing the platform from destabilizing. In order to ensure the stability of the platform, the electrical system is required to regulate the flow of the two motors.

The electrical control system is mainly controlled by Delta touch screen + PLC. The main functions are controlling the proportional amplifier, controlling the opening and closing of the electromagnetic relief valve, and controlling the red and green lights. The core is the proportional amplifier, through two proportional amplifiers can realize the load-sensitive proportional multi-way valve opening change. The schematic diagram of the electrical system is shown in Fig. 4.

The hardware module of the electrical control system consists of a power system module, a single motor start module, an electromagnetic relief valve module, a proportional amplification module, and a touch screen module.

(1) The power system module consists of 60V32Ah lead battery, 60–220 V, 1.5 kW inverter, air switch, rotary switch. The whole module provides 220 V AC power for the electrical control system.
(2) The single motor start module consists of a single motor, contactor, thermal relay, air switch, and PLC.
(3) The electromagnetic relief valve module consists of an electromagnetic relief valve and an intermediate relay.
(4) The proportional amplification module consists of PLC, switching power supply, DA conversion module, and air switch.

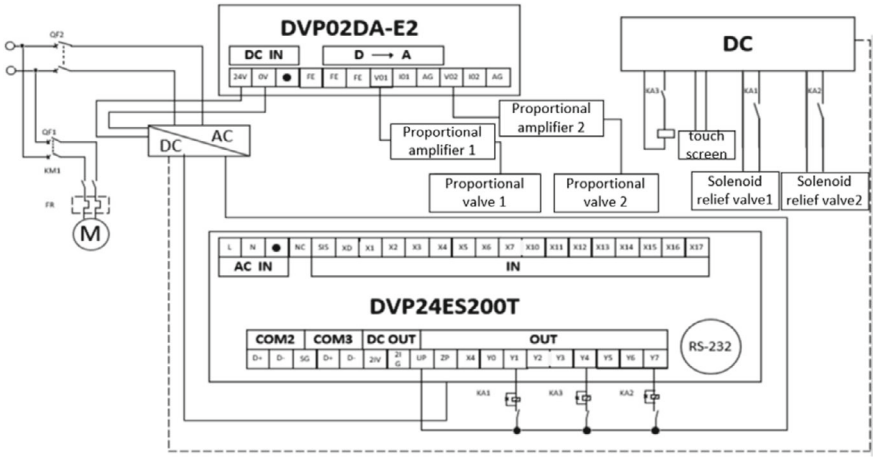


Fig. 4 Electrical schematic diagram of the carrier platform

Table 3 Delta PLC logic controller I/O address allocation tables

Input device	Motor start button	Motor stop button	Platform standby button	Mode switching button	Forward button
Corresponding input point number	M0	M1	M2	M3	M4
Input device	Turn left button	Right turn button	Back button	Spin in place button	
Corresponding input point number	M10	M11	M5	M14	
Output device	Low Pressure Mode Solenoid Relief Valve	High pressure mode solenoid relief valve	Motor contactor	Red light	Green light
Corresponding output point number	Y1	Y2	Y4	Y5	Y6

(5) The touch screen module consists of a touch screen, 24 V switching power supply. The software system consists of Delta PLC and the system software in the touch screen. The I/O address assignment table for PLC control is shown in Table 3.

The electrical principle of the carrier platform is shown in Fig. 4.

5 Sample Test Records

Based on the above 3D modeling and analytical calculations, the machining and fabrication of the test prototype was carried out. Subsequently, a traveling test was conducted on the mobile carrier platform, which was operated forward, backward, left and right.

The experimental process is as follows: first start the power supply unit, the motor will start, drive the hydraulic pump rotation. After the use of the hands of the touch screen control Delta PLC, press the forward button on the touch screen, moving the transport platform forward, while recording the pressure representation, forward process pressure representation shown in Fig. 5. After that, manipulate the Delta touch screen in turn, control the carrier platform backward, left turn, right turn three processes, respectively, to record the number of pressure. The backward process pressure is shown in Fig. 6, the left turn process pressure is shown in Fig. 7, and the right turn process pressure is shown in Fig. 8. Finally, the mobile transportation

platform successfully completed the four basic actions, and the pressure representation is maintained stable, are not more than the maximum pressure of the hydraulic power components, hydraulic control components and hydraulic actuators. The test results show that the design of the mobile transportation platform is successful. The pressure gauge data after speed stabilization was finally plotted in a table, and the test data is shown in Table 4.

Based on the observations in the above figure, the pressure gauge readings were 4.5 MPa during forward movement of the carrier platform, 4 MPa during backward movement, 6 MPa during left turn and 5.5 MPa during right turn.

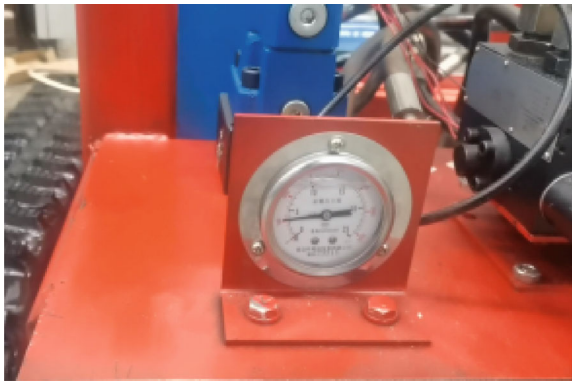


Fig. 5 Pressure gauge diagram for forward process



Fig. 6 Pressure gauge diagram during the retreat process

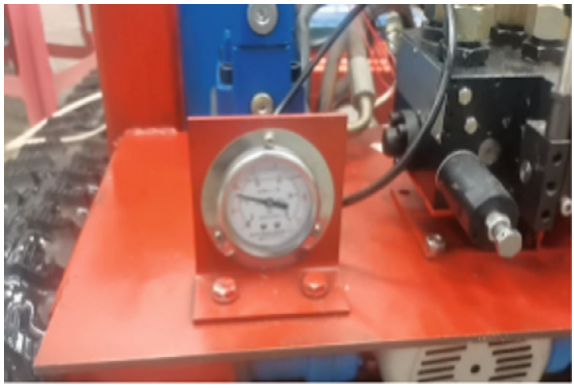


Fig. 7 Pressure gauge diagram during left turn process

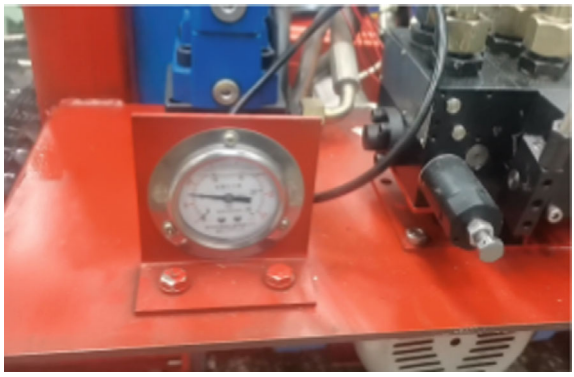


Fig. 8 Pressure gauge diagram during right turn process

Table 4 Pressure indicator data sheet

Operational state	Advance	Be back	Turn left	Turn right
Pressure indicator	4.5 MPa	4 MPa	6 MPa	5.5 MPa

6 Conclusion

Aiming at the problem that it is difficult to carry out dynamic tests on the traditional ground test bench. This paper carries out the following work research: firstly, the mechanical structure design of the mobile transportation track platform was developed. After that, the hydraulic system of the transportation platform was designed, and the feasibility of the system was verified. At the same time, the software and hardware parts of the electrical system of the transportation platform were designed. Finally, the fabrication and testing of the prototype were completed. This study can provide a reference program for the design of the transportation platform.

References

1. Li JF, Li G (2020) Review and prospect of energy, environment and climate change in China. Environ Sustain Dev 45(05):8–17
2. Ante K, Gojmir R, Sandro N (2020) Analysis of methods towards reduction of harmful pollutants from diesel engines. J Clean Prod 262(20):121105.1–121105.20
3. Ma J, Sun SZ, Rui HT et al (2018) Overview of academic research on road construction machinery in China-2018. China J Highw 31(06):1–164
4. Li YN (2023) Research on design and simulation of orchard wheel weeding machine. Shandong Jiaotong University
5. Fang SP, Wang NN, Yi KC et al (2017) Design and performance analysis of power system for pure electric tractor. J Chin Agric Mech 38(01):80–84
6. Wu WB, Han CY, Liang RX et al (2021) Design and test of a wheeled transporter driven by hub motor for hilly orchard and tea garden. Huazhong Agric Univ 40(03):286–294
7. Zhang XF (2023) Design and testing of self-following greenhouse wheeled transporters. Shenyang Agricultural University
8. Li C, Nie HY, Duan C et al (2022) Design and mode analysis of aeroengine transport vehicle. Univ Jinan 36(06):748–754
9. Yang Z, Ma CY, Dou YK et al (2022) Walking all-terrain hexapod robot. Chin Hydraul Pneum 1:109–116
10. Reiser D, Sehsah ES, Bumann O et al (2019) Development of an autonomous electric robot implement for intra-row weeding in vineyards. Agriculture 9(1):18

Open Access This chapter is licensed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License (http://creativecommons.org/licenses/by-nc-nd/4.0/), which permits any noncommercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if you modified the licensed material. You do not have permission under this license to share adapted material derived from this chapter or parts of it.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.



Research on Vehicle Center of Mass Measurement Method Based on Active Suspension



Jinming Chang, Dingxuan Zhao, Haoyu Jiang, Yunwang Ma, and Shiling Li

Abstract To address the issue of center of mass variation in heavy vehicles equipped with active suspension under different operating conditions, this study pro-poses a measurement method based on active suspension. The X and Y coordinates of the vehicle's center of mass are measured using a four-wheel weighing method, while the Z coordinate is determined using a weight response method. This weight response method compensates for the limitation of the four-wheel weighing method, which cannot measure the Z-direction center of mass coordinate. The measurement error for the X-direction center of mass coordinate obtained from the four-wheel weighing method is within 8 mm, and within 12 mm for the Y-direction. For the Z-direction coordinate, the weight response method yields a measurement error within 14 mm. Experimental results demonstrate that the proposed center of mass measurement method effectively addresses the control performance issues arising from center of mass variations, thereby offering potential improvements in the ride comfort and handling stability of vehicles with active suspension.

Keywords Centroid measurement · Active suspension · Four-wheel weighing method · Weight response method

J. Chang · D. Zhao (✉) · H. Jiang · S. Li
School of Mechanical Engineering, Yanshan University, Qinhuangdao, China
e-mail: zdx@ysu.edu.cn

State Key Laboratory of Crane Technology, Yanshan University, Qinhuangdao, China

Y. Ma
XCMG Xuzhou Heavy Machinery Co., Ltd., Xuzhou, China

© The Author(s) 2026

H. Saman et al. (eds.), *The 9th International Conference on Advances in Construction Machinery and Vehicle Engineering*, Lecture Notes in Mechanical Engineering,
https://doi.org/10.1007/978-981-96-5527-4_87

1 Introduction

The position of a vehicle's center of mass is a key parameter influencing its stability, handling, and safety performance. Accurate measurement of the center of mass coordinates is essential for vehicle design, performance optimization, and accident analysis. For vehicles equipped with active suspension control systems, the precision of the center of mass coordinates directly impacts the effectiveness of the active suspension, playing a decisive role in the ride comfort and handling stability of the vehicle. The active suspension control system utilizes sensors to obtain information about the vehicle's state and, in conjunction with hardware parameters such as the center of mass coordinates, calculates in real time the control forces and displacements required to achieve optimal driving conditions, thereby meeting the demands for ride comfort and handling stability.

Various methods can be employed to measure the center of mass coordinates of vehicles, which can be broadly categorized into static measurement methods and dynamic measurement methods. Static measurement methods include theoretical calculation methods, four-wheel weighing methods, and weight response methods. Dynamic measurement methods involve measuring the three-dimensional coordinates of the vehicle's center of mass using motion data collected during actual driving [1]. The theoretical calculation method determines the overall vehicle center of mass coordinates based on the mass and center of mass coordinates of components provided by the vehicle manufacturer, including the engine, cabin, chassis, and tires. The four-wheel weighing method measures the vertical loads on each wheel using four independent weighing plat-forms, combining the vehicle dimensions and torque balance equations to compute the center of mass coordinates in the X and Y directions; however, it cannot determine the Z-direction center of mass coordinates. Tomas Skrucany et al. re-searched the center of mass of vehicles under varying passenger loads using the four-wheel weighing method, confirming that the center of mass changes with load and further affects vehicle performance [2]. Patel et al. developed a device for measuring the mass and center of mass of irregular objects using the four-wheel weighing method, providing valuable insights for practical applications of theory [3]. Chien et al. applied the principles of the four-wheel weighing method in the design and implementation of a two-wheeled self-balancing vehicle based on load sensors, further validating the feasibility of this method. The weight response method addresses the limitations of the four-wheel weighing method by placing the vehicle on a platform with adjustable incline angles. By recording changes in wheel loads at various incline angles and combining mechanical analysis, the Z-axis coordinate of the center of mass can be inferred [4]. Feng conducted research in his thesis to estimate the center of mass height of articulated heavy vehicles, achieving significant findings [5]. Cui et al. studied random errors during the weight response method measurement process and proposed a method for measuring the center of mass location, reducing measurement errors and ensuring precision [6]. Guo proposed an angle-adjustable wheel load measurement method to address the issue of measurement precision loss due to platform measurement errors, providing

new insights for center of mass measurement in multi-axle vehicles [7]. Liu designed an automatic measurement system for the center of mass of two-wheeled vehicles based on the principles of the weight response method, and experimental results showed high measurement accuracy and promising application prospects [8]. Qiu et al. developed a measurement platform for the vehicle's center of mass location based on the weight response method, achieving precise measurement of the vehicle's center of mass [9]. Li et al. proposed an improved weight response method to address issues such as inaccurate calculation of tire load points and significant vehicle roll, resulting in precise center of mass measurements [10]. Synak et al. applied the weight response method to study changes in the vehicle's center of mass position under different load distributions, concluding that variations in load distribution significantly impact vehicle performance [11]. Dynamic measurement methods utilize motion data collected during actual vehicle operation to measure the three-dimensional coordinates of the center of mass. These methods typically employ inertial measurement units (IMUs), such as accelerometers and gyro-scopes, along with external positioning systems like GPS, in conjunction with vehicle kinematic models to compute the center of mass position. Jin et al. con-ducted research on vehicle state estimation using dynamic measurement methods, suggesting that data-driven estimations can enhance the performance of vehicle dynamics estimation systems [12]. Huang et al. conducted real-time estimation studies of the center of mass height of lightweight vehicles based on the instantaneous effective radius of the tires, demonstrating through simulation and experimental results that the designed estimator can accurately estimate the vehicle's center height in real-time [13]. Huang et al. proposed a real-time estimation method for the center of mass position of lightweight trucks based on a combination of the AKF and EKF methods, providing strong evidence for the practical application of dynamic measurement methods [14]. Existing research indicates that static measurement methods are structurally simple and convenient for measurements at fixed locations, while dynamic measurement methods allow for real-time center of mass coordinate measurements during vehicle operation but require numerous sensors, resulting in higher economic costs. Therefore, the choice of center of mass measurement method should depend on the specific operating conditions [15, 16].

The center of mass coordinates of heavy cargo vehicles vary with the loading mass and loading position, making both static and dynamic measurement methods unsuitable for measuring the center of mass coordinates. This study focuses on the measurement method for the center of mass coordinates of heavy cargo vehicles equipped with active suspension systems. Based on the observation that the loading mass and position primarily affect the center of mass coordinates of the suspended mass, we propose a method to measure the center of mass coordinates of the sprung and non-sprung masses separately, allowing for the subsequent calculation of the overall center of mass coordinates for the vehicle.

2 Theoretical Analysis

2.1 Measurement Principle

This study focuses on the measurement method of the center of mass for a two-axle crane equipped with an active suspension system. Active suspension vehicles can be divided into two components: the suspended mass and the non-sprung mass. During crane operations, the center of mass coordinates of the suspended mass are affected, while those of the non-sprung mass remain unchanged. The non-sprung mass and its center of mass coordinates are determined at the factory, so by measuring the suspended mass and its center of mass coordinates, the overall vehicle mass and its center of mass coordinates can be calculated. The measurement process for the center of mass coordinates of the suspended mass consists of measuring the coordinates in the X, Y, and Z directions. The X and Y coordinates can be obtained using the four-wheel weighing method, while the Z coordinate is deter-mined using the weight response method.

2.2 Calculation Model

Based on the working principle, the measurement of the center of mass coordinates for the suspended mass is divided into three directions: X, Y, and Z. A co-ordinate system OXYZ is established, with the intersection point of the vehicle's left–right geometric symmetry plane, the center plane of the first and last axles, and the ground serving as the origin O. The positive direction of the X-axis is to the right of the vehicle's forward direction, the positive direction of the Y-axis is aligned with the vehicle's forward direction, and the positive direction of the Z-axis is oriented upwards, perpendicular to the XOY plane.

The structure of the hydraulic actuator is shown in Fig. 1, and the force generated by the hydraulic actuator is as follows:

$$f = P_2A_2 - P_1A_1 \tag{1}$$

In the equation, f represents the output force of the actuator, while P_1 and P_2 denote the pressures in the rod chamber and the non-rod chamber, respectively. A_1 and A_2 refer to the effective areas of the rod chamber and the non-rod chamber, respectively.

When installed, the actuator forms an angle with the vertical, so the vertical support force of the actuator can be expressed as follows:

$$F = f \cdot \cos\beta \tag{2}$$

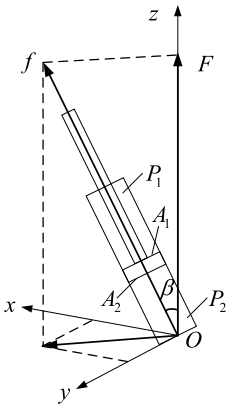


Fig. 1 Schematic diagram of actuator structure

In the equation, F represents the vertical support force of the actuator, and β denotes the angle between the actuator and the vertical.

The vertical support forces of all actuators are balanced by the gravitational force of the suspended mass, so the suspended mass can be expressed as follows:

$$m_s = \left(\sum_{i=1}^{m} F_i\right)/g \tag{3}$$

In the equation, m_s represents the suspended mass, m denotes the number of wheels, and g is the acceleration due to gravity.

The measurement principle for the center of mass coordinate value x_s of the suspended mass is illustrated in Fig. 2. In the XOZ plane, the point where the suspended mass on the right side (or left side) is hinged to the actuator serves as the moment center. The support force from the left actuator (or right actuator) acting on the suspended mass balances the moment created by the gravitational force of the suspended mass. Therefore, the center of mass coordinate value x_s in the vehicle coordinate system can be expressed as follows:

$$x_s = a \cdot \frac{m_s g - 2F_l}{2m_s g} \text{ or } x_s = a \cdot \frac{2F_r - m_s g}{2m_s g} \tag{4}$$

In the equation, a represents the distance between the left and right hinge points, $\boldsymbol{F}_l$ denotes the support force from all left-side actuators acting on the suspended mass, and $\boldsymbol{F}_r$ represents the support force from all right-side actuators acting on the suspended mass.

The measurement of the center of mass coordinate value y_s of the suspended mass is illustrated in Fig. 3. In the YOZ plane, the coordinate origin serves as the moment center. The support forces from the actuators on each axis acting on the suspended mass balance the moment created by the gravitational force of the suspended mass. Therefore, the center of mass coordinate value y_s in the vehicle coordinate system

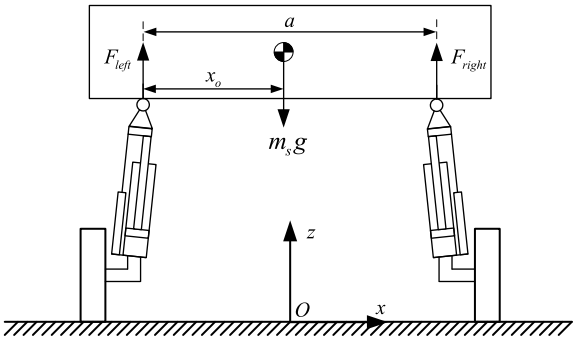


Fig. 2 Schematic diagram of X-axis coordinate measurement

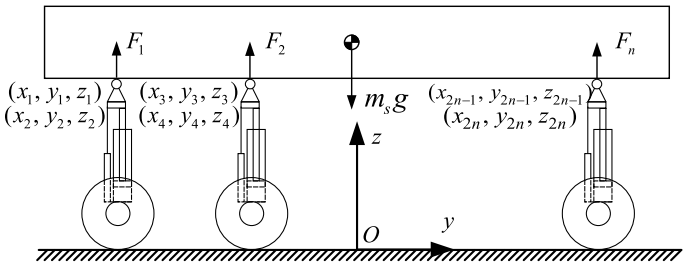


Fig. 3 Schematic diagram of Y-axis coordinate measurement

can be expressed as follows:

$$y_s = \left(-\sum_{j=1}^{n} M_j \right) / m_s g \tag{5}$$

In the equation, n represents the number of axles of the vehicle.

The measurement of the center of mass coordinate value z_s of the suspended mass is illustrated in Fig. 4. In the XOZ plane, the suspension actuators are adjusted to a low position while maintaining the suspended mass in a horizontal orientation. The right-side (or left-side) actuator is then extended by a specified length, and the adjusted angle of the suspended mass can be expressed as follows:

$$\alpha = \arcsin\left(\frac{\Delta L}{a} \right) \tag{6}$$

In the equation, α represents the attitude angle, and ΔL denotes the extension length of the actuator.

Based on the support force from the left-side (or right-side) actuator acting on the suspended mass before and after the attitude adjustment, along with the attitude angle and vehicle parameters, the distance of the center of mass of the suspended

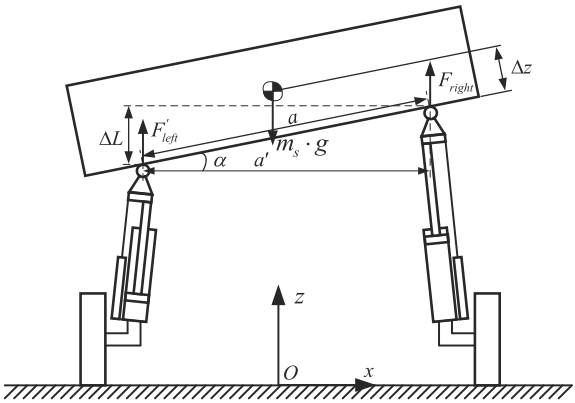


Fig. 4 Schematic diagram of Z-axis coordinate measurement

mass in the Z direction relative to the hinge point can be determined.

$$\Delta z = \frac{a(F_l' - F_l)}{m_s \times \tan(\alpha)} \text{ or } \Delta z = \frac{a(F_r' - F_r)}{m_s \times \tan(\alpha)} \tag{7}$$

In the equation, $\boldsymbol{F}_{\mathrm{l}}$ and $\boldsymbol{F}_{\mathrm{l}}{}'$ represent the support forces from all left-side actuators acting on the suspended mass before and after the attitude adjustment, respectively. $\boldsymbol{F}_{\mathrm{r}}$ and $\boldsymbol{F}_{\mathrm{r}}{}'$ denote the support forces from all right-side actuators acting on the suspended mass before and after the attitude adjustment, respectively.

The center of mass coordinate value z_{s} of the suspended mass in the vehicle coordinate system can be expressed as follows:

$$z_s = \Delta z + \frac{1}{m} \cdot \sum_{i=1}^{m} z_i \tag{8}$$

In the equation, z_{i} represents the Z-axis coordinate values of each hinge point in the vehicle coordinate system, and m denotes the number of actuators.

The non-sprung mass and its center of mass coordinates are determined at the factory, so the overall center of mass coordinates for the vehicle can be expressed as follows:

$$x = \frac{m_s \cdot x_s + m_n \cdot x_n}{m_s + m_n} \tag{9}$$

$$y = \frac{m_s \cdot y_s + m_n \cdot y_n}{m_s + m_n} \tag{10}$$

$$z = \frac{m_s \cdot z_s + m_n \cdot z_n}{m_s + m_n} \tag{11}$$

In the equation, m_n represents non-sprung mass, x_n, y_n, z_n are the coordinates of the non-sprung mass.

3 Experimental Comparison and Analysis

The active suspension two-axle crane used in this study is shown in Fig. 5, with the main parameters listed in Table 1. To validate the center of mass measurement method proposed in this paper, it is necessary to obtain the accurate center of mass coordinate values of the vehicle for experimental comparison. To achieve this, the experimental vehicle underwent center of mass coordinate measurements at the Jiawang Comprehensive Testing Ground, as illustrated in Fig. 5.

During the testing, the actuators were fixed at the specified positions, and the test platform was activated to repeat the measurement of the center of mass coordinates



Fig. 5 Experimental site diagram

Table 1 Main parameters of the vehicle

Test Times	Parameters
a	1320 mm
m	4
n	2
A1	0.00125 m^2
A2	0.00567 m^2
Non-sprung mass	1.2 t
X-value of non-sprung mass	0 mm
Y-value of non-sprung mass	28.32 mm
Z-value of non-sprung mass	635 mm

Table 2 Statistics of center of mass measurement results

Test times	X value (mm)	Y value (mm)	Z value (mm)
1	− 6	30.5	1848
2	− 4	28.8	1855
3	− 5	25.6	1850
4	− 3	29.3	1890
5	− 6	27.4	1865
Average	− 4.8	28.32	1861.6

five times. The measurement results are shown in Table 2, and the final center of mass coordinates are taken as the average of the five measurements.

The X and Y values of the center of mass for the suspended mass are measured using the four-wheel positioning method mentioned earlier. Therefore, it is necessary to obtain the pressures in the rod chamber and the non-rod chamber of each actuator to calculate the support forces exerted by each actuator on the suspended mass. To minimize the influence of friction between the hydraulic rod and the hydraulic cylinder on the measurement results, the actuators are approached from both the longest and shortest ends to the target position, and the actuator output force is taken as the average of the two values. The vehicle is parked on a level surface, and the experiment is repeated five times. The output forces of each actuator are shown in Fig. 6.

Based on the output forces of each actuator and the relationships between the center of mass coordinates of the suspended mass, the non-sprung mass, and the overall vehicle center of mass coordinates, the X and Y values of the vehicle's center of mass are calculated. The results from the five experiments are presented in Figs. 7 and 8.

From the experimental results, it can be concluded that when measuring the X-axis coordinate with the right-side hinge point as the moment center, the deviation of the third experimental result is slightly larger. Overall, the data obtained using the active suspension center of mass measurement method is relatively close to the data

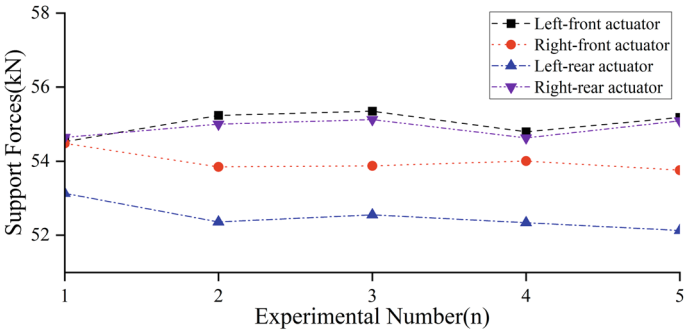


Fig. 6 Output force of the actuator

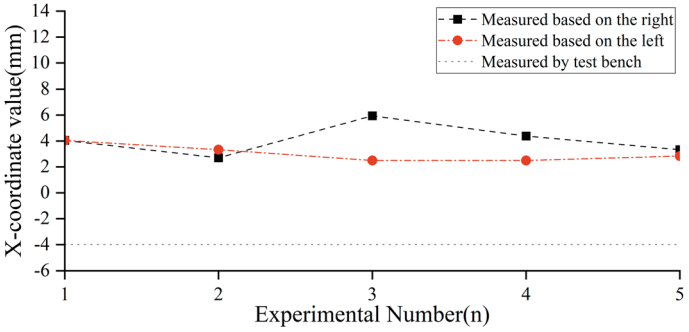


Fig. 7 Results of X-axis coordinate measurement

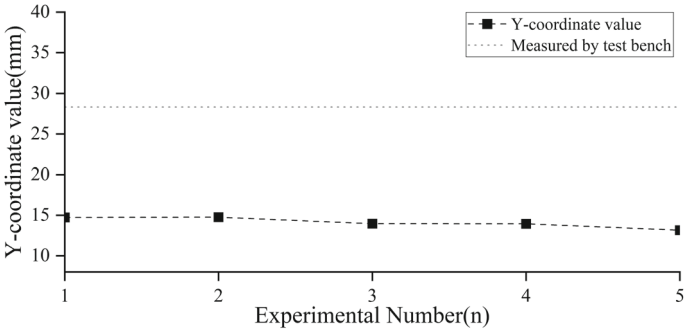


Fig. 8 Results of Y-axis coordinate measurement

measured by the test platform, with a deviation of approximately 8 mm. The Y-axis coordinate measurements are generally stable, with a difference of about 12.5 mm compared to the data from the test platform, which is within the allowable range of experimental error.

The Z value of the center of mass for the suspended mass is measured using the mass response method mentioned earlier. The vehicle is parked on a level surface, and the actuators are adjusted to a low position while maintaining the suspended mass in a horizontal orientation. The left-side actuator is kept fixed, and the pressures in the rod chamber and non-rod chamber of the left-side actuator are recorded. The right-side actuator is then extended by 180 mm, and the pressures in the rod chamber and non-rod chamber of the left-side actuator are recorded again. This process is repeated for five experiments. Similarly, five experiments are conducted with the right-side actuator fixed while following the same procedure. The results of the Z value for the center of mass of the suspended mass from both experimental setups are presented in Fig. 9.

From Fig. 9, it can be deduced that the results from the five experiments exhibit a high level of consistency, thereby validating the repeatability of the methodology proposed in this study. The measurements obtained with the right-side actuator fixed

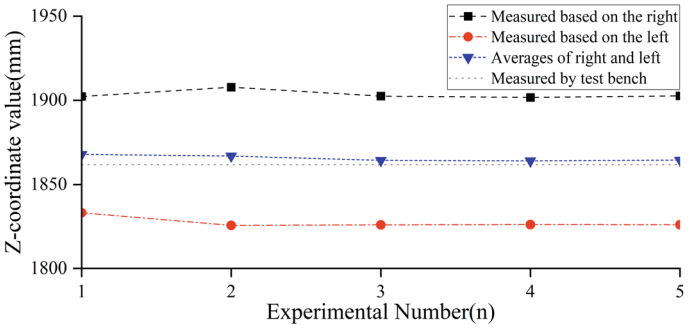


Fig. 9 Results of Z-axis coordinate measurement

consistently exceed those recorded by the test platform, whereas the measurements obtained with the left-side actuator fixed are consistently lower than the platform readings. To enhance the accuracy of the center of mass coordinates, the average of the two sets of measurement results can be calculated, with the final averaged value serving as the definitive measurement of the vehicle's center of mass.

4 Conclusion

This paper focuses on the active suspension two-axle crane, providing an in-depth analysis of the principles of vehicle center of mass measurement and pro-posing a measurement scheme based on active suspension vehicles. By obtaining the pressures in the various chambers of the active suspension actuators and their displacement data, and combining these with the vehicle's original parameters, the formula for the overall vehicle center of mass coordinates is derived and experimentally investigated. The conclusions are as follows:

(1) The four-wheel weighing method can accurately determine the center of mass coordinates of the vehicle in the X and Y directions, with experimental results showing a high degree of consistency with the test platform measurements, making it a viable method for measuring the center of mass of heavy vehicles with active suspension.
(2) The mass response method can measure the center of mass coordinate of the vehicle in the Z direction; however, it exhibits slightly larger measurement errors compared to the four-wheel weighing method. While there is some discrepancy between the mass response method results and the test platform measurements, the errors remain within an acceptable range.
(3) During the experiment, pressure fluctuations were observed in both the rod and non-rod chambers as the actuator reached the target position. Future work should focus on enhancing the stability of displacement control to mitigate measurement errors caused by these pressure fluctuations.

Acknowledgements The Joint Fund for Regional Innovation and Development of the National Natural Science Foundation of China under Grant (U20A20332). Hebei Province Natural Science Foundation Innovation Research Group Project (E2020203174).

References

1. Li W, Wang X (2022) Analysis and research on center of mass measurement methods for commercial vehicles. Autom New Power 5(05):61–63
2. Skrucany T, Synak F, Semanova S et al (2018) Detection of road vehicle's centre of gravity, pp 1–7
3. Patel S, Topiwala J (2017) Measuring a centre of gravity of an object using 4 load transducer method. Int J Eng Tech Res V6(1):210–214
4. Chien SY, Wang AS, Wong CC (2020) Design and implementation of two-wheeled self-balancing vehicle based on load sensors. In: 2020 international conference on system science and engineering (ICSSE)
5. Yang F (2005) On the estimation of center of gravity height of arbitrarily loaded articulated freight vehicle. Concordia University
6. Cui A, Guan S, Ling S et al (2021) Research on vehicle center of mass position measurement methods. Mech Eng 01:119–121
7. Guo Y (2019) Analysis of center of mass measurement methods for multi-axle vehicles. Heilongjiang Sci 10(14):70–71
8. Liu J (2023) Design and research of an automatic vehicle center of mass measurement system based on the Principle of Torque Balance. Modern Mach 05:40–44
9. Qiu Y, Xiong Y, Li J (2018) Research on a vehicle center of mass position measurement platform based on the weight reaction method. Agric Equipm Veh Eng 56(09):88–90
10. Li R, Zhao Y, Gu L et al (2024) Improved lateral tilt method for vehicle center of mass measurement. J Jilin Univ 1–8
11. Synák F, Nedeliaková E (2024) Changing the position of the vehicle's center of gravity as a result of different load distribution. Appl Sci 14(20):9328
12. Jin X, Yin G, Chen N (2019) Advanced estimation techniques for vehicle system dynamic state: a survey. Sensors 19(19):4289
13. Huang X, Wang J (2013) Center of gravity height real-time estimation for lightweight vehicles using tire instant effective radius. Control Eng Pract 21(4):370–380
14. Huang X, Wang J (2014) Real-time estimation of center of gravity position for lightweight vehicles using combined AKF–EKF method. IEEE Trans Veh Technol 63(9):4221–4231
15. Wittmer K, Sawodny O, Henning KU (2023) Model-based estimation of vehicle center of gravity height and load. J Dyn Syst Meas Contr 145(5):051001
16. Previati G, Gobbi M, Mastinu G (2008) Improved measurement method for the identification of the centre of gravity location and of the inertia tensor of rigid bodies. In: Proceedings of the ASME international design engineering technical conference and computers and information in engineering conference 2008(05):841–850

Open Access This chapter is licensed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License (http://creativecommons.org/licenses/by-nc-nd/4.0/), which permits any noncommercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if you modified the licensed material. You do not have permission under this license to share adapted material derived from this chapter or parts of it.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.



Design and Dynamic Performance Analysis of Rail Grinding on Metro Lines with Small Radius Curves



Haifu Xing

Abstract In response to the issue of severe rail wear in metro line, this paper designs the grinding rail profiles for the left and right rails of small radius curved track, then analyzes the wheel-rail contact and curve negotiation before and after grinding. The main approach is to perform personalized grinding on the outer rail curve and the inner rail curve. After grinding, the contact points in a better contact state. The equivalent conicity of the grinding rail has increased, from 0.085 to 0.104. After grinding, the contact patch area increases, and the contact stress is reduced, indicating that grinding effectively decreases the contact stress and improves the wheel-rail contact relationship. When designed the grinding rail profile, the curve passing performance is enhanced, with a reduction in both the wheel axle lateral force and the derailment coefficient, which is beneficial for reducing wheel-rail wear.

Keywords Merto lines · Small radius curves · Rail grinding · Wheel-rail contact · Curve passing performance

1 Introduction

With increased railway transportation speed and volume, wheel-rail contact fatigue, wear, and noise issues have become increasingly prominent. Rail grinding is an effective maintenance method, that improves the geometric characteristics of wheel and rail, reduces lateral forces and wear, lowers the derailment coefficient, and thus enhances the dynamic performance of vehicles. Research progress has shown that personalized rail grinding schemes can optimize the rail profile in the wheel-rail contact area for specific line conditions.

Currently for the small radius curve rail wear and rail grinding-related research, Tao et al. [1] through the model and data combination of small radius curve rail wear analysis. Povilaitiene et al. [2] analyzed the impact of rail gauge on the small

H. Xing (✉)
Hangzhou Metro Operations Co., Ltd., Hangzhou, China
e-mail: xinghaifu@hzmetro.com

© The Author(s) 2026

H. Saman et al. (eds.), *The 9th International Conference on Advances in Construction Machinery and Vehicle Engineering*, Lecture Notes in Mechanical Engineering,
https://doi.org/10.1007/978-981-96-5527-4_88

radius curve rail side grinding research, Uhlmann et al. [3] studied the impact of rail sanding on the surface hardness of the rail surface for related research and analysis. Krishna et al. [4] considered the effect of rail grinding in the prediction of rail wear and analyzed the long-term evolution of rail wear. Mao et al. [5] proposed a scheme to design rail profile used the contact angle and the equivalent conicity. Wang et al. [6] used the response surface method to design the rail profile for curve tracks in heavy-haul lines, and the designed profile effectively improved the vehicle's curve passing performance. However, these methods are relatively complex in the rail profile design, and with little research on the design of the grinding profile. Analysis of Measured Wheel-Rail Profiles.

The Miniprof rail profile measuring instrument is used in the test in Fig. 1a, b and c are the measured rail profiles. The inner rail has severe wear mainly in the rail head area; while the outer rail has severe side wear and vertical wear, especially from the rail head to the gauge corner side, where the wear is severe. This severely affects the wheel-rail relationship and can have adverse effects on train operation.

2 Personalized Rail Profiles Design of the Curved Track

Design Guidelines: Due to severe rail damage and surface quality defects in this part, the wheel-rail profiles have changed. To prevent surface damage from progressing deeper, and improve the safety and comfort of the vehicle, the rail profile is optimized and grinding to maintain good contact conditions. The GFC method was used to design the rail profile. The worn rail and the grinding rail are shown in Fig. 2, where CHN60W-L and CHN60W-R represent the profiles before grinding, and CHN60G-L and CHN60G-R represent the profiles after grinding.

The GFC method [7] is as follows: The rail profile is divided into two areas. $f_1(x_1; \sigma_1, c)$ is the left section, $f_2(x_2; \sigma_2, c)$ is the right section. Let the left zone be β_1 is the variance in left section, $2\beta_1$ is for right section. β_2 is a factor to correct $f(x; \sigma, c)$.

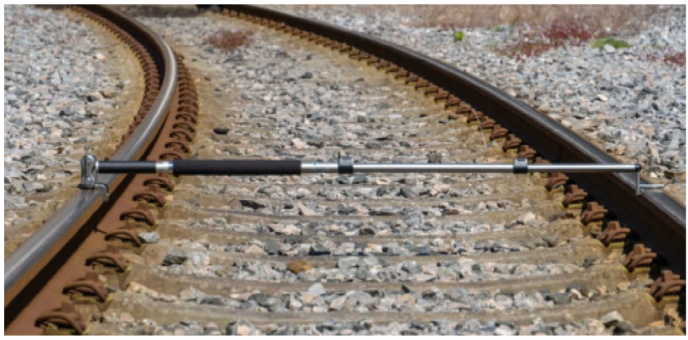
$$f_1(x_1; \sigma_1, c) = e^{\frac{-(x_1-c)^2}{2\sigma_1^2}} \quad x_1 \leq c \tag{1}$$

$$f_2(x_2; \sigma_2, c) = e^{\frac{-(x_2-c)^2}{2\sigma_2^2}} \quad x_2 \geq c \tag{2}$$

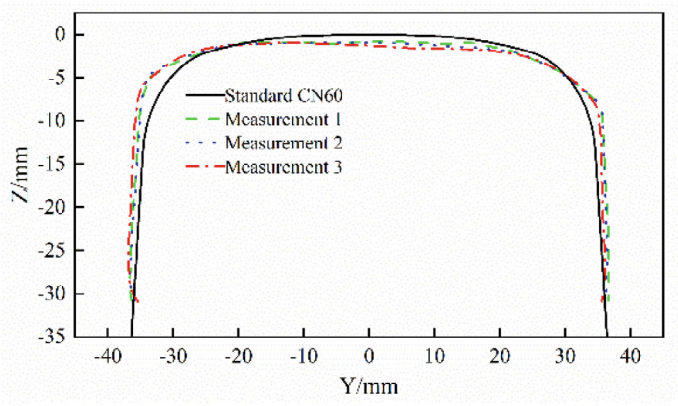
$$\sigma_2 = 2\sigma_1 = 2\beta_1 \tag{3}$$

$$f(x; \sigma, c) = f_1(x_1; \sigma_1, c) + f_2(x_2; \sigma_2, c) = f_1(x_1; \sigma_1, c) + f_2(x_2; \alpha_1\sigma_1, c) \tag{4}$$

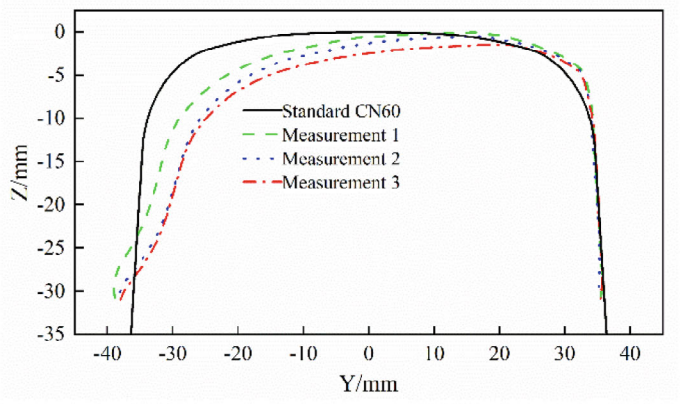
$$y_2 = \beta_2 f(x; \sigma, c) \tag{5}$$



(a)



(b)



(c)

Fig. 1 Rail Profiles test: **a** Miniprof wheel tester; **b** Inner side rail profiles; **c** Outside rail profiles

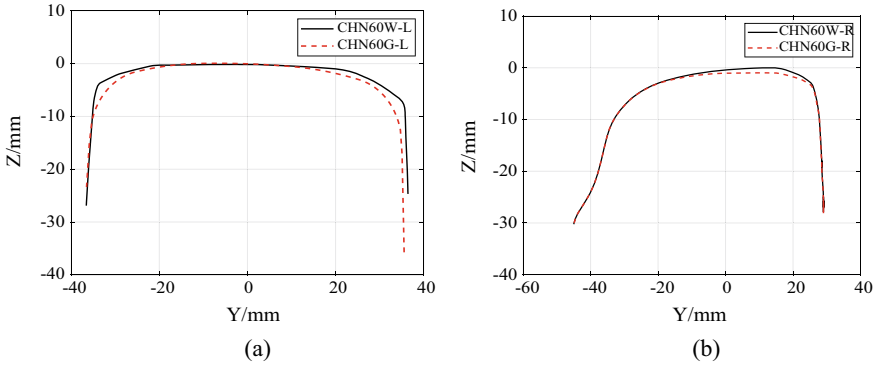


Fig. 2 Rail Profiles test: **a** Inner side rail profiles; **b** Outside side rail profiles

3 Analysis of Wheel-Rail Contact Relationship Before and After Design

The geometric situation of wheel-rail contact is shown in Fig. 3. Figure 3a, b represent contact positions before rail grinding, with contact positions taken at intervals of 0.5 mm. Figure 3c and d represent contact positions after rail grinding. The main analysis focuses on the contact points under lateral shifts ranging from −12 mm to 12 mm. Before rail grinding, the wheel-rail contact positions are relatively concentrated, with the inner rail mainly distributed on the rail shoulder and the outer rail mainly on the rail head. Additionally, there is a phenomenon of contact point jumping in the middle. After grinding, the wheel-rail contact relationship has been improved. Compared with before grinding, the contact points are less concentrated, which can reduce the contact pressure on the rails, and the contact points are more even, resulting in a better contact condition. Figure 3e shows the equivalent conicity curve; the equivalent conicity after grinding is increased by 3 mm. Figure 3f shows the wheel diameter difference. Overall, these improvements enhance the comfort and safety of train operation.

Figure 4 illustrates the diagram of contact patches on the rails before and after grinding. Before grinding, the maximum area of the contact patch on the left rail is 91.8216 mm^2, and after grinding, it increases to 101.825 mm^2. For the right rail, the maximum area of the contact patch before grinding is 64.5244 mm^2, and after grinding, it increases to 68.3049 mm^2. The increase in contact patch area after grinding indicates a reduction in contact stress, demonstrating that grinding effectively decreases contact stress and improves the wheel-rail contact relationship.

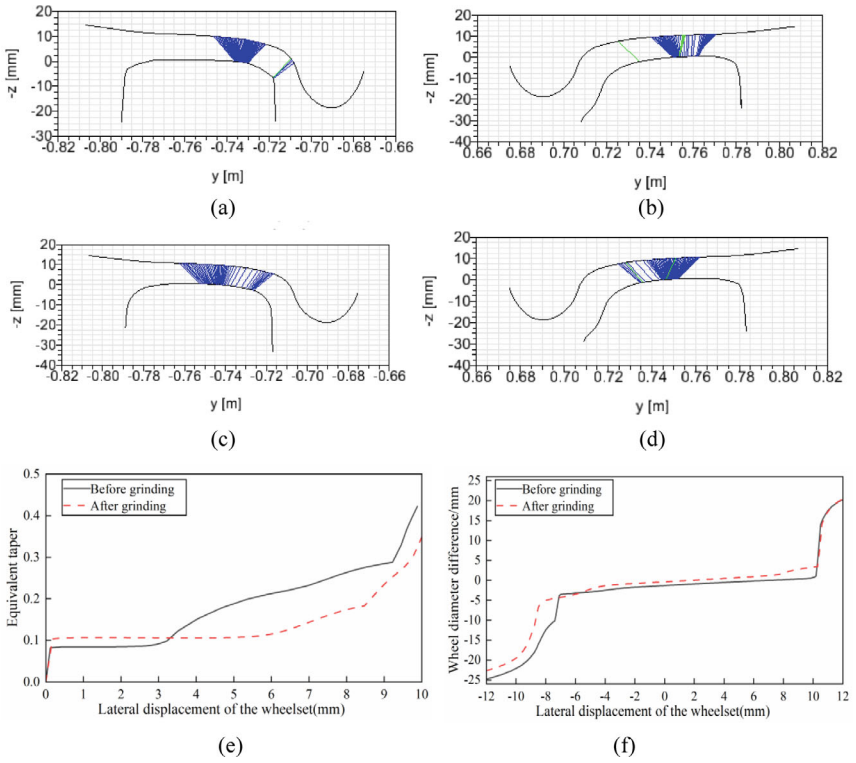
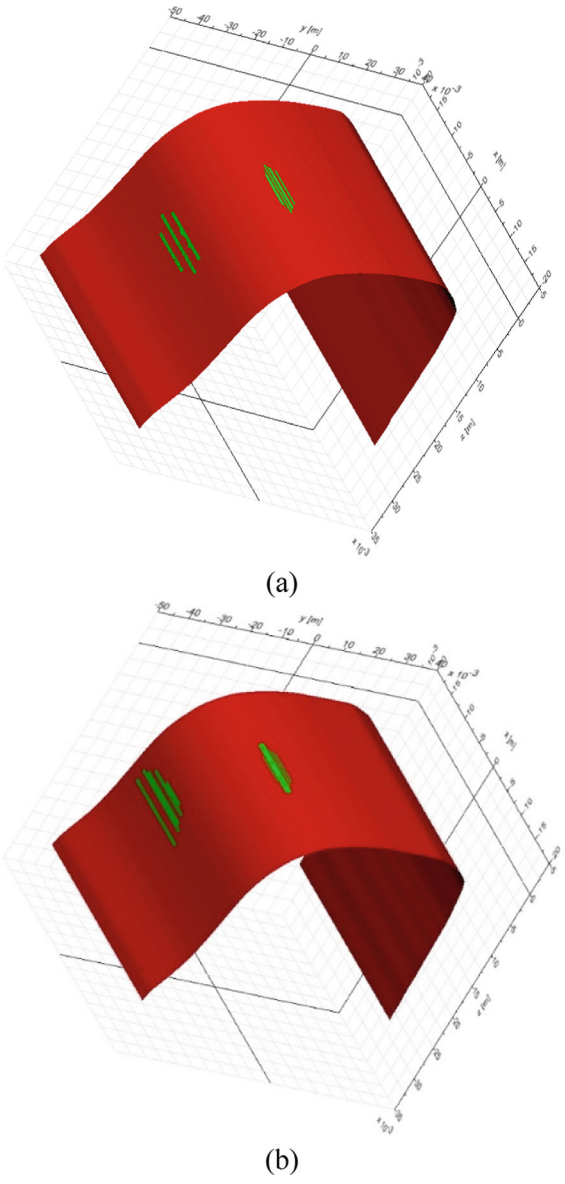


Fig. 3 Wheel-rail contact relationship **a** Contact points of worn left rail; **b** Contact points of worn right rail; **c** Contact points of left desin rail; **d** Contact points of right desin rail; **e** Equivalent conicity; **f** Wheel diameter difference



(a)

(b)

Fig. 4 Contact patches on the rail **a** Worn right rail; **b** Design right rail

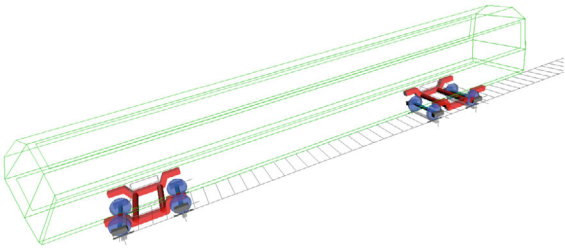


Fig. 5 Vehicle dynamics model

4 Dynamic Response Before and After Grinding

4.1 Vehicle Dynamic Model

The dynamic model of China'2s metro vehicles was established in Simpack, which uses LM tread profiles. The model primarily consists of 1 carbody, 2 bogies, 4 wheelsets, and 8 axleboxes. Considering the effects of the primary and the secondary suspension. The metro dynamic model is established in SIMPACK presented in Fig. 5. The model considers the nonlinear contact relationship between wheels and rails as well as the vehicle suspension system.

4.2 Curve Passing Performance

This study simulates two actual railway metro lines to investigate the relevant indicators of curve passing performance. Line 1 (consisting of 100 m straight + 130 m easing curve + 312.982 m circular curve + 130 m easing curve + 100 m straight, with a radius of 400 m, superelevation of 74 mm, and a speed of 60 km/h) is based on actual measured lines. The calculation results are shown in Fig. 6. Under the curved condition, the wheel axle lateral force and derailment coefficient are both within the safety limits. Moreover, after grinding, the wheel axle lateral force and derailment coefficient have both decreased, indicating that the curve passing safety performance is better and more stable, which meets the expected standards of grinding.

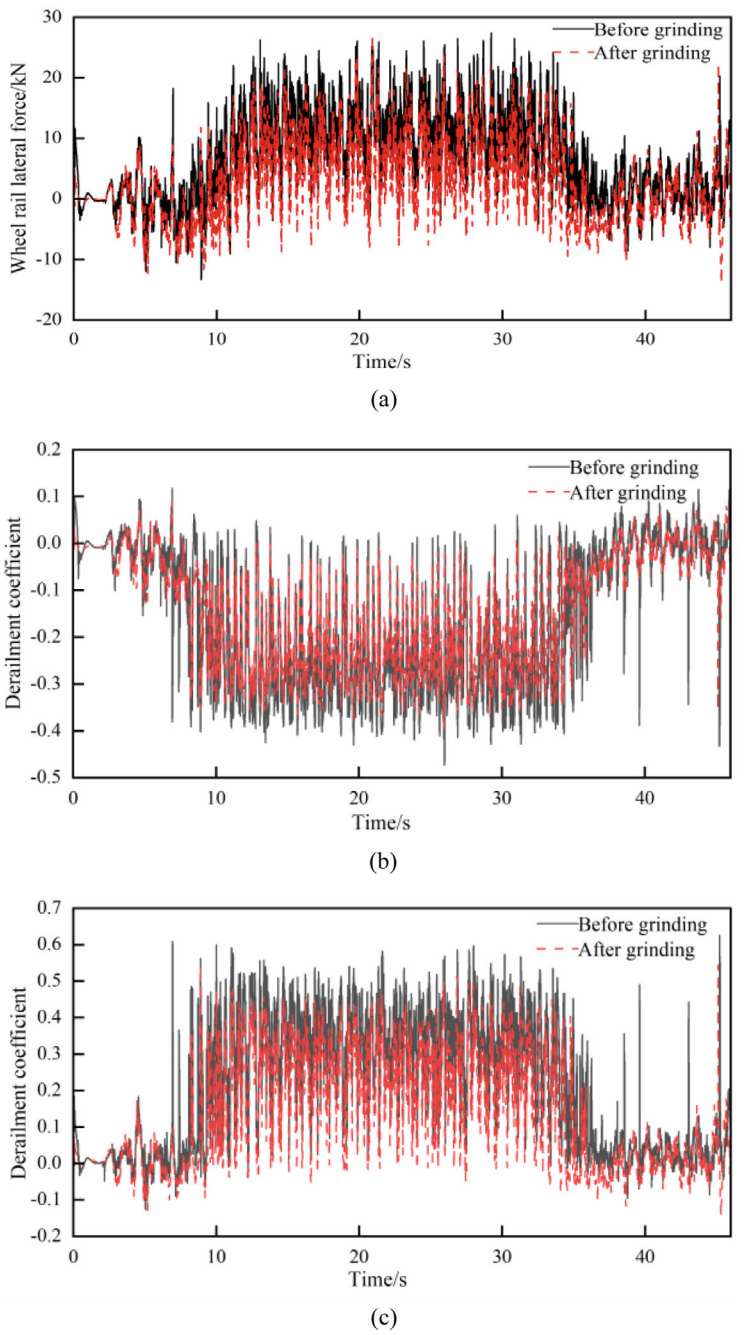


Fig. 6 Curving performance of Line 1: **a** Lateral force on the wheel axle; **b** Derailment coefficient on the left wheel; **c** Derailment coefficient on the left wheel

5 Conclusion

As the issue of severe wear on rails in small-radius curved tracks in metro lines, this paper designs the grinding rail profiles.

(1) After personalized grinding, the wheel rail contact interaction was improved. Compared to before grinding, contact points was more even, the equivalent conicity after grinding was increased in 3 mm.
(2) The contact patch area increased after grinding, which resulted in a reduction of contact stress. The grinding effectively decreased the contact stress and improved the wheel and rail relationship.
(3) After grinding, both the wheel axle lateral force and the derailment coefficient were reduced, indicating that the safety performance of curve passing was better and more stable, meeting the expected standards of grinding.

References

1. Gao T, Wang Q, Yang K (2021) Estimation of rail renewal period in small radius curves: a data and mechanics integrated approach. Measurement 185:110038
2. Povilaitiene I, Kamaitis IZ, Podagelis I (2006) Influence of gauge width on rail side wear on track curves. J Civ Eng Manag 12(3):255–260
3. Uhlmann E, Lypovka P, Hochschild L, Schröer N (2016) Influence of rail grinding process parameters on rail surface roughness and surface layer hardness. Wear 366:287–293
4. Krishna VV, Hossein-Nia S, Casanueva C, Stichel S (2020) Long term rail surface damage considering maintenance interventions. Wear 460:203462
5. Mao X, Shen G (2017) An inverse design method for rail grinding profiles. Veh Sys Dyn 55(7):1029–1044
6. Wang J, Chen S, Li X, Wu Y (2016) Optimal rail profile design for a curved segment of a heavy haul railway using a response surface approach. Proc Inst Mech Eng F: J Rail Rapid Transit 230(6):1496–1508
7. Qi Y, Dai H, Song Y, Ye Y, Li D (2024) Optimisation design of grinding rail profile of metro lines for small radius curved track based on the GFC method: a case study. Veh Syst Dyn 62(8):1929–1951

Open Access This chapter is licensed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License (http://creativecommons.org/licenses/by-nc-nd/4.0/), which permits any noncommercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if you modified the licensed material. You do not have permission under this license to share adapted material derived from this chapter or parts of it.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.



Advanced Vehicle and Robot Control and Optimisation Technologies

Modal Analysis and Experimental Study of Wheeled Special Vehicle Body



Wenshuo Li, Mingming Dong, and Weiwei Lv

Abstract As a large box-bearing structure, the body of the special vehicle must not only bear various loads, but also have a certain protective capacity. As an important carrier in the whole life cycle of special vehicles, the structural design and test mode analysis of body-in-white have an important impact on the performance and reliability of special vehicles. In order to obtain the dynamic characteristic parameters of the body and verify its structural rationality, the calculation mode and test mode analysis of a wheeled special vehicle body-in-white were taken as the research object. The reliability of the finite element model is verified by the comparison and analysis of the modal simulation and the experimental results.

Keywords Special vehicles · Finite elements · Modal testing

1 Introduction

The body of a special vehicle is usually a large box-bearing structure, which is excited by multiple sources such as road surface, power transmission, suspension device, and task load, and the working conditions are extremely harsh. In order to meet the multi-body dynamics requirements of wheeled special vehicles driving on land, and at the same time meet the design requirements of protection, vehicle life cycle reliability, vibration characteristics, etc., the design of special vehicle body structure is very important. In general, in vehicle modal research, vibration problems need to focus on the control of low-order modal parameters [1, 2]. Structure-borne noise is usually affected by local modal parameters [3], and the extraction of modal parameters can be carried out by finite element calculation modes or test modes. Modal analysis usually uses two methods to compare and verify each other.

At present, there are many modal studies on small and medium-sized car bodies such as passenger cars, buses, and commercial vehicles [3–7], but few modal studies

W. Li · M. Dong · W. Lv (✉)
School of Mechanical Engineering, Beijing Institute of Technology, Beijing, China
e-mail: lvweiwei@bit.edu.cn

© The Author(s) 2026

H. Saman et al. (eds.), *The 9th International Conference on Advances in Construction Machinery and Vehicle Engineering*, Lecture Notes in Mechanical Engineering,
https://doi.org/10.1007/978-981-96-5527-4_89

on large load-bearing special bodies. In this paper, the dynamic characteristics of the box-bearing body structure of a certain type of wheeled special vehicle are studied, the experimental mode is studied on the basis of the calculation mode analysis of the body-in-white, and the extracted modal characteristic parameters provide theoretical support for improving the dynamic response characteristics of the vehicle, and the finite element model verified by the experimental mode can be used for the optimal design of the body structure.

2 Modal Simulation

2.1 Modal Analysis Theory

Depending on the material and structural characteristics, the body-in-white generally adopts the following undamped multi-degree-of-freedom vibration system [8]

$$M\ddot{x} + Cx = 0 \tag{1}$$

where: M is the mass matrix; C is the system stiffness matrix; x is the displacement array; $\ddot{x}$ is an array of accelerations.

According to the vibration theory, the free vibration of any elastomer can be decomposed into a series of superpositions of simple harmonic vibrations, so that

$$x = X\sin(\omega + \varphi) \tag{2}$$

where: X is the amplitude; ω is the natural frequency; φ is the phase angle.

Substituting Eq. (2) into Eq. (1), we can obtain

$$(K - \omega^2 C)X = 0 \tag{3}$$

When the system vibrates freely, there must be a displacement, then there is a non-zero solution in Eq. (3), then the coefficient determinant should be equal to zero, i.e.:

$$|K - \omega^2 C| = 0 \tag{4}$$

If the system is an nth-order vibration system, the n eigenvalues (frequencies) and eigenvectors (mode shapes) of Eq. (4) can be solved.

2.2 Finite Element Model Establishment and Solving

In order to save time and expense in the analysis, only critical structures are usually considered, so the model is simplified: non-load-bearing components and some decorative parts are ignored; Removed some of the smaller diameter process holes; Removed some transition fillets.

The body is mainly made of sheet metal stamping parts after welding, and the thickness of the material is small, so the plate and shell elements are selected for the discretization of the structure, the geometric model is meshed with a unit size of 50 mm, and the mesh quality is controlled in the process of division. The body structure is mainly assembled by welding, and rigid connections are used to simulate seam welding.

The body material is defined as structural steel, the elastic modulus is Ex = 2.1 × 1011 Pa, the Poisson's ratio is μ = 0.3, and the density is ρ = 7.85 × 103 kg/m^3. The finite element model of the whole vehicle is shown in Fig. 1, and the entire body structure is discretized into 519,500 elements, with a total of 1,183,272 nodes.

The finite element model is imported into the Ansys Modal Solver Module, and since it is a free-mode solution for the whole vehicle, no constraints are imposed. The solution method adopts the block Lanczos method, which uses a sparse matrix solver, which can efficiently deal with large symmetric eigenvalue problems, and is suitable for complex models such as solid elements, shell elements, and beam elements.

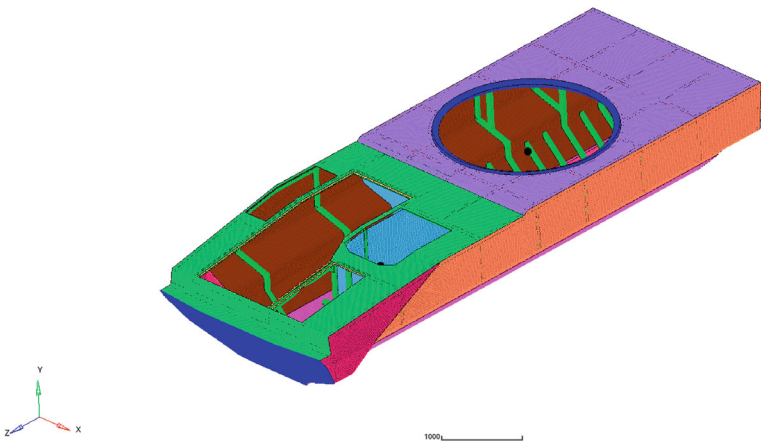


Fig. 1 Finite element model

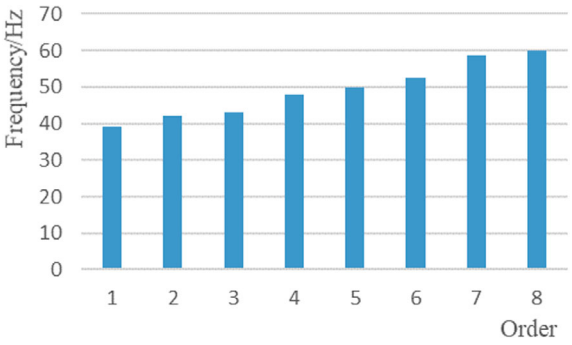


Fig. 2 Frequencies of the first 8 free modes

Table 1 Mode shapes of the first 8 free modes

Order	Frequency (Hz)	Mode shapes
1	39.101	The first order torsion of the whole vehicle
2	42.225	There is noticeable vibration near the bezel of the cooling fan mounting on the top of the car
3	43.256	First-order vibration in the center of the head-mounted panel
4	47.945	There was noticeable vibration on the roof on the side of the turret base
5	49.924	The coupling of first-order bending with second-order torsion
6	52.538	Transverse power compartment bulkhead vibrates
7	58.808	Longitudinal power compartment bulkhead vibrates
8	60.035	Longitudinal power compartment bulkhead vibrates

2.3 Simulation Results

The free mode natural frequencies and mode shapes of the body structure were analyzed by the Ansys software modal solver module, and the frequencies and mode shapes of the first 8th order modes were calculated (the first six rigid body modes were eliminated), as shown in Fig. 2 and Table 1.

3 Modal Testing

3.1 Test System and Measurement Point Layout

The test measurement and analysis system consists of the following three parts. Test excitation system: excitation Dongling ESD-045 and B&K 4825, and GF-10 power amplifier; Response pickup system: including PCB three-axis accelerometers

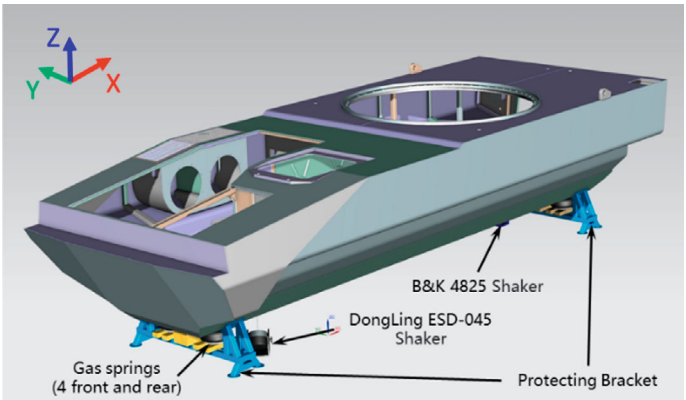


Fig. 3 Schematic diagram of the layout of the excitation points

(10), force sensors (2) and data acquisition front-end LMS SCADAS (sampling rate 204.8 Hz), modal analysis and processing system: software LMS Test.lab 17A.

The body modal test adopts the free mode method, the body structure is supported in the Z direction through 4 air springs, which is used to simulate the free boundary conditions, the natural frequency of the support is about 2 Hz, and the support points are located at the bottom of the front and rear of the car respectively. The auxiliary bracket is protected from contact, as shown in Fig. 3.

The position of the measurement point should accurately obtain the modal formation, and the spatial aliasing should not occur due to the loss of modal information, so after comprehensive consideration, the sensors are arranged on the main longitudinal beam and stiffener on the outside of the vehicle body to avoid the weak plate surface. In addition, based on the results of pre-test and finite element analysis, the selected measurement points should avoid the node position as much as possible on the premise that the main shape of the body structure can be outlined. A total of 411 measurement points are arranged in the whole vehicle, as shown in Fig. 4.

3.2 Data Acquisition

First, the body is supported by air springs (inflation pressure of 4 atmospheres, total load capacity of 12 tons). Then the excitation points are arranged, the exciter and the force sensor are fixed with glue, and the power amplifier is connected. Then, the response point is arranged, the acceleration measurement point is planned according to the body model, marked on the actual vehicle, and the acceleration sensor is fixed at the marked position through the magnetic base. During the test, the sine swept signal is selected for excitation, and the excitation method adopted is two-point excitation. Due to the limited number of accelerometers, mobile accelerometers are required to measure in batches.

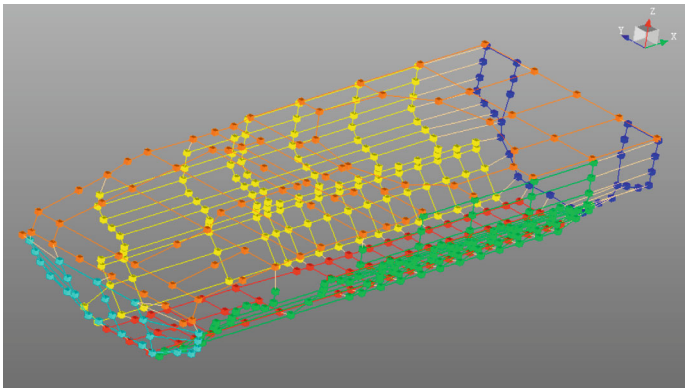


Fig. 4 Distribution of measurement points

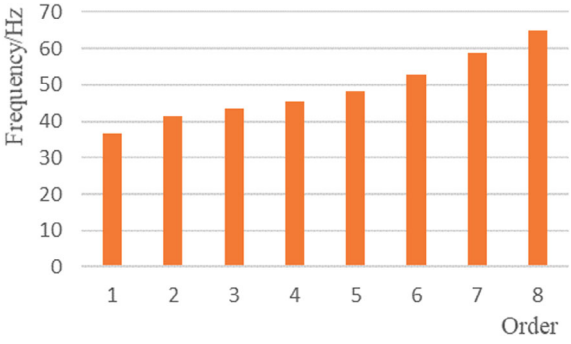


Fig. 5 Frequencies of the first 8 free modes

3.3 Data Processing

The Least Squares Complex Index (LSCE) method of LMS Test.Lab modal analysis software was used to identify the modal parameters of each order, as shown in Fig. 5 and Table 2.

4 Comparison of Simulation and Test Results

The first 8 elastic modes were compared with the experimental results, as shown in Table 3 and Fig. 6. It can be seen that the relative error of the first eight modes is within the acceptable range, and the maximum error occurs in the eighth order, which is 7.7%.

In general, the confidence level of the error between the simulation analysis results and the experimental analysis results is within 10%, indicating that the reliability of

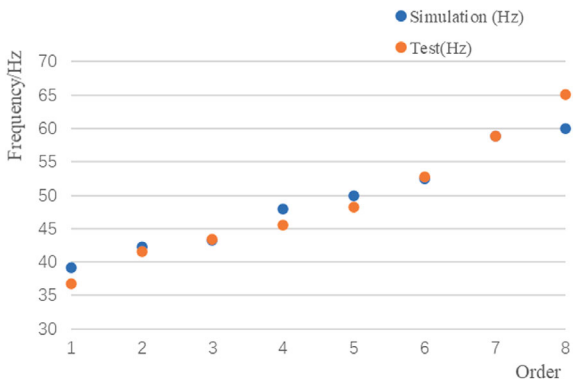
Table 2 Mode shapes of the first 8 free modes

Order	Frequency (Hz)	Mode shapes
1	36.76	The XY plane first-order bending is coupled to the torsional mode, and the main deformation comes from the front body
2	41.63	The front and tailgate are deformed greatly, and the amplitude on both sides of the body is obvious
3	43.37	The amplitude of the center of the front roof plate and the rear of the turret base is obvious, which may affect the stability of the turret
4	45.55	The center of the front roof has a noticeable amplitude
5	48.22	The rear of the body is deformed with a matchbox, and the stiffness of the four corners of the YZ plane needs to be strengthened
6	52.74	Second-order matchbox coupling with respiratory modality
7	58.84	There is no noticeable vibration
8	65.05	The middle part of the roof is deformed, which may affect the stability of the turret

Table 3 Comparison of simulation and experimental results

Order	Simulation (Hz)	Test (Hz)	Relative error (%)
1	39.101	36.76	6.4
2	42.225	41.63	1.4
3	43.256	43.37	0.3
4	47.945	45.55	5.3
5	49.924	48.22	3.5
6	52.538	52.74	0.4
7	58.808	58.84	0.1
8	60.035	65.05	7.7

Fig. 6 Comparison of simulation and experimental results



the calculation results is good [9]. As can be seen from Table 3, the relative error between the simulated values of the first 8 elastic modes and the experimental values is not more than 10%.

Considering that the error may be due to the fact that the subframe and opening and closing parts have been installed, and the chassis has been coated with chassis armor and cushioning materials, which affects the accuracy of the body-in-white mode.

5 Conclusion

Through the comparative analysis of the finite element mode and the test mode of the body-in-white, the agreement between the test mode and the calculated mode is good. The natural frequency error is within the acceptable range, and the mode shape results are in good agreement, which verifies the correctness of the simulation results of the finite element model established in this paper, and can be used to use the body-in-white finite element model for other simulation calculations and optimization design.

According to the simulation and test results of the modalities, the following suggestions are put forward for the optimization of the body structure:

(1) Thicken the roof and front panels to improve their stiffness and reduce deformation, prevent it from resonating at low excitation frequencies resulting in excessive amplitude, thereby improving board reliability and reducing potential structure-borne acoustic radiation.
(2) The rear part of the tailgate and turret seat should be reinforced to prevent excessive vertical vibration radiations from affecting turret stability.
(3) The first-order bending and torsional modes are both above 30 Hz, and the integrated stiffness is considered to be good from the engineering point of view, and the excitation frequency of the vibration source such as engine and road surface needs to be decoupled from the natural frequency of the key modes of the vehicle.

Of course, there are still deficiencies in the work done in this paper, mainly including two aspects: first, the construction of the finite element model is not precise enough, so there is a gap with the actual car body; Second, the selection of measurement points in the modal test is not dense enough, and it is difficult to accurately present the mode shapes of each part of the body. Although the resulting error will not have much impact on the subsequent structural optimization work in this paper, the above problems should be fully considered if the higher modal analysis accuracy requirements are met.

References

1. Shengqin L, Xinyuan F (2020) Study of structural optimization design on a certain vehicle body-in-white based on static performance and modal analysis. Mech Syst Signal Process 135:106405
2. Abdullah NAZ, Fouzi MSM, Sani MSM (2020) Computational modal analysis on finite element model of body-in-white structure and its correlation with experimental data. Int J Automot Mech Eng 17(2):7915–7926
3. Asri MNAM, Abdullah NAZ, Sani MSM (2021) The effect of modal properties of crash box structures with trigger mechanisms towards the crashworthiness by using finite element analysis. J Mech Eng Sci 15(3):8459–8468
4. Daojie J, Yanhai X, Pengxing Z, Shengya L, Keqi L (2018) Research on the influence of BIW solder joint simulation on modal analysis of commercial vehicles. J Guangxi Univ (Nat Sci Ed) 43(3):916–925
5. Xiaoping X, Xu H, Guodong C, Changjiang Z (2010) Modal analysis of commercial vehicle cab's body-in-white. Hunan Daxue Xuebao/J Hunan Univ Nat Sci 37(5):24–30
6. Changsheng L, Changchen Y, Yue Hao C, Yixiu W, Chuanhsing H, Yiting L (2021) Design and analysis of a light electric vehicle. Mech Sci 12(1):345–360
7. Guosheng Z, Guosong Z, Ye C, Jie J (2011) Modal analysis and local stiffness optimization for body-in-white of car. J Highw Transp Res Dev (English Ed) 5(2):137–141
8. Hao C, Chihua L, Zhien L, Cunrui S, Yi S, Menglei S (2020) Structural modal analysis and optimization of SUV door based on response surface method. Shock Vib 2020(1):1–12
9. Ward H (2001) Theory and experiment of modal analysis (Bai HT, Guo JZ, trans). Beijing Institute of Technology Press, Beijing

Open Access This chapter is licensed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License (http://creativecommons.org/licenses/by-nc-nd/4.0/), which permits any noncommercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if you modified the licensed material. You do not have permission under this license to share adapted material derived from this chapter or parts of it.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.



Optimization of Burner Temperature Control Based on Genetic Algorithm



Zhiming Zhang, Wengang Yan, Haiying Cheng, Tiejian Wang, and Fei Liu

Abstract In response to issues of poor control accuracy and insufficient stability in pulverized coal burner control systems caused by nonlinear factors, this paper proposed a fuzzy PID control strategy deeply optimized based on genetic algorithms (GA), leveraging GA's unique global convergence and parallel optimization capabilities. The parameters of the fuzzy PID controller were optimized using a genetic algorithm, enabling rapid tuning of the system's weight factors, thus significantly enhancing its control performance and achieving precise control of burner temperature. MATLAB was used for simulation experiments, with comparative analysis against conventional PID and fuzzy PID control methods. The study shows that system optimization of fuzzy PID through genetic algorithms can effectively reduce errors and improve the control effect of the burner temperature control system. The optimized fuzzy PID controller demonstrates faster response speed, smaller overshoot, and a shorter time to reach steady state compared to conventional PID and fuzzy PID.

Keyword Aggregate dried pulverized coal burner · Fuzzy PID control · Genetic algorithm · Temperature control system

Z. Zhang · W. Yan · T. Wang · F. Liu
Inner Mongolia University, Hohhot, China

H. Cheng (✉)
ChangAn University School, Xi'an, China
e-mail: 45004629@qq.com

Engineering Research Center of Highway Construction Equipment and Technology, Ministry of Education, Xi'an, China

© The Author(s) 2026

H. Saman et al. (eds.), *The 9th International Conference on Advances in Construction Machinery and Vehicle Engineering*, Lecture Notes in Mechanical Engineering,
https://doi.org/10.1007/978-981-96-5527-4_90

1 Introduction

The coal powder burner for aggregate drying achieves heating and drying of aggregates by controlling the flame temperature during the combustion process, and the effectiveness of temperature control directly affects the quality of the aggregate drying [1]. In the past, combustion equipment was often operated based on workers' experience, resulting in poor temperature control accuracy, low thermal efficiency, and high pollutant emissions. Modern optimized control systems are able to control the combustion temperature more precisely [2], improving efficiency and shortening the time to reach a steady state, resulting in lower pollutant emissions [3].

In the rapidly advancing era of technology, traditional PID control methods often fail to achieve the desired effect [4]. Therefore, some scholars have attempted to integrate fuzzy algorithms with traditional PID control strategies to achieve control over fuzzy PID. Liu et al. [5] proposed a fuzzy PID control algorithm, which was studied and compared through simulation experiments of boiler temperature control systems to explore the differences between traditional PID and fuzzy PID control algorithms. The research revealed that fuzzy PID controllers demonstrate superior performance and efficacy in actual operation. To reduce the temperature fluctuation of the adhesive on a jet dispensing machine, Deng et al. [6] designed a fuzzy PID temperature system controller; the study demonstrated that the new system's stability is superior to that of the classic PID controller. Guan and Pan [7] introduced the concept of distributed reasoning, combined with traditional fuzzy self-tuning PID control, to design an optimized fuzzy controller for better control of superheated steam temperature in boilers. The study showed that this method effectively improves overshoot and settling time and adapts well to changes in operating conditions. Fuzzy PID control combines the advantages of traditional PID control and fuzzy control, exhibiting high robustness and adaptability, allowing precise control in environments lacking accurate mathematical models and effectively handling certain complex nonlinear systems. However, due to the inherent blindness and potential errors during manual parameter tuning, issues such as errors and system overshoot may arise during fuzzy PID parameter adjustment, making it difficult to achieve optimal system performance [8, 9].

In this paper, the genetic algorithm (GA) method was adopted for the optimal design of the temperature control system of the coal powder burner for aggregate drying. The parameters of the fuzzy PID controller were optimized through the genetic algorithm, allowing the system's weighting factors to be quickly tuned, thereby achieving precise control of the burner's temperature. The output response curve of the burner temperature control system was obtained through MATLAB, and a comparative analysis with traditional PID control methods was conducted. The research shows that the GA-based fuzzy PID controller demonstrates faster response speed, smaller overshoot, and the ability to reach a steady state in a shorter time during the system regulation process.

2 Model of the Coal Powder Burner Temperature Control System

The aggregate drying process evaporates the internal moisture of wet aggregates using the heat generated by the coal powder burner to reduce the moisture content of the aggregates. In the temperature control system of the coal powder burner, both the coal feed rate and the primary air volume are key parameters. An excessive supply of coal powder leads to incomplete combustion, increasing the concentration of coal powder in the combustion chamber and causing problems such as slagging, which in turn increases the load and wear of the equipment, affecting its stability and lifespan. On the other hand, an insufficient supply of coal powder results in incomplete combustion and a decrease in burner temperature, thereby affecting the drying effect of the aggregates. Primary air provides the necessary oxygen for the volatile components in coal powder, promoting complete combustion. When the moisture content of the aggregates is relatively high, increasing the primary air volume can raise the flame combustion temperature. Conversely, when the moisture content of the aggregates is relatively low, reducing the primary air volume can decrease the heat released by the flame's combustion.

The flame temperature of the coal powder burner is detected by a high-temperature thermocouple and then converted into a standard voltage signal. After receiving and evaluating the signal, the CPU transmits it to the inverter, thereby controlling the speed of the coal mill and the primary air volume, achieving the goal of controlling the coal feed rate to optimize the combustion efficiency of the coal powder. The flame temperature control diagram is shown in Fig. 1.

In the aggregate drying process, the coal feed rate was found to affect the drying temperature. Both the proportional and inertia components can approximately describe the coal feeding process. Moreover, because the characteristic of the controlled object in the simulation experiment is temperature, to reduce the complexity of the simulation experiments, secondary factors such as reverse heat transfer were ignored, and a second-order inertia with delay was used for description. The transfer function of the simulation experiment is:

$$G_1 = \frac{K_1}{(T_1 s+1)(T_2 s+1)} e^{-\tau s} \tag{1}$$

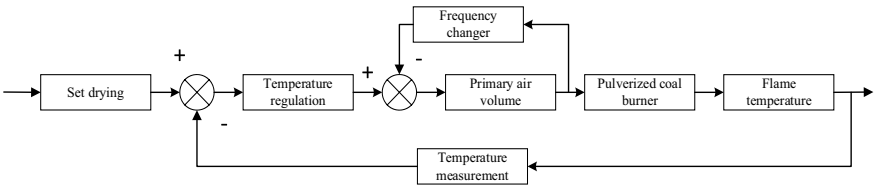


Fig. 1 Aggregate drying temperature control chart

In the equation, K_1 is the system gain, τ represents the delay time s denotes the complex frequency domain variable, and T_1 and T_2 are time constants.

In this paper, the flame temperature of the coal powder burner was determined using a parameter model identification method. Through the coal powder transport process and the combustion process, the transfer function for the flame temperature of the aggregate drying coal powder burner is approximately determined as:

$$G_s = \frac{5}{0.6s^2+4.4s+1} e^{-20s} \tag{1}$$

3 System Control Methods

3.1 Conventional PID Control

The conventional PID controller, as a common controller, operates based on feedback control of the system. By continuously comparing the error between the feedback signal and the desired target, a control output is generated to adjust the system's behavior. Its main task is to adjust the control output according to the size of the error, its accumulation, and rate of change, achieving control over aspects such as system stability and response performance.

The PID controller adjusted the output of the control system based on the error $e(t)$ between the setpoint and the actual output value of the system, through proportional §, integral (I), and derivative (D) operations, to achieve control over system deviations. In PID control, the controller weights and sums the proportional, integral, and derivative terms according to different weighting coefficients to obtain the final control output $u(t)$, expressed as:

$$u(t) = K_p\left[e(t) + \frac{1}{T_i}\int e(t)\mathrm{d}t + T_d\frac{\mathrm{d}e(t)}{\mathrm{d}t}\right] \tag{3}$$

In the equation, K_p is the proportional gain T_i is the integral time, and T_d is the derivative time.

3.2 Fuzzy PID Control

Fuzzy PID control was described as a method that combines fuzzy logic with a traditional PID controller to address systems that are nonlinear, complex, or highly uncertain. The fuzzy PID controller processes inputs, outputs, and rules through fuzzification, which replaces the need for precise mathematical models and parameter tuning required by traditional PID controllers. In this study, the input variable for the designed fuzzy controller is the flame temperature of the pulverized coal burner,

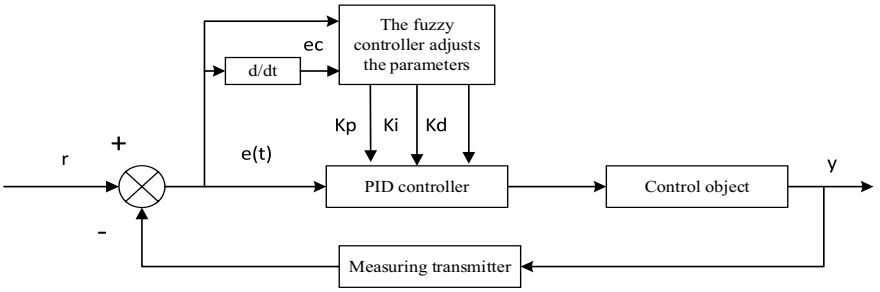


Fig. 2 Block diagram of fuzzy autotuning

while the output variable is the primary air volume of the pulverized coal burner. The block diagram of the fuzzy self-tuning principle is shown in Fig. 2.

During the dynamic adjustment process of the flame temperature control system, high precision was not required, thus a two-dimensional fuzzy controller was chosen to control the system. According to the self-tuning principles of PID parameters, a fuzzy controller with two inputs and three outputs was selected to control the burner temperature control system, as shown in Fig. 3.

The fuzzy universe of discourse for the aforementioned variables was set as $\{-6, 6\}$, and each variable's fuzzy universe of discourse was divided into multiple fuzzy sets to describe the possible states of the variables. The linguistic values for each variable are {NB, NM, NS, ZO, PS, PM, PB}.

According to the principles of coverage, sensitivity, stability, and robustness of fuzzy PID, triangular membership functions were used to describe the fuzzy subsets to achieve precise control and robustness optimization of the system. The expression is as follows:

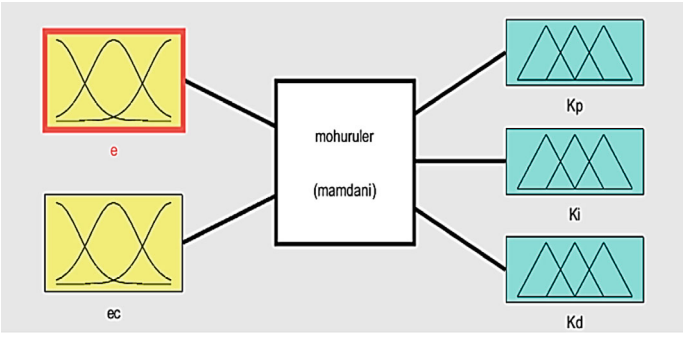


Fig. 3 A two-input, three-output fuzzy controller

$$u(x,a,b,c) = \begin{cases} 0 \; x < a \\ \frac{x-a}{b-a} \; a \le x \le b \\ \frac{c-x}{c-b} \; b \le x \le c \\ 0 \; x > c \end{cases} \tag{4}$$

In the expression $u(x, a, b, c)$ is the output variable, and a, b, c are the parameters of the triangular membership function.

The formation of fuzzy rules was typically based on expert knowledge and experience, which can be obtained through trial and error or expert insights. In a fuzzy PID controller, fuzzy rules are used to determine the behavior of the fuzzy controller and adjust PID parameters based on input variables and their corresponding fuzzy sets, as well as the fuzzy sets corresponding to the output variables. Given that real-time operational control requires compliance with fuzzy reasoning, the design of the fuzzy rule tables for ΔK_p, ΔK_i, ΔK_d was conducted by consulting the principles of tuning parameters in PID controllers, as shown in Table 1. The fuzzy rule surface plot is shown in Fig. 4.

Table 1 ΔK_p, ΔK_i and ΔK_d fuzzy rules

E	*ec*						
	NB	NM	NS	ZO	PS	PM	PB
NB	PB/BN/PB	PB/NB/NS	PM/NM/NB	PM/NM/NB	PS/NS/NB	ZO/ZO/NM	ZO/ZO/PS
NM	PB/NB/PS	PB/NB/NS	PM/NM/NB	PS/NS/NM	PS/NS/NM	ZO/ZO/NS	NS/ZO/ZO
NS	PM/NB/ZO	PM/NM/NS	PM/NSNM	PS/NS/NM	ZO/ZO/NS	NS/PS/NS	NS/PS/ZO
ZO	PM/NM/ZO	PM/NM/NS	PS/NS/NS	ZO/ZO/NS	NS/PS/NS	NM/PM/NS	NM/PM/ZO
PS	PS/NM/ZO	PS/NS/ZO	ZO/ZO/ZO	NS/PS/ZO	NS/PS/ZO	NM/PM/ZO	NM/PB/ZO
PM	PS/ZO/PB	ZO/ZO/PS	NS/PS/PS	NM/PS/PS	NM/PM/PS	NM/PB/PS	NB/PB/PB
PB	ZO/PB/PB	ZO/ZO/PM	NM/PS/PM	NM/PM/PM	NM/PM/PS	NB/PB/PS	NB/PB/PB

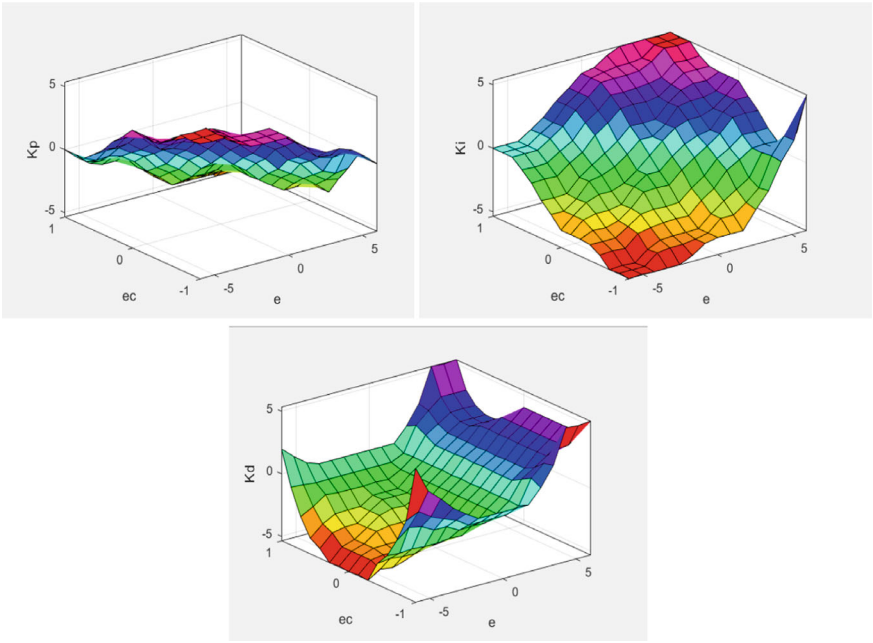


Fig. 4 Blur regular surface diagrams

4 GA Optimization of Fuzzy PID

4.1 Basic Principles of Genetic Algorithms

The genetic algorithm, as a type of optimization algorithm that has garnered significant attention, was inspired by the evolutionary processes in natural life systems. This algorithm aims to simulate critical aspects of biological genetics, such as inheritance, mutation, and natural selection, in order to achieve deep exploration and precise optimization of the solution space for complex problems.

The key to implementing the genetic algorithm lies in constructing a fitness function to evaluate the quality of each chromosome in the population. With this fitness function, all chromosomes in the current population can be assessed, allowing the most advantageous chromosomes to be selected and determining whether these chromosomes are eligible to continue participating in the reproduction process of subsequent generations. Meanwhile, to maintain the diversity of the population, a series of genetic operations are performed on the selected specific chromosomes. Through iterative improvements to the population, the algorithm ultimately approaches the optimal parameter combination.

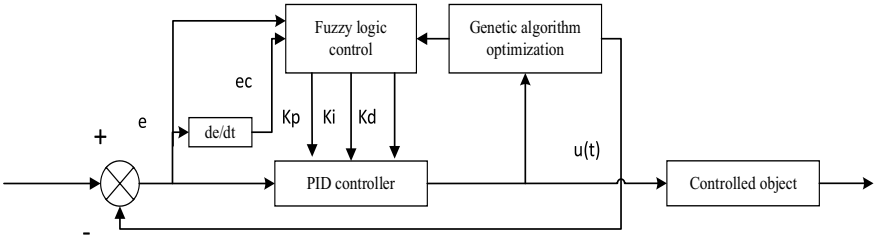


Fig. 5 Fuzzy PID controller optimized based on genetic algorithm

4.2 Control Flow of Genetic Algorithms

The genetic algorithm first initializes the population. By constructing a fitness function, candidate solutions are continuously evaluated and improved. By selectively retaining superior individuals and introducing new genetic information, the algorithm gradually approaches the optimal solution to find the best individual, thereby achieving excellent control performance. The GA-optimized fuzzy PID controller is shown in Fig. 5.

The control flow of the genetic algorithm has the following key steps:

Parameter encoding is the process of converting the feasible solution space into the chromosome search space. Common encoding methods include binary, real number, and integer encoding. Considering the characteristics of multiple fuzzy control variables and chromosome genes, integer encoding was used in this study to effectively represent and optimize the parameters. Meanwhile, to facilitate optimization, the membership function was chosen as an isosceles triangle.

Assuming the vertex coordinates of a triangle are constant, half of its base width (denoted as L) is taken as the parameter to be optimized. That is, there are 21 parameters to be optimized for the membership function, encoded as {2222222 2222222 2222222}, which represent the input variables ΔI, ΔI_c, and the control variable U. These are the encodings of the triangular membership function L of each fuzzy subset. Encoding was also applied to the quantization factors (K_e,K_c) and the proportional factor (K_u). These parameters were encoded into binary form, constructing a two-dimensional matrix, which was then flattened to form the gene sequence of the chromosome. Thus, the gene length of an individual chromosome is:

$$7 \times 7 + 3 \times 7 + 3 = 73$$

The fitness function calibrates the genetic algorithm, which is a process of searching for the optimal solution. The fitness objective function chosen in this study is:

$$J = \int_{n}^{0} t \cdot e^2(t) \mathrm{d}t \tag{5}$$

In this equation, t represents time and e represents error.

The characteristic of this fitness function is that it focuses on the error during the process, with the initial error being relatively ignored. The fitness function is defined as:

$$F = 1/J \tag{6}$$

The difference in the probability of individuals being selected in the population is referred to as selection pressure, which is positively correlated with the optimization function of the genetic algorithm. The magnitude of the selection pressure influences the optimization capability of the genetic algorithm. Therefore, a linear scaling method was used in this study to calibrate the objective function, as shown below:

$$F(x) = af(x) + b \tag{7}$$

$$b = a(-f_{\min}(x) + \varsigma) \tag{8}$$

In the equation, $F(x)$ is the fitness function, a and b are the determined coefficients, $f(x)$ is the objective function, $f_{\min}(x)$ is the minimum value of the objective function, x is the chromosome, and ς is a small number.

Selection During the selection process of the offspring, individuals with the highest fitness were produced. Therefore, in the selection process of each generation, the best individuals of the current generation are retained and passed on to the next generation. This selection method is called elite selection, and its purpose is to collect results more effectively and find the global optimization more quickly. The probability of an individual being selected is shown below:

$$P_i = \frac{F_i}{\sum_{i=1}^{M} F_i} \tag{9}$$

In the equation, F_i is the fitness of the individual, and M is the population size.

Crossover Parent individuals were selected, and crossover points were randomly determined on their chromosomes. The crossover operation is performed on the chromosomes of two parent individuals, and the newly generated offspring individuals are replaced into the population.

$$\begin{aligned} \text{Parental sample}: P_1 &= X = (x_1, x_2 \ldots x_N) \\ P_2 &= Y = (x_1, x_2 \ldots x_N) \\ \text{After crossover}: Z_1 &= aX + (1-a)Y \\ Z_2 &= aY + (1-a)X \end{aligned}$$

Mutation Individuals requiring mutation operations were selected from the population, and mutation points were randomly determined in their chromosomes for mutation operation. By updating the mutated individuals, they were replaced into

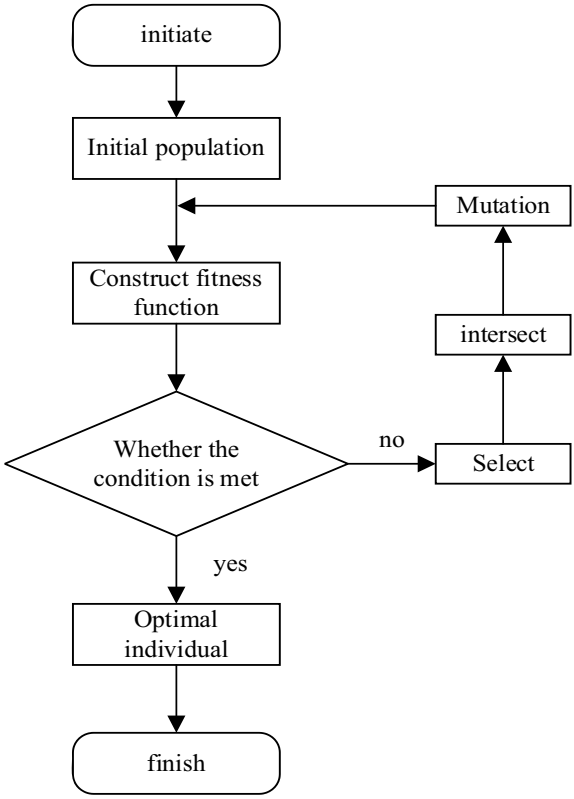


Fig. 6 Flow chart of PID control parameters optimization based on genetic

the population. This operation helps maintain the diversity of the population and introduces new gene combinations, thereby aiding in global search.

Through the above analysis, the PID control parameters GA optimization flowchart is obtained as shown in Fig. 6.

5 Simulation Analysis

In order to test the control effect of the improved genetic algorithm fuzzy PID controller parameters on the burner temperature control system, MATLAB was used to conduct simulation experiments on the burner temperature control system and compare it with the conventional PID and fuzzy PID control methods. The simulation structure of the burner temperature control system is shown in Fig. 7.

The initial parameters of the genetic algorithm are as follows: the initial population size M was set to 10, the number of generations G was set to 20, the crossover probability was set to 0.7, and the mutation probability was set to 0.001.

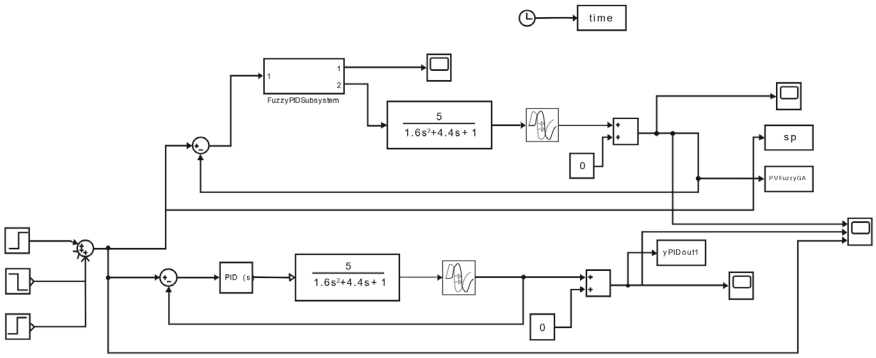


Fig. 7 Simulation structure diagram of burner temperature control system

The simulation curve graphs of conventional PID, fuzzy PID, and GA fuzzy PID were shown in Fig. 8. It can be seen from Fig. 8 that the regulation time of conventional PID was relatively long, which would lead to an inability to quickly stabilize the flame, adversely affecting the drying quality of the aggregates. At the same time, the conventional PID control system exhibited excessive overshoot, prolonging the time to reach system steady state, and resulting in larger fluctuations, significantly impacting the experimental results. Fuzzy PID effectively improved the system's response time, shortening the time to reach steady state. However, the expert scoring of the weight factors in the fuzzy PID controller had subjectivity, making it unable to perform parameter tracking adjustments, resulting in slower adjustment speed. GA fuzzy PID, as an improved control strategy based on fuzzy PID control, can quickly respond to control demands, reduce overshoot, and shorten the time for the system to reach steady state. From the simulation results, GA fuzzy PID shows a clear advantage in control effectiveness compared to conventional PID and fuzzy PID, significantly reducing the system's adjustment time and enhancing the overall system performance.

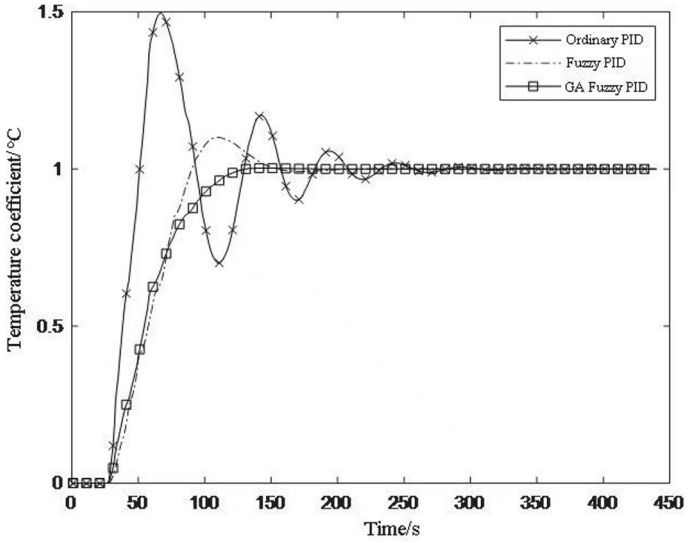


Fig. 8 Simulation curve

6 Conclusion

Compared to the conventional PID and fuzzy PID control algorithms, fuzzy PID control optimized by genetic algorithms demonstrated faster feedback response, lower system overshoot, and a quicker ability to reach a stable state during the system regulation process, ensuring accuracy and stability in the temperature control system for aggregate drying pulverized coal combustion. This provides a feasible method for designing the burner temperature controller. However, there are still shortcomings in this study. In future research, it is hoped that a threshold alarm device can be introduced so that when the burner's temperature is too high or too low, an alarm is immediately triggered to control the temperature promptly, ensuring the temperature stability of the combustion equipment.

With the in-depth research on PID control and the rapid development of intelligent optimization algorithms, the introduction of genetic algorithms into fuzzy PID control not only avoids excessive reliance on expert experience for blind optimization but also effectively optimizes the uncertain weight factors in fuzzy PID controllers. It fully exploits the advantages of fuzzy PID controllers in control, enabling a more flexible response to various complex working conditions and requirements, thereby expanding its application range.

Funding Project Science Program of Inner Mongolia Autonomous Region (2020GG0269).

References

1. Li XJ, Cheng HY, Hu ZY et al (2021) Application of particle swarm optimization fuzzy PID in burner temperature control. Mech Sci Technol 40(02):276–280
2. Shen P, Duan XG (2018) Fuzzy PID double-layer parameter tuning and its application in curing oven. J Chongqing Univ 41(01):78–87
3. Wu J, Zhang H (2024) Fuzzy PID control of two-coil MR brake optimized by genetic algorithm. Automot Eng 46(03):526–535
4. Zhihan Z, Di B (2022) Optimization of improved PID control strategy based on genetic algorithm. J Phys: Conf Ser 2417(1):012025
5. Liu H, Liu MC (2018) Simulation research on boiler temperature control system based on fuzzy PID. Autom Instrum 33(04):20–25
6. Deng G, Feng Z (2019) Fuzzy PID based temperature controller for jetting dispenser. IOP Conf Ser: Mater Sci Eng 677(5):052083
7. Guan CL, Pan WJ (2018) Simulation of optimal control of superheated steam temperature in thermal power plant boiler. Comput Simul 35(10):145–149
8. Zhao X, Jun L (2022) Research on temperature control of liposome high pressure homogenizer based on genetic algorithm optimization PID. Proc Comput Sci 208(5):330–337
9. Xie X, Xu W, Zhu HZ et al (2023) Fuzzy PID-controlled screw feeding metering system based on GA optimization. J Hubei Univ Technol 38(04):29–33

Open Access This chapter is licensed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License (http://creativecommons.org/licenses/by-nc-nd/4.0/), which permits any noncommercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if you modified the licensed material. You do not have permission under this license to share adapted material derived from this chapter or parts of it.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.



Research on Anti-rocking Control of Wind Turbine Crane Based on Input Shaping Technology



Xuyang Cao, Fengzhan Xu, Xiaochun Tie, and Dan Wang

Abstract A rigid-flexible coupling dynamics model is proposed for the wind turbine crane's truss boom system, considering the large slenderness ratio of the structure and the complexity of the operating environment. This model simultaneously addresses the vibration of the truss boom and the swing of the suspended load. The system is modeled using an equivalent spring-mass-damping system, a suspension system, elastic two-force bars, and an instantaneous wind model to simulate the elastic vibrations of the truss boom, the swing of the suspended object, the rope dynamics, and the wind load. The rigid-flexible coupling dynamics of the truss boom-lifting system is established using Lagrange's equations, and dynamic response analysis is conducted in SIMULINK. Based on the input shaping technique, an input shaper is designed for the target wind turbine crane. An anti-sway control simulation platform is also built in SIMULINK for experimental testing. Results show that the ZVD_2 input shaper provides an effective anti-shaking control for large flexible truss arm systems.

Keywords Wind turbine crane · Rigid-flexible coupling · Lagrange equation · SIMULINK · Input shaping

1 Introduction

In the field of wind power, lattice boom crawler cranes have an irreplaceable role. In the process of wind turbine installation, the working environment is harsh, and the skewing of the lifting process is inevitable, which seriously affects the efficiency and safety of the lifting operation; therefore, studying effective anti-sway control technologies is crucial for the safe operation of the crane.

X. Cao · F. Xu (✉)
Dalian University of Technology, Dalian, China
e-mail: xfz0314@163.com

X. Tie · D. Wang
Taiyuan Heavy Industry Co., Ltd., Taiyuan, China

© The Author(s) 2026

H. Saman et al. (eds.), *The 9th International Conference on Advances in Construction Machinery and Vehicle Engineering*, Lecture Notes in Mechanical Engineering,
https://doi.org/10.1007/978-981-96-5527-4_91

To more accurately capture the coupling effect between the flexible jib's vibration and the swing of the suspended load, Yan et al. [1] considered the geometric nonlinear effect of the jib; Park et al. [2] modeled the jib as an elastomer using the finite element method, which provides a more realistic representation of the boom's vibration and the coupling effect of the pendant swing, though the equations are complex. Currently, the research on anti-rocking control mainly focuses on bridge gantry cranes. Mohammed et al. [3] and Abdullah et al. [4] proposed an enhanced input shaping control scheme for anti-rocking control of bridge cranes; jib cranes cause the spatial yawing of the hoisted object during the slewing motion, and the difficulty of its anti-rocking control is higher than that of bridge gantry type. Kawada et al. [5] and Yan et al. [6] designed variable-gain PD controller and dual-mode input shaper.

Input shaping technology is simple, reliable, and widely used in crane anti-sway control. This paper establishes a crane's dynamic model using the equivalent system method and Lagrange's equation. Based on the system's vibration characteristics, an input shaper is designed, and a SIMULINK-based anti-sway control simulation platform is developed to verify the algorithm's effectiveness through simulation experiments.

2 Crane System Dynamics Model

The wind turbine crane's truss boom-lifting system is modeled as a rigid-flexible coupled dynamics system, as illustrated in Fig. 1 [7].

In spring-mass damped system, the position vector of the mass point in $o - xyz$ is

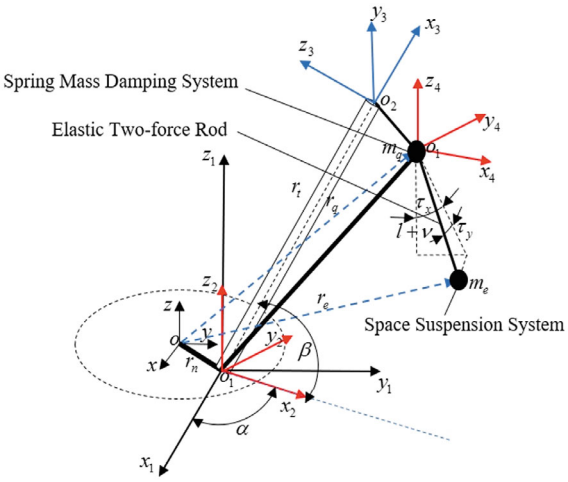


Fig. 1 Rigid-flexible coupling dynamics model

$$r_q = R_a[R_b(r_t + u) + r_n] \tag{1}$$

where, R_a and R_b are the coordinate transformation matrices; u represents the three-dimensional vibration displacement of the mass point m_q at time t in the $o_2 - x_3y_3z_3$ coordinate system.

The first order derivative of Eq. (1) with respect to time gives the velocity of the mass m_q in the coordinate system $o - xyz$ as

$$\dot{r}_q = R_a\Omega_\alpha[R_b(r_t + u) + r_n] + R_aR_b\Omega_\beta(r_t + u) + R_aR_b\dot{u} \tag{2}$$

where $\Omega_\alpha = \dot{\alpha}(t)\Phi_\alpha$, $\Omega_\beta = \dot{\beta}(t)\Phi_\beta$, Φ_α, Φ_β are the skew-symmetric matrices corresponding to the slewing angular velocity and luffing angular velocity of the truss arm.

The equivalent spring-mass-damped system's kinetic energy is given by

$$T_q = \frac{1}{2}\dot{r}_q^{\mathrm{T}} m_q \dot{r}_q = \frac{1}{2}\begin{bmatrix}\dot{\alpha} & \dot{\beta} & \dot{u}^{\mathrm{T}}\end{bmatrix}\begin{bmatrix} M^q_{\alpha\alpha} & M^q_{\alpha\beta} & M^q_{\alpha u} \\ M^q_{\beta\alpha} & M^q_{\beta\beta} & M^q_{\beta u} \\ M^q_{u\alpha} & M^q_{u\beta} & M^q_{uu} \end{bmatrix}\begin{bmatrix}\dot{\alpha} \\ \dot{\beta} \\ \dot{u}\end{bmatrix} \tag{3}$$

The dissipation energy and elastic potential energy are

$$F_{cq} = \frac{1}{2}\dot{u}^{\mathrm{T}} C_q \dot{u} \quad U_q = \frac{1}{2}u^{\mathrm{T}} K_q u \tag{4}$$

where, C_q, K_q are the damping coefficient matrix and stiffness coefficient matrix.

Suspension system, the vectorial diameter of the mass point in $o - xyz$ is

$$r_e = R_a[R_b(r_t + u) + \Gamma_e + r_n] \tag{5}$$

where Γ_e represents the position vector of the mass point m_e in the follower coordinate system $o_1 - x_4y_4z_4$.

The first order derivative of Eq. (5) with respect to time gives the velocity of the mass m_e in the coordinate system $o - xyz$ as

$$\begin{aligned}\dot{r}_e = {} & R_a\Omega_\alpha[R_b(r_t + u) + \Gamma_e + r_n] + R_aR_b\Omega_\beta(r_t + u) \\ & + R_aR_b\dot{u} + R_aV_l\dot{\Psi} + R_aV_J\dot{\omega}\end{aligned} \tag{6}$$

where, V_l is the matrix describing the velocity of the wire rope; V_J is the matrix describing the pendant deflection; Ψ is the rope deformation; ω describes the pendant deflection angle in both directions.

The kinetic energy of the space suspension system is

$$T_e = \frac{1}{2}\left[\dot{\alpha}\ \dot{\beta}\ \dot{u}^{\mathrm{T}}\ \dot{\Psi}^{\mathrm{T}}\ \dot{\omega}^{\mathrm{T}}\right]\begin{bmatrix} M^e_{\alpha\alpha} & M^e_{\alpha\beta} & M^e_{\alpha u} & M^e_{\alpha\Psi} & M^e_{\alpha\omega} \\ M^e_{\beta\alpha} & M^e_{\beta\beta} & M^e_{\beta u} & M^e_{\beta\Psi} & M^e_{\beta\omega} \\ M^e_{u\alpha} & M^e_{u\beta} & M^e_{uu} & M^e_{u\Psi} & M^e_{u\omega} \\ M^e_{\Psi\alpha} & M^e_{\Psi\beta} & M^e_{\Psi u} & M^e_{\Psi\Psi} & M^e_{\Psi\omega} \\ M^e_{\omega\alpha} & M^e_{\omega\beta} & M^e_{\omega u} & M^e_{\omega\Psi} & M^e_{\omega\omega} \end{bmatrix}\begin{bmatrix} \dot{\alpha} \\ \dot{\beta} \\ \dot{u} \\ \dot{\Psi} \\ \dot{\omega} \end{bmatrix} \tag{7}$$

The dissipated energy, potential energy of the suspension system is

$$F_{ce} = \frac{1}{2}\dot{\omega}^{\mathrm{T}} C_{\omega}\dot{\omega} \tag{8}$$

$$U_e = m_e g\left[(l_t + u_x)\sin\beta + u_z\cos\beta - (l+v)\cos\tau_x\cos\tau_y\right] + \frac{1}{2}k_g v^2 \tag{9}$$

In the formula, C_w is the matrix of system damping coefficient, c_w is the damping coefficient of pendulum of suspended object; k_g is the stiffness coefficient of wire rope.

Crane Design Code GB/T 3811-2008 [8] provides that the wind load of the crane in working condition is

$$P_w = \frac{1}{2}C_w\rho_w v_w^2 A \tag{10}$$

where C_w is the wind force coefficient; ρ_w is the air density; v_w is the instantaneous wind speed; A is the crane's effective windward area.

Consider the air friction resistance of a single mass in a space suspension system

$$F_{\mu} = 2\mu_g M^e_{\omega\omega}\dot{\omega} \tag{11}$$

Taking the generalized coordinates $q_i = \left[\alpha(t)\ \beta(t)\ (u(t))^{\mathrm{T}}\ (\Psi(t))^{\mathrm{T}}\ (\omega(t))^{\mathrm{T}}\right]^{\mathrm{T}}$, Combining the above equations, the flexible body dynamics equations of the truss arm-lifting system can be obtained by using Lagrange's equations as

$$M\ddot{q}_i + C\dot{q}_i + Kq_i = Q_{out} + Q_v \tag{12}$$

where, M is the system mass coefficient matrix; C is the system damping coefficient matrix; K is the system stiffness coefficient matrix; Q_{out} is the system external load matrix; Q_v is the system inertia force.

3 Input Shaper

For a typical second-order system, when n input shaper pulses of amplitude A_i are applied at time t, the unit impulse response of the system is

$$y(t) = \sum_{i=1}^{n} \frac{A_i \omega_n}{\sqrt{1-\xi^2}} e^{-\xi\omega_n(t-t_i)} \sin\left(\omega_n\sqrt{1-\xi^2}(t-t_i)\right) \tag{13}$$

Without the input shaper, the response of the system t_i moment under the action of the unit pulse signal is

$$c(t) = \frac{\omega_n}{\sqrt{1-\xi^2}} e^{-\xi\omega_n(t-t_i)} \sin\left(\omega_n\sqrt{1-\xi^2}(t-t_i)\right) \tag{14}$$

The residual vibration ratio of the system is

$$V(\omega_n, \xi) = \frac{y(t)}{c(t)} \tag{15}$$

$V(\omega_n, \xi) = 0$ It is the residual vibration constraint equation of the system, and a single residual vibration constraint equation is not enough to complete the design of the input shaper, which needs to be combined with the pulse amplitude constraint equation, the time-optimal constraint equation, and the robustness constraint equation to jointly design the input shaper that meets the requirements.

The pulse amplitude and time lag of the single mode ZVD shaper are

$$\begin{bmatrix} A_i \\ t_i \end{bmatrix} = \begin{bmatrix} \frac{1}{1+2k+k^2} & \frac{2k}{1+2k+k^2} & \frac{k^2}{1+2k+k^2} \\ 0 & \frac{T}{2} & T \end{bmatrix} \quad (i = 1, 2, 3) \tag{16}$$

In the formula, $k = \mathrm{e}^{-\frac{\xi\pi}{\sqrt{1-\xi^2}}}$, $T = \frac{2\pi}{\omega_n\sqrt{1-\xi^2}}$.

The dual-mode input shaper is obtained by the convolution operation of two single-mode input shapers [9].

$$\begin{bmatrix} A_i \\ t_i \end{bmatrix} = \begin{bmatrix} A_i^{\omega_1} \\ t_i^{\omega_1} \end{bmatrix} * \begin{bmatrix} A_i^{\omega_2} \\ t_i^{w_2} \end{bmatrix} = \begin{bmatrix} A_i^{\omega_1} * A_i^{\omega_2} \\ t_i^{\omega_1} + t_i^{\omega_2} \end{bmatrix} \tag{17}$$

4 Calculated Example Analysis

4.1 Rigid-Flexible Coupled Dynamics Model Parameters

The target wind power crane truss boom length of 130 m, the analysis of the initial pitch angle of 75 degrees through the rotational moment of inertia equivalent, deformation equivalent and free attenuation vibration method can be determined by the system parameters, as shown in Table 1.

Table 1 Simulation results (average values)

$k_x/\mathrm{N}\cdot\mathrm{m}^{-1}$	$k_y/\mathrm{N}\cdot\mathrm{m}^{-1}$	$k_z/\mathrm{N}\cdot\mathrm{m}^{-1}$	m_q/kg	ξ_q	ω_{nq}/Hz	ξ_e	ω_{ne}/Hz
7.33×10^6	3.57×10^5	1.4×10^5	7.27×10^5	0.08	0.873	0.02	0.06

This paper uses an average wind speed of 9.35 m/s at grade five [10].

4.2 Simulation Platform Construction

The simulation platform is built without considering the coupling influence of the driving signals, and the velocity is directly used as the initial input signal of the system. At this time, Eq. (12) is degraded to a six-degree-of-freedom system, and only the three-way elastic vibration of the truss arm, the two-way deflection angle of the suspended object and the elastic deformation of the wire rope are considered. The driving velocity signal in the solution process is simulated using the cosine function, as shown in Fig. 2.

4.3 Analysis of Anti-shaking Control

In the actual wind power crane lifting operations, for safety reasons there is generally only a single condition operation, this paper takes the most complex slewing condition as an example to study the input shaping technology on the wind power crane lattice boom—lifting system anti-shaking control effect. At the time of analysis, the speed of the slewing drive signal in the uniform phase is $v_\alpha = \pi/100$ rad/s, the wire rop has a length of $l = 50$ m, and the mass of the lifted object is $m_e = 50t$.

From the curve of Fig. 3, it can be seen that ZVD_2 shaper has better suppression effect on the y-direction and z-direction vibration of the truss arm, and the y-direction vibration has the most obvious suppression effect, and the largest value of the y-direction vibration of the boom under the driving of the original signal is 758.44 mm, and the largest value of the y-direction vibration after the shaping of ZVD_2 is 157.33 mm. From the curve of Fig. 4, it can be seen that the ZVD_2 shaper has better suppression effect on the swing angle of both the lifted variable plane and the slewing plane. The maximum positive and negative swing angles are 4.135° and 3.470°, and the maximum positive and negative amplitude plane swing angles are 1.105° and 1.041°. After ZVD_2 shaping, the maximum positive and negative swing angles are 0.047° and 0.015°, and the maximum positive and negative amplitude plane swing angles are 1.041°. The maximum positive and negative swing angles are 0.047° and 0.015°, the maximum positive and negative luffing plane swing angles are 0.014° and 0.030°, and the maximum swing angles of the lifted object's slewing plane and luffing plane are 1.14% and 2.71% of the unshaped ones, respectively; it

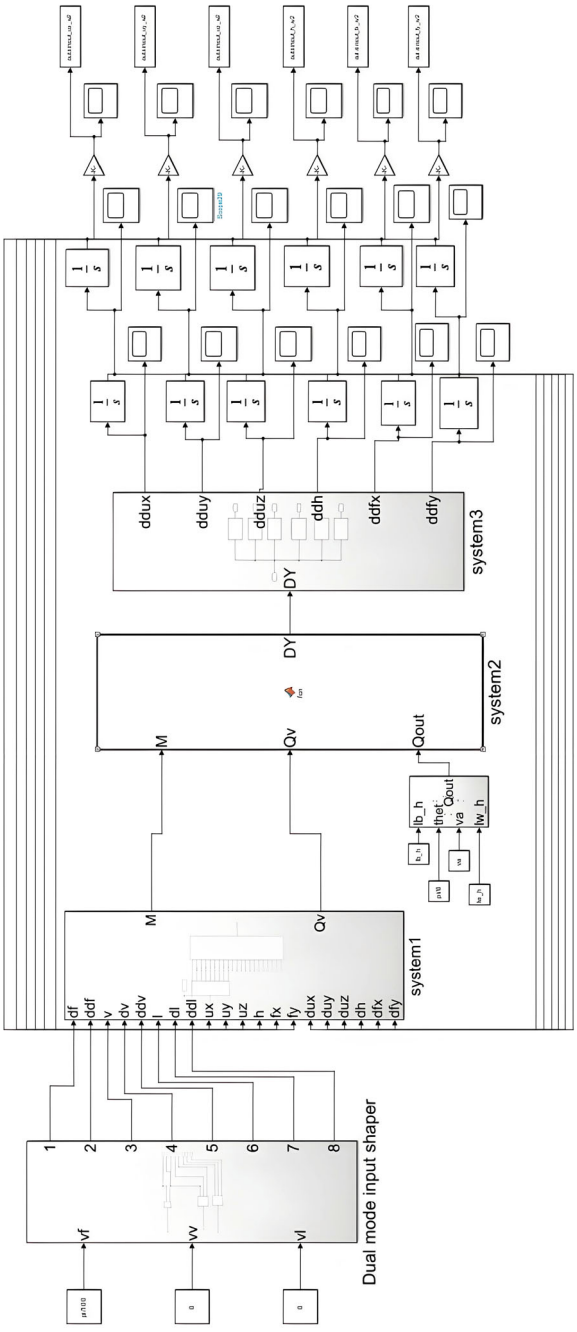


Fig. 2 Anti-sway control simulation platform

can be seen that the ZVD_2 shaper based on the modes of the truss arm and the lifted object's luffing modes has a better anti-rocking control effect on the truss arm-lifting system of the wind turbine cranes.

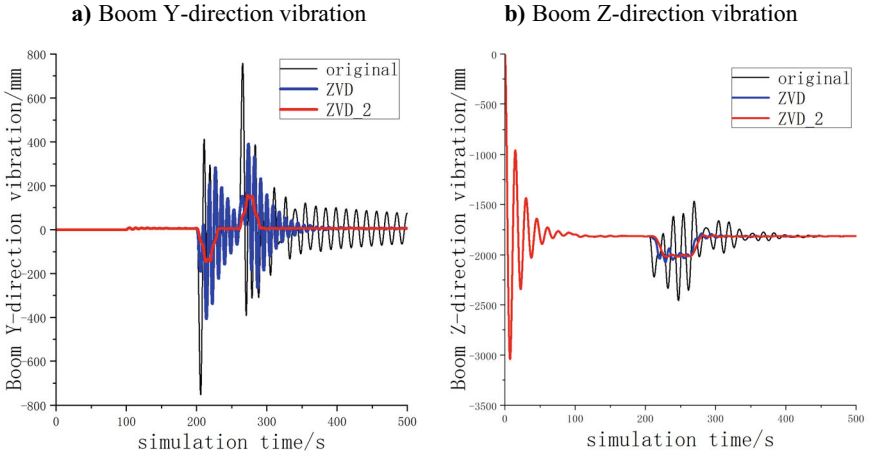


Fig. 3 Vibration response of the arm frame before and after shaping

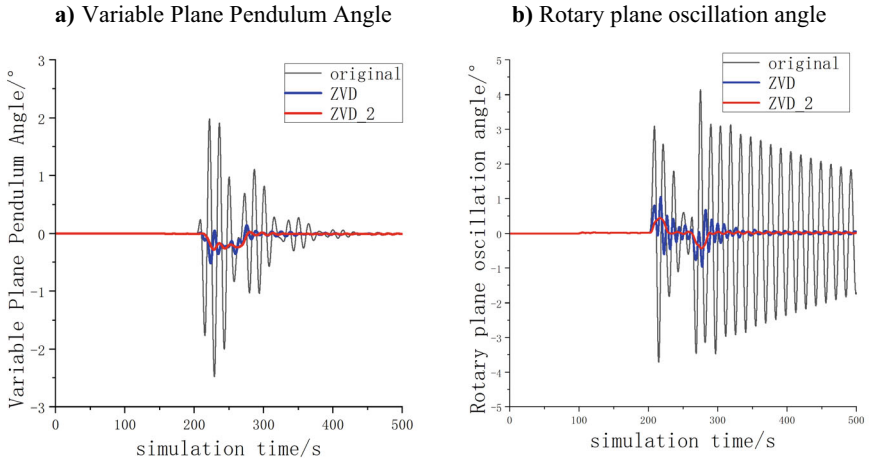


Fig. 4 Angle of deflection of the suspended object before and after straightening

5 Conclusion

This paper proposes a rigid-flexible coupling dynamics model for the truss boom-suspension system, using an equivalent spring-mass-damping system, a space suspension system, elastic two-force bars, and a transient wind model. These elements are employed to simulate the elastic vibration of the truss boom, the swing of the suspended load, the rope's elasticity, and the wind loads acting on the system. Based on the vibration characteristics, the ZVD and ZVD_2 input shapers are designed. The system's mathematical model is integrated into a SIMULINK-based anti-sway control simulation platform. Due to the coupling effects between the truss boom's elastic vibration and the suspended object's spatial deflection, the ZVD_2 input shaper, which takes into account both the boom's vibration modes and the suspended object's deflection modes, provides superior anti-rocking performance for the wind turbine crane.

Acknowledgements This work is supported by National Natural Science Foundation of China under Grant No.52275088 and the Fundamental Research Funds for the Central Universities under Grant (No. DUT22LAB506, No. DUT22LAB507).

References

1. Yan S, Peng J, Liu Z et al (2018) Dynamic model of rigid-flexible coupling for slewing lifting of large flexible lifting arm system. Vib Shock 37(05):175–179+201
2. Park K, Cha J, Lee K (2011) Dynamic factor analysis considering elastic boom effects in heavy lifting operations. Ocean Eng 38(10):1100–1113
3. Mohammed A, Altuwais H, Alghanim K (2023) An optimized shaped command of overhead crane nonlinear system for rest-to-rest maneuver. J Eng Res 11(4):548–554
4. Abdullah M, Khalid A, Masood AT (2020) An adjustable zero vibration input shaping control scheme for overhead crane systems. Shock Vibr
5. Kawada K, Yamamoto T, Mada Y (2005) Gain scheduled PD sway control of a lifted load for a mobile crane. Control Intell Syst 33(1):48–54
6. Yan S, Peng J, Wang S et al (2020) Active control of vibration in rotary lifting multibody systems taking into account rigid-flexible coupling effect. J Appl Mech 37(06):2455–2462+2698–2699
7. Cao X, Xu G, Hu Y et al (2023) Rigid-flexible coupling dynamics analysis of boom-hoisting system of wind power crane. In: International conference on advances in construction machinery and vehicle engineering. Springer, Singapore, pp 3–20
8. National standard GB/T3811-2008 crane design code. Beijing, Beijing Lifting and Transportation Machinery Design and Research Institute, 2012-01-01
9. Guo Z, Zhang J, Zhang P (2023) Research on the residual vibration suppression of delta robots based on the dual-modal input shaping method. Actuators 12(2):84
10. GB/T 14560-2022 (2022) Crawler cranes. China Standard Press, Beijing

Open Access This chapter is licensed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License (http://creativecommons.org/licenses/by-nc-nd/4.0/), which permits any noncommercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if you modified the licensed material. You do not have permission under this license to share adapted material derived from this chapter or parts of it.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.



Active Suspension Anti-rollover Control Strategy for Wind Turbine Blade Mountain Transport Vehicles



Yiying Wang, Zhong Wang, Yuchen Wang, Mingyang Wang, and Yuhao Jiang

Abstract Large wind turbine blades can be transported using a specialized truck called the Wind Turbine Blade Mountain Transporter (WTBMT). However, during transit, the vehicle's body and surrounding conditions make it more vulnerable to rollover issues brought on by impacts from the road. The goal of this research project is to overcome the rollover issue by modifying the Active Suspension System's (ASS) control force. It is based on the repetitive nature of the WTBMT uphill tooling operation. In this work, we establish 2-DOF and 4-DOF ASSs, and based on them, we suggest a feedback-feedforward control technique that combines LQR control and iterative learning (ILC) control to avoid sudden changes in the sprung mass displacement. Ultimately, the simulation analysis demonstrates that the suggested approach not only solves the jitter issue and makes the ASS learnable, demonstrating its efficacy and producing superior outcomes.

Keywords Iterative learning control · Active suspension systems · LQR control · Rollover prevention

1 Introduction

The wind farm approach road has significant longitudinal and transverse slopes, a low road surface grade, and unstable conditions. In addition, the wind turbine blades are large and heavy [1], the transportation process's center of mass is high and subject to wind and other external environmental factors, etc., and rollover accidents involving WTBMTs occur occasionally. As a result, improving the WTBMT's rollover stability has become a hot topic for both domestic and international research. Much work has been done on rollover prevention, such as differential braking anti-rollover control system [2], anti-roll bar anti-rollover control system [3], active steering anti-rollover control system [4, 5], and active/semi-active suspension anti-rollover control system

Y. Wang (✉) · Z. Wang · Y. Wang · M. Wang · Y. Jiang
Yanshan University, Qinhuangdao, Hebei, China
e-mail: yiyingwang@stumail.ysu.edu.cn

© The Author(s) 2026

H. Saman et al. (eds.), *The 9th International Conference on Advances in Construction Machinery and Vehicle Engineering*, Lecture Notes in Mechanical Engineering,
https://doi.org/10.1007/978-981-96-5527-4_92

[6]. In order to restrict the rollover motion while the vehicle is steered, Nguyen devised a fuzzy control approach in [7] that was applied to a stabilizer bar model based on the stability and responsiveness of the system. However, the effectiveness of the control model and algorithm has not been thoroughly studied. To increase straight-line acceleration and cornering stability, Feng et al. [8] suggested active roll control, four-wheel independent braking, and coordinated front and rear traction control systems. While this integrated chassis control [8] guarantees handling stability during cornering, it is not easily transferable to heavy-duty, multi-axle vehicle studies. Since vehicle rollover typically occurs within seconds of sudden steering or encountering obstacles, vertical control of actuators is considered a direct solution to address rollover issues. Many researchers have selected the active suspension anti-rollover control as a control mechanism that provides direct vertical forces, to various vehicle locations [3, 9].

During WTBMT transport procedure, each tire's suspension height must be manually adjusted in order to bring the semi-trailer's overall level. This includes checking to see if the WTBMT complete truck is horizontal and whether any of the tires have a noticeable off-ground overhang. Furthermore, following the wind farm's location, the WTBMTs must be transported up the mountain multiple times [10] and typically must be operated three times to create a generator with a single rotor and three blades. But according to earlier research, there are no findings about how to increase the rollover stability of the WTBMT by designing controllers with iterative learning algorithms and taking into account the properties of the active suspension system. In order to accomplish active suspension rollover control of WTBMT under unstructured road circumstances, this study suggests a new LQR feedback-feedforward ILC-based control technique based on the suspension model for the particular operational scenario of WTBMT.

2 ASS Physics and Mathematical Modeling

A 2-DOF ASS is typically utilized in control studies of suspension systems (see Fig. 1a). The differential equation of motion for the 2-DOF ASS can be expressed as follows using Newton's second law:

$$\begin{cases} m_s\ddot{z}_s + c_s(\dot{z}_s - \dot{z}_u) + k_s(z_s - z_u) + F_a = 0 \\ m_u\ddot{z}_u - c_s(\dot{z}_s - \dot{z}_u) - k_s(z_s - z_u) + k_t(z_u - z_r) - F_a = 0 \end{cases} \tag{1}$$

Table 1 displays the letters' meanings. In order to convert Eq. (1) into a state space model Eq. (2), define the state variable $\mathrm{X} = [x_1, x_2, x_3, x_4]^{\mathrm{T}}$, where $x_1 = z_s - z_u$, $x_2 = \dot{z}_s$, $x_3, = z_u - z_r$, and $x_4 = \dot{z}_u$ represent the suspension travel, sprung mass velocity, wheel deflection, and unsprung mass velocity, respectively.

$$\dot{X} = AX + BU \tag{2}$$

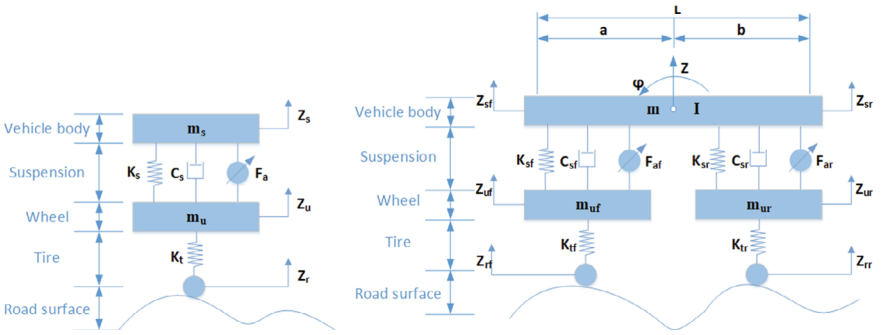


Fig. 1 Active suspension physical model

where, $A = \begin{bmatrix} 0 & 1 & 0 & -1 \\ -k_s/m_s & -c_s/m_s & 0 & c_s/m_s \\ 0 & 0 & 0 & 1 \\ k_s/m_u & c_s/m_u & k_t/m_u & -c_s/m_u \end{bmatrix}$; $B = \begin{bmatrix} 0 & 0 \\ 0 & 1/m_s \\ -1 & 0 \\ 0 & -1/m_u \end{bmatrix}$; $U = \begin{bmatrix} \dot{z}_r \\ F_a \end{bmatrix}$.

A twelve-degree-of-freedom vehicle dynamics model should be developed for the five-axis WTBMT of Shandong Shiyun Co. The vehicle dynamics model can be simplified to a 4-DOF ASS (see Fig. 1b) for the study of lateral tilt stability in this paper. This allows for the response of the 4-DOF ASS to the vibration characteristics of the vehicle's vertical and lateral tilts, as well as the balance of forces in the vertical direction of the wheels. Based on Newton's law of dynamics, the differential equations for the motion of the 4-DOF ASS can be established as follows:

Table 1 Model parameters of the active suspension system

Symbols	Description	Values	Units
m_s	Vehicle body mass	234	kg
m_u	Vehicle wheel mass	40	kg
k_s	Suspension spring stiffness	104	kN/m
c_s	Suspension damping coefficient	6160	N/(s · m^{-1})
k_t	Wheel stiffness	40	kN/m
m	Vehicle body mass	690	kg
I	Moment of inertia (mechanics)	1222	Kg · m^2
$m_{uf} = m_{ur}$	Vehicle wheel mass	40.5	kg
$k_{sf} = k_{sr}$	Left and right suspension spring stiffness	17	kN/m
$c_{sf} = kc_{sr}$	Damping coefficient of left and right suspension	1500	N/(s · m^{-1})
$k_{tf} = k_{tr}$	Left and right wheel stiffness	192	kN/m
a = b	Body center of mass value left–right axis distance	1	m

$$\begin{cases} m\ddot{z} + c_{sf}\left(\dot{z}_{sf} - \dot{z}_{uf}\right) + k_{sf}\left(z_{sf} - z_{uf}\right) + c_{sr}(\dot{z}_{sr} - \dot{z}_{ur}) + k_{sr}(z_{sr} - z_{ur}) + F_{af} + F_{ar} = 0 \\ I\ddot{\psi} - ac_{sf}\left(\dot{z}_{sf} - \dot{z}_{uf}\right) - ak_{sf}\left(z_{sf} - z_{uf}\right) + bc_{sr}(\dot{z}_{sr} - \dot{z}_{ur}) + bk_{sr}(z_{sr} - z_{ur}) - aF_{af} + bF_{ar} = 0 \\ m_{uf}\ddot{z}_{uf} - c_{sf}\left(\dot{z}_{sf} - \dot{z}_{uf}\right) - k_{sf}\left(z_{sf} - z_{uf}\right) + k_{tf}\left(z_{uf} - z_{rf}\right) = 0 \\ m_{ur}\ddot{z}_{ur} - c_{sr}(\dot{z}_{sr} - \dot{z}_{ur}) - k_{sr}(z_{sr} - z_{ur}) + k_{tr}(z_{ur} - z_{rr}) = 0 \end{cases} \tag{3}$$

Table 1 displays the letters' meanings. Define the state vector $X = \left[z_{sf}, z_{uf}, z_{sr}, z_{ur}, \dot{z}_{sf}, \dot{z}_{uf}, \dot{z}_{sr}, \dot{z}_{ur}\right]$ and apply the following steps to convert Eq. (3) into a state space model:

$$\dot{X} = AX + BU + \Gamma W \tag{4}$$

Let $k_1 = I + a^2m = I + b^2m$, $k_2 = I - abm$, $k_3 = Im$, we have $A = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}$; $B = \begin{bmatrix} B_1 \\ B_2 \end{bmatrix}$; $U = \begin{bmatrix} F_{af} \\ F_{aR} \end{bmatrix}$; $\Gamma = \begin{bmatrix} \Gamma_1 \\ \Gamma_2 \end{bmatrix}$; $W = \begin{bmatrix} z_{rf} \\ z_{rr} \end{bmatrix}$.

3 Controller Design

3.1 Road Excitation Input

One of the suspension system's inputs, road surface excitation (see Fig. 2), which can have a major influence on the vehicle very quickly. The real operation of the WTBMT's serves as the basis for the definition of the periodic trapezoidal impact road surface excitation model, as follows:

$$z_r(t) = \begin{cases} 0, & 0 \ll t < 3 \\ \frac{h}{3}(t-3), & 3 \ll t < 6 \\ h, & 6 \ll t < 9 \\ \frac{h}{3}(t-9), & 9 \ll t < 12 \end{cases} \tag{5}$$

3.2 LQR-Based Feedback-Feedforward ILC Controller

The configuration of the ILC control to reliably lower the body roll angle during recurrent uphill operations of a WTBMT is covered in this section. In order to attain extremely accurate tracking of the intended trajectory, the ILC control enhances the control inputs by correcting for faults. The iterative control expression of PD type is:

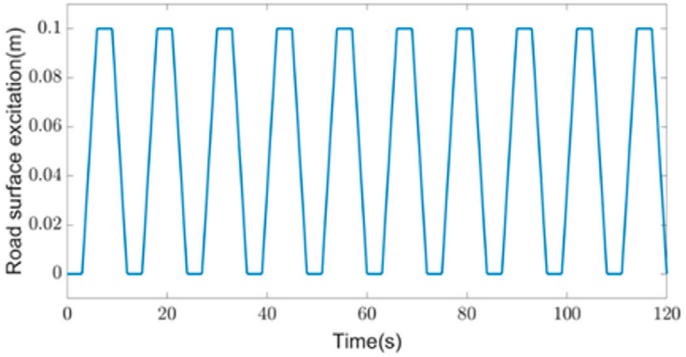


Fig. 2 Periodic trapezoidal impact pavement excitation modeling

$$U_{k+1}(t) = U_k(t) + K_p e_k(t) + K_d[e_k(t) - e_k(t-1)] \tag{6}$$

where $e_k(t)$ is the tracking error of the kth iteration and K_p and K_d are the proportional and differential coefficients, respectively. The controller iteratively learns that its desired trajectory is $y_d = z_{sd} = 0$, its output is $U_{k+1}(t) = F_A$.The asymmetry of the system control effects in the hydraulic actuator of the suspension system, and K_P, K_d will typically be adjusted larger to achieve a higher convergence speed, which raises the system control error even more.While LQR considers the system's output at each iteration $z_{s_k}(t)$ to account for oscillations in the ILC algorithm. We introduce the performance metric J:

$$\mathrm{J} = \frac{1}{2}\int_0^t \left(x^{\mathrm{T}} Q x + u^{\mathrm{T}} R u\right)\mathrm{d}t \tag{7}$$

where the weight matrices are denoted by Q and R. The feedback regulator can be obtained as:

$$\mathrm{u} = -\mathrm{Kx} = -\left(R^{-1}B^{\mathrm{T}}P\right)x \tag{8}$$

LQR control and ILC control are combined in the feed-forward-feedback control strategy. While LQR ensures the system's flexibility to outside disturbances and system jitter, ILC ensures the system's tracking performance at each iteration. Below is the control rule for the proposed LQR-based feed-back-feedforward ILC controller(LQR_ILC controller):

$$\begin{cases} U_{k+1}(t) = U^f_{k+1}(t) + U^b_{k+1}(t) \\ U^f_{k+1}(t) = U^f_k(t) + K_p e_k(t) + K_d[e_k(t) - e_k(t-1)] \\ U^b_{k+1}(t) = -K x_{k+1} \end{cases} \tag{9}$$

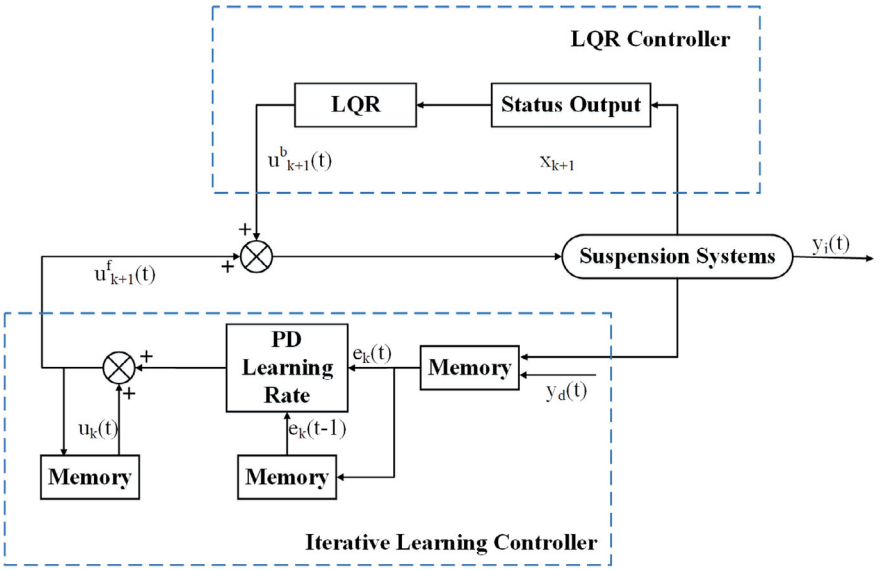


Fig. 3 Proposed LQR-based feedback-feedforward ILC control block diagram

In summary, the block diagram of the LQR_ILC control logic proposed in this paper is shown in Fig. 3.

4 Simulation Result Analysis

This section includes a simulation examination of the MATLAB/Simulink designed controller to confirm the efficacy of the proposed LQR_ILC controller. Table 1 displays the 2-DOF and 4-DOF ASS characteristics.

Passive suspension (no control), PID control, LQR control, ILC control and LQR_ILC control are applied and analyzed comparatively.

(1) C1: Passive suspension system (control force input from suspension actuator absent)

(2) C2: The controller gain $\left(K_p = 100,\ K_i = 800,\ K_d = 0\right)$ has been chosen by the ASS using PID control. The controller is configured to:

$$\mathrm{u}(\mathrm{t}) = K_p e(t) + K_i \int e(t)\mathrm{d}t + K_d \frac{\mathrm{d}}{\mathrm{d}t} e(t) \tag{10}$$

(3) C3: The weighting matrix for the ASS with LQR control has been chosen (Q = diag{$1 * 1e10$, $1 * 1e8$, 1, 1}, R = 1), and the controller is represented by the Eq. (8).

(4) C4: Eq. (6) represents the controller for the ASS utilizing ILC control; specifically, $K_p = 12000$, $K_d = 0$.
(5) C5: ASS with LQR_ILC control, the controller is denoted as Eq. (9).

4.1 Comparative 2-DOF ASS Simulation Results

The simulation results of the ASS actuator under various control strategies are displayed in Figs. 4, 5, and 6, and a periodic trapezoidal impact pavement excitation mod-el is employed in this subsection (see Fig. 2).

The displacement of the spring-loaded mass is displayed in Fig. 4. Repeated inputs had no effect on the spring-loaded mass displacements of the unlearned passive, PID, and LQR controls; the passive control had the biggest displacement fluctuations at

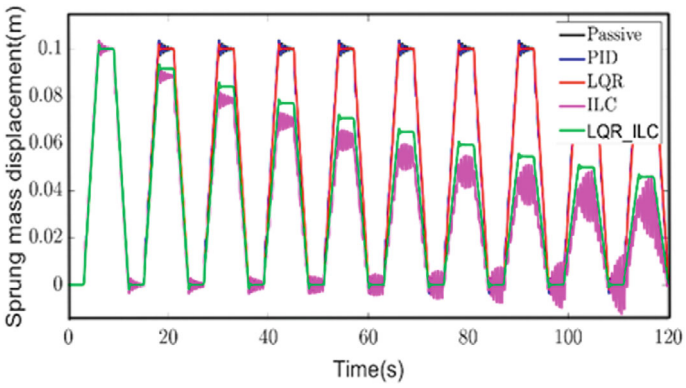


Fig. 4 Sprung mass displacement

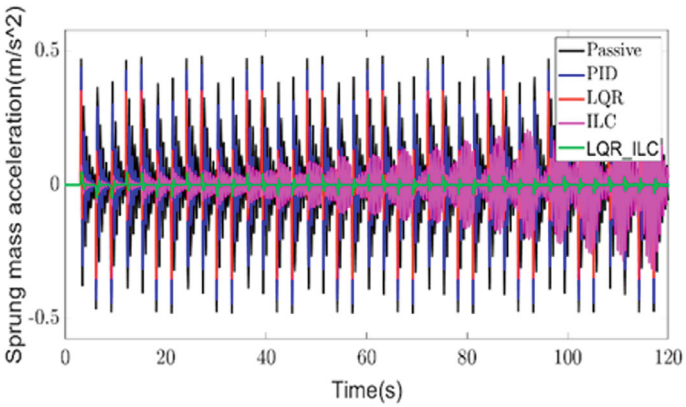


Fig. 5 Sprung mass acceleration

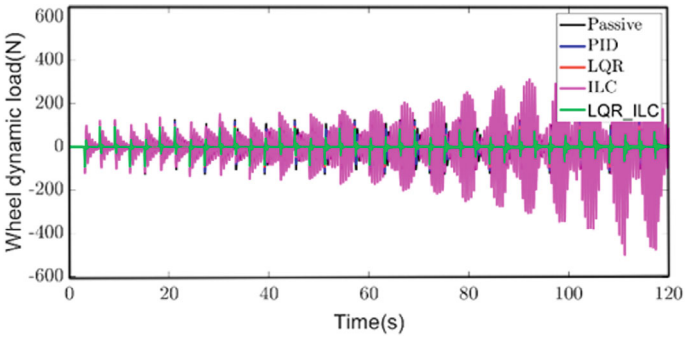


Fig. 6 Wheel dynamic load

each shock encounter, followed by the PID control, and the LQR control had the smallest displacement fluctuations. After a few repetitions, the ILC control alone exhibits significant output fluctuations despite being trained for the input pavement excitation. The proposed LQR_ILC control has a better control effect on the repetitive inputs, with the increase of the number of iterations of the spring loaded mass displacement gradually decreases, the convergence speed although compared to the unique ILC control is reduced, but in the tenth iteration, that is, in the third operation of the wind turbine blade mountain transshipment fleet, the lateral inclination angle has been reduced by more than 50%, and the suppression of fluctuation effect is significant, in general, the control effect is the best.

Figures 5 and 6 illustrate the acceleration of the sprung mass and the dynamic wheel load, respectively. Once again, the LQR_ILC control produces the optimal control outcomes. In conclusion, all of the performance index curves exhibit improved control and accurate response, demonstrating that the LQR_ILC control outperforms other techniques in suppressing system jitter, learning from repeated inputs, and converges more quickly.

4.2 Comparative 4-DOF ASS Simulation Results

The periodic trapezoidal impact pavement excitation Eq. (5), where, is the right wheel pavement excitation input of the 4-DOF ASS. The excitation input for the left wheel surface is 0. Figure 7 displays the ASS actuator's simulation results under the passive, LQR_ILC control approach.

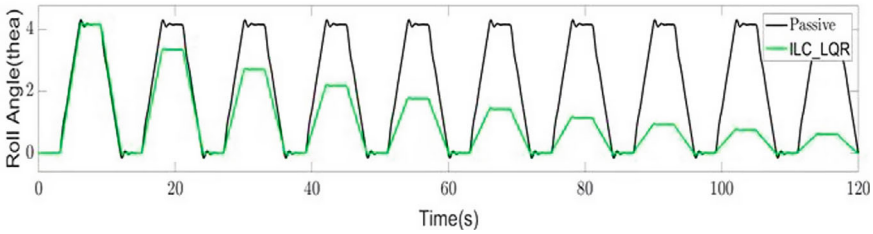


Fig. 7 Vehicle body tilt angle

5 Conclusion

For the actual driving situation of a wind turbine blade mountain transporter, a feedback-feedforward ILC control method based on LQR is proposed in this paper. LQR control and ILC control are modularized and designed to form a feedback-feedforward controller, which is used to continuously reduce the displacement change of the spring-loaded mass and subsequently improve the lateral stability of the vehicle. whereby the entire vehicle dynamics model is condensed into a 4-DOF ASS in order to ease the controller design process. Furthermore, LQR control is implemented based on ILC control to efficiently suppress system jitter and enhance control precision. Feedback and feedforward can operate separately and mutually reinforce one another. The efficiency of the control method suggested in this research is confirmed by establishing the ASS and analyzing the control impacts of various controllers through simulation comparison. The following effort is to further test the control performance of the control strategy under difficult situations by introducing the dynamics of tractors and wind turbine blades.

Appendix

Regarding parameters listed in the text that aren't clearly valued or specified, please refer to the Baidu network disk link:

https://pan.baidu.com/s/1Sfx3jo72qvHxrT5Fh8VzyQ

Extraction code: 7fvd.

References

1. Macioszek E (2020) Oversize cargo transport in road transport—problems and issues. SJSUTST 108:133–140
2. Feng J, Chen S, Qi Z (2020) Coordinated chassis control of 4WD vehicles utilizing differential braking, traction distribution and active front steering. IEEE Access 8:81055–81068

3. Nguyen TA (2021) Improving the stability of the passenger vehicle by using an active stabilizer bar controlled by the fuzzy method. Complexity 20. 6569298
4. Chen C, Zhang J, Zheng H (2024) A modified handling stability control strategy for steer-by-wire vehicles based on variable steering ratio and active front steering control. Proc Inst Mech Eng D J Automobile Eng 0(0)
5. Liang W, Ahmac E, Khan MA et al (2021) Integration of active tilting control and full-wheel steering control system on vehicle lateral performance. Int J Automot Technol 22:979–992
6. Konieczny J, Sibielak M, Rączka W (2020) Active vehicle suspension with anti-roll system based on advanced sliding mode controller. Energies 13(21):5560
7. Nguyen TA (2024) Establishing a novel adaptive fuzzy control algorithm for an active stabilizer bar with complex automotive dynamics model. Ain Shams Eng J 15(1):102334
8. Feng J, Gao Z, Guo B, Du C (2024) Comprehensive chassis control strategy based on coordination of traction/braking distribution and active roll control. Electronics 13(14):2731
9. Qingshuo H, Yang L, Sha Y, Zhuopei L, Xing Z, Chenyu Z (2024) Research on anti-rollover intelligent suspension control based on dynamic bayesian network. In: 2024 IEEE 19th conference on industrial electronics and applications (ICIEA), Kristiansand, Norway, pp 1–6
10. Breuning L, Isaza AC, Gawlick J, Kerekeš A, Hamacher T (2024) Combined photovoltaic and wind power plant planning for the production and transportation of liquefied green hydrogen: a case study of Egypt. Int J Hydrogen Energy 49:150–165. ISSN: 0360-3199

Open Access This chapter is licensed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License (http://creativecommons.org/licenses/by-nc-nd/4.0/), which permits any noncommercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if you modified the licensed material. You do not have permission under this license to share adapted material derived from this chapter or parts of it.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.



Automatic Calibration and Variable Scale Segmentation Algorithm of Driveless Dumping Truck Based on Mining Area



Boxua Zhang, Zeyu Liu, Chenhui Li, DaWei Li, and Dawei Jiang

Abstract At present, the ground of open-pit mines is rugged and bumpy, often accompanied by more dust and soil slopes. In order to obtain more front view of the driveless dumping truck perception system, the installation angle of the lidar needs to be changed during the tooling process. In this paper, an algorithm for automatic calibration and variable scale segmentation of lidar is proposed. First, the driveless dumping truck is parked on a wide ground to ensure that the road ahead and the current position of the driveless dumping truck are at the same level. The current pitch angle and roll angle of the lidar are obtained by analyzing the characteristic normal vector in a square area 10 m in front of the trucl. A symmetrical object is placed directly in front of the lidar, and its centroid deviation is calculated to obtain the yaw angle of the lidar. After the corrected point cloud is obtained, the point cloud data is segmented with variable scale. The point cloud is divided into fine-grained grid and coarse-grained grid, and ground and non-ground point clouds are obtained through granular transformation, plane fitting, and neighborhood smoothing. The scheme proposed in this paper can effectively solve the situation of poor ground environment, strong turbulence, and frequent occurrence of soil slopes and smoke in the mining area, so as to ensure the safe operation of driveless dumping truck in the mining area.

Keywords Unmanned technology · Driveless dumping truck · Variable scale · Granular transformation · Mining area

B. Zhang (✉) · Z. Liu · C. Li · D. Li · D. Jiang
Beijing Institute of Mechanical Equipment, Beijing, China
e-mail: zhangboxuanFred@163.com

© The Author(s) 2026

H. Saman et al. (eds.), *The 9th International Conference on Advances in Construction Machinery and Vehicle Engineering*, Lecture Notes in Mechanical Engineering,
https://doi.org/10.1007/978-981-96-5527-4_93

1 Introduction

1.1 Autonomous Haulage System

The driveless dumping truck uses the lidar for the obstacle detection task of environment perception [1], it needs to consider the hardware performance, development cycle, data cost and other issues. Traditional point cloud processing algorithms can be divided into calibration algorithms, segmentation algorithms, and detection and tracking algorithms. This paper focuses on discussing the first two steps.

1.2 Lidar Calibration

Lidar calibration can be divided into two methods: manual calibration and self-calibration. The user needs to correct the six degrees of freedom of the lidar. The translation calibration usually uses the manual calibration method, the distance is relatively certain, and the calibration result is accurate. The angle calibration is the focus of the calibration work, and the angle error will directly affect the subsequent detection effect. The angle error is divided into three directions: pitch, roll, and yaw. These three angles are tough to be measured directly by tools, and have a great influence on the accuracy of the final output result. Therefore, a lidar angle calibration is introduced here. Compared with other methods, this method has lower requirements on the calibration environment and tools, and is easy to implement.

1.3 Ground Segmentation

The commonly used ground segmentation algorithms include plane grid method, plane fitting method and ray method. The grid method usually establishes a plane grid (or multi-layer grid or three-dimensional voxel) according to the settled size, then projects the original point cloud into their respective grids, and extracts features from the point set in each grid, such as average height, maximum height, height difference, density, etc. [2]. Its problems: when there are few lidar harnesses, such as the road data collected by the 16 line lidar in the following figure, On the ground 20 m away from the front of the vehicle, there are few laser points that can be respond, and there is usually only one laser harness hitting the obstacles. If the height feature is used in the grid for ground filtering, low obstacles will be easily filtered out as ground points. The plane fitting algorithm [3] is to randomly select 3 points each time to determine a plane equation ($ax + by + cz = d$) within the set iteratio; Bring all points into the plane equation in turn, and make judgment according to the set distance threshold. If it is within the threshold range, it is considered to be the inner point of the plane, otherwise it is the outer point. Count the number of inner points after traversing

all points; The plane equation with the largest number of interior points within the iteration is the ground equation. The interior points under the ground equation are the ground point clusters, and the exterior points are the obstacle point clusters. The problem of this method is that the model-based method is used to segment the ground on an ideal road surface, and the effect is acceptable. However, when the ground is not an absolute plane, the ground equation calculated by [4] is also the same, When the front ground point set is taken as the obstacle point, the ray method reduces the three-dimensional space of the point cloud (x, y, z) to the two-dimensional grid of (D, z) [4–6]. At the same time, different sector point cloud division strategies are adopted to divide the grid segments and bin, and the local fitting line is used to judge whether it is a ground point.

2 Methodology

2.1 Lidar Calibration

The so-called lidar calibration is to calculate the difference between the lidar's own coordinate system and the coordinate system we need (such as the body coordinate system), so as to obtain the lidar detection results in the vehicle coordinate system, which is convenient for subsequent perception calculation. This error value can be divided into rotation angle error and translation error. The translation error can be measured by laser rangefinder and other measuring tools, and has little impact on the final output accuracy. The important thing is the angle error, which can be divided into three directions: pitch, roll and yaw. These three angles are not convenient to be measured directly by tools, and have a great impact on the accuracy of the final output results. Therefore, a lidar angle calibration method is introduced here. Compared with other methods, this method has lower requirements on the calibration environment and tools, and is easy to implement. Generally speaking, single lidar calibration is divided into manual calibration and automatic calibration.

The main steps for manual calibration of a single lidar are:

(1) find a place with large and flat ground and place a regular obstacle in front of the lidar;
(2) convert the origin point cloud and display the calibrated point cloud;
(3) manually adjust pitch, yaw and roll angles and x Offset, y Offset and other parameters by observing the calibrated point cloud.

Main steps for automatic calibration of a single lidar:

(1) find a place with large and flat ground and place a regular obstacle in front of the lidar;
(2) determine the approximate coordinates of the flat ground (within what range is the horizontal and vertical coordinates of the flat ground);
(3) calculate the normal vector of the range;

(4) using the normal vector to determine the pitch angle PTICH and roll angle (atan) of the lidar;
(5) determine the approximate position of the obstacle, and determine the yaw angle of the lidar through the atan values of the X and Y coordinates of the obstacle;
(6) x_Offset and y_Offset and height manual calibration.

Singular value decomposition (SVD) is used to solve the lidar calibration in this paper. It is known that the lidar data is a three-dimensional point cloud, and a large number of ground points fit the plane equation:

$$ax + by + cz = d \tag{1}$$

In order to conveniently obtain the distance from the plane to the origin, it is necessary to add constraints:

$$a^2 + b^2 + c^2 = 1 \tag{2}$$

The center of gravity coordinates of the three-dimensional point cloud are obtained, and it is obvious that the plane must pass through the center of gravity coordinates, that is:

$$a\overline{x} + b\overline{y} + c\overline{z} = d \tag{3}$$

Subtract (3) from (2), we get,

$$a(x_i - \overline{x}) + b(y_i - \overline{y}) + c(z_i - \overline{z}) = 0 \tag{4}$$

Written in matrix form:

$$\begin{bmatrix} x_1 - \overline{x}\ y_1 - \overline{y}\ z_1 - \overline{z} \\ x_2 - \overline{x}\ y_2 - \overline{y}\ z_2 - \overline{z} \\ x_3 - \overline{x}\ y_3 - \overline{y}\ z_3 - \overline{z} \end{bmatrix} \begin{bmatrix} a \\ b \\ c \end{bmatrix} = 0 \tag{5}$$

The covariance matrix is:

$$C = \sum_{i=1:|s|} (\mathrm{s_i} - \hat{s})(\mathrm{s_i} - \hat{s})^T \tag{6}$$

where $\mathrm{S_i}$ represents the three-dimensional coordinates of the point cloud and is the mean value of all points. Let matrix A be C.

Set column vector X:

$$\mathbf{X} = \begin{bmatrix} a \\ b \\ c \end{bmatrix} \tag{7}$$

Get:

$$\mathbf{AX} = 0 \tag{8}$$

When fitting, the lidar point cloud is always not on the plane, and the sum of the minimum distances from the point cloud to the plane is required to be as small as possible, so the fitting objective function:

$$\min \|\mathbf{AX}\| \tag{9}$$

When:

$$\|\mathbf{X}\| = 1 \tag{10}$$

Singular value decomposition is performed on A to obtain:

$$\mathbf{A} = \mathbf{UDV}^{\mathrm{T}} \tag{11}$$

where D is a diagonal matrix, U and V are unitary matrices. According to the properties of unitary matrix,

$$\|\mathbf{AX}\| = \left\|\mathbf{UDV}^{\mathrm{T}}\mathbf{X}\right\| = \left\|\mathbf{DV}^{\mathrm{T}}\mathbf{X}\right\| \tag{12}$$

where $\mathrm{V}^{\mathrm{T}}\mathrm{X}$ is a column matrix, and because the diagonal element of D is singular value, the last diagonal element is the minimum singular value if and only if

$$\mathbf{V}^{\mathrm{T}}\mathbf{X} = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \tag{13}$$

then $\|\mathbf{AX}\|$ be a minimum, and

$$\mathbf{X} = \begin{bmatrix} a \\ b \\ c \end{bmatrix} = \mathbf{V}\begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \tag{14}$$

where (a, b, c) is the normal vector of the plane $\mathbf{n}$.

2.2 Variable Scale Segmentation Algorithm

The idea of this paper is a ground detection method of unmanned mine car based on area search. Firstly, the target area is rasterized, and the point cloud in the grid is

preliminarily filtered through the elevation information to obtain the candidate point set; Then, the candidate point set is fused, the candidate points are arranged in the grid, and the ground fitting is carried out in the fused large-scale grid, the matching degree of the neighborhood ground and the grid ground to the candidate points in the grid is compared, and the ground equation of the grid is determined; Furthermore, after the ground equation of each grid is obtained, the ground equation of each grid is smoothed. For the grid that has not been successfully fitted, the neighborhood ground equation is used to solve the approximate equation of the grid; Finally, the distance of all input point clouds to the local grid ground equation and the neighborhood grid ground equation is calculated, and whether it is less than the preset distance threshold is analyzed to determine whether it is a ground point.

Step1. Initial extraction of candidate points based on fine mesh.

Traditional ground detection algorithms often segment the whole ground of the point cloud in the driving area. Another method is to grid the whole area and segment the ground of different grids. However, these methods have large cumulative errors, that is, in a large area, the distribution range of point cloud elevation values is larger, which makes it impossible to effectively distinguish low obstacles, resulting in false segmentation. In this paper, firstly, the target area is divided into fine grids, and the square grid is obtained. The lowest elevation point in the grid is counted as the ground reference point. By expression,

$$\left|Z_n - Z_{p0}\right| < \lambda \quad (n = 0, 1, 2, \ldots, n) \tag{15}$$

Judge whether the distance between the current point and P0 is less than the threshold λ, If it is less than, the candidate point set will be pushed. If it is greater than, it will be judged as a non-ground point.

Step2. Ground segmentation based on large area mesh.

(1) More complete ground point information can be obtained from the candidate point set filtered by the fine grid, but there will be misjudgments (such as ground pits or large obstacles). Therefore, the small grids are fused and redistributed, and the candidate points in the small grids are arranged in the corresponding large-area grids. Sort the large-area grids, calculate the distance between the center point of each grid in the lidar coordinate system and the lidar, and take the nearest grid as the initial fitting grid, through Eq. 1,

Fit the ground points in the grid, iterate through the number of iterations determined by Eq. 3, and continuously update the number of interior points until the interior points meet the confidence point pf = p • pall, the iteration ends, and the plane parameters are output

$$N > \log(1 - p)/\log(1 - p_0) \tag{16}$$

(2) After the ground fitting equation of the nearest grid is obtained, the elevation value at the center point of the grid is calculated. When the next grid is reached, the elevation value at the center point of the neighborhood grid is used as the neighborhood information to filter the candidate points in the current grid twice and filter

out the non ground points in the candidate points. That is, the candidate points of the current grid are filtered twice through the effective center point height in the neighborhood grid, that is, through Eq. 15,

$$(A \cdot z_1 + B \cdot z_2 + \ldots, H \cdot z_n)/n \quad (n <= 8) \tag{17}$$

Complete secondary filtration. Repeat step 1 for the candidate points that have completed the filtering to obtain the plane equation of the current mesh. For the plane equation that does not meet the requirements, search the effective plane equation of the neighborhood grid, and calculate the fitness of the candidate points in the current grid to the effective plane equation of the neighborhood grid. When the fitness reaches the threshold, the effective plane equation of the neighborhood grid will replace the plane equation of the grid.

(3) Continue to repeat steps 1–2 until all meshes obtain effective ground fitting equations (or are judged as invalid meshes), and the ground segmentation of large-area meshes is completed.

Step3. Region based ground smoothing and searching.

For each large-area grid, the ground equation fitted by the grid itself is only applicable to the interior of the grid or its individual neighborhoods. However, in real life, the ground is continuous. In order to eliminate the interference of non ground points, the ground parameters of the current grid are smoothed through the neighborhood grid. Four neighborhood grids connected with the current grid are taken, and the ground parameters of the current grid are smoothed through Eq. 16,

$$A_f = (4 \cdot A_0 + A_{left} + A_{right} + A_{up} + A_{down})/8 \tag{18}$$

Is the smoothed plane equation parameters, and other equation parameters are obtained by the same method.

After obtaining the smoothing equation parameters of each large-area grid, determine whether the current point belongs to the ground point by calculating the distance from all points in the current grid to the plane of the current grid and comparing the distance with the distance threshold, and eliminate the point cloud that does not belong to the ground point to complete the ground segmentation.

3 Experimental Verification

Firstly, the calibration algorithm is verified. The calibration equipment we use is Ouster64 line mechanical lidar, with Jetson AGX Xavier as the computing unit. The above equipment is arranged on the MTE108 mining truck of the LGMG.

Take the 5 m*5 m area ground point cloud 10 m in front of the unmanned mine card, and ensure that there is no slope change between the front ground and the ground where the unmanned mine card is located. Verify the detection effect of the

pitch angle, roll angle and yaw angle detected by the calibration algorithm, and the detection accuracy is shown in Table 1.

As shown in the above table, the pitch angle, yaw angle and roll angle obtained by the calibration algorithm are within ± 2% of the manual calibration value, which has a good calibration effect. After the actual use process verification, it can also be proved that the calibrated lidar has high accuracy in the subsequent segmentation tasks. The calibration algorithm is used for the self-calibration effect of the pitch angle of the unmanned mine car as shown in Fig. 1. The blue in the figure is the point cloud after correction, the color is the point cloud before correction,

Next, we will verify the ground detection performance of the variable-scale segmentation algorithm. Using RSBP 32 line mechanical blind-filling lidar as data input, the calculation unit is Jetson AGX Xavier, and the above equipment is also arranged on the MTE108 mining truck of the LGMG. The RSBP lidar is placed at the rear of the unmanned vehicle, 1.6 m from the horizontal ground, to detect the ground point cloud of the unmanned vehicle when reversing. The unmanned vehicle is backed by a 3-degree slope, which is a standard flat-slope-flat road section. The existing algorithm [7] and the algorithm in this paper are used to detect the ground behind the unmanned vehicle, and the detection results are as follows:

As can be seen from the above Figs. 2, 3, 4, and 5, in a typical mining scene, that is, a flat-slope-flat road section, the unmanned mining truck can filter out the ground points on the same plane and a certain slope through the existing algorithm. However, the ground point cloud on the slope was not completely filtered out. The variable-scale segmentation algorithm can segment the ground point cloud of different slopes and different road sections. It has high robustness and environmental adaptability, and can meet the scenes with many soil slopes and bumpy road sections in mining areas.

Table 1 Detection accuracy of calibration algorithm

Pitch ground-truth	Pitch predict	Roll ground-truth	Roll predict	Yaw ground-truth	Yaw predict
10.51	10.3	6.37	6.33	15.32	15.2
8.44	8.41	3.21	3.20	10.44	10.38
5.32	5.30	0.59	0.60	5.46	5.2



Fig. 1 Effect of lidar pitch angle calibration

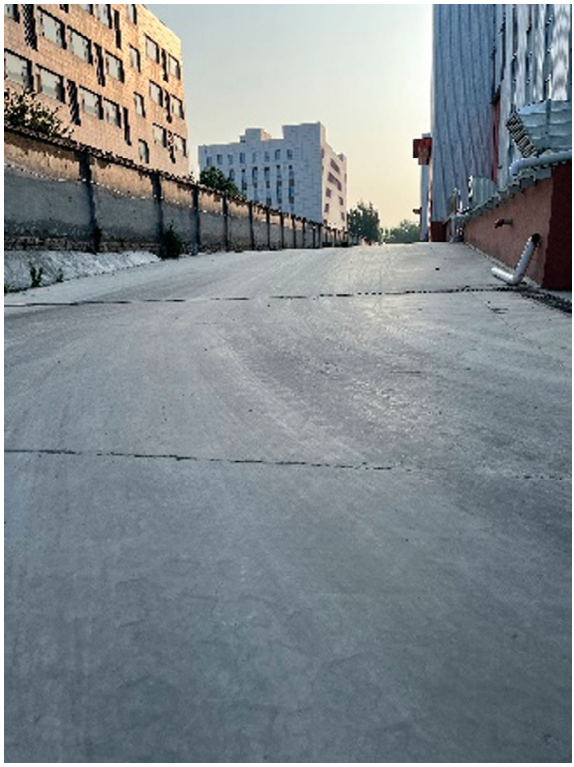


Fig. 2 Scene map of the rear slope of the unmanned vehicle

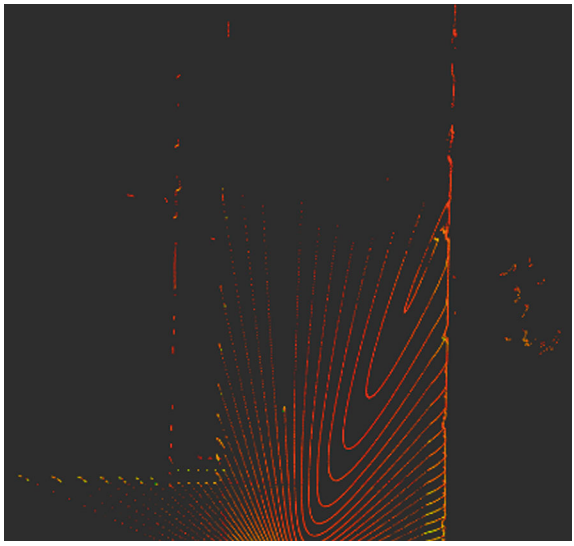


Fig. 3 Raw laser point cloud map

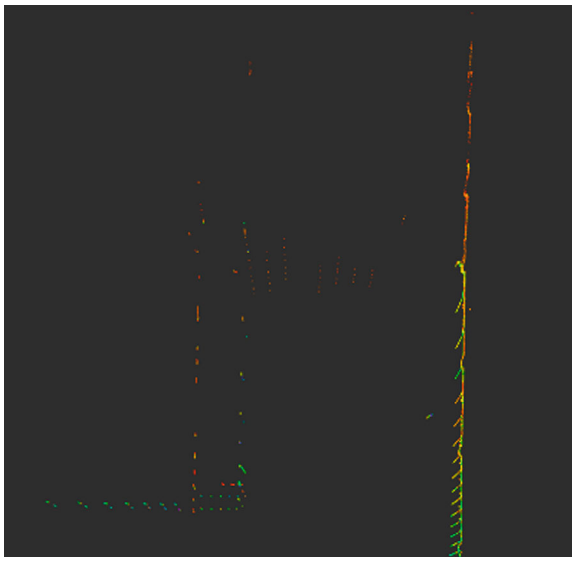


Fig. 4 Effect map of ground segmentation in the paper

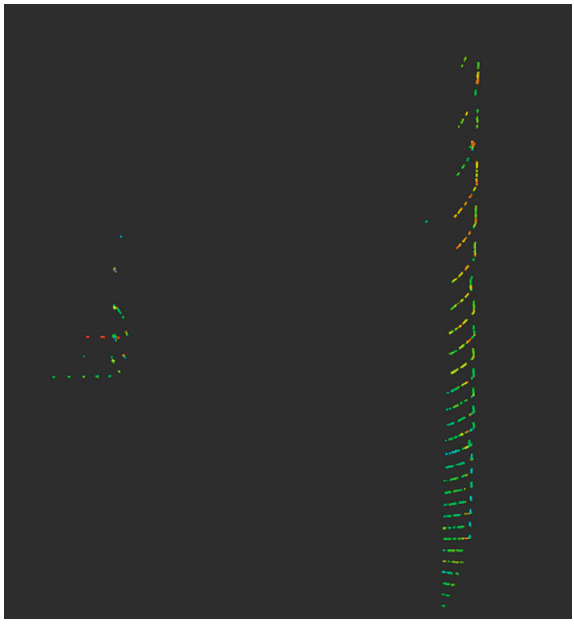


Fig. 5 The ground segmentation effect map in this article

4 Conclusion

It can be seen from the above experiments that the lidar calibration algorithm has high calibration accuracy, and the operation is relatively simple in the actual application process, the site requirements are not high, and the reproducibility is strong. The tester checks the lidar status at any time and makes immediate corrections for the lidar angle deviation. Significantly reduces lidar calibration time. However, in specific scenarios, this algorithm has certain limitations. For example, at the intersection of uphill and downhill, if the assumption of horizontal consistency in the area in front of the vehicle is broken, it will lead to a decrease in calibration accuracy. Therefore, it is necessary to consider how to avoid automatic calibration bias in specific scenarios in the future.

The variable-scale segmentation algorithm can segment lidars of different scenes and different linear shapes, and has strong versatility. For typical Z-shaped slopes, the ground point cloud can still be detected well. In the environment with more smoke, dust and slopes inside the mining area, the high segmentation accuracy of this segmentation algorithm can effectively retain the characteristics of various obstacles and ensure the stable driving of the unmanned vehicle in the mining area.

References

1. Pereira LG, Fernandez P, Mourato S, Matos J, Mayer C, Marques F (2021) Quality control of outsourced LiDAR data acquired with a UAV: a case study. Remote Sensing
2. Bekey GA (2008) Springer handbook of robotics. In: Siciliano B, Khatib O (eds) IEEE robotics & automation magazine, vol 15, issue 03, pp 110–110
3. Zermas D, Izzat I, Papanikolopoulos N (2017) Fast segmentation of 3D point clouds: a paradigm on LiDAR data for autonomous vehicle applications. In: 2017 IEEE international conference on robotics and automation (ICRA), May 29–June 3, 2017, Singapore. IEEE, New York, pp 5067–5073
4. Qiu JY, Lai JZ, Li ZM et al (2020) A lidar ground segmentation algorithm for complex scenes. Chin J Sci Instrument 41(11):244–251
5. Li QQ, Zhang L, Mao QZ et al (2014) Motion field estimation for a dynamic scene using a 3D LiDAR. Sensors 14(09):16672–16691
6. Zhou Y, Wang D, Xie X et al (2014) A fast and accurate segmentation method for ordered LiDAR point cloud of large-scale scenes. IEEE Geosci Remote Sens Lett 11(11):1981–1985
7. Himmelsbach M, Hundelshausen FV, Wuensche HJ (2010) Fast segmentation of 3D point clouds for ground vehicles. In: Intelligent vehicles symposium, IEEE

Open Access This chapter is licensed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License (http://creativecommons.org/licenses/by-nc-nd/4.0/), which permits any noncommercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if you modified the licensed material. You do not have permission under this license to share adapted material derived from this chapter or parts of it.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.



Effect of MHD Control on Turbulent Boundary Layer Separation Flow in Compression Corner



Zhaoyu Wu, Mingsong Ding, and Weizhong Dong

Abstract In order to study the effect of MHD control on turbulent boundary layer separation of compression corners, the actuator model is developed and verified with a typical example. A numerical simulation of MHD control on the boundary layer separation of the compression corner is carried out to study the influence of various magnetic induction intensities on it. The results show that the numerical simulation results of the calculation method established in this paper are consistent with the experimental results, and boundary layer separation of compression corners is alleviated by applying MHD control. Under different magnetic induction intensities, the size of the separation zone decreases when magnetic induction intensity increases.

Keywords MHD · Compression corner · Boundary layer separation · Numerical simulation

1 Introduction

The interaction between the shock and boundary layer is the basic flow characteristic of compression corners. To reduce boundary layer separation of compression corner turbulence, MHD control has become a new method in recent years. In 2006, Zaidi proposed a type of MHD excitation method called snowplow [1]. The experimental study found that the gas velocity increased by 50–200 m/s under the action of Lorentz force in the boundary layer. In 2008, Saito carried out an experimental study on the MHD control of the separation zone of the front stage [2]. The size of the separation zone was successfully reduced by about 10% in the tunnel. However, the interaction of turbulence and magnetic field is not considered in the paper.

Z. Wu (✉)
Northwest Institute of Nuclear Technology, Xian, Shanxi, China
e-mail: wzy98@mail.ustc.edu.cn

M. Ding · W. Dong
China Aerodynamics Research and Development Center, Mianyang, Sichuan, China

© The Author(s) 2026

H. Saman et al. (eds.), *The 9th International Conference on Advances in Construction Machinery and Vehicle Engineering*, Lecture Notes in Mechanical Engineering,
https://doi.org/10.1007/978-981-96-5527-4_94

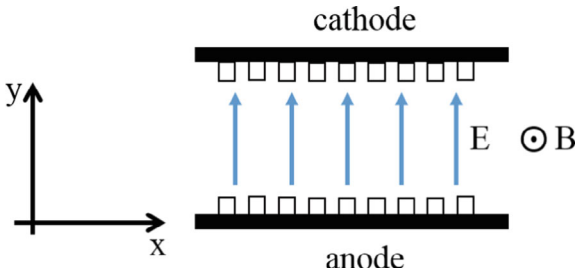


Fig. 1 MHD plasma actuator model

It follows that research on the MHD control of the boundary layer separation in compression corner turbulence needs to be further studied. Hence, the numerical simulation method and program of low magnetic Reynolds number three-dimensional magnetohydrodynamics turbulent flow are established in this paper. A typical model of a two-dimensional 24° compression corner is selected to carry out the research on MHD control to suppress boundary layer separation flow at the compression corner, and the effect of different magnetic induction intensities on the size of the separation zone of the compression corner is considered.

2 Actuator Model

Figure 1 is the MHD plasma actuator model. By using DC electrode discharge [3], the air between the anode and cathode is ionized to form plasma. Lorentz force is formed along the x-axis in the ionization region after applying a magnetic field. The flow field at the bottom of the separation zone is accelerated so as to reduce the size of the separation zone.

3 Code Validation

Hartmann flow is one of the classical basic flows in MHD [4]. It describes the complete flow between two parallel plates under the action of a normal magnetic field: Lorentz force and viscous stress. As shown in Fig. 2, the distance between parallel plates h is set to 0.01 m. The length of the plates L is 0.2 m, and the plate length is 20 times the distance between plates to reduce the interference of flow near the wall. A uniform magnetic field with magnetic induction intensity B_0 is applied perpendicular to the plate along the y-axis. The whole flow field is set to constant conductivity σ_0. There is a fixed pressure gradient $Dp = -\frac{\partial p}{\partial x}$ in the incoming direction to drive the fluid movement. The component of velocity in the x-axis is obtained through theoretical analysis:

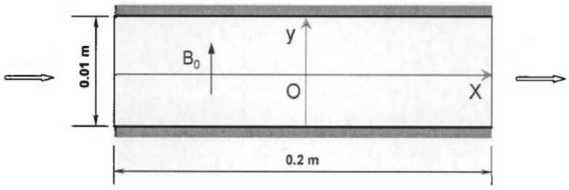


Fig. 2 Hartmann flows on a flat plate

$$u(y) = U_{\max}\left[1 - \frac{\cosh(Ha \cdot 2y/h)}{\cosh(Ha)}\right] \tag{1}$$

where $U_{\max} = -\frac{1}{\sigma_0 B_0^2} \cdot \frac{\partial p}{\partial x}$ is the maximum speed and $Ha = \frac{B_0 h}{2} \cdot \sqrt{\frac{\sigma_0}{\mu}}$ is the Hartmann number.

Flow conditions are $Ma_\infty = 0.01$, $p_\infty = 101956.0$ Pa, $\rho_\infty = 1.225$ kg/m^3, $Dp = 10$ Pa, and $\sigma_0 = 800$ S/m. The periodic boundary is adopted for flow velocity. The range of the Hartmann number is adjusted through the change in magnetic induction intensity.

Figure 3 compares the numerical results of the velocity type $u(y)/U_{mean}$ of the pipeline profile with the distribution of the exact solution, where U_{mean} is the average velocity in the y-axis. Figure 3 shows the fully developed laminar flow velocity type under different Ha numbers, and the numerical simulation is completely consistent with the accurate solution. The magnetic induction intensity is zero where Ha = 0, and there is no effect of the magnetic field. The fluid is balanced under the interaction of driving pressure and viscous stress. The magnetic induction intensity increases with the increase of the Ha number, and the velocity type flattens. The velocity near the central area of the flat plate basically remains constant when Ha > 10.

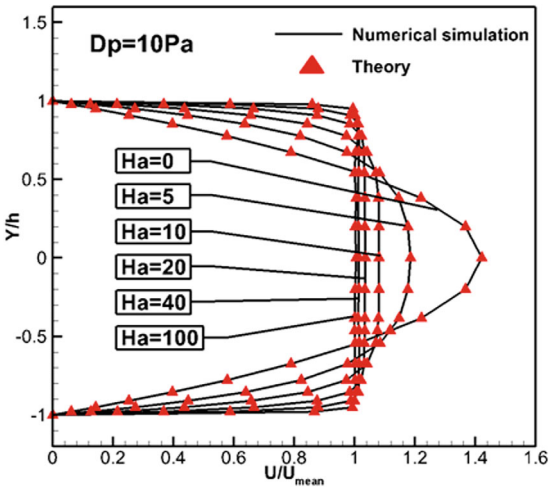


Fig. 3 Comparison between calculation and theoretical results under different Ha numbers

4 Simulation Results and Discussion

According to the experiment of shock boundary layer separation carried out by Settles [5], the calculation model selected in this paper is the same two-dimensional 24° compression corner model so as to facilitate the verification of the accuracy of the program. As shown in Fig. 4, the inflection point of the compression corner is set as the origin, and a section of flat plate is set in front of it. The length of the plate is 1.226 m, and the length of the inclined plane is 0.8 m. Flow conditions are $Ma_\infty = 2.85$, $p_\infty = 23219.10$ Pa, $T_\infty = 99.83$ K, Reynolds number $\mathrm{Re}_\infty = 7.3 \times 10^7$/m, and incoming boundary layer thickness $\delta = 3.5$ mm.

4.1 Flow Field Structure Without Control

Figure 5 shows the comparison between the numerical solution of wall friction coefficient distribution and test data, as well as the literature results [6]. The wall friction coefficient is less than zero at the corner, indicating that this area is a separation zone.

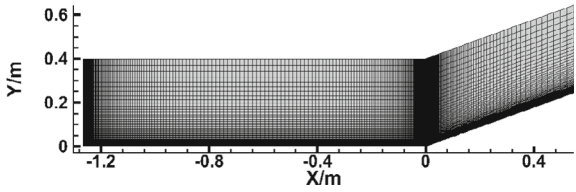


Fig. 4 Calculation grid of compression corner

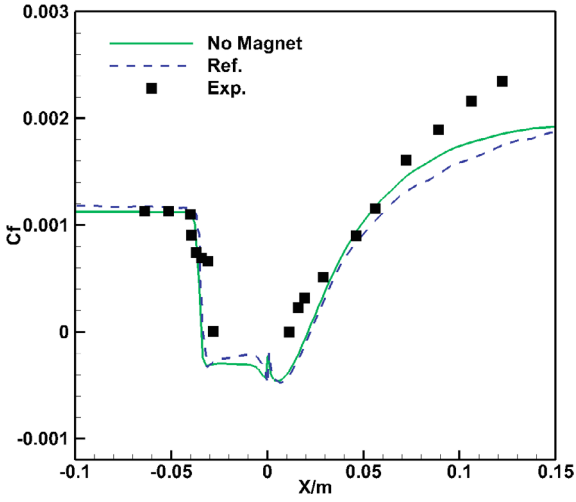


Fig. 5 Wall friction coefficient without control

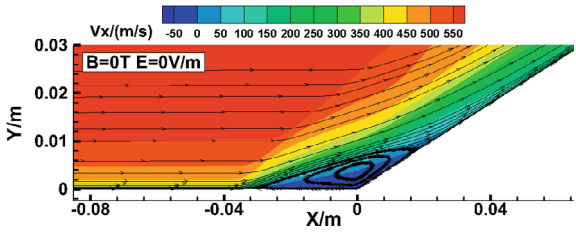


Fig. 6 Streamlined distribution of the compression corner without control

Figure 6 shows the streamlined distribution of compression corners without MHD control. Under the action of the reverse pressure gradient, a low-speed separation zone appears near the corner after the incoming flow collides with the inclined plane. The separation zone spans both sides of the corner with a length of 0.0573 m and a thickness of 0.006 m at the corner, accompanied by separation shock, expansion wave, reattachment shock, and other structures.

4.2 Effect of Magnetic Induction Intensity

MHD control is applied to the separation zone of the compression corner. Figure 7 shows that the MHD control area is set in a zone of -0.04 m $< x <$ 0.03 m with a height of 0.001 m, plus a 500 V/m uniform electrical field perpendicular to the compression corner. Cathode and anode electrodes are arranged under the wall of the compression corner. Magnetic induction intensity B is set to 0.5, 1.0, and 2.0 T along the positive direction of the z-axis. Uniform conductivity is set as the value of 800 S/m.

Figure 8 shows the streamlined distribution of the separation zone of the compression corner under different magnetic induction intensities. After MHD control is applied, the boundary layer is accelerated by Lorentz force, and the size of the separation zone is reduced. With the increase of magnetic induction intensity, the size of the separation zone decreases gradually, and the influence of restraining separation is more obvious.

Figure 9 further shows the wall friction coefficient distribution of compression corners under various magnetic induction intensities. The area with a wall friction coefficient less than zero is significantly reduced after MHD control is applied. The

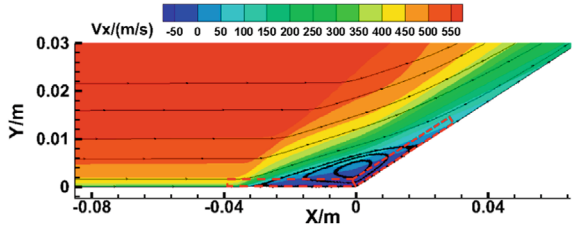


Fig. 7 MHD control area

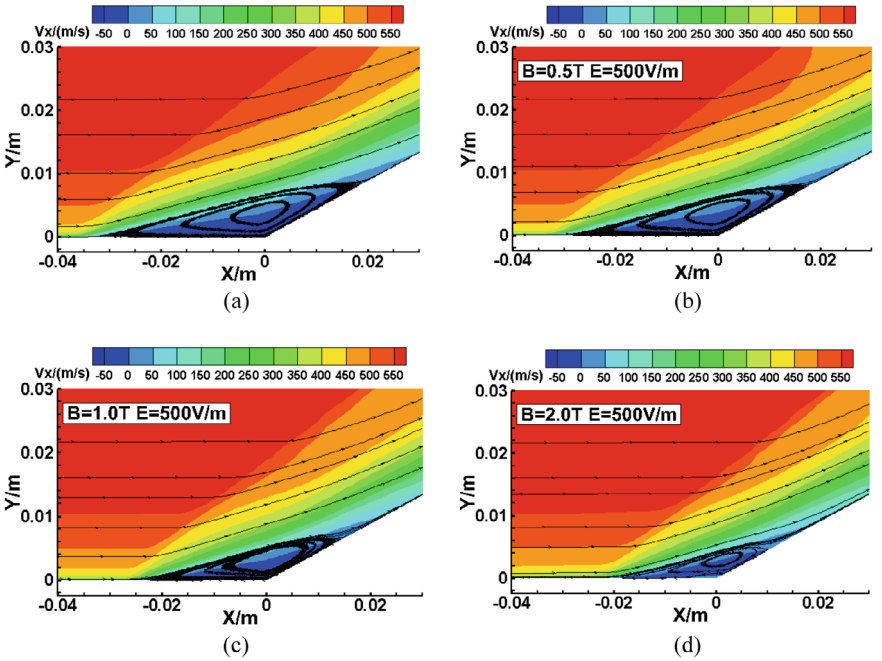


Fig. 8 Streamlined distribution of the separation zone under various magnetic induction intensities: **a** B = 0 T, **b** B = 0.5 T, **c** B = 1.0 T, **d** B = 2.0 T

points of separation and reattachment are close to the corner. When B = 1.0 T, the size of the separation zone is about 0.0407 m, which is reduced by about 30%. Wall friction coefficient under three kinds of magnetic induction intensity has an obvious upward trend near the corner. This indicates that MHD control is most effective in the area near the corner, and the reverse pressure gradient in the corner area is easier to suppress than in other areas.

5 Conclusions

Based on the results and discussions presented above, the conclusions are obtained as below:

(1) The computation of the Hartmann flow is in good agreement with the literature's results.
(2) The numerical calculation of the two-dimensional 24° compression corner model shows that the size of the separation zone of the compression corner is reduced after MHD local control is applied. Compared with the case of no MHD control, the size of the separation zone is reduced by about 30% when magnetic

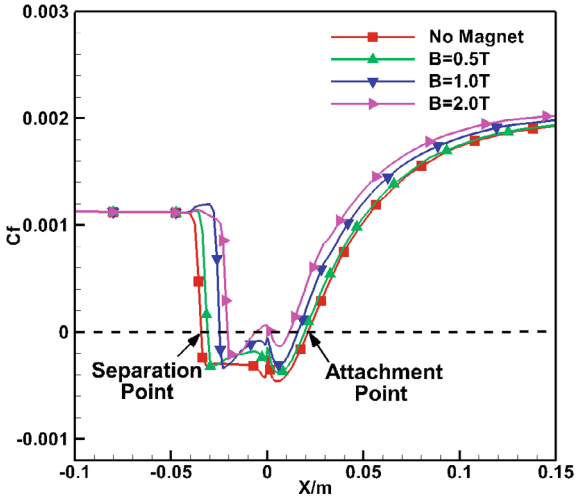


Fig. 9 Wall friction coefficient under different magnetic induction intensities

induction intensity is 1.0 T, which improves the boundary layer separation of the compression corner.

(3) It is found that the effect of MHD control is better with an increase in magnetic induction intensity.

References

1. Zaidi SH, Smith T, Macheret S et al (2006) Snowplow surface discharge in magnetic field for high-speed boundary layer control. In: 44th AIAA aerospace sciences meeting and exhibit. Reno, Neveda, p 2006
2. Shinya S, Keisuke U, Kenji K et al (2008) Boundary layer separation control by MHD interaction. In: 46th AIAA aerospace sciences meeting and exhibit. Reno, Neveda, p 1091
3. Xu HM, Jia PY, Ran JX et al (2021) Investigation on discharge mode of an argon plasma jet array with a single electrode excited by a direct-current voltage. J Nanchang Univ (Nat Sci) 45(5):455–462
4. Zhao QJ, Mao BQ, Bai XH et al (2022) Effect of transverse magnetic field on magnetogasdynamic flow and heat transfer characteristics in insulated/conductive circular tubes. Acta Phys Sin 71(16):164702
5. Luo SC, Liu J, Hu SC et al (2024) Effect of magnetohydrodynamic control on the self-starting ability of hypersonic inlets. Sci Sin-Phys Mech Astron 54(3):234711
6. Feng YY, Li ZL, Ji ZL (2021) LES study of microbubble drag reduction on flat-plate turbulent boundary layers. Energy Conserv Technol 39(4):305–311

Open Access This chapter is licensed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License (http://creativecommons.org/licenses/by-nc-nd/4.0/), which permits any noncommercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if you modified the licensed material. You do not have permission under this license to share adapted material derived from this chapter or parts of it.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.



Research on the Characteristics of Electro-Hydraulic System of Crane Pressure Flow Composite Control



Pengfei Yang, Luhong Yue, and Jing Yang

Abstract Aiming at the problems of low flow control accuracy, slow response and large energy loss in existing crane load-sensitive systems, a crane pressure-flow composite control electro-hydraulic system based on pressure/flow mode switching and bilinear interpolation flow control strategy is proposed. According to the working conditions, the pump and valve are combined with pressure and flow sensitivity. During micro-movement, the flow compensator dynamically adjusts the valve opening according to the real-time pressure difference before and after the valve port to achieve precise flow control. During rapid movement, the valve port is fully opened and the hydraulic pump adjustment system is used. flow. This article establishes the AMESim simulation model of the system and builds a pump-valve collaborative test platform. The control characteristics of the constructed system under constant pressure difference and varying flow rates were investigated through simulation and experimental analysis. Additionally, the system's behavior was examined under different pressure differences while maintaining a constant flow rate. The research results show that when the pressure-flow composite control electro-hydraulic system is in micro-flow ($\leq$ 70 L/min), the fluctuation of the flow during the dynamic change of the load is not greater than 10.98%, which can effectively suppress the disturbance of the flow caused by the load change. When the present pressure difference of the pump control subsystem increases, the system responds quickly and has high dynamic performance. When the present pressure difference decreases, the system flow control accuracy can be improved.

Keywords Hydraulics · Load sensitivity · Pressure and flow compound control · Pump valve cooperation · BODAS controller

P. Yang · L. Yue · J. Yang (✉)
College of Mechanical and Vehicle Engineering, Taiyuan University of Technology, Taiyuan, China
e-mail: yangjing001@tyut.edu.cn

© The Author(s) 2026

H. Saman et al. (eds.), *The 9th International Conference on Advances in Construction Machinery and Vehicle Engineering*, Lecture Notes in Mechanical Engineering,
https://doi.org/10.1007/978-981-96-5527-4_95

1 Introduction

Load-sensitive systems are commonly used in the current field of cranes and construction machinery, taking advantage of their proportional flow distribution, high speed and stiffness, and small impact of load on flow. Kong et al. [1, 2] studied the dynamic characteristics of load-sensitive systems and analyzed the impact of feedback pipeline parameters on the dynamic characteristics of load-sensitive systems. Zhang et al. [3–5] studied the pre-valve compensation system, analyzed the flow-pressure relationship curve of the load-sensitive valve during the compression process, and analyzed the working principle of the load-sensitive system. Gao et al. [6–8] studied the LUDV system and analyzed the flow distribution characteristics of hydraulic oil through the multi-way valve and the impact of the structural parameters of the multi-way valve on flow distribution. As the requirements for positioning accuracy and low energy consumption of construction machinery continue to increase, load-sensitive systems using machine-hydraulic pressure feedback loops face problems such as slow dynamic response, weak robustness, high energy consumption, and inaccurate flow control. The shortcomings become more prominent.

With the continuous development of electronic control technology and sensor technology, it is possible to replace the original machine-hydraulic servo control with electronic control. Finzel et al. [9] proposed a hydraulic control system using an electric proportional multi-way valve in the form of pre-valve compensation, and conducted a preliminary theoretical discussion on the energy saving and responsiveness of a large flow system. Long et al. [10–12] proposed the principle of electro-hydraulic proportional flow control. By combining the internal feedback valve of the open-loop system with the electronic closed-loop control valve of the closed-loop system, the flow rate through the valve port is independent of the load.

At present, the above-mentioned electro-hydraulic flow control method is only suitable for independent control of proportional valves or electro-hydraulic proportional pumps, and has not yet realized the combined control of pressure and flow in hydraulic systems. Dujia et al. [13] proposed a collaborative pressure-flow composite control system for crane pump valves based on the principle of electronic pressure compensation. This paper proposes a crane pressure-flow composite control electro-hydraulic system based on the control strategy of pressure/flow mode switching and bilinear interpolation flow control strategy, and builds a crane pressure-flow composite control electro-hydraulic system test platform, and conducts experiments on the same set pressure The control characteristics of different flow rates under different pressure differences and the control characteristics of the same flow rate under different set pressure differences were simulated and experimentally analyzed.

2 System Principles and Control Strategies

The crane pressure-flow composite control electro-hydraulic system includes multiple components, such as electro-hydraulic proportional pumps, electro-hydraulic proportional multi-way valves, pressure sensors and actuators. Compared with the traditional anti-flow saturation load-sensitive system, this system has made some structural improvements, such as eliminating the pressure compensator behind the proportional multi-way valve and using an electro-hydraulic proportional variable pump to replace the load-sensitive pump. At the same time, pressure sensors are installed before and after the valve of the multi-way reversing valve for pressure detection. Functionally, the control unit can calculate the valve control signal in real time based on the present flow rate and the pressure difference in front of and behind the valve to control the opening of the main valve core. The pressure sensor detects the pressure in front of and behind the valve in real time. The controller can calculate the pump control signal in real time based on the present pressure difference and the pressure in front of and behind the valve to control the displacement of the electro-hydraulic proportional pump. The principle of the crane pressure-flow composite control electro-hydraulic system is shown in Fig. 1.

During the working process, the pressure sensor transmits the monitored pump outlet pressure and load pressure to the controller. The pressure controller adjusts the displacement of the electro-hydraulic proportional pump according to the feedback pressure changes to ensure that the pump outlet pressure always maintains a ratio higher than the maximum. The load pressure is higher by a fixed value Δp_0 to achieve load sensitivity. The controller calculates the valve control signal based on the set flow

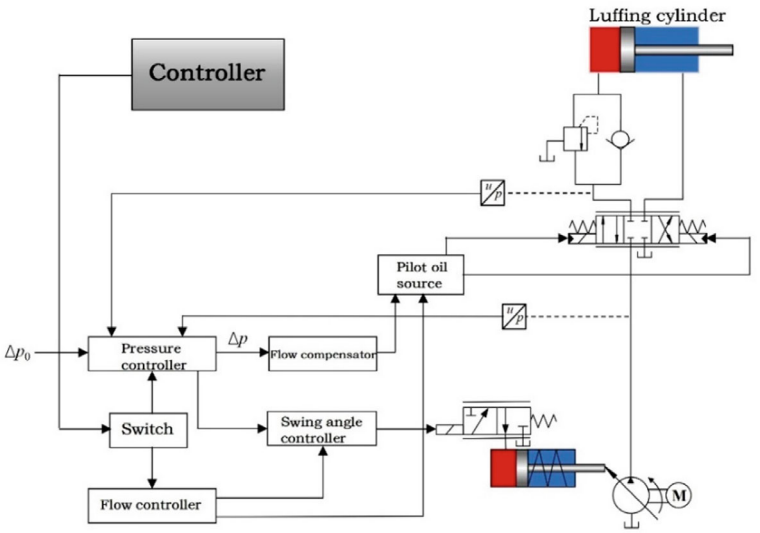


Fig. 1 Principle diagram of electro-hydraulic system for crane pressure-flow composite control

rate and the current real-time pressure difference to control the spool displacement of the multi-way valve. The specific flow compensation principle is as follows:

The flow formula of hydraulic oil through the multi-way valve port is:

$$Q = C_d A \sqrt{\frac{2\Delta p}{\rho}} \tag{1}$$

In the formula: Q is the flow rate through the multi-way valve port; A is the multi-way valve port area; Δp is the valve port pressure difference, ρ is the hydraulic oil density, and C_d is the flow coefficient.

Assume that when the load changes at a certain moment, the pressure difference between the valve port changes to Δp, and the flow rate through the valve port is Q. At this time, the difference ΔQ between the actual flow rate Q through the valve port and the initial flow rate Q_0 is calculated as follows:

$$\Delta Q = C_d A \sqrt{\frac{2}{\rho}} (\sqrt{\Delta p} - \sqrt{\Delta p_0}) \tag{2}$$

Changing the multi-way valve port area A can make the flow rate Q through the valve port return to Q_0.when the load changes. At this time it should be:

$$\Delta A = A(1 - \sqrt{\frac{\Delta p_0}{\Delta p}}) \tag{3}$$

In the main valve of a multi-way valve, the flow area and the valve core displacement usually have a linear relationship, which can be obtained:

$$\Delta x = x\left(1 - \frac{\sqrt{\Delta p_0}}{\sqrt{\Delta p}}\right) \tag{4}$$

It can be calculated from this that when the pressure difference at the multi-way valve port deviates from the present value Δp_0, in order to ensure the stability of the flow through the multi-way valve port, the position of the multi-way valve spool needs to be changed by Δx.

The working modes of the crane luffing system mainly include rapid movement mode, micro-motion mode and descent mode. In the fast motion mode, focusing on energy efficiency characteristics, the multi-way valve is fully opened, and the flow into the actuator is controlled by controlling the swash plate angle of the electro-hydraulic proportional pump, that is, the volume control method is adopted. In the inching mode, attention is paid to the flow control accuracy, and the multi-way valve flow control method is mainly used. The electro-hydraulic proportional pump works in the pressure control mode control strategy, that is, the throttling control method; in the descending mode, the current crane mainly uses external control. Balance valve

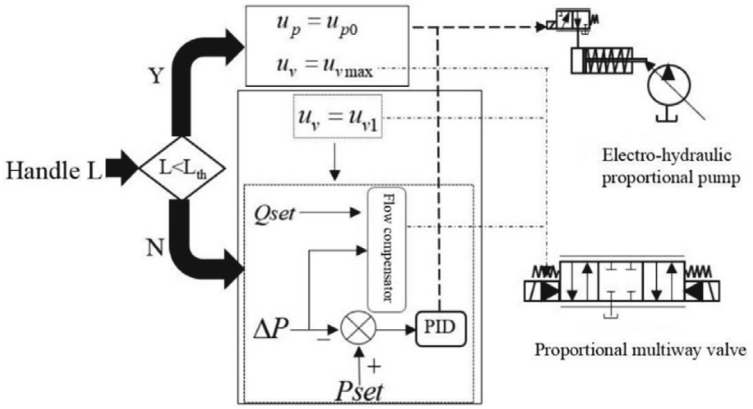


Fig. 2 Mode switching control strategy

to control the lowering speed of the actuator. The control strategy flow of the system is shown in Fig. 2.

During rapid movement, the multi-way valve port is fully open. The flow through the multi-way valve is proportional to the swing angle control signal of the electro-hydraulic proportional pump. When the handle is switched to the inching mode, the electro-hydraulic proportional pump works in the pressure control mode, which implements pressure closed-loop control based on the pressure feedback of the outlet pressure sensor, so that the outlet pressure of the electro-hydraulic proportional pump is always higher than the maximum load pressure by a fixed value Δp_0. Through the calculation of the flow compensator, each opening angle of the handle will obtain a uniquely determined multi-way valve spool displacement Δx under the current pressure difference Δp. Figure 3 shows the displacement curve of the flow area and valve core of the multi-way valve luffing joint valve port of a certain type of crane.

As shown in Fig. 3, when the valve core displacement is 2–5 mm and the valve core displacement is 6–11 mm, the valve core displacement and the flow area have

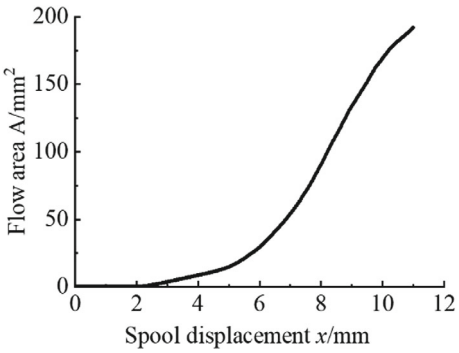


Fig. 3 Typical orifice flow area and valve core displacement curve

a linear relationship respectively. In the principle of flow compensation, the default C_d is a constant. However, the flow coefficient C_d of the U-shaped throttle notch is closely related to the opening of the valve port. When the valve port is at a small opening, the flow coefficient C_d is close to 1; when the valve port opening is in the middle area, the flow coefficient C_d is maintained between 0.7 and 0.8; when the valve port is close to fully open, the flow coefficient C_d increases again to 1 [14]. Flow compensation cannot be performed directly through the flow formula. A large number of static characteristic tests of valve opening-pressure difference-flow are required to obtain the relationship between (Q, $\sqrt{\Delta p}$, U) and establish flow compensation based on bilinear interpolation. device [15]. The flow compensator first needs to locate the four closest points to (Q, $\sqrt{\Delta p}$), obtain their corresponding spool displacement values U_{11}, U_{12}, U_{21} and U_{22} respectively, and then perform linear interpolation operations. The specific interpolation process is as follows:

First perform two horizontal interpolation calculations,

$$U_{x1} = \frac{\sqrt{\Delta p_a} - \sqrt{\Delta p_2}}{\sqrt{\Delta p_1} - \sqrt{\Delta p_2}} U_{11} + \frac{\sqrt{\Delta p_a} - \sqrt{\Delta p_1}}{\sqrt{\Delta p_2} - \sqrt{\Delta p_1}} U_{21} \tag{5}$$

$$U_{x2} = \frac{\sqrt{\Delta p_a} - \sqrt{\Delta p_2}}{\sqrt{\Delta p_1} - \sqrt{\Delta p_2}} U_{12} + \frac{\sqrt{\Delta p_a} - \sqrt{\Delta p_1}}{\sqrt{\Delta p_2} - \sqrt{\Delta p_1}} U_{22} \tag{6}$$

Then perform longitudinal interpolation calculation,

$$U_x = \frac{q_a - q_2}{q_1 - q_2} U_{x1} + \frac{q_a - q_1}{q_2 - q_1} U_{x2} \tag{7}$$

Through transverse and longitudinal interpolation calculations, the displacement of the multi-way valve core corresponding to the current target flow rate and the current valve port pressure difference can be obtained U_x.

3 Establishment of Simulation Model of Pressure-Flow Composite Control Electro- Hydraulic System

By analyzing the existing M7 series load-sensitive hydraulic system proportional multi-way valve, the AMESim simulation model of the multi-way valve was successfully established. The specific model is shown in Fig. 4. After removing the pressure compensator, the part of the pressure compensator can be replaced by a volume chamber [13].

The model of the electro-hydraulic proportional pump is more complex. The variable mechanism of the electro-hydraulic proportional pump adopts a mechanical-hydraulic servo system of a four-way valve-controlled cylinder. This system can be approximately regarded as a third-order linear system. Therefore, the step response characteristics of the electro-hydraulic proportional pump can be measured, and

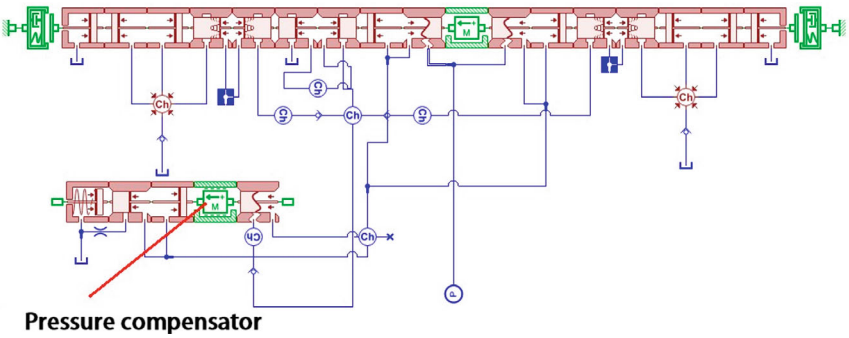


Fig. 4 AMESim model of proportional multi-way valve

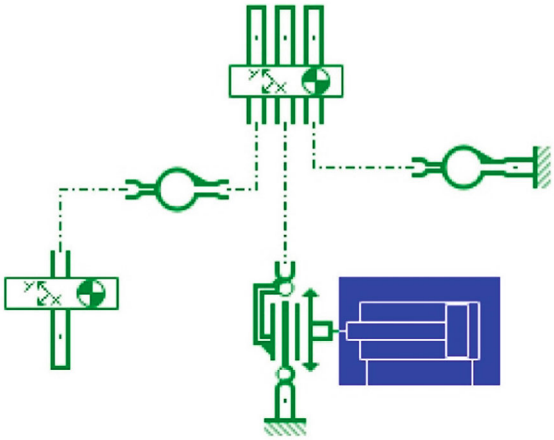
the dynamic response model of the electro-hydraulic proportional pump can be established by using the third-order transfer function to fit the step response curve [16].

The luffing mechanism of the crane is used to change the angle of the boom, thereby increasing the working range of the crane. It can be modeled using the 2D plane mechanical library of AMESim software. The AMESim model of the crane luffing mechanism is shown in Fig. 5.

For other hydraulic components such as three-way proportional pressure reducing valves and pilot pumps, a simulation model is established through the sample parameters of the corresponding components.

The simulation model of the crane pressure-flow composite control electro-hydraulic system based on the control strategy of pressure/flow mode switching and the bilinear interpolation flow control strategy is shown in Fig. 6.

Fig. 5 AMESim model of crane luffing mechanism



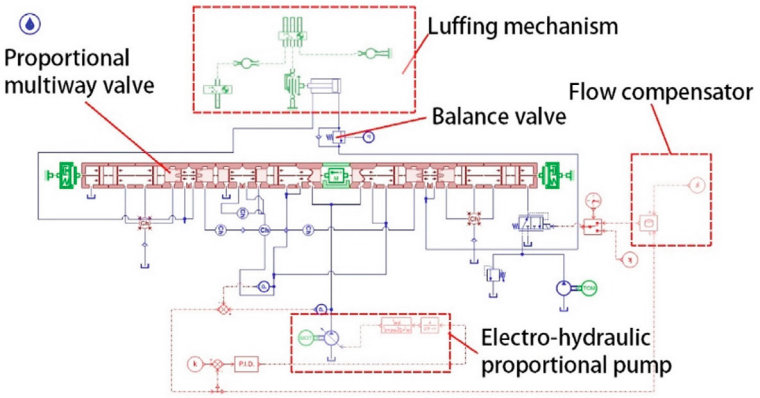


Fig. 6 Pressure-flow composite control electro-hydraulic system AMESim model

4 Establishment of Test Platform for Pressure-Flow Composite Control Electro-Hydraulic System

4.1 Construction of Pressure-Flow Composite Control Electro-Hydraulic System Test Platform

According to the working principle of the pressure-flow composite control system of the crane luffing mechanism, a test platform was built as shown in Fig. 7. The test platform consists of an electro-hydraulic proportional pump power source L11vo115L, a three-term asynchronous motor LMP-250M-4, and a Rexroth M7 series proportional It consists of multi-way valve (luffing joint), three-way proportional pressure reducing valve PPCD05-NG, proportional relief valve, fuel tank, pressure sensor M5134-C2501X-050BG, BODAS RC28-14/30 controller and pipelines and other auxiliary components. Some test device parameters are shown in Table 1.

4.2 Controller Selection and Control Algorithm Writing

The system controller uses BODAS RC28-14/30 controller, which shows excellent control effects in various complex environments. It has been widely used in engineering machinery systems, especially in the field of mobile machinery, and has achieved remarkable results [17].

BODAS RC28-14/30 has 9 analog voltage input terminals and 28 proportional output terminals. The proportional output terminals are all power level outputs without the use of amplifiers. According to the system principle and control strategy,



(a) (b)

1-Proportional multiway valve; 2-Electro-hydraulic proportional pump; 3-BODAS controller; 4-Boarding position.

Fig. 7 Test bench for pressure-flow composite control system

Table 1 Experimental device parameters

Test device	Value
Electro-hydraulic proportional pump displacement/cc/r	115
Three asynchronous motor speeds/r/min	1480
Proportional multi-way valve diameter/mm	22
Three-way proportional pressure reducing valve maximum pressure/bar	30
Proportional relief valve maximum pressure/bar	280

this experiment requires 3 analog signal input terminals and 2 PWM signal output terminals. The output port assignments of the controller are shown in Table 2.

According to the IEC61131-3 standard, the controller is programmed using codesys programming software and supports six programming languages: CFC, ST, SFC, IL, FBD and LD. Moreover, these six programming languages can also be nested into each other to complete the writing of complex programs [17]. This experiment uses three programming languages: CFC, ST and SFC.

According to the control strategy, this test can be divided into fast motion mode and pressure flow control mode. In order to simplify the test process, the test does not conduct the test of the fast moving part and starts directly from the pressure flow control mode.

Table 2 Controller output port assignment

Signal name	Type	Measuring range	Pin
Hydraulic pump outlet pressure	Analog input	0.5~4.5 V	164
Load pressure	Analog input	0.5~4.5 V	166
Flow through multi-way valve	Analog input	0~10 V	231
Valve control signal	PWM Output	0~700 mA	153
Pump control signal	PWM Output	0~700 mA	177

First, the valve core opening of the multi-way valve is controlled to a small fixed value, and the outlet pressure of the electro-hydraulic proportional pump is always higher than the maximum load pressure by a fixed value Δp_0, realizing a pressure closed loop. Then, when the system is stable, the flow compensation calculation of the controller is used to obtain the spool displacement of the multi-way valve required when the current flow is set. Finally, the current value required by the three-way proportional pressure reducing valve is obtained through conversion, and then the multi-way valve is controlled. The displacement of the road valve spool is used to compensate for the flow.

In order to implement the logic in the control algorithm, SFC language can be used to establish a sequential logic block diagram. Then use CFC language to divide each sequential logic block diagram into different functional blocks to establish functional logic. After that, write this function block program in ST language inside different function blocks.

For some complex algorithms, the program can be written with the help of Simulink software. First, build the control algorithm in Simulink software. Then, access the in and out interfaces, connect them to the input and output ports of the control algorithm respectively, and use the Create Subsystem command to build the subsystem to achieve discretization. Finally, the PLC Code toolkit in Simulink is used to convert the control a lgorithm into the ST language of codesys and generate the subsystem.exp file. Then, use the project-import function of codesys to import subsystem.exp into codesys to generate the Function Block function block [18].

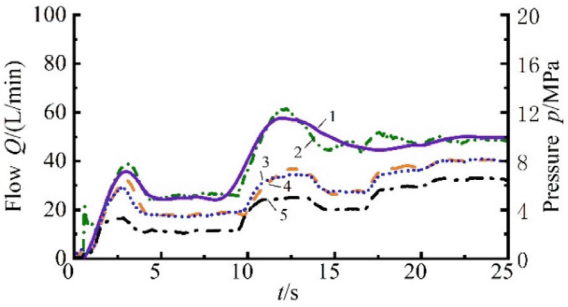
In the control algorithm of the pressure-flow composite control hydraulic system, the programming of the bilinear interpolation table is relatively complicated, and it is written using the above method.

5 Simulation and Experimental Analysis of Pressure-Flow Composite Control Electro-Hydraulic System

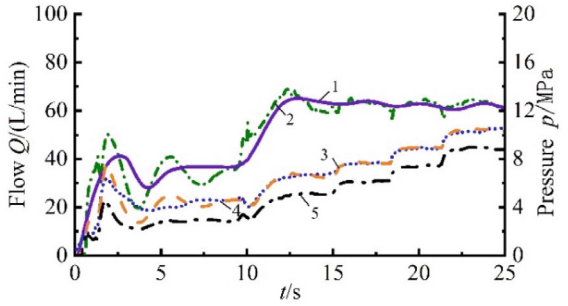
5.1 Analysis of Control Characteristics of Different Set Flow Rates

The maximum outlet pressure of the hydraulic pump in the crane pressure-flow composite control electro-hydraulic system is set to 28 MPa, and the target flow rates of the system are set to 50 L/min, 60 L/min, and 70 L/min respectively. From 0 to 10 s, according to the present flow rate and set pressure difference, the valve opening of the multi-way valve is opened to 4.5 mm, 4.75 mm and 4.9 mm respectively. The electro-hydraulic proportional pump is set to the pressure closed-loop control mode to ensure the pump's The outlet pressure is 1.5 MPa higher than the load pressure, achieving load sensitivity. After 10 s, the load pressure set by the loading relief valve begins to change, and the system uses the flow compensator to compensate according to the set flow rate. This stage runs until 25 s. The output flow, hydraulic

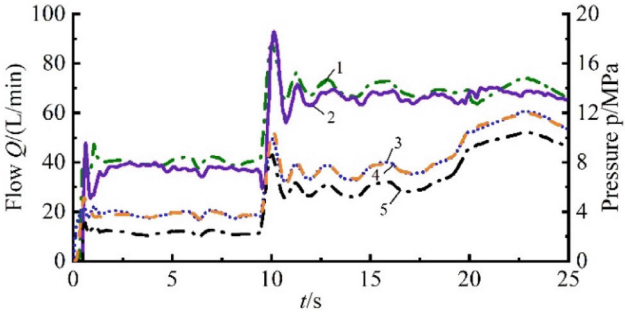
pump outlet pressure and load pressure curves of the pressure-flow composite control electro-hydraulic system are shown in Fig. 8.



(a) 50 L/min Test and simulation curves of target flow rate



(b) 60 L/min Test and simulation curves of target flow rate



(c) 70 L/minTest and simulation curves of target flow rate

1—Test flow; 2—Simulated traffic; 3—Test hydraulic pump outlet pressure; 4—Simulation pump outlet pressure; 5—Load pressure.

Fig. 8 Test and simulation curves for target flows

Table 3 Control characteristics of different target flows

Q/(L/min)	K_P	t_r/s	σ%
50	0.05	1.51	15.8
60	0.02	2.12	4.36
70	0.08	0.238	26.1

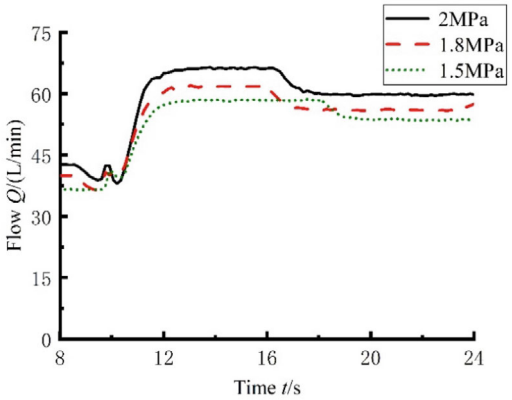
From the analysis of Fig. 8, it can be seen that the electro-hydraulic proportional pump is in the pressure control mode, and its outlet pressure changes with the change of the load pressure and maintains the pressure difference at 1.5 MPa, achieving load sensitivity. The flow rate through the multi-way valve port within 0–10 s is slightly less than the target flow rate, about 70% of the target flow rate. 10–25 s The system uses the flow compensator to compensate the valve port opening of the multi-way valve so that the flow through the multi-way valve port reaches the target flow value.

The electro-hydraulic proportional pump adopts the pressure control mode in the micro-motion working mode of the pressure-flow composite control electro-hydraulic system to ensure that the outlet pressure of the pump is always a certain value higher than the maximum load pressure. At this time, a PID control algorithm is needed to implement the above functions. Due to the influence of factors such as the compressibility of oil and the nonlinearity of the electro-hydraulic proportional control system, the determination of KP, KI and KD control parameters is determined by a large number of simulations and experimental optimization. In the above three sets of experiments, the integral and differential constants are the same, KI is 0.6, KD is 0.01, and the proportional coefficient KP is different. The proportional coefficient Kp, rise time tr and overshoot σ% when the target flow rate Q is at 50 L/min, 60 L/min and 70 L/min are shown in Table 3.

Because the pressure-flow composite control method can modify the control parameters according to the system working conditions, the crane pressure-flow composite control electro-hydraulic system has higher flow control accuracy and better load adaptability.

5.2 Analysis of Control Characteristics of Different Set Pressure Differences

It can be known from the principle of the crane pressure-flow composite control electro-hydraulic system that for the same set flow rate, when the set pressure difference becomes smaller, the opening of the proportional multi-way valve will increase. Because the dynamic response of the proportional multi-way valve is higher than that of the electro-hydraulic proportional pump, the adjustment range of the proportional multi-way valve can be increased by reducing the set pressure, thereby improving the resolution of flow control, but reducing the set pressure difference. At the same time, it will cause the problem of insufficient system power.



1—2 Mpa Set test flow rate for differential pressure; 2—1.8 Mpa Set test flow rate for differential pressure; 3—1.5 Mpa Set test flow rate for differential pressure.

Fig. 9 Test flow curve under different set pressure difference

In the study, the set pressure differences of 1.5 MPa, 1.8 MPa and 2 MPa were selected for the test. The target flow rate is selected as 60 L/min, the load pressure is set as 5 MPa, and the PID control parameters are: $K_p = 0.01$, $K_i = 0.4$, $K_d = 0.01$ and run for 30 s. The output flow Q of the electro-hydraulic system under pressure-flow composite control is shown in Fig. 9.

It is calculated that under the set pressure differences of 1.5 MPa, 1.8 MPa and 2 MPa, the variances of the flow Q through the multi-way valve are: 5.35256, 6.42106 and 8.72645 respectively. The rise times of the second stage are 1.237 s, 1.226 s, and 1.03 s respectively.

It can be seen that when the set pressure difference increases, the rise time of the flow rate through the multi-way valve port decreases, and the power of the electro-hydraulic proportional pump increases; but the flow control accuracy decreases. When the set pressure difference decreases, the rise time of the flow rate through the multi-way valve port increases, and the dynamic response of the system decreases, but the flow control accuracy increases.

6 Conslusion

Taking the crane luffing system as the research object, by analyzing the electro-hydraulic flow control method, a crane pressure-flow composite control strategy was proposed, a crane pressure-flow composite control electro-hydraulic system test platform was built, and experiments and simulation studies were conducted. The research results show:

When the pressure-flow composite control electro-hydraulic system is at a micro-flow rate (≤70 L/min), the fluctuation of the flow rate during the dynamic change of the load is not greater than 9.35%. The impact of load changes on the pump control subsystem can be effectively suppressed by changing the closed-loop control parameters of the pump control subsystem. The disturbance of flow further improves the flow control accuracy and the load adaptability of the system.

By reasonably setting the present pressure difference of the pump control subsystem, the flow control accuracy and dynamic characteristics required by the system under current working conditions can be matched.

Owing to constraints in experimental conditions, the study focused on a hydraulic system featuring the coordinated pressure and flow control of a single actuator pump valve. Future research will explore flow distribution among multiple actuators.

Acknowledgements The authors gratefully acknowledge the financial support of Research on key technologies of excavator distributed independent electro-hydraulic control system (2021YFB2011905), and the Shanxi Province Basic Research Project (20210302123200).

References

1. Jiao Z, Wang Z, Liu X et al (2022) Multi-parameter load sensing pump model simulation and flow rate characteristics research. Chin J Aeronaut (English Edition) 296–308
2. Kong X (2005) Dynamic of load-sensing pump with pipe effect considered. J Eng Des 219–222
3. Tian X, Stump P, Vacca A et al (2021) Power-saving solutions for pre-compensated load-sensing systems on mobile machines. Trans ASABE 1435–1448
4. Zhang L, Wang L, Li D et al (2015) Principle analysis of the load sensing system with the front valve compensation. Chin Hydraul Pneum 63–67
5. Chang Y, Zhou J (2022) Analysis of pre-valve compensation flow anti-saturation technology in mine load sensing system. Mach Tool Hydraul 129–132
6. He Q, Hao P, Zhang D (2008) Modeling and parameter estimation for hydraulic system of excavator's arm. J CentL South Univ Technol (English Edition) 382–386
7. Li D, Zhu S, Zhao Y et al (2022) Electro-hydraulic control system for automatic anti-sticking of coal mine drilling rig. J Chongqing Univ 114–124
8. Gao L, Liu J, Li J et al (2013) Research of multi-way valve in working device of hydraulic excavator. Mach Tool Hydraul 149–150
9. Finzel R, Helduser S, Jang DS (2009) Electro-hydraulic control systems for mobile machinery with low energy consumption//ICFP2009, Hangzhou
10. Long Q, Lin T (1995) Theory and implementation of new principle of electro-hydraulic proportional charge control. Trans CSAM 83–87
11. Qiao J, Feng Z, Zhang Z (2021) Design of low speed fine flow electro hydraulic servo control system. Ind Control Comput 2021:26–27
12. Qian S, Zhang W, Li Q (2021) Flow control design and analysis of electro-hydraulic proportional valve based on fuzzy PID. Chin J Constr Mach 512–517
13. Du J (2021) Energy efficiency characteristics analysis for crane hydraulic system of pump-valve coordinated pressure-flow composite control. Taiyuan University of Technology, Taiyuan, pp 33–34
14. Wang A, Kuang L, Zhang X et al (2018) A study on flow coefficient of combined throttling groove in spool valves. J Xi'an Jiaotong Univ 110–117

15. Wang H, Huang J, Quan L et al (2018) Characteristics of proportional flow valve based on bilinear interpolation control strategy. J Mech Eng 287–296
16. Wen C, Li Y, Pan Q (2022) Simulation analysis and experiment of electric proportional bent-axis piston pump. Chin Hydraul Pneum 34–36
17. Chen N (2019) Reaserch on bodas based leafy harvester control system. Shanghai Jiao Tong University, ShangHai, pp 41–50
18. Li Z, Yue L, Yang J (2023) Research on the characteristics of independent control system with variable speed load sensitivity for import and export journal of Beihang University 3132–3144

Open Access This chapter is licensed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License (http://creativecommons.org/licenses/by-nc-nd/4.0/), which permits any noncommercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if you modified the licensed material. You do not have permission under this license to share adapted material derived from this chapter or parts of it.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.



Dynamic Path Planning Algorithm for Unmanned Inspection Vehicle Based on Improved DWA



Zhiguo Zhao, Siyu Wang, Wen Xiao, Rong Zhou, Zhen Xu, Dong Xie, Xiaokang Wan, and Yeqin Wang

Abstract Aiming at the problem that DWA algorithm is easy to falls into the local optimal solution, which cannot reach the target point smoothly, or even collide with obstacles, an improved dynamic path planning method for unmanned patrol vehicles based on DWA algorithm is proposed. Based on the original evaluation function, a gap evaluation subfunction for the judgment of unknown obstacles is added to predict the motion trend of dynamic obstacles in advance, and a new driving path is planned to avoid obstacles, to improve the obstacle avoidance ability of unmanned patrol vehicles. Moreover, the DWA algorithm is based on the node of the global path as the moving direction, so that the local path can fit the global path. Avoid situations where you cannot reach the target point. The results show that compared with the traditional DWA algorithm, the improved DWA algorithm proposed in this paper can predict the motion trend of dynamic obstacles in advance, plan a new driving path to avoid dynamic obstacles and solve the problem that the DWA algorithm reaches the target point due to the existence of locally optimal solutions.

Keywords Path planning · Dynamic window approach · Dynamic obstacles · Unmanned inspection vehicles

1 Introduction

As a kind of mobile robot with autonomous navigation ability, unmanned inspection vehicle is widely used in iron and steel plants, nuclear power plants, power plants, chemical plants, unmanned logistics and park inspection and other fields [1]. The unmanned inspection vehicle with the functions of autonomous movement, autonomous obstacle avoidance, autonomous navigation, and automatic data

Z. Zhao (✉) · S. Wang · W. Xiao · R. Zhou · Z. Xu · D. Xie · X. Wan · Y. Wang (✉)
School of Transportation Engineering, Huaiyin Institute of Technology, Huai'an, China
e-mail: zhaozg@hyit.edu.cn

Y. Wang
e-mail: wangyeqin@hyit.edu.cn

© The Author(s) 2026

H. Saman et al. (eds.), *The 9th International Conference on Advances in Construction Machinery and Vehicle Engineering*, Lecture Notes in Mechanical Engineering,
https://doi.org/10.1007/978-981-96-5527-4_96

upload can replace human beings to perform efficient and accurate inspection tasks in complex, changeable, and dangerous environments.

As a key technology for positioning and navigation of unmanned patrol vehicles, path planning algorithms can be divided into global path planning and local path planning. Global path planning focuses on the optimality and efficiency of the overall path, such as Dijkstra's algorithm [2]; A* algorithm [3]; local path planning focuses more on real-time and dynamic obstacle avoidance, such as Dynamic Window Approach (DWA) [4]; TEB algorithm [5].

Global path planning algorithms such as A* and RRT are usually suitable for static environments, while dynamic environment path planning requires local motion planning algorithms. Due to the good performance of the DWA algorithm in static environments, but also the poor performance in dynamic environments due to the lack of consideration of obstacle velocities, most scholars have chosen to improve upon the DWA algorithm.

To better realize path planning in a dynamic environment, a dynamic path planning method for unmanned inspection vehicles based on the improvement of the DWA algorithm is proposed to improve the obstacle avoidance ability of unmanned inspection vehicles by adding a gap evaluation subfunction for judging unknown obstacles based on the original evaluation function, which is used for predicting the motion trend of the dynamic obstacles in advance and replanning a new driving path to avoid the obstacles.

2 DWA Algorithm

The Dynamic Window Approach (DWA) is a local path-planning algorithm used for the autonomous navigation of unmanned patrol vehicles. It searches for the best speed instruction during path planning and transforms the position constraint problem into a speed constraint problem to avoid unknown obstacles.

2.1 Principles of the DWA Algorithm

The implementation principle of DWA algorithm is based on the dynamics model and environmental information of the unmanned vehicle to obtain the speed and pose information of the unmanned vehicle in its current state. A dynamic window is searched in the speed-turn space. For each speed instruction within the range of speed variation in the dynamic window, its motion trajectory is simulated. The motion trajectory is evaluated by the evaluation function, and the speed, distance proximity to target points and obstacles of each motion trajectory are evaluated and scored. Based on the score, the speed instruction of the optimal trajectory is selected as the next moving speed of the unmanned inspection vehicle. Judge whether the target point is reached by the trajectory, if it is reached, the speed and position information

of the optimal trajectory is transferred to the chassis of the unmanned inspection vehicle; if the target point is not reached, the sampling and selection of the speed is carried out again.

2.2 Chassis Speed

The speed of each wheel in the chassis of the unmanned inspection vehicle can be calculated by the encoder value, and the overall motion speed of the chassis can be obtained by knowing the wheel speed. Through the speed analysis of each wheel and chassis during the movement of McNamee wheel chassis, its omnidirectional kinematics model is shown in Fig. 1. Speed was analyzed, as shown in Fig. 2.

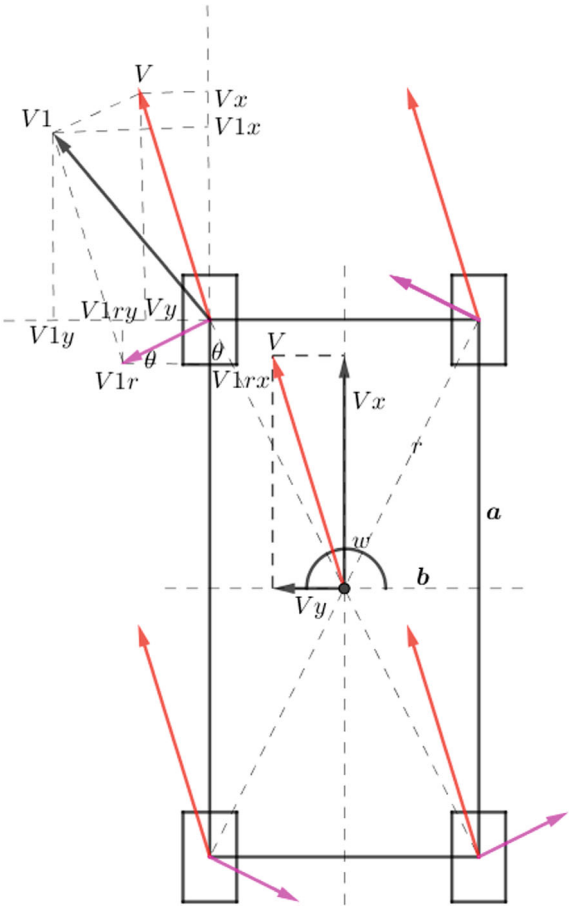


Fig. 1 Omnidirectional kinematic model

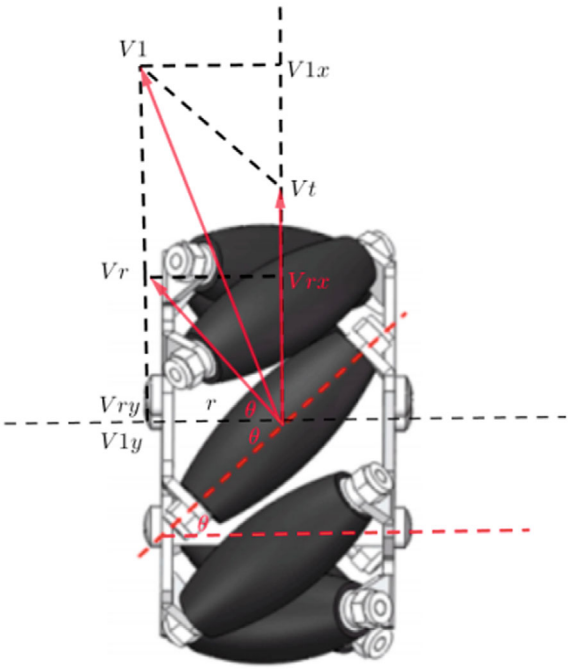


Fig. 2 Velocity vector analysis

In Fig. 1, the motion speed of the chassis is controlled by the linear speed V and the angular speed w, and the speed of each wheel can be broken down into two components, namely the linear speed and the angular speed components of the chassis. Vector v_r is generated when the roller of the McNamee wheel is driven, not generated by power, and its value cannot be determined. Speed analysis of the front left wheel of the chassis. It is known that r is the distance from the center point of the chassis of the unmanned inspection vehicle to the center point of each McNamee wheel, that is the radius of the arc. The distance between the front and rear wheels of the chassis is 2a, and the distance between the left and right wheels is 2b. Using the trigonometric function relationship, the proportional relationship of each side can be determined:

$$V_1 = \begin{bmatrix} V_{1x} \\ V_{1y} \end{bmatrix} = \begin{bmatrix} V_x - V_{1rx} \\ V_y + V_{1ry} \end{bmatrix} = \begin{bmatrix} V_x - bw \\ V_y + aw \end{bmatrix} \tag{1}$$

In Fig. 2, it is known that the wheel contains two degrees of freedom, tangential and lateral. The tangential speed component V_t is provided by the motor, and the lateral speed component V_r is generated when the roller is driven, so the wheel speed V_1 can be represented by V_t and V_r. Using the constant relationship between the components, the relationship between the motor drive wheel linear velocity V_t and the overall chassis velocity V and w can be found, Thus, the relationship between $V_{t1,}$

V_{t2}, V_{t3}, V_{t4} and the overall linear and angular velocity of the chassis is as follows:

$$V_{t1} = V_x - V_y - (a+b)w \tag{2}$$

$$V_{t2} = V_x + V_y + (a+b)w \tag{3}$$

$$V_{t3} = V_x + V_y - (a+b)w \tag{4}$$

$$V_{t4} = V_x - V_y + (a+b)w \tag{5}$$

Equations (2)–(5) are the inverse kinematics formula of McNamee wheel chassis, and the inverse transformation is the forward kinematics formula of McNamee wheel omnidirectional chassis. The transformation process is as follows:

$$V_x = \frac{V_{t1} + V_{t2} + V_{t3} + V_{t4}}{4} \tag{6}$$

$$V_y = \frac{-V_{t1} + V_{t2} + V_{t3} - V_{t4}}{4} \tag{7}$$

$$-V_{t1} + V_{t2} - V_{t3} + V_{t4} = 4(a+b)w \tag{8}$$

$$w = \frac{-V_{t1} + V_{t2} - V_{t3} + V_{t4}}{4(a+b)} \tag{9}$$

The forward kinematics is obtained by solving the motion speed of the chassis through the speed of each McNamee wheel. Convert the above formula into matrix form as follows:

$$\begin{bmatrix} V_x \\ V_y \\ w \end{bmatrix} = \frac{1}{4} \begin{bmatrix} 1 & 1 & 1 & 1 \\ -1 & 1 & 1 & -1 \\ -\frac{1}{a+b} & \frac{1}{a+b} & -\frac{1}{a+b} & \frac{1}{a+b} \end{bmatrix} \begin{bmatrix} V_{t1} \\ V_{t2} \\ V_{t3} \\ V_{t4} \end{bmatrix} \tag{10}$$

2.3 Evaluation Function

The evaluation function of the dynamic window method is the core of the local path planning algorithm. When the speed of the unmanned patrol vehicle is sampled, the motion trajectory will be generated in the dynamic window according to the sampling speed. The existence of the evaluation function is to score all the motion trajectories in the dynamic window based on the evaluation criteria. The evaluation function of DWA local path planning algorithm is:

$$G(v, w) = \sigma\left[\alpha.heading(v, w) + \beta.dist(v, w) + \gamma.velocity(v, w)\right].iscollide(v, w) \tag{11}$$

where σ is the smoothing coefficient, α, β, γ are the evaluation function coefficients, *heading*(*v*,*w*) is the azimuth rating function, *dist*(*v*,*w*) is the obstacle distance rating function, and *velocity*(*v*,*w*) is the speed rating function.

The normalization of evaluation function can make the value range of evaluation function fixed in a finite interval, avoid too large or too small evaluation value, and effectively improve the stability and convergence of the algorithm.

$$normal_head(i) = \frac{head(i)}{\sum_{i=1}^{n} head(i)} \tag{12}$$

$$normal_dist(i) = \frac{dist(i)}{\sum_{i=1}^{n} dist(i)} \tag{13}$$

$$normal_velocity(i) = \frac{velocity(i)}{\sum_{i=1}^{n} velocity(i)} \tag{14}$$

3 Dynamic Path Planning Method Based on Improved DWA Algorithm

When the path planning of unmanned patrol vehicles is based on DWA algorithm, the algorithm may have a local optimal path, so it cannot reach the target point, as shown in Fig. 3.

Based on the same starting point and the target point, Dijkstra's algorithm successfully planned a global optimal path, but DWA algorithm was easy to fall into the local optimal solution, failing to successfully reach the target point and even collision with obstacles. Therefore, an improved dynamic path planning method based on DWA algorithm was proposed. Dynamic obstacle prediction is added to the original evaluation function to improve the obstacle avoidance ability of unmanned patrol vehicles, and the DWA algorithm is based on the moving direction of the nodes of the global

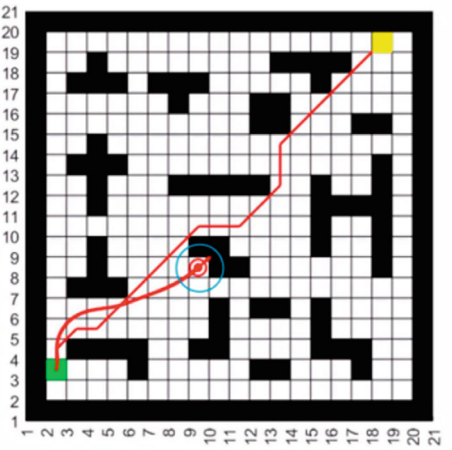


Fig. 3 DWA localized path planning diagram

path, so that the local path fits the global path, and the situation of failing to reach the target point is avoided.

3.1 Overall Idea of the Improved Algorithm

First, the speed of the unmanned patrol vehicle is obtained by sampling the information from inertial measurement unit, odometer and LIDAR parameters. The speed of the dynamic obstacle is sampled by continuously firing laser around the LIDAR probe. By judging whether there is a possibility of collision between the unmanned patrol vehicle and the dynamic obstacle, the judgment basis is that if the Angle formed by the direction vector of the unmanned patrol vehicle and the dynamic obstacle is less than 180°, then there is a possibility of collision, the collision point is predicted, and the traveling speed and direction of the unmanned patrol vehicle are optimized. The dynamic window method is used to randomly generate several driving paths, and each path is evaluated by evaluation function. The final path generated when the unmanned patrol vehicle traverses to the endpoint is judged. If it approaches the global path and does not collide with dynamic obstacles, it is selected as the optimal path to end the path planning.

3.2 Improvement of the Evaluation Function

When the unmanned inspection vehicle carries out path planning in complex environments, the weight β of the known obstacle clearance evaluation subfunction $\beta.dist(v,w)$ will be reduced, at which time the vehicle is unable to avoid obstacles

in time, so the unknown obstacle clearance evaluation subfunction *iscillide*(*v*,*w*) is added to be used for predicting dynamic obstacles in the path planning process, to improve the obstacle avoidance ability of the unmanned inspection vehicle.

The improved rating function is:

$$G(v, w) = \sigma\left[\alpha.heading(v, w) + \beta.dist(v, w) + \gamma.velocity(v, w)\right].iscollide(v, w) \tag{15}$$

Normalized smoothing of the rating function:

$$normal_head(i) = \frac{head(i)}{\sum_{i=1}^{n} head(i)} \tag{16}$$

$$normal_dist(i) = \frac{dist(i)}{\sum_{i=1}^{n} dist(i)} \tag{17}$$

$$normal_velocity(i) = \frac{velocity(i)}{\sum_{i=1}^{n} velocity(i)} \tag{18}$$

$$normal_iscollide(i) = \frac{iscollide(i)}{\sum_{i=1}^{n} iscollide(i)} \tag{19}$$

Take the predicted collision point P as the origin to establish the plane right-angle coordinate system, and use the translation property to make the collision point P as the end point and the start point of the unmanned inspection vehicle's speed vector $\vec{V}_c$ and the dynamic obstacle's speed vector $\vec{V}_b$, respectively, and use the triangular vector rule to get the combined speed vector $\vec{V}_s$; select the optimal speed instruction through several trajectories of the dynamic window and send it to the bottom disk, and then calculate the speed of the four wheels in the bottom disk of the unmanned inspection vehicle, and realize the optimization of the traveling direction by controlling the rotational speed. The rotation speed of the four wheels in the bottom tray of the unmanned inspection vehicle is calculated, and the optimization of the traveling direction is realized by controlling the rotation speed. The specific algorithm is as follows:

$$\begin{bmatrix} \omega_A \\ \omega_B \\ \omega_C \\ \omega_D \end{bmatrix} = \frac{1}{R} \begin{bmatrix} 1 - 1 - (x_1 + x_2) \\ 11(x_1 + x_2) \\ 11 - (x_1 + x_2) \\ 1 - 1(x_1 + x_2) \end{bmatrix} \begin{bmatrix} V_{cx} \\ V_{cy} \\ \omega \end{bmatrix} \tag{20}$$

where w_A, w_B, w_C, w_D are the rotational speeds of the corresponding chassis wheels, respectively.

4 Analysis of Physical Simulation and Experimental Results

To verify the feasibility of the improved algorithm, this paper compares the global path planning algorithm and the improved DWA algorithm from multiple dimensions in the 20 × 20 raster map built.

The unmanned inspection vehicle local path planning is simulated based on the robot motion model, the obstacle radius safety domain value R = 0.6 is set, and the weights *a*, β, and γ are set to (0.3, 0.5, and 0.2), respectively, and the parameters of the improved DWA algorithm simulation are shown in Table 1.

The improved DWA local path planning algorithm takes the nodes provided by the Dijkstra global path planning algorithm as the goal points at the next moment, the more the number of nodes and the longer the path, the less efficient the algorithm is. As shown in Fig. 4, the number of nodes is reduced by smoothing and optimizing the path of the global path planning algorithm, which makes the path of the improved DWA algorithm shorter and more relevant to the global path.

The simulation results after smoothing and optimizing the global optimal path of Dijkstra's algorithm are shown in Table 2. The smoothed optimization process results in fewer global path nodes, fewer steering times, and shorter path lengths.

The global path after smoothing optimization decreases the number of nodes, total angle of inflection, and path length by 36.4%, 36.4%, and 3.86%, respectively, compared with the pre-optimization. The reduction in the number of nodes can effectively reduce the local path length of the improved DWA algorithm. The simulation results of the improved DWA algorithm in a 20 × 20 raster map are shown in Fig. 5, where the black part is the known static obstacles, and the lower-left and upper-right corners are the start and target points of path planning, respectively.

The gray grid in Fig. 5 is the newly added static obstacles and the static obstacles are arranged on the globally optimal path. The yellow grids are dynamic obstacles moving uniformly from top to bottom and from left to right. The simulation results of the improved DWA algorithm are shown in Table 3.

The improved DWA algorithm can avoid dynamic obstacles by predicting the movement trend of dynamic obstacles in advance and planning a new driving path. Therefore, the path length is increased, but the unknown static and dynamic obstacles

Table 1 Simulation parameters of DWA algorithm

Parameters	Numerical value	Unit (of measure)
Maximum line speed	2.0	m/s
Maximum angular velocity	0.4	rad/s
Accelerations	0.01	m/s^2
Rotational acceleration	1.0	rad/s^2
Speed resolution	0.28	m/s
RPM resolution	0.01	rad/s

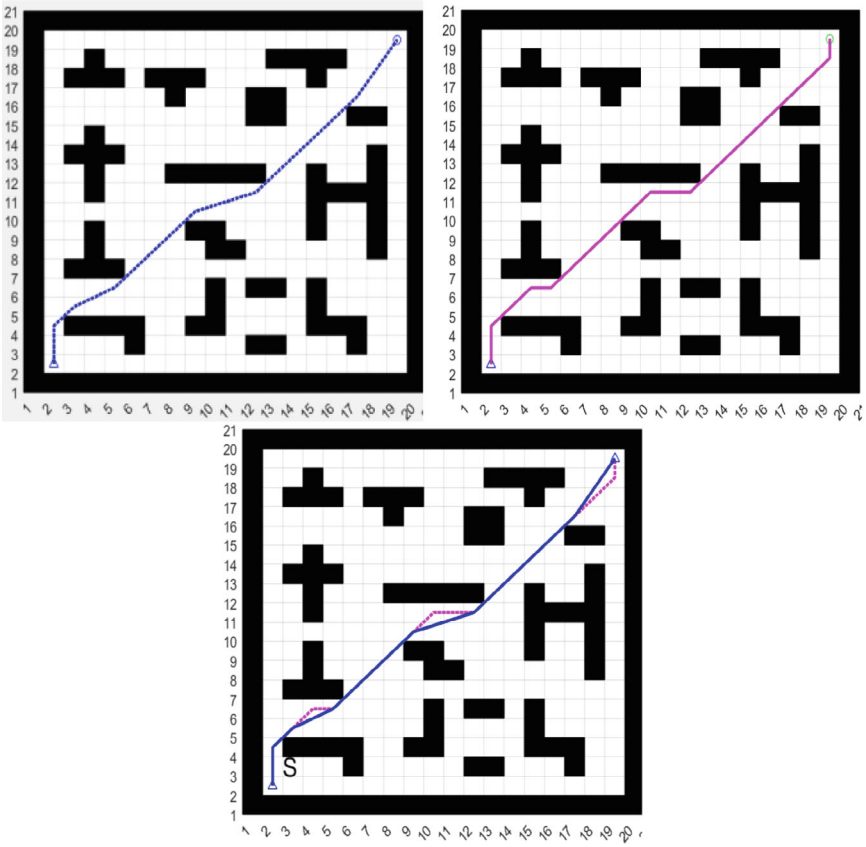


Fig. 4 Global path smoothing optimization

Table 2 Simulation results of global path smoothing optimization

	Number of nodes (pc)	Total turning angle (°)	Path length (m)
Pre-optimization	11	495	25.796
One-stop optimization	9	405	25.146
Secondary optimization	7	315	24.799

can be effectively avoided, as shown in the green circle in the figure, which improves the obstacle avoidance ability of the algorithm. In addition, the local path moves in the direction of the node provided by the global path, such as the red round node in the figure, which solves the situation that the DWA algorithm reaches the target point due to the existence of a locally optimal solution.

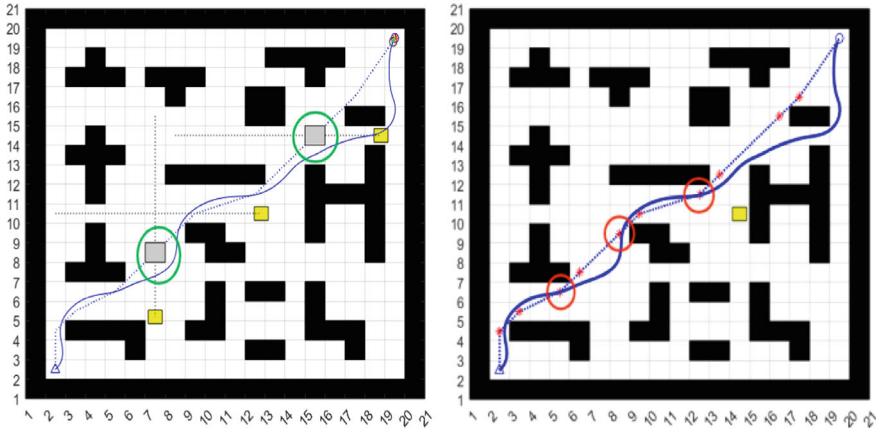


Fig. 5 Localized path planning diagram of the improved DWA

Table 3 Simulation results of the improved DWA algorithm

	Path length	Whether the target point is reached or not	Obstacle avoidance or not
Dijkstra	25.7990	Be	Be
DWA	26.9386	Clogged	Be
Improved DWA	27.7407	Be	Be

5 Conclusion

In this paper, we optimize and improve the traditional DWA local path planning algorithm, propose a dynamic path planning method based on the improvement of DWA algorithm, and introduce a new evaluation sub-function in its evaluation function, which is used to improve the unmanned inspection vehicle's obstacle avoidance ability to obstacles in unknown environments, and let the node of the DWA algorithm based on the global path be the moving direction, to avoid not being able to reach the target point. The experimental results show that the improved algorithm can improve the obstacle avoidance ability and the planned path can fit the global path better. Future research can further explore and develop these problems, mainly including the following two points: (1) Environmental adaptability can be improved by adjusting the weight combination of the dynamic window method, and corresponding optimization can be carried out according to the actual distribution of obstacles. (2) Machine learning methods such as reinforcement learning can be applied to path planning to achieve more intelligent and adaptive path planning.

References

1. Liu L, Wang X, Yang X, Liu H, Li J, Wang P (2023) Path planning techniques for mobile robots. Review and prospect. Expert Syst Appl
2. Shu-Xi W (2012) The improved Dijkstra's shortest path algorithm and its application. Procedia Eng 29:1186–1190
3. Erke S, Bin D, Yiming N, Qi Z, Liang X, Dawei Z (2020) An improved A-Star based path planning algorithm for autonomous land vehicles. Int J Adv Robot Syst 17(5)
4. Li T, Ya R, We G, Su L (2019) Local path planning algorithm for blind-guiding robot based on improved DWA algorithm. In: 2019 Chinese control and decision conference (CCDC), pp 6169–6173
5. Wu J, Ma X, Peng T, Wang H (2021) An improved timed elastic band (TEB) algorithm of autonomous ground vehicle (AGV) in a complex environment. Sensors 21(24)

Open Access This chapter is licensed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License (http://creativecommons.org/licenses/by-nc-nd/4.0/), which permits any noncommercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if you modified the licensed material. You do not have permission under this license to share adapted material derived from this chapter or parts of it.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.



Simulation Analysis of Lifting Equipment for Underwater Robot



Xin Wang, Xinyu Li, Tianchuan Pan, Hao Li, Sisi Zhu, and Zetao Li

Abstract In order to solve move-in and move-out of a container for an underwater robot, a lifting equipment is proposed and designed. The mainly focuses in this paper are its working principle, structural design, structural finite element analysis and action simulation. Multi-level planar truss is designed and integrated with the container so as to ensure the distance to the outside of the container and all back inside of the container. Longitudinal movement by 2nd level and 3rd level motors set on the correspond planar truss, and transverse movement by electric hoist set on the 3rd level beam can cover all space of the container. According to finite element analysis, the structure meets the strength and stability requirements. The motion simulation of the lifting equipment is carried out and shows its feasibilities of all kinds of movements. After that, a test prototype is manufactured and physical load tests are carried out to further verify its reliability and practicability as well as its theoretical model.

Keywords Underwater robot · Lifting equipment · Lifting plan · Finite element analysis

1 Introduction

With the rapid development of national economy and science and technology, the construction of hydropower stations in China has also been developed rapidly [1]. From the time the hydropower station is put into use, the concrete surface of the dam

X. Wang (✉) · Z. Li
Dalian University of Technology, Dalian, China
e-mail: wangxbd21@163.com

X. Li · S. Zhu
China Yangtze Power Co., Ltd., Yichang, China

T. Pan · H. Li
Dalian YILIY Construction Machinery Co., Ltd., Dalian, China

© The Author(s) 2026

H. Saman et al. (eds.), *The 9th International Conference on Advances in Construction Machinery and Vehicle Engineering*, Lecture Notes in Mechanical Engineering,
https://doi.org/10.1007/978-981-96-5527-4_97

will appear with the use of time, such as erosion, dissolution, cavitation, freezing and thawing and other diseases [2], and even cracks and other problems, so the work of underwater defect detection of the dam is urgently needed [3, 4]. Then, underwater detection robot becomes the research focus. The lifting equipment in this paper is designed for underwater robots to move-in and move-out containers. Underwater robots are usually transported to the hydroelectric dam by transport vehicles or container carriers [5], and in order to facilitate the entry and exit of the container, the design of the lifting equipment integrated inside the container is one of the effective ways. So, comprehensively compared several ones, the lifting scheme placed on the top of the container is determined. Through the finite element simulation and motion simulation, the proposed scheme is feasible. Furthermore, through the load test, the effectiveness of the proposed lifting equipment is verified. It provides favorable support conditions for the underwater robot into the water smoothly for inspection.

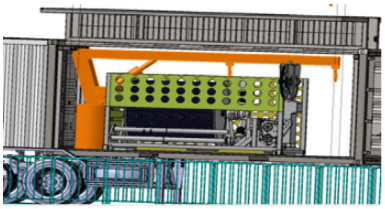
2 Designing Scheme of Lifting Equipment

The lifting system realizes to transport an underwater robot between different hydropower stations, and lifting and transfer the underwater robot between dam and water surface when arriving the hydropower station [6–8]. For integrated the lifting with the transportation, a lorry crane may be selected, as shown in Fig. 1 [9, 10]. In this way, the underwater robot and all auxiliary components, such as umbilical cables and drum, can be placed on the open carriage of the crane when transporting. Arriving at the dam, the crane directly lifts the underwater robot to the water. It is easy to implement transportation and lifting of underwater robot, but sometimes it is limited to near to the dam, thus affecting its applicability. In addition, the lifting rope length of lorry crane is limited, so it is difficult to be applied to dams with a height of up to 100 m.

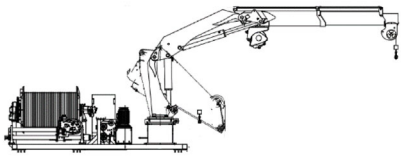
For easy transportation and integration of the control room, container transportation can be considered. The inside of the container can be divided into a manual

Fig. 1 Lorry crane





(a) Upper part of lorry crane with straight boom in front of a container



(b) Upper part of lorry crane with folding boom at the end of a container

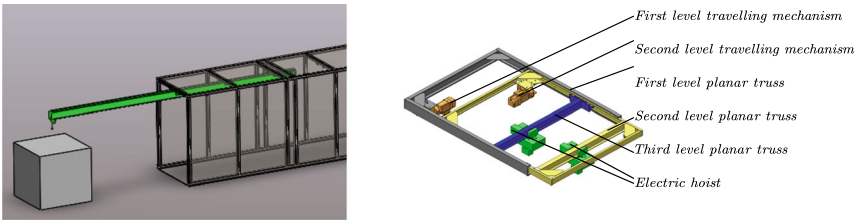
Fig. 2 Upper part of lorry crane inside the container

operational room and a storage room. In the operational room, operators can monitor and control the underwater robot. While the storage room can be applied to store the underwater robot and its accessories. Moreover, the container occupies a small space and can be conveniently placed along the dam.

Because of the robot's large size and self-weight, in order to smoothly move the underwater robot into and out of the container, a lifting equipment is required. In current available products, it is usually to place upper part of a lorry crane in the front of or at the end of the container, as shown in Fig. 2. But its slewing structure and jib take up a large space, so it is difficult to perform the role of lorry crane.

From above, the storage space of the container and the functional requirements of access for the underwater robot are further analyzed. Underwater robots and the drum wrapped by umbilical cables occupy larger space, and are obviously fixed on the ground of the container. Then the top space of the container is cleared, so we can integrate the lifting equipment on the top. In order to realize in and out of the container, the lifting equipment should have a telescopic function. When not in use, it can be all retracted back to the container, and when putting in use, it can be extended to the container outside. According to the structural style of lifting machinery, two kinds of structures can be adopted, one is telescopic boom with box section, and the other one is multi-level planar truss. Due to compact dimensions, the telescopic boom can be placed in the center axis of the container, but its working range cannot cover the entire space in the container, as shown Fig. 3a). However, the multi-level planar truss can be set up with transverse and travelling mechanisms to realize the coverage of the whole space of the container, and can realize the space demand of the container outside. According to above analysis, the multi-level planar truss is selected as the structure of the lifting equipment.

In order to realize more working space outside the container, three-level planar truss is adopted, as shown in Fig. 3b). The first level planar truss is fixed with the steel frame of the container, and the second level planar truss inside the first level one, can be moved longitudinally along the container by a motor. The third level one is a flanged beam, and can be moved longitudinally by a motor inside the second level ones. One electric hoist is installed and can move transversely on the third level ones. So, the movements of the two planar trusses and electric hoist can cover the whole space of the container, and the underwater robot can be placed in any position



(a) Telescopic boom with box section **(b)** Multi-level planar truss

Fig. 3 The structure style of lifting equipment placed on top of the container

of the container. In order to move the underwater robot smoothly, the single lifting point is changed to double lifting points, and one additional set of third level flanged beam and electric hoist is to be set up.

3 Finite Element Analysis of Lifting Equipment Structure

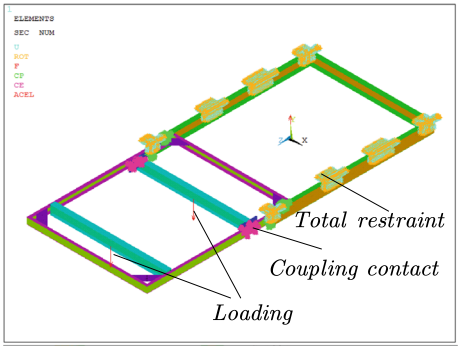
In order to meet its load carrying capacity and stability, the structure of the lifting equipment is analyzed by finite element method.

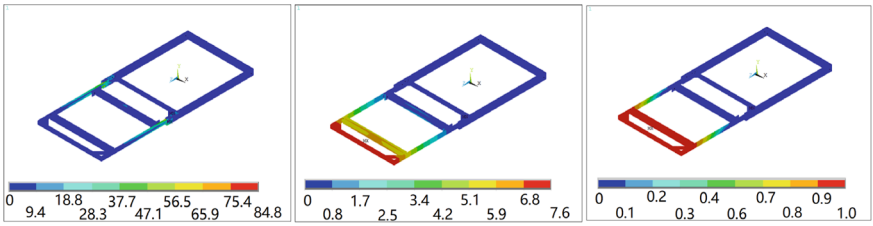
The structural finite element model is built according to the state that the second level planar truss is fully extended and the flanged beam is located at the end of the second level ones. The structure is composed of steel plates, so the finite element model is simulated by shell element 181. Sime small chamfers and fillets and narrow process edges of the structure are ignored. Triangles and quadrilaterals are selected to mesh the model.

Static displacement constraints are applied at the connections between the first level planar truss and the container steel frame, and coupling constraints are applied at the contacts between the second level planar truss and the first level planar truss, also between the second level planar truss and the flanged beam. The applied load is 400 kg self-weight of the underwater robot and its packing box, and the load factor 1.25 is also considered. So the load is total 500 kg, and located two positions which are the midspan and end span of the flange beam. The built finite element model is shown in Fig. 4.

The distributions of stress, displacement and stability are shown in Figs. 5, 6 and Table 1. It is shown that the maximum stresses of two load positions are 84.8 MPa and 115.4 MPa respectively, and all in overlap location both the second level planar truss and first level ones, and not more than the allowable stress 227 MPa. The maximum displacements of two load positions are 7.6 mm and 10.2 mm respectively, and all in the second level planar truss outside. The first eigenvalues of two load positions are 67.2 and 58.2 respectively, and do not lost their stabilities. Therefore, the designed structure of the lifting equipment meets the working requirements.

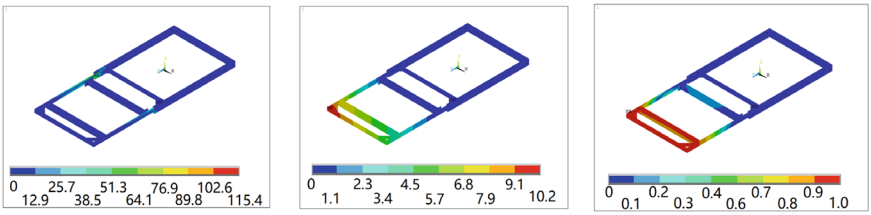
Fig. 4 Finite element model





(a) Stress cloud **(b)** Displacement cloud **(c)** Stability cloud

Fig. 5 Finite element cloud when the load is located in the middle of the span



(a) Stress cloud **(b)** Displacement cloud **(c)** Stability cloud

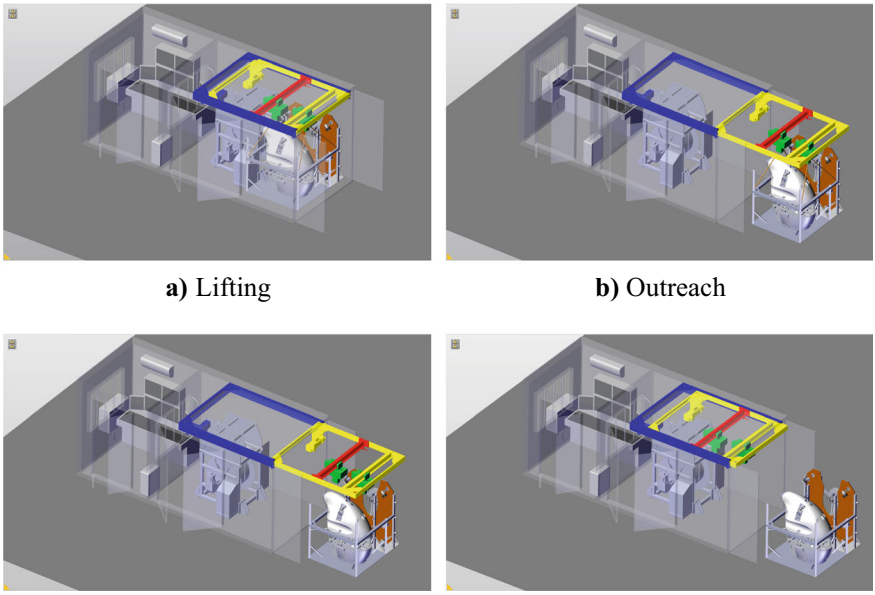
Fig. 6 Finite element cloud diagram when the load is located at the span end

Table 1 Analysis results

Load position at the flanged beam	Max. stress (MPa)	Max. displacement (mm)	First eigenvalue
Mid span	84.8	7.6	67.2
End span	115.4	10.2	58.2

4 Motion Simulation of the Lifting Equipment

In order to verify its correctness, the lifting equipment motion is simulated by Solidworks software, mainly including 5 steps. Firstly, the flanged beam and the electric hoists are moved so as to connect the riggings with the lifting points of the underwater robot box. Secondly, the underwater robot and its packing box are lifted slowly and smoothly by the two electric hoists as a whole, and left from the container ground. Thirdly, they are moved out of the container by the first level motor. Fourthly, after arriving at the target position, they are slowly lowered to the ground through the electric hoist. Finally, the lifting points are disconnected, and the lifting equipment returns back into the container through the first level motor, and the electric hoist hooks are returned to the original positions. The simulation screenshots, as shown in Fig. 7, show that the movements of the lifting equipment meet the functional requirements of the underwater robot to move-in and move-out the container.



a) Lifting **b)** Outreach

c) Lowering **d)** Back to the container

Fig. 7 Motion simulation



a) Prototype **b)** Load test at midspan **c)** Load test at end span

Fig. 8 Lifting equipment prototype and load test

Table 2 Comparison of test and calculation results

Test loads	Load (kg)	Vertical displacement of the second level planar truss outstretched ends (mm)		Error (%)
		Test value	Calculation value	
1.25 times rated load at midspan	500	8.5	7.6	− 10.6
1.25 times rated load at end span	500	11.3	10.2	− 9.7

5 Load Test

According to the design scheme of the lifting equipment, the prototype was made, as shown in Fig. 8, and the midspan load test and end span load test were carried out. Also the motion performance of each mechanism was tested. Due to the limit test conditions, only the structural displacements of the two load tests when the secondary level planar truss outstretched of the container. The test results compared with calculation results are shown in Table 2. It is shown that the calculation results are slightly smaller than the test results. The reasons of different results may be related to the ideally modelling and the possible errors of actual manufacture and installation. However, the two values are close. Therefore, the reasonableness and effectiveness of the design scheme is further verified through the test. Next, the integrated test with the underwater robot will be carried out.

6 Conclusion

In this paper, for the lifting work of underwater robots, lifting equipment with multi-level planar truss on the top of a container is comprehensively compared and selected. The feasibility of the truss structure is confirmed through finite element simulation and motion simulation, and the rationality and the effectiveness of the lifting equipment are further verified through prototype and load tests.

Acknowledgements This research was supported by the National Key Research and Development Program of China (Project No. 2022YFB4703405).

References

1. Xu P, Chen M, Kai Y et al (2023) Research progress of underwater robot for dam body inspection in large hydropower station. J Tsinghua Univ (Nat Sci Ed) 63(07):1032–1040
2. Chen Q, Ding Y, Jiang T (2023) Analysis of hydroelectric power plant and dam inspection technology based on unmanned underwater robot. Robot Ind 6:94–98
3. Hirai H, Ishii K (2019) Development of dam inspection underwater robot. In: International conference on artificial life & robot. Alife Robot Corp Ltd, International Steering Community; Chinese Association Artificial Intelligence, pp 528–531
4. Kramarchuk A et al (2020) The research bearing capacity of crane beams for possible establishment of bridge crane on them. In: Lecture notes in civil engineering, pp 202–210
5. Crane: a versatile truck-mounted tool. Public Works 119(12):62–63 (1988)
6. Li C, Wang Y, Xu X (2022) Design of boom and telescopic mechanism of truck-mounted crane. Constr Mach 6:108–111
7. Sugimoto H, Moriya Y, Ogasawara T (2017) Underwater survey system of dam embankment by remotely operated vehicle. Haeundae, Busan, Korea
8. Zhang L, Jiang D, Li Y (2024) Design and application of a multi-lift truck-mounted crane. Constr Mach 55(03):110–114+10
9. Xing (2022) Design of boom and telescopic mechanism of truck-mounted crane. Constr Mach 6:108–111
10. Li C, Zhang Y, Lian D (2021) Optimisation of crane structure design based on ANSYS. Times Autom 12:111–112

Open Access This chapter is licensed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License (http://creativecommons.org/licenses/by-nc-nd/4.0/), which permits any noncommercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if you modified the licensed material. You do not have permission under this license to share adapted material derived from this chapter or parts of it.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.



Analysis of Steering Control of a Five-Axle All-Terrain Crane Considering Body Rolling



Tingting Zhang, Xin Wang, Xifeng Zhang, and Shuai Song

Abstract Aiming at the problem of poor steering stability of five-axle all-terrain cranes, a 3-DOF model was established on the basis of the classical 2-DOF model, taking into account the body's lateral inclination movement. A 5-axis steering model was built in Adams to verify the accuracy of the 3-DOF model. Based on the two control strategies of first-axis corner feedforward control and optimal control based on vehicle state feedback and proportional feedforward, the two control strategy models are established in Matlab/Simulink, and the effectiveness of the control strategies is verified through the angular step test. The simulation results show that the optimal control is the most effective, the steady state value of the center of mass lateral deflection angle is close to zero, and the vehicle is consistent with the driving direction, which enhances the anti-skidding ability; compared with the feed-forward control, the vehicle traverse angular velocity under the optimal control is basically the same as that of the first-axis steering; minimized lateral inclination under optimal control ensures good crane overtures.

Keywords Multi-axis steering · 3-DOF model · Feedback control · Simulation

1 Introduction

Five-axle all-terrain cranes have a large steering radius, poor maneuverability and flexibility, poor stability when driving, and are prone to tail-swinging and other phenomena. In order to improve its maneuvering performance, at present, the domestic mostly use mechanical steering and electronic control combination, the method can improve maneuverability and reduce tire wear [1].

T. Zhang · X. Wang (✉)
Dalian University of Technology, Dalian, China
e-mail: wangxbd21@163.com

X. Zhang · S. Song
Taiyuan Heavy Industry Co., Ltd., Taiyuan, China

© The Author(s) 2026

H. Saman et al. (eds.), *The 9th International Conference on Advances in Construction Machinery and Vehicle Engineering*, Lecture Notes in Mechanical Engineering,
https://doi.org/10.1007/978-981-96-5527-4_98

Li [2] and others designed an optimized tracking control strategy to reduce transient understeer and oversteer under vehicle steering conditions by controlling the steering angle of the rear axle, which improves the steering stability of the vehicle; Zhang [3] proposed to link the steering models of multi-axle and four-wheeled vehicles and designed a static-free tracking controller based on a 2-DOF vehicle steering model, which improves the safety and maneuvering stability of all-terrain cranes; Shet et al. [4] proposed an adaptive single-input fuzzy logic control based on Lyapunov's stability theorem. The proposed control scheme is robust to the effects of parameter uncertainty, exogenous noise and perturbations. The overall control scheme of Lyapunov's theory ensures the global asymptotic stability of an autonomous vehicle; Sanchez et al. [5] proposed a novel extremum-seeking controller for stabilizing the tracking trajectory of an all-drive four-wheeled skid-steer vehicle, and verified the effectiveness of the proposed design by comparing the sliding-mode based control action with the conventional state feedback formulation.

This paper takes five-axle all-terrain crane as the research object, establishes the 3-DOF model of all-wheel steering by considering the crane's lateral motion, and compared with Adams simulation to verify the reliability of the 3-DOF model. Based on the two control strategies of proportional feed-forward control of the first axis corner and optimal control based on the crane state feedback and proportional feed-forward, the two control strategy models are established in Matlab/Simulink, and the control effect of the control strategy on the five-axis all-terrain crane is verified through the angular step test.

2 Five-Axis All-Terrain Crane Steering Model

2.1 2-DOF Model

According to the automobile theory, the 5-axis all-terrain crane can be simplified as 2-DOF model, considering only the lateral motion along the y-axis and the transverse pendulum motion around the z-axis, as shown in Fig. 1.

$$\begin{cases} mv(\dot{\beta}+\omega_r) = \sum_{i=1}^{3} k_i\left|\beta+\frac{l_i\omega_r}{u}-\delta_i\right| + \sum_{i=4}^{5} k_i\left|\beta-\frac{l_i\omega_r}{u}-\delta_i\right| \\ I_z\dot{\omega}_r = \sum_{i=1}^{3} k_i l_i\left|\beta+\frac{l_i\omega_r}{u}-\delta_i\right| - \sum_{i=4}^{5} k_i l_i\left|\beta-\frac{l_i\omega_r}{u}-\delta_i\right| \end{cases} \tag{1}$$

where k_i is the lateral deflection stiffness; β is the center of mass lateral deviation angle; and ω_r is the cross-swing angular velocity.

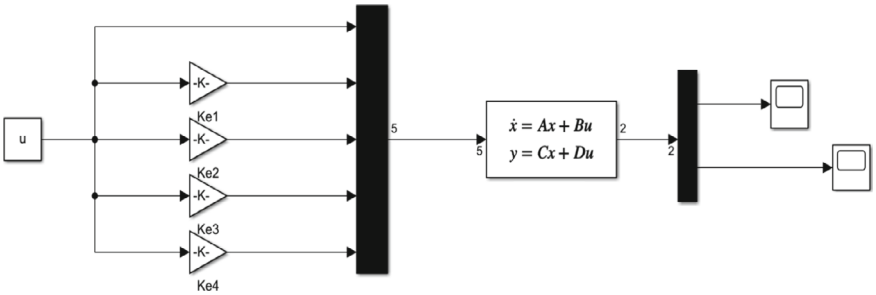


Fig. 1 2-DOF model

2.2 3-DOF Model

Although the 2-DOF model can study the basic characteristics of steering, it can only reflect the longitudinal motion and transverse pendulum motion when steering, when it comes to the crane rollover, it is necessary to take into account the body tilting motion, so it is necessary to establish 3-DOF model that includes the body tilting to reflect the steering characteristics of the crane. Based on the traditional 2-DOF model, 3-DOF model is established by considering the body roll [6], as shown in Fig. 2.

Assuming a constant longitudinal speed of the Crane, linear tire side deflection characteristics, the dynamics equations of the 3-DOF model are:

$$\begin{cases} (m-m_s)(\dot{v}_y + v_x\omega_r) + m_s(\dot{v}_y + v_x\omega_r - h_s\ddot{\phi}) = \sum_{i=1}^{5} F_{yi} \\ I_z\dot{\omega}_r + I_{xz}\ddot{\phi} = \sum_{i=1}^{3} l_i F_{yi} - \sum_{i=4}^{5} l_i F_{yi} \\ I_x\ddot{\phi} + I_{xz}\dot{\omega}_r - m_s h_s(\dot{v}_y + v_x\omega_r - h_s\ddot{\phi}) = m_s h_s g\sin\phi - K\phi - C\dot{\phi} \end{cases} \tag{2}$$

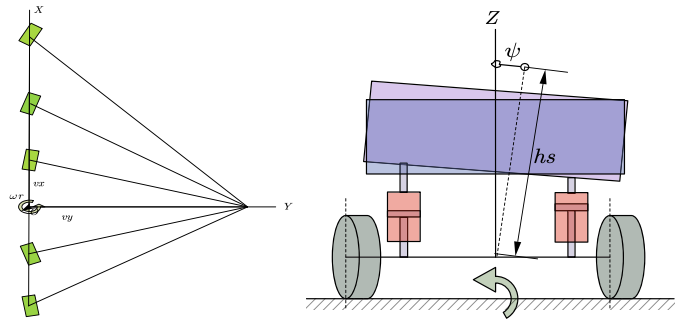


Fig. 2 3-DOF model

where, F_y is the lateral force on the tire; ϕ is the angle of lateral inclination, and g is the acceleration of gravity; δ is the angle of rotation of the first wheel.

2.3 Comparative Analysis of Adams Dynamics

A five-axis steering system model is established in Adams, as shown in Fig. 3, and the angular step simulation is carried out on the steering system model, given that the angle input of the first axis is 0.25 rad, the speed of the crane is 5 m/s, and the simulation time is 10 s. After obtaining the simulation results, the data is exported, and the comparison with the theoretical results is carried out, which verifies the accuracy of the 3-DOF model. The structural parameters used for the simulation are shown in Table 1.

The basic parameters measured are the center of mass lateral deflection angle, swing angular velocity and body lateral inclination angle, as shown in Fig. 4.The 3-DOF model compared to the 2-DOF model takes into account the crane's lateral

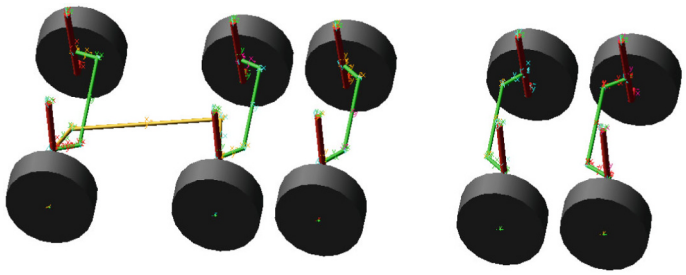
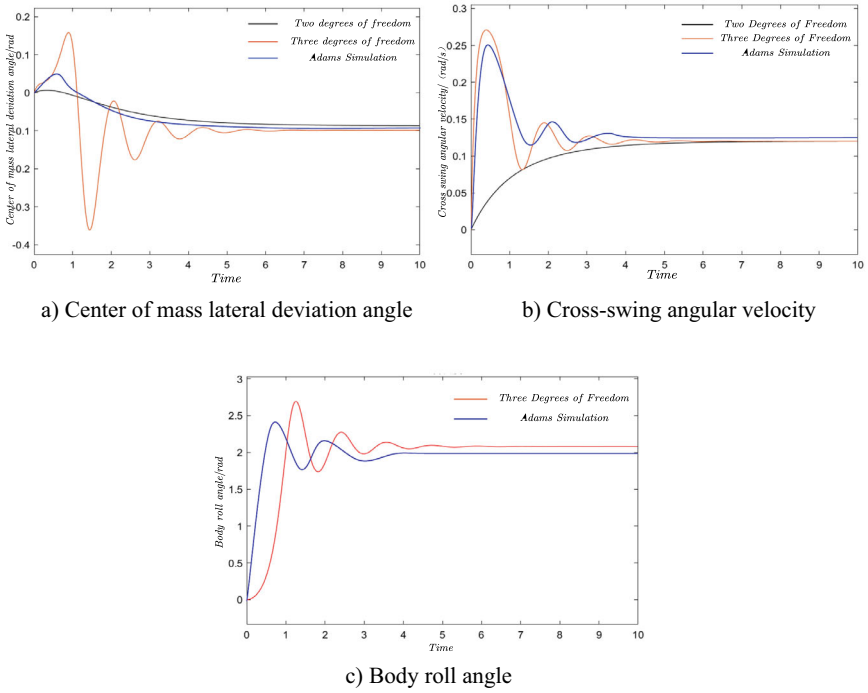


Fig. 3 Adams 5-axis steering model

Table 1 Structural parameters

Categorization	Name	Numerical value	Work unit
Crane structure	Total mass	86	t
	Wheelbase	2.65/1.65/1.65/2.9	m
	Equivalent sway stiffness	58,000	N/rad
	Equivalent sway damping	0.2–0.4	/
Spring loaded mass	Total mass	62	t
	Moment of inertia about the x-axis	410,000	kg/m^2
	Moment of inertia about z-axis	930,000	kg/m^2
	Moment of inertia around x and z axes	780,000	kg/m^2
	Lateral tilt center to center of mass distance	2.2	m



a) Center of mass lateral deviation angle

b) Cross-swing angular velocity

c) Body roll angle

Fig. 4 Comparison pictures

movement during steering. From Fig. 4, it can be seen that the deviation among the theoretical curves (2-DOF model and 3-DOF model) and the simulation curves is small, and the trend is consistent, so the established steering 3-DOF model of five-axis all-terrain crane is reliable.

3 Control Strategy Analysis

In order to ensure that the five-axis all-terrain crane has good stability and steering when steering, two control strategies based on first-axis cornering feed-forward control and optimal control based on vehicle state feedback and front-wheel feed-forward, are established. In Fig. 5a, the transient fluctuation of the center-of-mass lateral deflection curve under the 3-DOF model is large, and the amplitude of the curve fluctuation can be optimized by two control strategies.

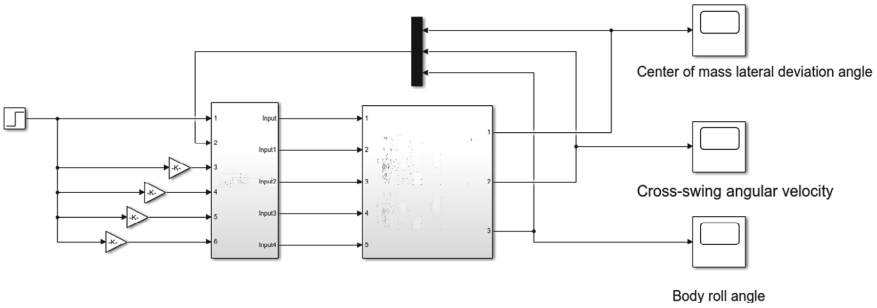


Fig. 5 State feedback control model

3.1 Proportional Control Based on the First Axis Angle

In the crane studied in this paper, the first two axles are mechanical linkage-driven steering, and the rear three axles are electronically controlled steering [7], so the proportionality coefficients of the angle of the second axle relative to the first axle are determined according to the geometry, and the proportionality coefficients of the rear three axles are determined according to the formula. Let the proportional relationship between the angle of the rear three axes and the angle of the first axis be:

$$\delta_i = G_i \delta_1 (i = 3, 4, 5) \tag{3}$$

To determine the proportionality coefficient G_i, it is necessary to determine L_1 first, and the following determines L_1 based on the zero-mass lateral deviation control strategy, which is based on the five-axis all-wheel steering linear 2-DOF model. When the whole crane reaches the steady state, it satisfies $\beta = \dot{\beta} = \dot{\omega}_r = 0$, which can be obtained by substituting into Eq. 4:

$$L_1 = \frac{a_{11}a_{22}(L_{12} + L_{13} + L_{14} + L_{15}) - a_{12}a_{21}(L_{12} + L_{13} - L_{14} - L_{15})}{5a_{11}a_{22} - a_{12}a_{21}} \tag{4}$$

$$\begin{bmatrix} a11 & a12 \\ a21 & a22 \end{bmatrix} = \begin{bmatrix} \frac{2(\sum_{i=1}^{5} ki)}{mv} & \frac{2(\sum_{i=1}^{3} l_i k_i - \sum_{i=4}^{5} l_i k_i)}{mv^2} - 1 \\ \frac{2(\sum_{i=1}^{3} l_i k_i - \sum_{i=4}^{5} l_i k_i)}{I_z} & \frac{2(\sum_{i=1}^{3} l_i^2 k_i - \sum_{i=4}^{5} l_i^2 k_i)}{I_z v} \end{bmatrix} \tag{5}$$

Combined with Ackermann's theorem, this gives:

$$G_3 = \frac{L_1 - L_{13}}{L_1} G_4 = \frac{L_1 - L_{14}}{L_1} G_5 = \frac{L_1 - L_{15}}{L_1} \tag{6}$$

Substituting L_1 into Eq. 6 gives the feedforward scale factor.

3.2 State Feedback Control

The value of pendulum angular velocity under the first axis steering is set as the ideal value. The state feedback control strategy [8] for five-axis steering is described as follows: a set of optimal steering angles is calculated to optimize the dynamic stability during five-axis steering under the conditions that the crane speed, cornering angle and crane parameters are determined. That is, the side deviation of the center of mass is close to zero, the traverse angular velocity can reach the ideal traverse angular velocity, and the value of the body roll angle is minimized.

Control objectives are converted to mathematical language for description [9]: the smaller the individual quantities in the output vector $Y = \left(\beta\ \omega_{\mathrm{r}}\ \phi\right)^T$, the better, that is to say, the dynamic stability performance index to reach a very small value, the performance index expression is designed according to Eq. 7, expressed as.

$$J = \frac{1}{2}\int_0^\infty \left(Y^T QY + U^T RU\right)dt = \frac{1}{2}\int_0^\infty \left(XQ_{\mathrm{d}}\mathrm{X} + U^T RU\right)dt \tag{7}$$

where U is the optimal control matrix, $U = -KX$;K is the matrix of optimal feedback coefficients, $K = -R^{-1}B^T L$;P is a solution of the Riccati equation.

The Riccati matrix equation is given below [10]:

$$LA + A^T L - LBR^{-1}B^T L + Q = 0 \tag{8}$$

The Riccati equation is solved for the matrix K. The optimal steering axis compensation angle is:

$$\delta_c = -KX = -K\left[\beta\ \omega_r\ \phi\right]^T \tag{9}$$

4 Control Simulation Analysis

In the actual turning of the crane, only the first axis tire steering, the rest of the axis does not steer the case, the tires work in the straight line area, the stability of the crane is the best [11]. As shown in Fig. 5 is the state feedback control model. In order to analyze the control effect, the proportional feedforward control model and the optimal control model based on vehicle state feedback and front wheel feedforward are established and simulated together with the front wheel steering model. Given that the first axis angle input is 0.25 rad and the crane speed is 5 m/s, the control effect of the two control strategies on the crane steering is analyzed and compared with the first axis steering condition. As shown in Fig. 6.

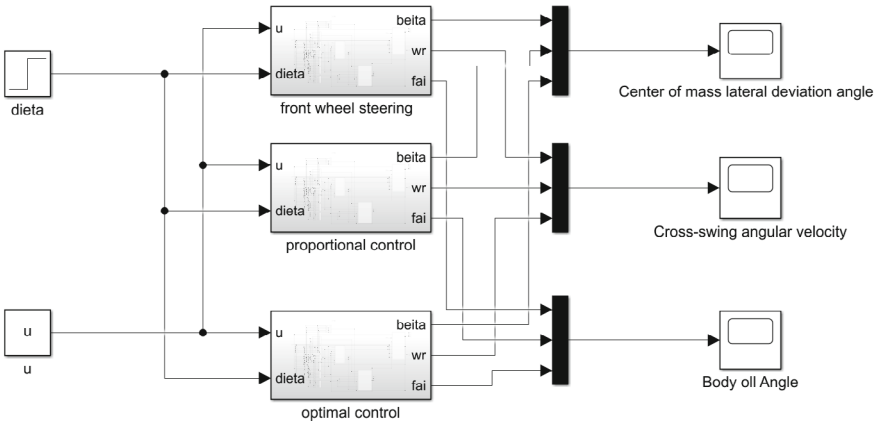
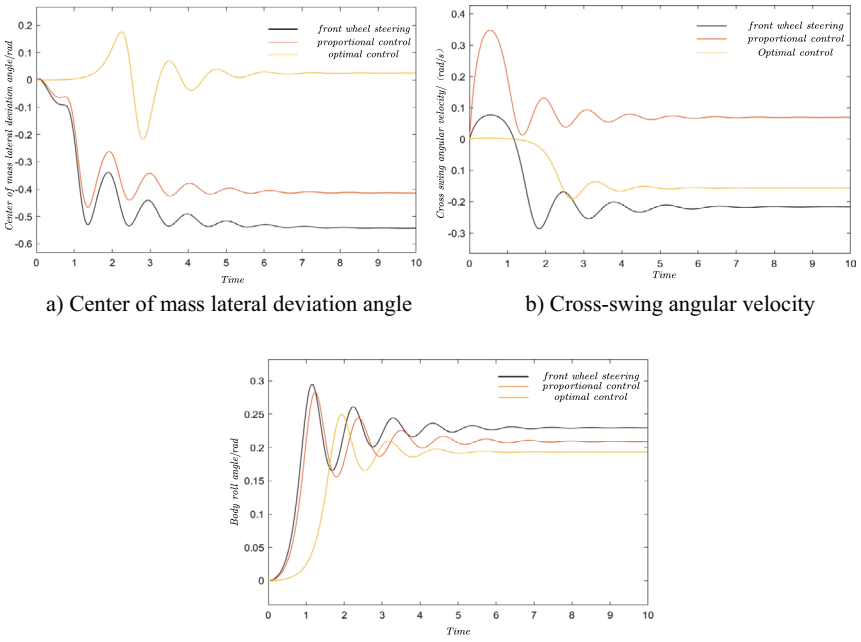


Fig. 6 Simulation model

In Fig. 7a, it can be seen that the transient fluctuation of the center-of-mass lateral deflection curve under the 3-DOF model is large, and the magnitude of the curve fluctuation is reduced by the two control strategies. From Fig. 7a, it can be seen that the center-of-mass lateral deflection under optimal control has the best effect, and its steady-state value is close to zero, which achieves the control objective and can balance the trajectory and stability, followed by the center-of-mass lateral deflection under proportional control, which is better than that under the first axis steering. From Fig. 7b, it can be seen that the cross-swing angular velocity under the optimal control is -0.17 rad/s, which is close to the cross-swing angular velocity of 0.22 rad/s under the first axis steering, and the steering sensitivity is better, which reduces the driver's driving burden. From Fig. 7c, it can be seen that the minimum sway angle under optimal control is 0.18 rad, which basically satisfies the control effect, followed by the sway angle under proportional control is better than that under first axis steering.



a) Center of mass lateral deviation angle

b) Cross-swing angular velocity

c) Body roll angle

Fig. 7 Comparison pictures

5 Conclusion

Based on the classical 2-DOF model, the 3-DOF model of the crane is established by considering the body tilt, on the basis of which the optimal controller based on crane state feedback is designed, and the first axis angular step input condition is selected, so that the optimal control is compared and analyzed with the first axis steering and proportional control in order to validate the control effect of the controller.

Comparative analysis results show that the crane under the optimal control based on state feedback can make the crane's center of mass lateral deflection angle closer to zero, and the crane stability is better; and the swing angle velocity can track the ideal model well, which ensures that the driver can keep the steering feeling better, and the lateral inclination angle during steering is minimized, which ensures a good attitude of the crane body in the steering process, and it is not easy to roll over, and the control effect is better than the previous wheel steering and proportional control, which greatly improves the maneuvering stability of the whole crane. The control effect is better than the previous wheel steering and proportional control, which greatly improves the maneuvering stability of the crane.

Acknowledgements This research was supported by Taiyuan key core technology research unveiled list of marshal projects.

References

1. Chen X (2018) Research on steering characteristics analysis and multi-objective cornering control of multi-axle steering vehicles. Jilin University
2. Li Y, Gao X, Li Z et al (2006) Optimal tracking of multi-bridge dynamic steering system for vehicle-mounted self-propelled artillery. J Jilin Univ (Eng Ed) S2:36–40
3. Zhang X (2011) Simulation and experimental research on steering performance of all-terrain crane. Jilin University
4. Shet MR, Lakhekar VG, Iyer CN et al (2024) Robust fuzzy quasi-SMC-based steering control of autonomous vehicle subject to parametric uncertainties and disturbances. Int J Autom Technol 1–19
5. Sanchez HA, Poznyak A, Chairez I (2024) Extremum seeking control for the trajectory tracking of a skid steering vehicle via averaged sub-gradient integral sliding-mode theory. Robot Auton Syst 174104609
6. Bai X, Wang J, Hou Q et al (2021) Research and analysis of multi-axle steering control strategy for large engineering vehicles. Times Autom 18:14–16
7. Xie C, Fan Y, Yang J et al (2022) Research on optimal control of four-wheel steering automobile based on state feedback. Mech Des Manuf 1:102–106
8. Fan Y, Zi H, Xie C et al (2021) Handling stability analysis of a four-wheel steering vehicle. J Taiyuan Univ Sci Technol 42(2):128–131+137
9. Lhoussain HE, Mehdi EM, Vidas Ž et al (2023) A novel robust adaptive neuro-sliding mode steering controller for autonomous ground vehicles. Robot Auton Syst 170
10. Deng Z, Yi Q, Gao W et al (2022) Simulation study of closed-loop LQR control for four-wheel steering vehicles. Mech Des Manuf 1:20–25
11. Xu F (2020) Research on control strategy and test of multi-mode wire-controlled steering system for emergency rescue vehicles. Jilin University

Open Access This chapter is licensed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License (http://creativecommons.org/licenses/by-nc-nd/4.0/), which permits any noncommercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if you modified the licensed material. You do not have permission under this license to share adapted material derived from this chapter or parts of it.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.



Brief Introduction of Hybrid Electric Vehicle Research



Shengqi Yao, Xiaoshan Yao, Yilu Liu, Chenfei Zheng, Weiqiang Zhan, Jie Ma, Ziliang Zhu, and Guang Yang

Abstract This paper explains the background and significance of researching hybrid vehicles, and summarizes and analyzes the current research status of external hybrid vehicles. It also analyzes the components, arrangement forms, and performance characteristics of the mainstream hybrid vehicle drive system structures (series hybrid system, parallel hybrid system, and hybrid system).

Keywords Hybrid electric vehicle · Research · Overview · Power system structure

1 Research Premises

1.1 Background

The automobile industry is closely related to People's Daily life, especially in People's Daily travel plays a key role. Although the automobile industry has created a lot of convenience and benefits for our life, it has also brought us some adverse effects. The consumption of oil by automobile has caused the economy to over-rely on fossil resources, and the global oil supply is gradually less than the demand, which leads to the conflict between People's Daily consumption and energy production becomes more acute. However, oil resources are very important strategic resources for any country, and it is related to the national energy security of each country. According to the report of relevant research institutions, in 2015, China's dependence on foreign oil resources has reached 60.6%, exceeding 60% for the first time,

S. Yao (✉)
Detroit Green Technology Institute (DGTI), Hubei University of Technology, Wuhan, China
e-mail: 19145720@qq.com

X. Yao · Y. Liu · W. Zhan · J. Ma · Z. Zhu · G. Yang
Air Force Early Warning Academy, Wuhan, China

C. Zheng
GAC R&D Center, Guangzhou, Guangdong, China

© The Author(s) 2026

H. Saman et al. (eds.), *The 9th International Conference on Advances in Construction Machinery and Vehicle Engineering*, Lecture Notes in Mechanical Engineering,
https://doi.org/10.1007/978-981-96-5527-4_99

and its oil import volume ranks second in the world. In 2016, China imported more oil than the United States, ranking first in the world. In 2017, China imported 420 million tons of crude oil, and its dependence on foreign countries reached 68.6%, or nearly 70% [1]. Table 1 shows China's crude oil import volume and the degree [2] of external dependence in the past ten years. OPEC predicts that the gap between China's oil demand and production will continue to widen in 2020, and the dependence on foreign imports will continue to rise. Transportation is a never-ending operation process, according to relevant data show that transportation oil consumption accounts for 61% of the world's total oil consumption, accounting for 28% of the total terminal energy consumption, is one of the most energy consumption industries.

Up to now, China produces over 20 million cars per year, car ownership is still increasing, not only consumes a lot of oil resources and mineral resources every year, but also causes countless exhaust emissions. The United Nations environmental protection organization shows that a large part of the pollutants in the atmosphere is produced by the combustion of automobile fuel, including CO, HC, NOx, PM and other harmful substances, seriously endangering people's physical and mental health and living environment. Since entering the twenty-first century, the degree of air pollution in China has been increasing day by day, and the continuity of more serious haze weather in most cities, $PM_{2.5}$ has gone off the charts in many places, has caused widespread concern from all walks of life [3]. Therefore, in [4] order to solve the problem of the increasing shortage of oil resources, get rid of the dependence on foreign oil resources, slow down the supply pressure of energy, achieve energy security, while reducing vehicle exhaust emissions, reduce environmental damage, the development of clean energy vehicles is the most effective way.

Table 1 China's crude oil import volume and dependence on imports from 2008 to 2017

Year	The import volume of crude oil (in billion tons)	Dependence on foreign countries (%)
2008	1.79	48.00
2009	2.04	51.17
2010	2.39	53.72
2011	2.52	55.11
2012	2.71	56.42
2013	2.82	57.39
2014	3.10	59.60
2015	3.26	60.60
2016	3.81	65.40
2017	4.20	68.60

1.2 Meaning

Facing serious energy and environmental problems, the development of new energy vehicles with zero (low) emissions and low noise has become an urgent need for the healthy development of cities. At the same time, the development plan of the automobile industry issued by The State Council has further clarified the development direction [5] of the automobile industry in the future.

Pure electric vehicle is the most ideal new energy model, it can achieve the ideal pollution-free state, is an environmentally friendly means of transportation. Moreover, pure electric vehicles simply use electricity as the driving energy, which is rich in electric energy reserves in China and lower in cost compared with oil resources, and can also solve the problem of load imbalance between day and night in the power grid. However, the core of pure electric vehicles—three electric technology (motor, electronic control, battery) still need to break through, especially battery technology, the current energy density of the battery is low, can not reach as long as the fuel vehicle driving range and the cost is still high, resulting in electric vehicles can not be introduced to the market [6] in large quantities. Therefore, traditional fuel vehicles can not be completely replaced by pure electric vehicles in a short period of time, in order to make the traditional fuel vehicles gradually transition to pure electric vehicles, hybrid vehicles have now become the focus of wide concern of enterprises, scientific research institutions and universities. Hybrid electric vehicles have the advantages of long driving range of fuel vehicles and low emissions of electric vehicles, which is in line with daily travel and market demand. After unremitting efforts in recent years, hybrid electric vehicles have basically formed three fixed power drive structures (series hybrid system, parallel hybrid system, hybrid hybrid system), these three structures have their own characteristics, but the manufacturing cost is generally high [7].

2 Current Research Status

Under the major situation of energy and environmental problems, the research and development of clean energy vehicles and the acceleration of the transformation and upgrading of the traditional automobile industry have become the challenges that the development of the global automobile industry needs to shoulder. Japan, Europe and the United States and other developed places, due to the early start, the research and development and manufacturing of hybrid electric vehicles has done a lot of work, and has a richer technical reserve [8] than other countries.

Toyota of Japan has the most mature hybrid technology in the world, and has also been recognized by the market, occupying a large number of global market shares. The current market of the third generation of Toyota Prius, the THS II system will be several advantages of hybrid vehicles to a new height. THS II system is composed of engine, generator, motor and planetary gear train, by the vehicle control strategy

to complete the power distribution work, determine the engine, generator, motor working state, by the planetary gear train to complete the vehicle power coupling. The output of the engine and the motor through the planetary gear train speed regulation without conflict, to ensure the smoothness of the power output. This is the hybrid technology so far the pursuit of engine energy conversion rate of the best scheme, the current third-generation Toyota Pruis 100 km fuel consumption is only 4.3 L.

GM Automobile Company in the United States has made deep progress on BSG micro-mixing and parallel mixing technology. The Lacrose Eco-Hybrid is equipped with a 2.4 L in-line four-cylinder engine and a nickel-metal hydride battery-powered motor assist system. The motor auxiliary system can assist the vehicle to drive under high load, relieve engine pressure and reduce fuel consumption. The comprehensive fuel consumption of the hybrid model is 8.3 L per 100 km, which is 16% lower than the original car.

In addition, Volkswagen, BMW, Honda, Nissan and so on are vigorously researching hybrid power technology. Among them, the United States and Japan mainly develop gasoline hybrid technology, the main object of research and development in the United States is large-displacement high-end models, Japan takes economic passenger cars as the main research models, while the hybrid vehicles developed in Europe are mainly mild diesel hybrid technology, and diesel and some low-displacement vehicles are the main research models. However, most of these hybrid models have complex technology or structure, high manufacturing costs, and there is still a certain distance from being introduced to the market in a large area.

In 2001, the state listed the research and development of new energy vehicles as one of the major research projects at the national level, with sustainable development as the strategic orientation, put forward the "three horizontal and three vertical" development strategy [9] based on models and common technologies. Through the close cooperation of enterprises, research institutions and universities as well as the state's planning and development and policy incentives, the development of China's hybrid electric vehicles has achieved [10] phased results.

Byd Auto Company launched BYD Qin 100 hybrid electric vehicle, composed of 1.5 T engine with electric motor DMII dual-mode hybrid system, the total power can reach 295 horsepower, the total torque can reach 479 Nm, the power parameters are very impressive. What's more, the Qin uses BYD's self-developed lithium iron phosphate battery, which can be charged directly by household power and can be fully charged in 5 h. At the same time, the dual-mode hybrid system also has an efficient braking energy recovery system, and the recovered energy can travel about 15 km [11].

Roewe e550 model, independently developed by SAIC, is equipped with a 1.5 L gasoline engine, an Integrated Starter Generator (ISG) Motor and a Traction Motor (TM) motor, and matched with EDU intelligent variable speed system. It can flexibly switch the driving mode. And put forward the hybrid power technology of "double core, three core, eight mode": "double core" is manifested as two power source engine and drive motor; "Three core" refers to its gasoline engine, ISG start and power generation unit, TM traction motor three nuclear power; "Eight modes" means that it has eight operating modes. The vehicle has a combined fuel consumption of

2.3L per 100 km, a maximum speed of 200 km/h, and an acceleration time of 0–100 km/h of 10.5 s. In addition, FAW, Dongfeng, Changan, GAC and Chery all have their own hybrid products. In recent years, due to the guidance of national policies, coupled with the unremitting efforts of enterprises, research institutions and universities, China's hybrid automobile industry has been rapid development. However, at present, China still has a certain gap with Japan and the United States in the key technologies and product quality of hybrid electric vehicles. The national finance is not perfect, the industrial chain is not concentrated enough, the lack of supporting facilities, and the low market acceptance are the main problems restricting the development of China's hybrid automobile industry. In order to rapidly promote the strategic transformation of the traditional automobile industry to new energy vehicles, and realize the industrialization and commercialization of hybrid electric vehicles, it needs the encouragement and strong support of national policies. At the same time, domestic automobile enterprises should strengthen their independent innovation capabilities, invest more research and development efforts, and be more committed to cultivating their own brands to further shorten the gap [12] with foreign countries.

3 Analysis of Power System Structure

3.1 Series Hybrid Power System

The structure of the series hybrid power system is shown in Fig. 1. Its essential structural characteristics are that the engine and the transaxle are relatively independent without direct connection, and the energy transfer must be completed through the middle drive motor. The motor controller connects the generator and the drive motor, and completes the power balance of the vehicle through the adjustment of the power battery pack. When the vehicle is running in high load condition, the power generated by the engine can not meet the driving power of the vehicle, and at the same time, the power battery output additional electric energy to supplement the needs of the vehicle; When the vehicle in the parking condition or other low load conditions, the engine generated power is greater than the vehicle's demand power, then the motor controller will control the generator to store additional electrical energy in the power battery [13, 14]. Since the engine power is calculated by the vehicle at a certain stable speed, but in the actual road conditions, the vehicle can not always run at a certain stable speed. When there is a gap between the power required to drive the motor and the power generated by the generator, the motor controller will change the working state of the generator according to the actual situation. The series hybrid power system has the following structure and control characteristics:

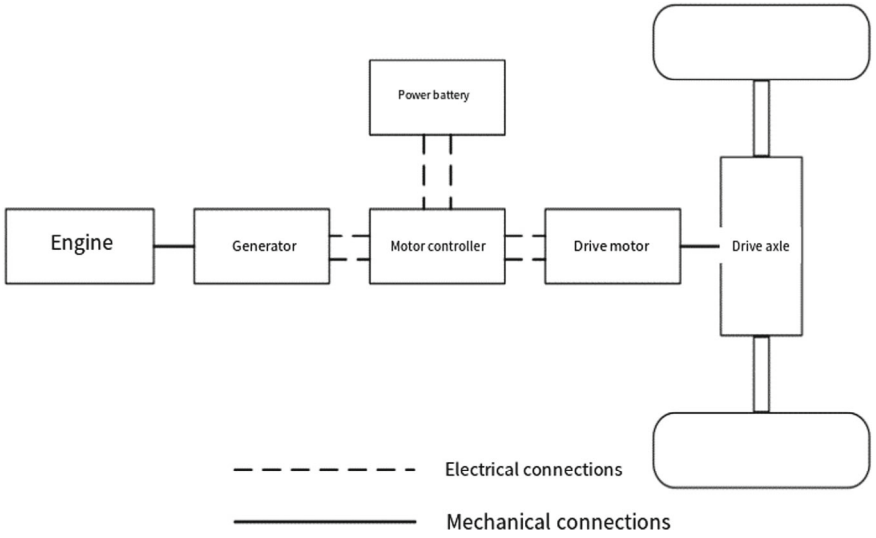


Fig. 1 Series hybrid power system

① The structure arrangement and control method of the series hybrid power system are easy and easy to realize. At the same time, the engine and the drive motor are relatively independent, there is no connection of mechanical parts, and it is easy to arrange the structure of the vehicle.
② The engine and the drive axle are relatively independent, so the working state of the engine is not interfered with the road conditions, but also easy to adjust the working range of the engine, which can make the engine always work stably in the low fuel consumption area, thus showing good economic performance.
③ Because the power battery can adjust the power balance between different power sources at any time, the low-power efficient engine can be selected as far as possible when matching the engine, as long as the power required under a certain stable speed can be met.
④ The engine does not need to directly drive the vehicle, and the drive motor is relatively independent, indicating that there are basically no restrictions on the selection of the engine, so you can try to choose efficient prime movers, such as: Stirling engine, high-speed gas turbine and so on.
⑤ Because the power battery needs to adjust the power balance between the power source at any time, while maintaining the battery life as much as possible to avoid the phenomenon of abnormal working of the battery, the power battery is required to have a larger capacity, increasing the vehicle weight and manufacturing costs.

The mechanical energy output of the engine must go through two energy conversion in order to provide driving force for the vehicle, so the utilization of energy in the series hybrid power system is poor. Series hybrid power system in frequent start, braking and low-speed operation of urban conditions has significant advantages, can

make the engine in the best working area for a long time, so the power system is suitable for buses, urban buses.

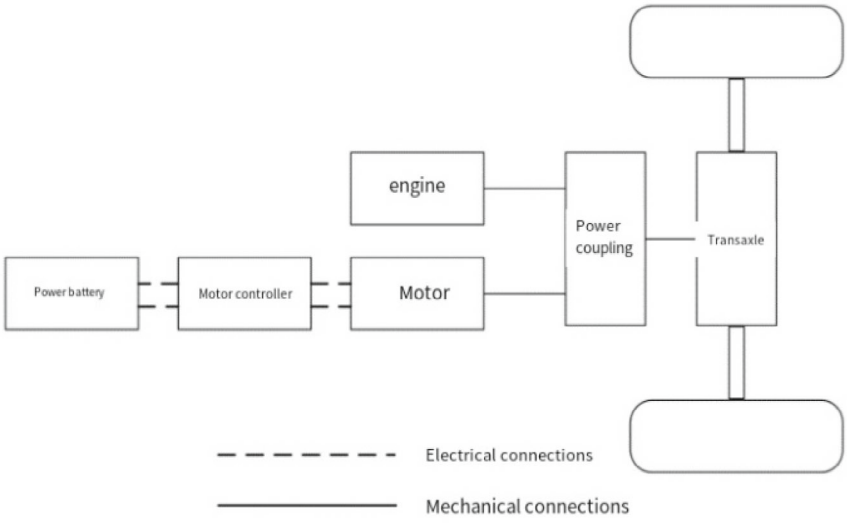
3.2 Parallel Hybrid Power System

The parallel hybrid power system can be simply understood as an additional set of independent electric drive system installed on the power system of the traditional fuel vehicle. Its structure is shown in Fig. 2. The engine and motor can provide driving force for the vehicle separately under different working modes, and can also provide driving force for the vehicle jointly through the power coupling device. Generally, the engine is the main power source in the parallel hybrid power system, which meets most of the driving needs.

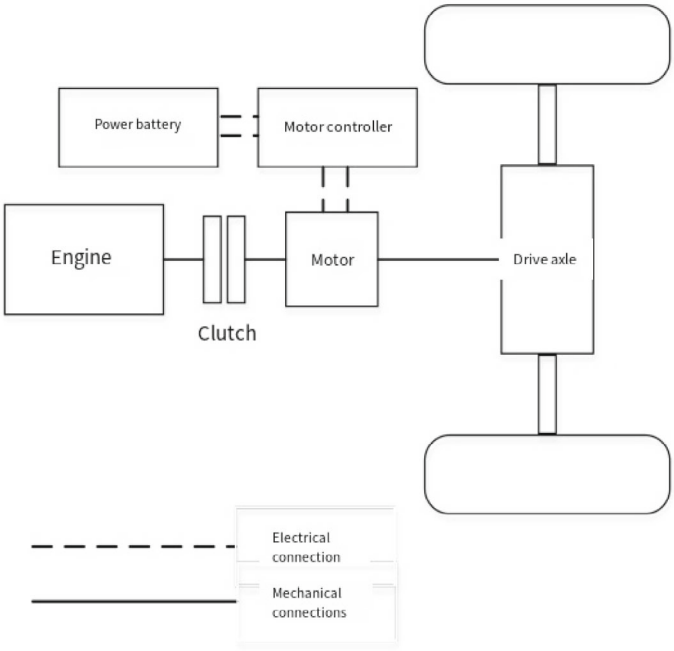
The motor plays a regulating role for the auxiliary power source, so the motor has two working states, on the one hand as a driving motor, and on the other hand as a generator. In the parallel hybrid power system, the engine and the motor often need a separate mechanical device for different power source coupling [15, 16].

In the starting or low load conditions, the motor provides driving force for the vehicle alone, the engine does not provide driving force, to avoid the engine running in the low efficiency zone; When the vehicle accelerates and climbs, the motor auxiliary drive shares part of the load of the engine and regulates the working interval of the engine; When the vehicle is cruising at high speed, after the engine provides the power required by the vehicle, the additional power can be output to the motor to charge the battery. The parallel hybrid power system has the following structural and control characteristics [17]:

① The engine can drive the vehicle alone in a specific mode, without the conversion process of mechanical energy to electric energy, reducing energy loss, so the energy utilization rate of the parallel hybrid system is higher than that of the series hybrid system.
② In the parallel hybrid system, the motor is mainly an auxiliary power source, and also plays a role in speed regulation of the engine, so it can choose a low-power motor as far as possible to save costs.
③ In the parallel hybrid system, although the engine provides the main driving force, the engine can timely supplement the battery with electric energy, so it is not necessary to match the power battery with such a large capacity as the series hybrid system.
④ The working state of the engine is greatly affected by external road conditions, especially in complex and changeable working conditions, the engine may often work in the inefficient area, resulting in the fuel economy and emissions of the parallel hybrid system than the series hybrid system is worse.
⑤ In the hybrid mode, the power coupling device is needed to couple different power sources, so the structure of the parallel hybrid system is more complex than the series hybrid system, and the requirements for the control unit are higher.



(a) Dual shaft parallel



(b) Single axis parallel connection

Fig. 2 Parallel hybrid power system

The parallel hybrid system has a variety of drive modes, which has great advantages in stable driving under medium and high load conditions. At the same time, in the urban conditions of low load, zero emissions can be achieved through pure electric driving mode, but this has higher requirements for the drive motor and power battery.

3.3 Hybrid Hybrid System

The hybrid hybrid power system integrates the characteristics of the above two hybrid power systems, and its power system structure is shown in Fig. 3. On the one hand, the power generated by the engine directly drives the vehicle through the mechanical transmission components, on the other hand, it is used to generate electric energy to provide to the driving motor, and then drives the vehicle through the power coupling device. Generally, the hybrid hybrid electric vehicle works with a series hybrid system when driving at low speed, and a parallel hybrid system [18] when driving at high speed.

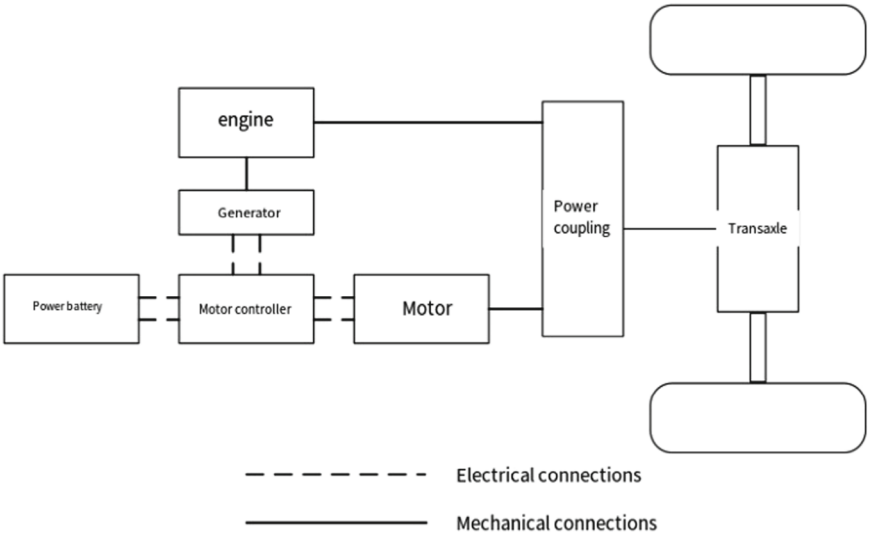


Fig. 3 Hybrid hybrid system

4 Conclusion

This article believes that hybrid vehicles, which combine the long range of gasoline vehicles and the low energy consumption and emissions of electric vehicles, are the future trend with a lot of industry investment. The article has a certain analysis of the structure of series-connected, parallel-connected, and hybrid-connected hybrid power systems, but the future development trend of different hybrid power systems needs further research.

References

1. Pang W (2018) In 2017, China's crude oil import exceeded 400 million tons of natural gas production and import double liters. China News Network (19 Mar 2018). http://www.ebnew.com/info/601150508.html
2. Wu T (2017) Beijing-Tianjin-Hebei Autumn and winter pollution control: Beijing $PM_{2.5}$ concentration decreased by 25% year-on-year. China Youth Network (25 Aug 2017). http://news.youth.cn/gn/201708/t20170825_10583726.htm
3. Overview of China's crude oil production and marketing in 2017 and analysis of crude oil dependence on foreign countries. China Industry Information Network (2017) (01 Sep 2017). http://www.chyxx.com/industry/201709/556736.html
4. Zhou J (2024) Research on the development status and prospect of new energy vehicles. Automot Knowl 24(07):23–25
5. Cuan Q, Chen L (2023) Some thoughts on promoting the transformation and upgrading of China's automobile industry. Shanghai State-owned Assets 7:65–68
6. Xue Y, Chen C (2023) Research on the status quo and development trend of new energy vehicle battery types. Times Automob 19:106–108
7. Jin M, Pan P, Wang F et al (2012) Analysis of hybrid passenger car demonstration and promotion products in China. Automot Eng J 2(06):400–406
8. Liu J (2023) Research on development problems and strategies of hybrid electric vehicle. Automot Test Rep 12:7–9
9. Zhen Z (2011) To achieve "three vertical and three horizontal" key technological breakthroughs scientific and technological innovation to drive China's new energy vehicle strategic emerging industry formation. China Sci Technol Wealth 7:26–28
10. Mou L, Zou Y, Jiang H (2024) Research status, existing problems and future development trend of new energy vehicles. Automot Mainten Tech 16:81–82
11. Shen J (2017) Ease your anxiety test drive 2017 BYD Qin 100. Car Home (Original test drive), 17 May 2017
12. Li X (2024) Development status and prospect trend of new energy vehicles in China. People's Forum 9:76–79
13. Cai L, Zhang X (2011) Match and optimization of series hybrid bus powertrain parameters. Appl Mech Mater 121–126:2710–2714
14. Xu M, Zhang Y-J (2019) Development of powertrain technology for hybrid electric vehicles in China. J Automob Saf Energy Conserv 15(03):269–294
15. Luo Y (2022) Research on torsional vibration control method of transmission system of parallel hybrid electric vehicle. Fujian University of Science and Technology, p 1, 16 Dec 2021 to 15 Jan 2022
16. Gao JP, Wei YH, Liu ZN et al (2013) Matching and optimization for powertrain system of parallel hybrid electric vehicle. Appl Mech Mater 341–342:423–431

17. Wu X, Fu J, Sun Z (2021) Research on power system parameter matching of single-axis parallel hybrid electric vehicle. Automot Pract Technol 46(22):89–91
18. Yin Y (2023) Analysis of energy management control strategy for hybrid electric vehicle. Automot Test Rep 8:143–145

Open Access This chapter is licensed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License (http://creativecommons.org/licenses/by-nc-nd/4.0/), which permits any noncommercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if you modified the licensed material. You do not have permission under this license to share adapted material derived from this chapter or parts of it.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.



Research on Injury to Hybrid III 50th Percentile Dummy Based on FMVSS 208 Test

Qian Zhang and Youjun Wang

Abstract This study rigorously adheres to the FMVSS 208 standard by conducting 30° left and right offset collision tests, as well as direct frontal collision tests, on the same vehicle model under unbelted conditions. The primary objective is to accurately assess the vehicle's crashworthiness and meticulously monitor the extent of injuries sustained by the front-row passenger and the driver across various body regions. Test results reveal that vehicle damage is particularly severe in the 100% frontal collision scenario, with injury metrics for the front-row passenger dummy generally exceeding those of the driver dummy in this type of collision. Consequently, in real-world traffic accidents where a vehicle strikes a building or another vehicle at a specific angle, it is essential to prioritize emergency care and conduct thorough examinations of the head, chest, and knees of front-row passengers while also enhancing safety measures for front-row occupants.

Keywords FMVSS 208 · Restraint conditions · Hybrid III 50th percentile dummy · Dummy injury

1 Introduction

There is still limited research on angled collisions both domestically and internationally, and this area has not been widely covered by safety standards and regulations in various countries [1]. It is noteworthy that 30° oblique collisions significantly differ from common frontal collision tests in terms of collision modes, testing dummy configurations, and injury assessment standards [2]. Furthermore, there is a relative scarcity of collision testing practices under unbelted conditions in domestic studies [3], leading to a limited understanding among professionals in this field [4–6]. In light of this, this study closely relies on the testing procedures and assessment system

Q. Zhang (✉) · Y. Wang
China Automotive Technology and Research Center Automotive Test Center (China Tianjin) Co., Ltd., Tianjin, China
e-mail: 1970456950@qq.com

© The Author(s) 2026

H. Saman et al. (eds.), *The 9th International Conference on Advances in Construction Machinery and Vehicle Engineering*, Lecture Notes in Mechanical Engineering,
https://doi.org/10.1007/978-981-96-5527-4_100

of FMVSS 208, focusing on a specific model from a manufacturing platform explicitly designed for pure electric vehicles. By conducting real-vehicle collision tests, the research thoroughly analyzes the collected data. The investigation centers on the crashworthiness of the vehicle across three scenarios: left 30° oblique, right 30° oblique, and frontal collisions. It includes detailed examinations of acceleration responses at the lower section of the B-pillar on the collision side, assesses the extent of front panel intrusion, analyzes vehicle body deformation, and evaluates injury conditions at critical points on dummies, such as heads, necks, chests, and thighs. This study aims to provide a scientific foundation for reducing the risk of occupant injuries in automotive traffic accidents, ultimately enhancing the safety of vehicle occupants.

2 Hybrid III 50th Percentile Dummy and Its Evaluation Method

The Hybrid III 50th percentile dummy is widely regarded as a globally respected standard tool for crash testing, as depicted in Fig. 1. It simulates a scenario in which a U.S. male driver operates a vehicle at a speed of 20–25 mph (30–40 kph) when the vehicle collides with a rigid wall at a 30° oblique, incorporating two typical collision modes: left-side 30° and right-side 30°, as depicted in Fig. 2.

The evaluation of the Hybrid III 50th percentile dummy is conducted according to FMVSS 208 testing, as shown in Table 1.



Fig. 1 Hybrid III 50th percentile dummy

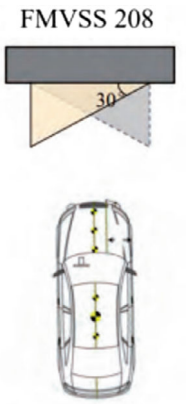


Fig. 2 e collision test based on FMVSS 208

Table 1 Injury criteria requirements for FMVSS 208

Inspection item		Standard requirement
Head injury criterion (HIC)		≤ 1000
3 ms Synthesis acceleration of chest		≤ 60 g
Chest deformation amount		≤ 76 mm
Thigh load (N)	Left leg	≤ 10,000
	Right leg	

3 Test Results and Analysis

This study concentrated on a specific vehicle type, employing a standardized Hybrid III 50th percentile dummy, and conducted three collision tests at a constant speed of 40 kph. The variations in weight among the test vehicles were meticulously controlled, with differences maintained within 10 kg. Furthermore, all dummies were unrestrained by seat belts to facilitate a comprehensive examination of the vehicle's crashworthiness and to assess the injury conditions across various body parts of the dummies.

3.1 Analysis of Crashworthiness of Vehicle Structure During Different Collision Tests

Acceleration Analysis at the Lower End of the B-Pillar on the Collision Side. The acceleration recorded at the lower end of the B-pillar serves as a critical parameter for assessing the acceleration experienced within the occupant compartment.

Figure 3 reveal that the variation trends of acceleration at the lower end of the left and right B-pillars show high similarity. In frontal collisions, the acceleration

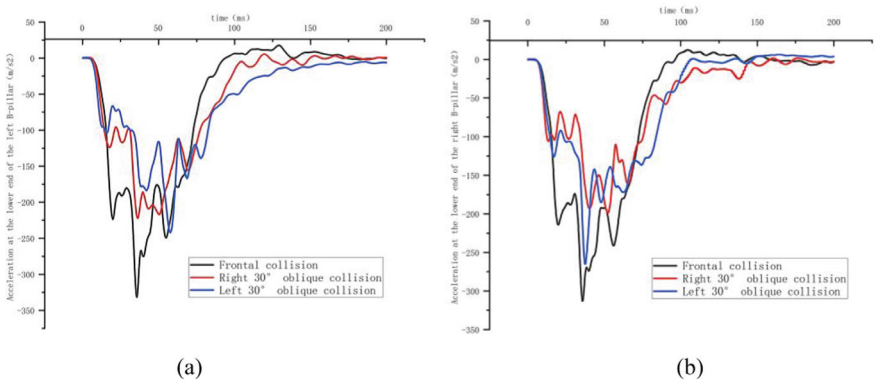


Fig. 3 Comparison of acceleration at the lower end of the left B-pillar and right B-pillar

of the B-pillar reaches its peak, indicating the maximum collision impact. Further comparison demonstrates that during the left 30° oblique collision, the acceleration of the right B-pillar is slightly higher than the corresponding value for the right 30° oblique collision, while the left B-pillar exhibits slightly lower acceleration. Given the high-strength characteristics of the B-pillar within the vehicle body system, it is not susceptible to displacement or significant deformation. Therefore, the acceleration variation curves of the B-pillars serve as effective indicators for assessing the overall stress conditions of the vehicle body and the extent of collision damage. It is inferred that the vehicle experiences the most intense impact force in frontal collisions, and correspondingly, the damage sustained by the vehicle body is also the most severe.

Displacement of the Steering Column. In the severe scenario of a frontal collision, the steering mechanism inevitably endures tremendous impact forces, as detailed in Table 2. During this process, the significant backward or upward movement of the steering wheel not only severely squeezes the survival space of the driver but also greatly increases the potential risk of injury to the driver's chest and head regions. An examination of the comprehensive data presented in Table 2 reveals that under the experimental conditions of left 30° oblique collision, right 30° oblique collision, and frontal collision, the displacement of the steering column in the X direction attains noteworthy levels, with the displacement during the frontal collision being the most prominent. The displacements during the left and right 30° oblique collisions are relatively close to each other. This data trend indicates that without seatbelt usage, frontal collisions not only cause more severe damage to the vehicle structure but also significantly increase the risk of injury to the driver.

Table 2 Displacement of the steering column in different tests

Name steering column center	Frontal collision	Right 30° oblique collision	Left 30° oblique collision
Nominal X	2677.297	2673.094	2676.239
Nominal Y	− 378.593	− 380.882	− 379.716
Nominal Z	1088.108	1086.891	1090.799
Measured X	2542.404	2596.472	2594.555
Measured Y	− 367.983	− 376.895	− 379.891
Measured Z	1067.900	1084.098	1084.742
Deviation X	− 134.893	− 76.622	− 81.684
Deviation Y	10.610	3.987	− 0.175
Deviation Z	− 20.208	− 2.793	− 6.057

3.2 Injury Analysis of the Hybrid III 50th Percentile Dummy in Different Collision Tests

Injury Analysis of Various Body Parts of the Hybrid III 50th Percentile Dummy in the Driver's Position

(1) Head

In the CNCAP evaluation system, HIC15 is a key indicator for measuring the severity of head injuries to the dummy, concentrating on the head impact within the first 15 ms following a collision. The head injury values obtained from this experiment are shown in Table 3. It is evident that when the driver is unrestrained by a seatbelt during a frontal 100% collision, the resultant acceleration of the driver dummy's head is markedly elevated, resulting in an extremely high HIC15 value, indicating the most severe head injury.

(2) Chest Synthetic Acceleration and Chest Compression

Figure 4 show that in the frontal 100% collision scenario, both the resultant acceleration and compression of the dummy's chest exhibit a high level. This phenomenon is primarily attributed to the high acceleration response of the vehicle's B-pillar and the significant displacement of the steering column during the frontal collision test. The combined effects of these two factors exacerbate the impact and injury sustained by the dummy's head and chest regions.

(3) Left Thigh Force and Right Thigh Force

Table 3 Head injury values for different test types

Test type	Frontal collision	Left 30° oblique collision	Right 30° oblique collision
Driver head HIC15 value	56	45	33

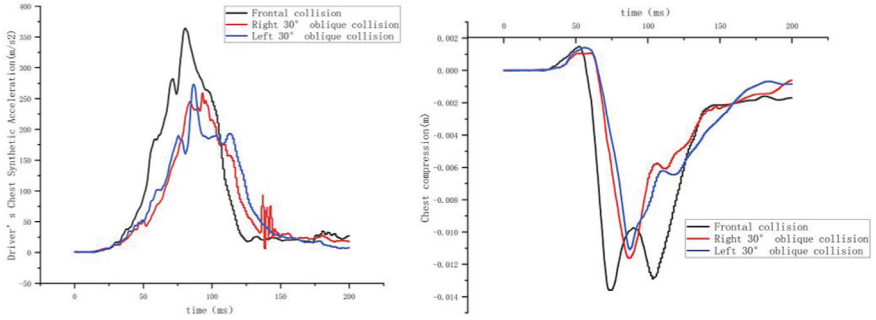


Fig. 4 Comparison curves of synthetic acceleration and chest compression at the chest

Comparative analysis of Fig. 5 indicate that in the frontal 100% collision test, the forces experienced by the left and right thighs of the dummy in the driver's position significantly exceed the corresponding values observed in the left 30° and right 30° oblique collisions.

Injury Analysis of the Hybrid III 50th Percentile Dummy in the Occupant Position

(1) Head

Table 4 and Fig. 6 represent that the head injury of the dummy in the occupant position is greater in the left 30° oblique collision test.

(2) Chest Synthetic Acceleration and Chest Compression

Figure 7 demonstrate that the chest injury of the dummy in the occupant position is greater in the frontal collision test.

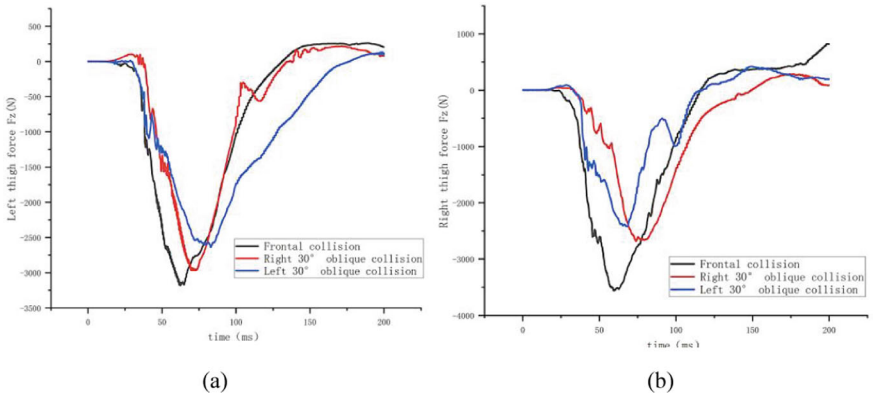


Fig. 5 Comparison curves of the left and right thigh force of the dummy in the driver's

Table 4 Head injury values for different test types

Test type	Frontal collision	Left 30° oblique collision	Right 30° oblique collision
Front row occupant head HIC15 value	159	188	129

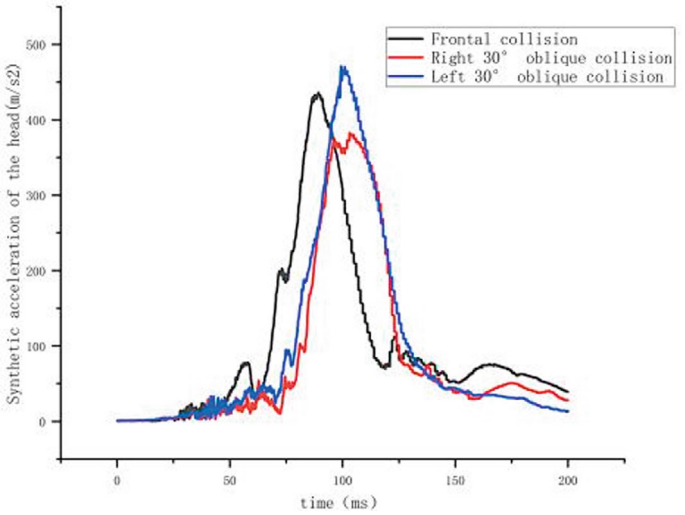


Fig. 6 Comparison curves of head injury for the dummy in the front-row occupant position

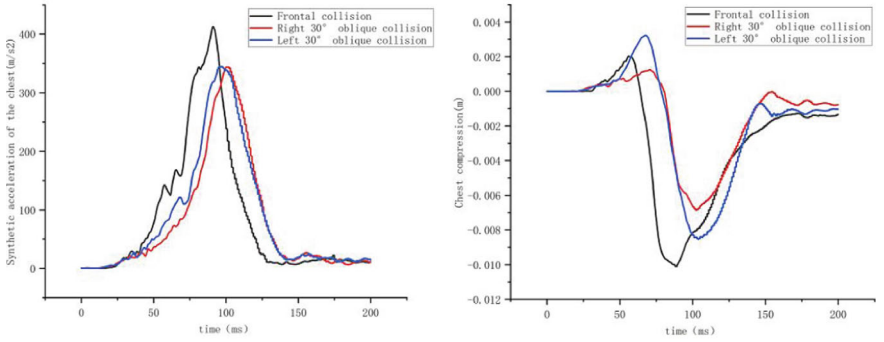


Fig. 7 Comparison curves of synthetic acceleration and chest compression at the chest

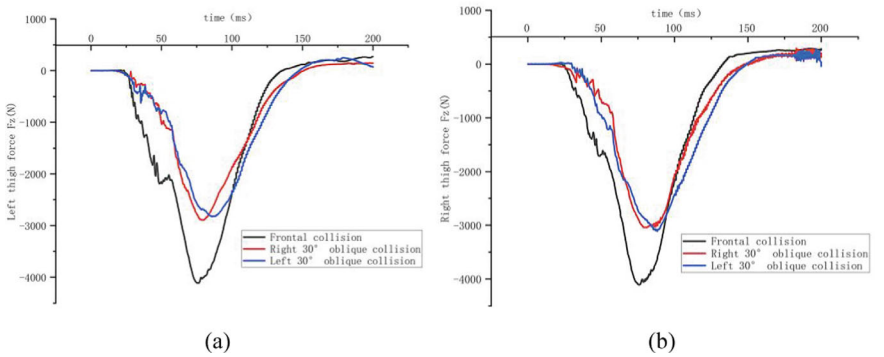


Fig. 8 Comparison curves of left and right thigh force for the dummy

(3) Left Thigh Force and Right Thigh Force

Figure 8 show that the left and right thigh forces of the dummy in the occupant position are relatively high during the frontal collision test.

3.3 Discussion on Injury Causes

Table 5 presents the injury values of the dummy in frontal collision tests conducted at various angles. In scenarios without seatbelt restraints, both front-row occupants, including the driver, exhibited considerable forward movement and upward leaning of the upper body during the collision. This motion led to an elevation of the buttocks. As a result, the dummy's head made contact with the airbag, triggering a peak head acceleration that highlights the substantial displacement and high impact acceleration on the head due to the lack of restraint. Additionally, the dummy's chest did not receive effective support due to the absence of seatbelt restraint. Consequently, when the chest of the driver-position dummy experienced dual impacts from the airbag and the steering wheel, the front-row occupant dummy's chest, after initially contacting the airbag, might subsequently collide with the vehicle's interior trim. This sequence of events resulted in a sharp impact force on the chest over a brief period, as indicated by a significant increase in synthetic acceleration and chest compression. Furthermore, the intense impact between the vehicle and a barrier during the frontal collision caused the displacement of the unrestrained dummy's hips and pelvic area, pulling the legs forward and significantly increasing the forces on both the left and right thighs.

Given that the three types of frontal angle collision tests were conducted under the protection of airbags, the head injuries suffered by the driver and front-row occupant dummies were relatively mild due to the effective cushioning provided by the airbags. However, during the frontal collision process, the overall movement of the front-row occupant dummy was quite pronounced, resulting in significant

Table 5 Injury values of dummies in frontal collision tests at various angles

Test type	Frontal collision		Left 30° oblique collision		Right 30° oblique collision	
Front row dummy	Driver	Occupant	Driver	Occupant	Driver	Occupant
Chest (3 ms)	36.49	40.87	26.28	35.03	24.98	34.83
Chest compression amount	13.65	10.11	11.04	8.58	11.64	6.86
Left thigh force	3.18	4.12	2.64	2.83	2.89	2.97
Right thigh force	3.56	4.12	2.43	3.11	2.68	3.04

head swaying along with the body. This led to an increase in both head acceleration and angular velocity, thus escalating the risk of head injury. Notably, regardless of whether it was a 100% frontal collision or a 30° oblique collision on either side, the test speed was set at 40 km/h, which falls within the low to moderate collision range, so the impact force was relatively low. At this speed, the protection offered by the airbags helped to mitigate head injuries for both the front-row occupant and the driver dummies. Additionally, both dummies' chests made contact with the airbags, preventing secondary hard impacts with the vehicle's interior trim, which helped reduce chest compression. However, the driver dummy's chest experienced relatively high collision impact forces during sequential contact with the airbag and the steering wheel, leading to a more significant compression of the chest. In the absence of seatbelt restraints, the front-row occupant and driver dummies' legs were more prone to moving forward and making secondary hard contact with the vehicle's interior trim. Notably, with more space in front, the front-row occupant dummy had an increased range of motion, which facilitated thigh collisions with the vehicle's interior trim and resulted in elevated thigh force values.

4 Conclusion

Through a series of frontal collision experiments, covering scenarios of 100% frontal impact and both left and right 30° angled collisions, we conducted an in-depth study of vehicle body and dummy injuries for a specific model from a manufacturing platform designed specifically for pure electric vehicles. The results showed that under conditions without seatbelt constraints, the damage to the vehicle body from a 100% frontal impact was particularly severe. This also led to significant forward displacement of the dummy in the front seats, increasing the likelihood of initial contact with the airbags and raising the risk of secondary hard collisions with interior components like the windshield. Consequently, this resulted in severe injuries to the heads, chests, and legs of front-seat passengers. While driver protection is generally comprehensive, this study reveals that the injury severity experienced by front-seat passengers in multi-angle frontal collisions is higher than that of the driver. Therefore, it is particularly urgent and essential to enhance awareness and implement measures

to improve the safety protection of front occupants. The research object of this paper is relatively single, and the data may be limited, and there may be some unconsidered variables or confounding factors that may have a potential impact on the results of the study, but have not been effectively controlled or analyzed. Future research should strive to expand the sample size and ensure that the sample is more representative to more accurately reflect the real situation and improve the generalizability and reliability of the research results.

References

1. Liu DC, Wang K, Zhang CJ (2021) Comparative analysis of frontal angle collision test methods for automobiles. Automot Accessories 19:4
2. Liu ML, Zhao W, Wu YQ (2022) Injury analysis of front row passengers in 18° frontal oblique collision test of a certain model. China Automot 3:11–14
3. Bai ZH, Wang YL, Liu L et al (2024) Adjustable occupant restraint system collision device: CN201610204273.7. CN107290151A
4. Hou FZ, Luan P, Li YM et al (2016) Experimental study on dummy injuries in frontal collisions of a certain model without wearing seatbelts. Eng Test 56(04):13–16
5. Yu CY (2007) Introduction to occupant protection in European and American commercial vehicle cabins: European ECE R29 vs. American FMVSS208 commercial vehicle news, vol Z3, p 22
6. Shang EY (2012) Research on dummy chest injuries in full width frontal collision tests. Automot Technol 3:18–21

Open Access This chapter is licensed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License (http://creativecommons.org/licenses/by-nc-nd/4.0/), which permits any noncommercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if you modified the licensed material. You do not have permission under this license to share adapted material derived from this chapter or parts of it.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.



Upper Limb Rehabilitation Robot Based on Planar Five-Bar Structure



Mingjie Cui, Huixiang Yang, Dawei Jiang, Jianwei Sun, and Tiecheng Ji

Abstract With the advancement in stroke rehabilitation, the application of upper limb rehabilitation robots for patient recovery is becoming increasingly popular. However, traditional rehabilitation robots face challenges such as complex control systems, high costs, and heavy structures, which hinder their widespread use in homes and small medical facilities. This paper introduces an innovative upper limb rehabilitation robot with a five-bar structure characterized by its flexible joint parallel configuration and user-friendly operation. This study conducts an in-depth analysis of the robot's kinematic properties and singular poses. Moreover, through the OpenSim simulation platform combined with Myoelectric Sensors data, we precisely investigate the dynamic changes in the activation levels of certain muscles and joint angles during rehabilitation. Experimental validations demonstrate that the robot can effectively promote moderate activation of key muscles, with all joint angles moving smoothly within their range of motion, indicating that the proposed device can provide effective rehabilitation training.

Keywords Upper limb rehabilitation · Parallel structure · OpenSim · Rehabilitation therapy

1 Introduction

In populations with neurological impairments, the loss of upper limb function has emerged as a prevalent issue [1–4], markedly diminishing the quality of life for patients. Given the trend of global population aging, the incidence of neurological disorders is expected to continue rising [5]. Concurrently, the shortage of therapists and caregivers further limits patients' opportunities to receive adequate rehabilitation

M. Cui · H. Yang · D. Jiang (✉) · J. Sun
Changchun University of Technology, Changchun, China
e-mail: jiangdawei@ccut.edu.cn

T. Ji
Jilin Electric Power Hospital, Changchun, China

© The Author(s) 2026

H. Saman et al. (eds.), *The 9th International Conference on Advances in Construction Machinery and Vehicle Engineering*, Lecture Notes in Mechanical Engineering,
https://doi.org/10.1007/978-981-96-5527-4_101

treatment. Therefore, exploring and promoting efficient methods for enhancing upper limb function recovery is of paramount urgency [6, 7]. The core objective of neurological rehabilitation focuses on strengthening muscle power and improving motor control, both of which are critical elements for patients to regain their daily living functions. Currently, task-oriented training models have been proven effective in improving functional outcomes due to their highly targeted nature and incorporation of repeated practice toward the desired functional recovery goals [8–10]. This type of training can facilitate neural plasticity and functional recovery following nerve injuries [11]. Robotic rehabilitation devices can effectively provide task-specific training, potentially addressing the gap in care needs [12].

End-effector robots have become vital tools for assisting patients in daily tasks due to their simple structure, ease of manufacturing and maintenance, user-friendly operation, cost-effectiveness, and broad applicability. These robots feature flexible mechanical designs, allowing for the use of rigid links or rope systems, which support precise movements in two-dimensional planes and three-dimensional spaces [13]. Particularly in the early stages of rehabilitation, planar end-effector robots are more suitable for patient use [14]. The initial MIT-Manus, which commenced its research and design in the 1990s [15], allows patients to grasp the handle at the end of the device. This innovative system facilitates two-dimensional rehabilitation training for the shoulder and elbow in a planar environment guided by the robot. The upper limb parallel end traction device (PETA) [16] utilized a specialized parallel five-bar mechanism as its driving system, featuring two coinciding active joints. Other rehabilitation devices have emerged as well. Henan University of Science and Technology designed a robot with a three-degree-of-freedom spatial structure [17]; however, issues related to noise levels have limited its applicability in extended rehabilitation sessions. The ULRR three-degree-of-freedom rehabilitation robot from the Beijing University of Technology [18] cleverly combined a four-bar mechanism with a lifting device to achieve three-dimensional motion. Meanwhile, Yanshan University took a different approach by developing a rehabilitation robot [19] that incorporates admittance control technology to ensure operational compliance. Still, its high cost and substantial size pose barriers to broader adoption. In contrast, parallel structures, particularly planar five-bar mechanisms, stand out with their simple, compact, and efficient design, presenting a more optimized approach for upper limb rehabilitation robots. Despite the added functionalities of existing upper limb rehabilitation robots, their complex mechanisms and control systems, high costs, and heavy structures still hinder their widespread application.

2 Structural and Kinematic Analysis

2.1 Structural Analysis

Parallel robots are ideal tools for performing rehabilitation tasks. Figure 1 displays an upper limb rehabilitation robot designed based on a five-bar structure, focusing on end-effector-guided rehabilitation training within a plane. This robot ingeniously employs four equal-length linkages connected in parallel to the end effector, enabling it to flexibly generate a variety of rehabilitation motion trajectories in the horizontal plane, accurately simulating daily activities. The rehabilitation robot features a lifting structure that adjusts the height of the work surface, making it easier for patients to set a suitable height for rehabilitation exercises, with an overall platform height ranging from 700 to 1200 mm. Regarding power drive, two high-performance Delta servo motors (from Taiwan) work in tandem with two Xinda (China) planetary gear reducers, effectively amplifying the output torque while reducing the speed, thus providing power support for the robot. Figure 1 provides a detailed illustration of the 3D model of the rehabilitation robot and its key components. In Fig. 1c, the servo motor and reducer are securely mounted to the profile frame via a fixed plate, with the coupling serving as the bridge for power transmission, transferring the motor's power to the linkages. Figure 1a illustrates the ingenious connection method between the linkages: the upper connecting block and the lower connecting block are securely joined together through the collaborative action of a fixed plate and bearings. Additionally, the fixed base plate encloses the lower connecting block, ensuring that the linkages maintain their rotational flexibility while preventing any vertical separation. Figure 1b focuses on the connection method of the end effector, where the inner shaft handle is directly secured to the linkage using screws, while the outer shaft handle is fixed to the inner shaft handle via a stepped shaft, allowing the handle to rotate independently. This design considers the dynamic changes in the angle between the patient's upper limb and the linkages during movement, providing the handle with an independent degree of rotational freedom and significantly reducing friction resistance during the rotation process through thrust bearings. Additionally, the rotating joints of the two linkages also utilize thrust bearings to minimize friction.

2.2 Kinematic Analysis

As shown in Fig. 2, A and C denote the drive pair connected to the two ends, B and D of the driving rod, respectively. The two supporting chains intersect at point P, which is designated as the output reference point. We established a fixed reference coordinate system $O - xy$, with its origin located at the midpoint of the line segment connecting A and C; the x-axis extends along the direction of AC, while the y-axis is oriented perpendicular to AC. Due to the structural symmetry, $AO = OC$,

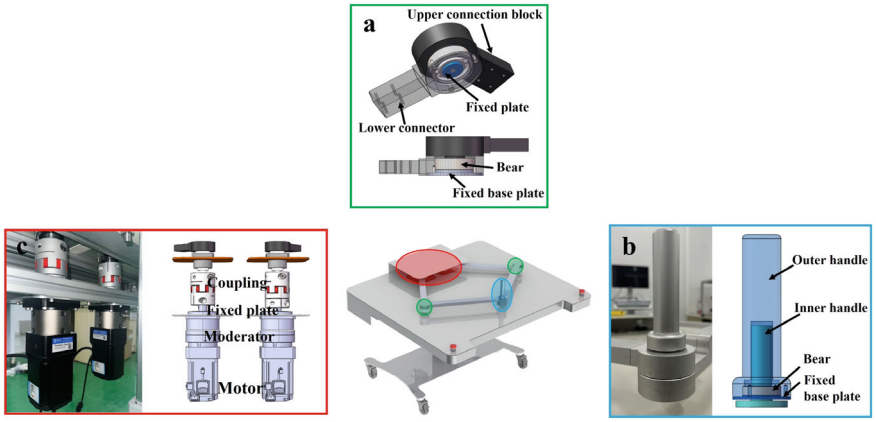


Fig. 1 Three-dimensional model of upper limb rehabilitation robot based on five-bar structure: **a** connection structure between legs; **b** grip structure; **c** driving structure

$AB = CD$ and $BP = DP$. For the length of the rod, $AO = OC = 100$ mm and $AB = CD = BP = DP = 500$ mm.

Inverse Kinematics. The position vectors of each point in the reference coordinate system are as follows:

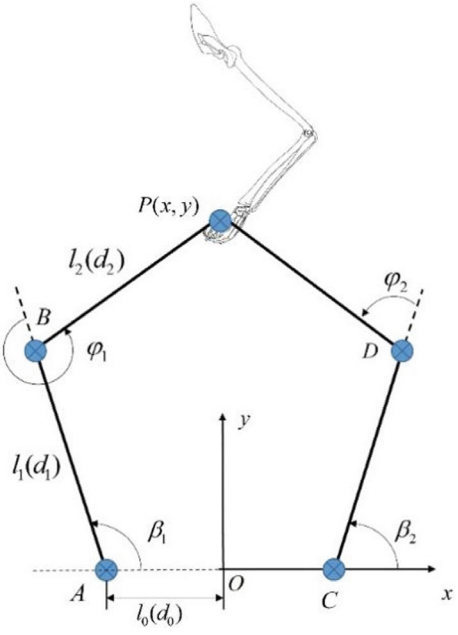


Fig. 2 Motion model of a five-bar upper limb rehabilitation robot

$$\begin{aligned}
\vec{a} &= \begin{bmatrix} -l_0 \\ 0 \end{bmatrix} \\
\vec{b} &= \vec{a} + l_1 \begin{bmatrix} \cos\beta_1 \\ \sin\beta_1 \end{bmatrix} \\
\vec{c} &= \begin{bmatrix} l_0 \\ 0 \end{bmatrix} \\
\vec{d} &= \vec{c} + l_1 \begin{bmatrix} \cos\beta_2 \\ \sin\beta_2 \end{bmatrix}
\end{aligned} \tag{1}$$

According to the constraint equation,

$$f_1: \quad \left\| \vec{p} - \vec{b} \right\| = l_2 \tag{2}$$

$$f_2: \quad \left\| \vec{p} - \vec{d} \right\| = l_2 \tag{3}$$

By squaring both sides of Eqs. (2) and (3), we obtain:

$$(x + l_0 - l_1\cos\beta_1)^2 + (y - l_1\sin\beta_1)^2 = l_2^2 \tag{4}$$

$$(x - l_0 - l_1\cos\beta_2)^2 + (y - l_1\sin\beta_2)^2 = l_2^2 \tag{5}$$

According to Eqs. (2) and (3), it can be obtained that:

$$\beta_1 = 2\tan^{-1}\frac{-A \pm \sqrt{A_1^2 + B_1 - C_1^2}}{C_1 - B_1} \tag{6}$$

$$\beta_2 = 2\tan^{-1}\frac{-A_2 \pm \sqrt{A_2^2 + B_2^2 - C_2^2}}{C_2 - B_2} \tag{7}$$

where

$$\begin{aligned}
A_1 &= 2yl_1 \\
B_1 &= 2l_1(x + l_0) \\
C_1 &= l_2^2 - l_1^2 - (x + l_0)^2 - y^2 \\
A_2 &= 2yl_1 \\
B_2 &= 2l_1(x - l_0) \\
C_2 &= l_2^2 - l_1^2 - (l_0 - x)^2 - y^2
\end{aligned} \tag{8}$$

The symbols ± in Eqs. (6) and (7) correspond to four distinct sets of inverse solution assembly structures. For the purposes of this study, we have selected the '±' set as the assembly configuration for the upper limb rehabilitation robot.

Passive Joint Angle. As shown in Fig. 3, the position of the output reference point p is known, and it can be obtained that:

$$L_{PA} = \sqrt{(x + l_0)^2 + y^2} \tag{9}$$

$$\varphi_1 = \pi + \cos^{-1} \frac{l_1^2 + l_2^2 - L_{PA}^2}{2l_1 l_2} \tag{10}$$

$$L_{PC} = \sqrt{(x - l_0)^2 + y^2} \tag{11}$$

$$\varphi_1 = \pi + \cos^{-1} \frac{l_1^2 + l_2^2 - L_{PA}^2}{2l_1 l_2} \tag{12}$$

Taking the time derivative of f_1 and f_2 in Eqs. (2) and (3), we obtain:

$$J_X \dot{\vec{p}} = J_q \dot{\vec{\beta}} \tag{13}$$

where p refers to the velocity vector and β represents the input velocity vector, defined as follows: $\dot{\vec{p}} = \begin{bmatrix} \dot{x} \\ \dot{y} \end{bmatrix}$ and $\dot{\vec{\beta}} = \begin{bmatrix} \dot{\beta}_1 \\ \dot{\beta}_2 \end{bmatrix}$.

$$J_x = \begin{bmatrix} l_0 + x - l_1 \cos\beta_1 & y - l_1 \sin\beta_1 \\ x - l_2 - l_1 \cos\beta_2 & y - l_1 \sin\beta_2 \end{bmatrix} \tag{14}$$

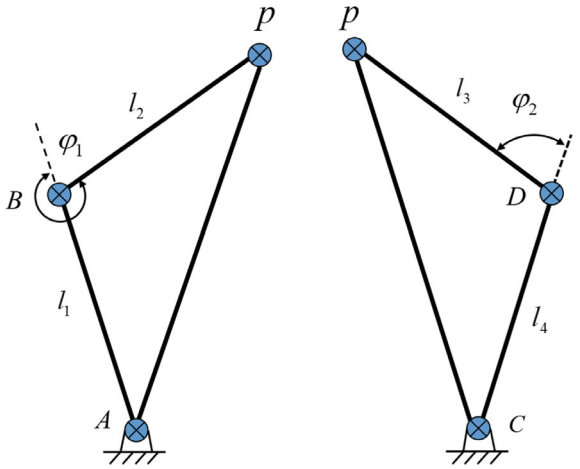


Fig. 3 Passive joint analysis diagram

$$J_q = \begin{bmatrix} l_1(y\cos\beta_1 - l_0\sin\beta_1 - x\sin\beta_1) & 0 \\ 0 & l_1(y\cos\beta_2 + l_2\sin\beta_2 - x\sin\beta_2) \end{bmatrix} \quad (15)$$

The robot's inverse kinematics Jacobian matrix J_I is given by:

$$J_I = J_q^{-1} J_x \quad (16)$$

3 Experimental Analysis

3.1 Singular Analysis

A singularity refers to a point in a mechanism's motion where its position cannot be determined or motion cannot occur, as the kinematic and dynamic performance of the mechanism in the neighboring area significantly deteriorates. Therefore, such points must be avoided or stayed away from as much as possible. For this device, singular poses mainly fall into serial singularity and parallel singularity. In the case of serial singularities, it is manifested explicitly as a specific co-linear arrangement between the rods, as illustrated in Fig. 4, where rod 1 overlaps with rod 2 or rod 3 overlaps with rod 4 along the same straight line.

As illustrated in Fig. 5, when rod 2 and rod 3 are collinear, it is referred to as a parallel singularity, and there are two such singular points. In the context of the two-degree-of-freedom five-bar mechanism, the alignment of rods 2 and 3 results in a situation where the mechanism remains inoperative regardless of the magnitude of the driving force applied. Consequently, the trajectory of point p is as follows:

$$\begin{cases} x = l_1(\cos\beta_2 + \cos\beta_1)/2 \\ y = l_1(\sin\beta_2 + \sin\beta_1)/2 \end{cases}$$

$$\beta_1 \in [0, \pi]$$

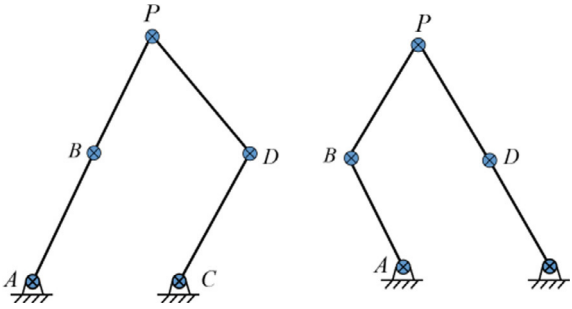
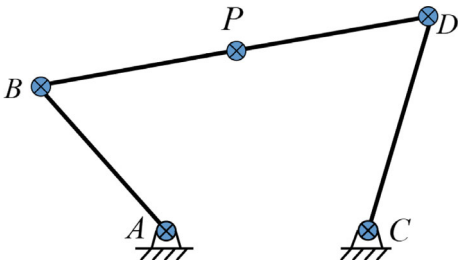


Fig. 4 Serial singularity

Fig. 5 Parallel singularity



$$\beta_2 = 2\tan^{-1}(k)$$

$$k = \left(-b \pm \sqrt{b^2 - 4ac}\right)/(2a) \tag{17}$$

$$a = 4l_2^2 + 4l_1 l_0 \cos\beta_1 + 4l_1 l_0 - 4l_0^2 - 2l_1^2 - 2l_1^2 \cos\beta_1$$

$$b = 4l_1^2 \sin\beta_1$$

$$c = 4l_2^2 + 4l_1 l_0 \cos\beta_1 - 4l_1 l_0 - 4l_0^2 - 2l_1^2 + 2l_1^2 \cos\beta_1$$

In the context of the two-degree-of-freedom five-bar mechanism, the occurrence of singular positions is primarily associated with parallel singularities, attributable to its distinctive motion characteristics. Nevertheless, it is possible for the dual-degree-of-freedom five-bar structure to circumvent these singular positions through the adjustment of the two input motions.

3.2 Determination of Robot Workspace

The design of the robot's accessible workspace is based on anthropometric standards. In the prescribed scenarios, the subjects were positioned 20 cm from the tabletop, taking into account the individualized arm length and relative distance from the torso. Based on the proportionality between height and each part of the upper limb, we assumed that the subject's height was 1.7 m, humerus length was 21.93 cm, upper arm length was 31.62 cm, and forearm length was 24.82 cm. As shown in Fig. 6, the coverage area of the working area of the upper limb on the horizontal plane was calculated using MATLAB.

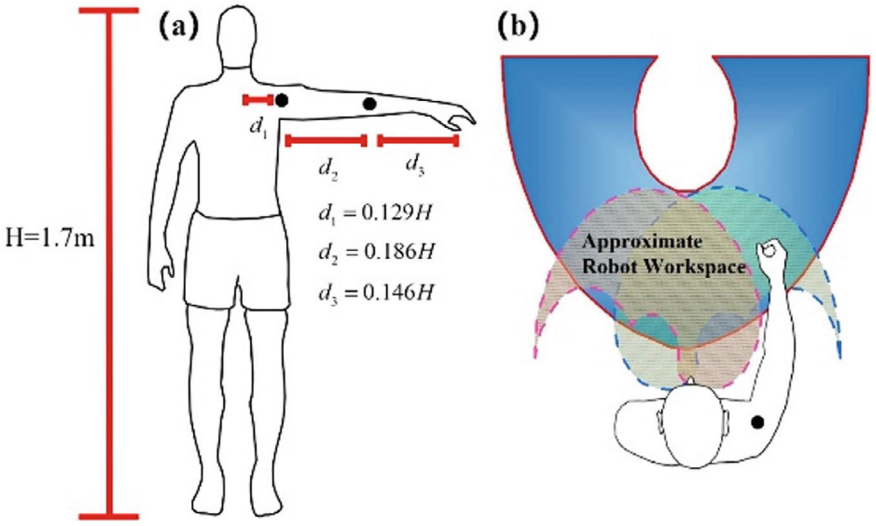


Fig. 6 **a** The structure of the device is designed to allow for a workspace, including the patient's arrival distance, to be used for measuring the human body; **b** use this standardized data to enable the robot's workspace to accommodate participants

3.3 Experimental Setup and Comparison of Results

The robot assists patients in using their residual muscle power for training through two modes: active-assisted and active-resistance, tailored for severe and moderate patients, respectively. In active-assisted mode, the robot features high-level and low-level active training controllers. The high-level controller applies the AAN principle and motor trajectory planning while regulating the low-level impedance controller. An example of trajectory planning includes a normal time-containing trajectory designed for arrival motion training.

$$\frac{x(t)-x_i}{x_d-x_i}=\frac{y(t)-y_i}{y_d-y_i}=8(t/\tau)^3-12(t/\tau)^4+5(t/\tau)^5 \tag{18}$$

where (x_i, y_i) and (x_d, y_d) represent the start and end point coordinates, t is the time parameter, and τ is the cycle time that regulates the interpolation duration. The 5th order polynomial form guarantees smooth transitions in position, velocity, acceleration, and jerk. Additionally, the high-level controller accounts for variations in patient motion, as illustrated by stretching exercises along the x-axis.

$$\begin{aligned} F_x &= \begin{cases} -k_x(x-x_{\text{ref}})-b_x\dot{x}, & x_{\text{th}}<x<x_{\text{ref}} \\ 0, & x<x_{\text{th}} \text{ or } x>x_{\text{ref}} \end{cases} \\ F_y &= \begin{cases} -k_y(|y|-|y_{\text{wall}}|)-b_y\dot{y}, & |y|>w_{\text{wall}} \\ 0, & |y|<w_{\text{wall}} \end{cases} \end{aligned} \tag{19}$$

F_x and F_y represent the robot's assistance and resistance in forward and vertical movements. The forward stiffness k_x is lower than the vertical stiffness k_y, allowing patients to exert more effort when moving forward. Elastic walls create a vertical deviation tolerance, facilitating patient exploration of the "tunnel" and aiding in motor relearning [20]. Robotic assistance activates only when patient movements exceed a defined threshold (x_{th}). This "virtual tunnel" and "assist threshold" strategy provides greater freedom in motor skill relearning compared to traditional error-based assistance, with the robot intervening only as necessary. This design integrates an adaptive strategy with a low-level impedance controller, demonstrating "on-demand assistance," which enhances patient engagement and feedback, proving more effective than passive training [21].

The low-level controller, based on impedance control [22], is illustrated in Fig. 7. Robot assistance depends on trajectory errors and velocities, where L represents robot kinematics, X is the end-effector position, J is the Jacobian matrix, and stiffness K and damping B are expressed as matrices. Two motor drives operate in current control mode, with torque reference Tm derived from the desired force F in joint space. Stability assessment is vital for patient safety, as previous research indicates that stiffness K must be limited to ensure stability, while virtual damping B enhances it. Although determining the maximum K is challenging, trials suggest it should be around 5000 N/m. Stiffness is maintained below 1000 N/m during rehabilitation training to ensure stability and patient safety.

The experimenter is 170 cm tall and weighs 70 kg. The initial point (288.630) corresponds to t = 0 on the circle. At the end time, t = 2π, the position returns to the initial point. Applying Eq. (18) for interpolation along the circular trajectory gives the following expression:

$$x(t) = x_i + (x_d - x_i)\left(8\left(\frac{t}{\tau}\right)^3 - 12\left(\frac{t}{\tau}\right)^4 + 5\left(\frac{t}{\tau}\right)^5\right)$$

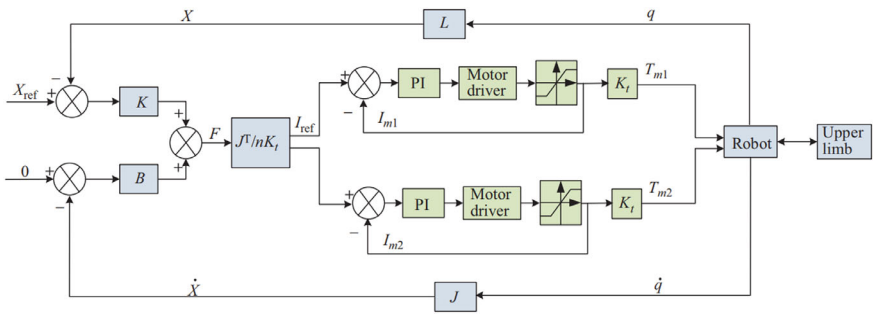


Fig. 7 Structure of the low level controller, where the joint torques are regulated through inner motor current control loop, K_t represents the motor torque constant, q = (θ_1, θ_2) is the joint space variable, and I_{m1} and I_{m2} represent the motor currents

$$y(t) = y_i + (y_d - y_i)\left(8\left(\frac{t}{\tau}\right)^3 - 12\left(\frac{t}{\tau}\right)^4 + 5\left(\frac{t}{\tau}\right)^5\right) \tag{20}$$

First, we affixed marker points to the upper limb of the subject, as illustrated in Fig. 8a. Subsequently, using the Qualisys three-dimensional motion capture system, we precisely captured and recorded the dynamic data of these markers during the passive movement of the subject, as shown in Fig. 8c. Next, we selected the Bimanual Upper Arm Model, which closely resembles human physiological structure, as the simulation model on the OpenSim platform. This model possesses matching degrees of freedom and integrates simulations of the major muscle groups, as depicted in Fig. 8b (muscles are hidden in the illustration). We made detailed scaling adjustments to the model's physical parameters based on the specific characteristics of each subject and added virtual markers corresponding to the marker locations on the subject's upper limb. Thereafter, we seamlessly integrated this customized Bimanual Upper Arm Model with the motion capture system, imported the previously collected motion data, and initiated the simulation process, as presented in Fig. 8d. Through this comprehensive approach, we could accurately simulate the activation levels of the muscles in various parts of the upper limb and the dynamic changes in joint angles during the subject's movements.

The detailed data presented in Fig. 9 shows that the initial posture of the shoulder joint is set to a flexion of 38°, internal rotation of 6°, and abduction of 23°. During the passive rehabilitation training, the upper limb's shoulder joint moves from an initial flexion of 40° to an extension of 66°, with a maximum flexion angle reaching 55°

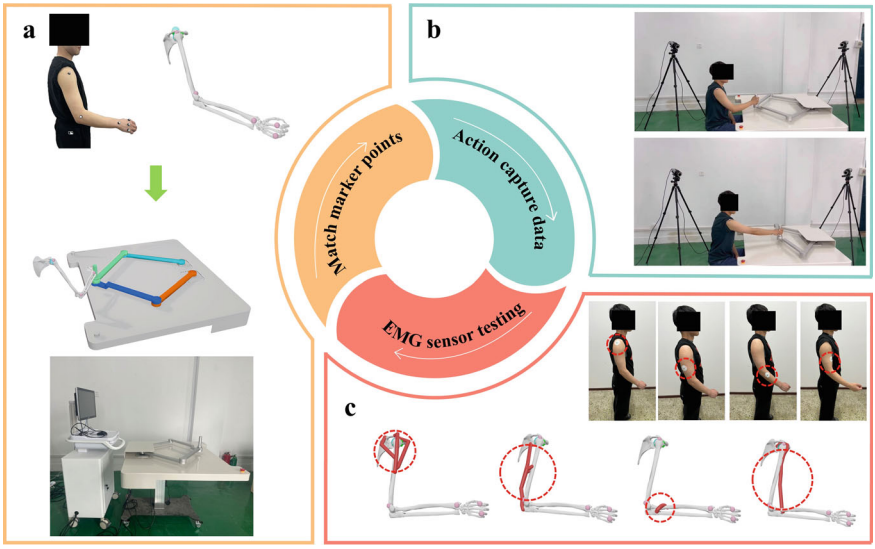


Fig. 8 Simulation analysis process: **a** location of subject marker points; **b** location of marker points in the model; **c** motion capture process; **d** simulation process

and a total range of motion of 121°. In terms of internal rotation, the shoulder joint transitions from a maximum of 19° of internal rotation to 28° of external rotation; the maximum adduction angle is 17°, while abduction extends to 44°. The elbow joint gradually changes from an initial 103°, flexing to a minimum of 23° and extending to a maximum of 116°. The initial angle for internal rotation is 44°, with a maximum of 48° and a minimum of 20° for internal rotation. The experimental results indicate that the movements of the shoulder and elbow joints are smooth, fulfilling the necessary motion range without significant fluctuations, which can provide rehabilitation training for patients.

The activation levels of the deltoid, triceps, supraspinatus, and biceps were evaluated during a predetermined circular motion trajectory using EMG sensors, as illustrated in Fig. 10. The experiments utilized a sampling rate of 3000 Hz with necessary preprocessing. To address noise in the raw EMG data, a 30–500 Hz bandpass filter was applied for continuous processing. Additionally, activation level data from the EMG sensors and OpenSim simulation outputs were smoothed using MATLAB functions. In addition, we smoothed the activation level data obtained from the EMG sensors and the activation output from OpenSim simulations using MATLAB functions.

Comparative analyses, as shown in Fig. 11, indicate that the deltoid and biceps muscles exhibit significant activation during passive rehabilitation training in elbow flexion. In contrast, the triceps brachii showed a relatively low level of activation. In contrast, when the elbow was extended, the activation level of the biceps decreased,

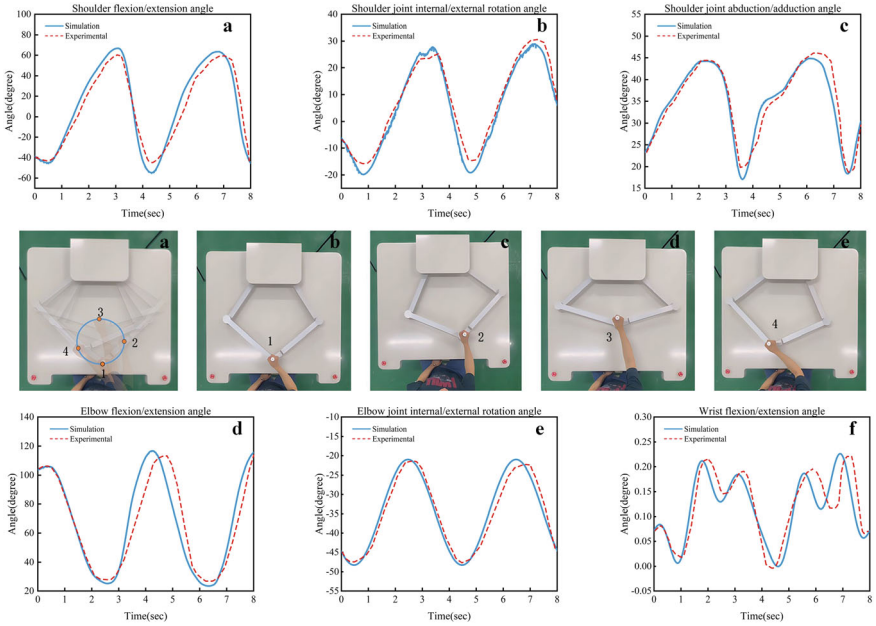


Fig. 9 Changes in joint angle

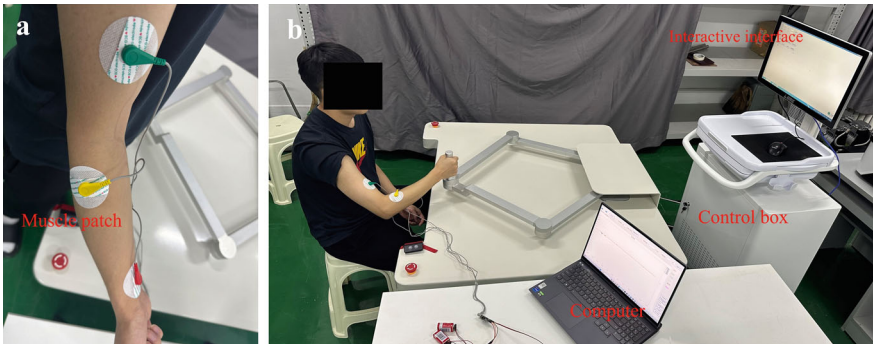


Fig. 10 Muscle strength sensor testing

whereas the activation level of the triceps increased significantly. The results indicated that the major muscle groups of the upper limb were effectively activated during exercise, which fully met the requirements of rehabilitation training. Further observation of the activation curves revealed that the simulated data were highly similar to the experimental data. The small deviation between the two may be due to the small differences between the marker positions of the upper limbs in the simulation model and those used on the patient. The device achieved the desired goal of providing rolling support for improved upper limb function through repetitive practice of the training paradigm [8–10]. It effectively addressed gaps in care needs [19]. Specifically, the intervention significantly enhanced upper limb function and facilitated close collaboration between the rehabilitation robot and clinicians.

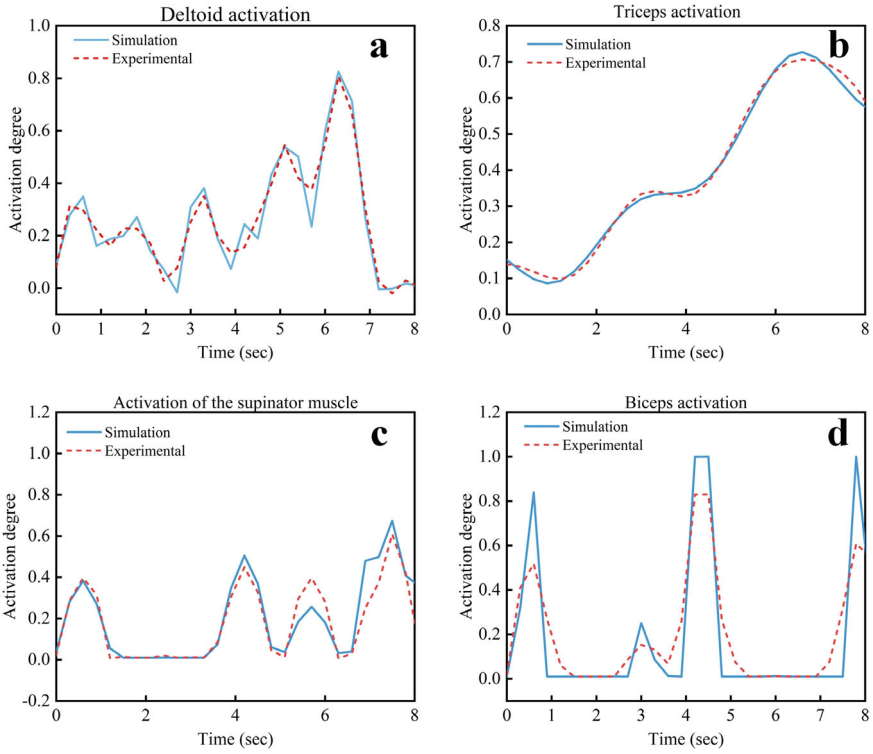


Fig. 11 Muscle activation levels: **a** Deltoid activation level; **b** Triceps activation level; **c** Supinator activation level; **d** Biceps activation level

4 Conclusion

This paper presents the design of a five-bar upper limb rehabilitation robot characterized by its parallel structure, compact layout, and aesthetic appearance. We conducted an in-depth analysis of the kinematic characteristics of the upper limb rehabilitation robot and thoroughly explored potential issues related to singular configurations. To validate the robot's rehabilitation effectiveness, we employed motion capture equipment to compare the actual motion trajectory with the planned trajectory under rehabilitation mode. Additionally, by comparing the data from Myoelectric Sensors with the muscle activation simulated by OpenSim software, we confirmed that the activation levels of the major muscles within the specified joint angle range can achieve the desired repetitive movements, thus effectively supporting rehabilitation goals. Looking ahead, we plan to expand the scope of our research by collecting and analyzing more training data. Furthermore, we will optimize the structure to make the five-bar upper limb rehabilitation robot even more compact while also incorporating interactive game modes to enhance the enjoyment of the rehabilitation process for patients.

Acknowledgements Research Project of the Jilin Provincial Department of Education on Scientific and Technological Research: Intelligent Elbow Rehabilitation Exercise Evaluation System (No.: JJKH20230755KJ).

References

1. Peers PV, Astle DE, Duncan J et al (2020) Dissociable effects of attention vs working memory training on cognitive performance and everyday functioning following fronto-parietal strokes. Neuropsychol Rehab 30(6):1092–1114
2. Wang YN, Liu XH (2023) Correlation and predictive effectiveness of subjective and objective balance function measurement in hemiparetic patients after stroke. Chin J Rehab Theory Pract 29(08):890–895
3. Yang JZ, Li D, Bei BS et al (2021) Associated mirror therapy enhances motor recovery of the upper extremity and daily function after stroke: a randomized control study. Neural Plast 2021:7266263
4. Wang CK, Liu LL, Liu XY et al (2021) Study on the rehabilitation effect of mirror therapy robotic rehabilitation on hand motor dysfunction during the subacute phase of stroke. Chin J Stroke 16(03):224–229
5. Yuanwen L, Mingyu Y, Jing L et al (2020) Effects of transcranial magnetic stimulation on the performance of the activities of daily living and attention function after stroke: a randomized controlled trial. Clin Rehabil 34(12):269215520946386
6. Richards L, Hanson C, Wellborn M et al (2008) Driving motor recovery after stroke. Top Stroke Rehabil 15(5):397–411
7. Veerbeek JM, van Wegen E, van Peppen R et al (2014) What is the evidence for physical therapy poststroke? A systematic review and meta-analysis. PLoS ONE 9(2):e87987
8. Sullivan KJ, Brown DA, Klassen T et al (2007) Effects of task-specific locomotor and strength training in adults who were ambulatory after stroke: results of the STEPS randomized clinical trial. Phys Ther 87(12):1580–1602
9. Hubbard IJ, Parsons MW, Neilson C et al (2009) Task-specific training: evidence for and translation to clinical practice. Occup Ther Int 16(3–4):175–189
10. Shepherd RB (2001) Exercise and training to optimize functional motor performance in stroke: driving neural reorganization? Neural Plast 8(1–2):121–129
11. Morrissey MC, Harman EA, Johnson MJ (1995) Resistance training modes: specificity and effectiveness. Med Sci Sports Exer 27(5):648–660
12. Maciejasz P, Eschweiler J, Gerlach-Hahn K, Jansen-Troy A, Leonhardt S (2014) A survey on robotic devices for upper limb rehabilitation. J Neuroeng Rehabil 11:1–29
13. Avizzano CA, Satler M, Cappiello G, Scoglio A, Ruffaldi E, Bergamasco M (2011) MOTORE: a mobile haptic interface for neuro-rehabilitation. In: Proceedings of the 2011 RO-MAN, pp 383–388
14. Hogan N, Krebs HI, Charnnarong J, Srikrishna P, Sharon A (1992) MIT-MANUS: a workstation for manual therapy and training. I. In: Proceedings of the IEEE international workshop on robot and human communication, pp 161–165
15. Zuo S, Li J, Dong M et al (2021) Optimum design and preliminary experiments of a novel parallel end traction apparatus for upper-limb rehabilitation. Front Mech Eng 1–21
16. Research on trajectory planning for a three-dimensional space end-traction upper limb rehabilitation robot. Chin J Rehabil Med 11:036 (2021)
17. Mingjie D, Wenpei F, Jianfeng L et al (2022) Patient-specific exercises with the development of an end-effector-type upper limb rehabilitation robot. J Healthc Eng 2022:4125606–4125606
18. Yan H, Wang HB, Chen P et al (2021) Research on optimal upper limb rehabilitation robot design with generalized shoulder joint. Acta Armamentarii 42(11):2491–2502

19. Lee BO, Saragih ID, Batubara SO (2023) Robotic arm used for upper limb rehabilitation after stroke: a systematic review and meta-analysis. Kaohsiung J Med Sci 39(5):435–445
20. Lo AC, Guarino PD, Richards LG et al (2010) Robot-assisted therapy for long-term upper-limb impairment after stroke. N Engl J Med 362(19):1772–1783
21. Krebs HI, Palazzolo JJ, Dipietro L et al (2003) Rehabilitation robotics: performance-based progressive robot-assisted therapy. Auton Robot 15:7–20
22. Hogan N (1984) Impedance control: an approach to manipulation. In: 1984 American control conference. IEEE, pp 304–313

Open Access This chapter is licensed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License (http://creativecommons.org/licenses/by-nc-nd/4.0/), which permits any noncommercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if you modified the licensed material. You do not have permission under this license to share adapted material derived from this chapter or parts of it.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.



Vehicle Trajectory Tracking and Optimization Based on Lateral and Longitudinal Control

Jie Tang, Zhiwen Zhang, and Jiyuan Zhang

Abstract The objective of this paper is to examine the issues of low accuracy and poor stability in vehicle trajectory tracking under a transverse controller, with a view to investigating the transverse and longitudinal control of vehicle trajectory tracking. Firstly, an MPC transverse controller based on a single-point pre-scan model is constructed. Secondly, a longitudinal controller with PID as the upper layer and vehicle dynamics as the lower layer is constructed. Once more, GA is applied to optimise the parameters of the MPC controller globally offline. Finally, a classical double-shift line trajectory tracking simulation model has been constructed using MATLAB and Carsim software. The results demonstrated that the average deviation of transverse displacement was reduced by 16.76%, and the average transverse angular velocity was reduced by 6.79%, under MPC+PID controllers relative to MPC control. Following the implementation of GA optimization, significant enhancements have been made to the performance of the MPC+PID controllers.

Keywords Model prediction · Inverse longitudinal dynamics · Genetic algorithm

1 Introduction

Trajectory tracking represents a fundamental aspect of autonomous driving technologies. Li et al. [1] proposes an improved artificial potential field method to accomplish the planning of trajectories by adding distance adjustment factors and so on. The corresponding constraint variables are set and the optimal objective function is established to construct the MPC controller to achieve the goal of trajectory tracking. Fan et al. [2] proposed a weighted matrix optimization method based on Improved Quantum Genetic Algorithm (IQGA) to solve the weight selection problem during

J. Tang · Z. Zhang (✉) · J. Zhang
College of Vehicle and Energy, Yanshan University, Qinhuangdao, China
e-mail: zhzhw@ysu.edu.cn

Z. Zhang
Hebei Key Laboratory of Special Carrier Equipment, Qinhuangdao, China

© The Author(s) 2026

H. Saman et al. (eds.), *The 9th International Conference on Advances in Construction Machinery and Vehicle Engineering*, Lecture Notes in Mechanical Engineering,
https://doi.org/10.1007/978-981-96-5527-4_102

trajectory tracking of unmanned wheeled tractor using Linear Quadratic Regulator (LQR). Simulation results show that the proposed IQGA has significant advantages over other algorithms in terms of performance. Model prediction control is a widely employed technique in the trajectory tracking of intelligent vehicles, due to its capacity to address multi-constraint and optimal value problems [3]. Gao et al. [4] developed a highly accurate nonlinear error vehicle model, which was subsequently linearized using Taylor's theorem and implemented within a MPC framework. The simulation results indicated an improvement in tracking accuracy. Wu et al. [5] focused on a four-wheeled, independently driven intelligent vehicle, incorporating vehicle dynamic stability into the constraints of the MPC controller. The simulations demonstrated superior tracking performance.

This paper explores vehicle trajectory tracking utilizing an MPC+PID approach for the purposes of transverse and longitudinal control. In Sect. 2, a horizontal controller based on the MPC controller is proposed; in the Sect. 3, a vertical controller is proposed; in the Sect. 4, the GA optimises the controller parameters; in the Sect. 5, the simulation is verified; and in the Sect. 6, the conclusion is drawn.

2 Lateral Controller Design

2.1 Vehicle Model

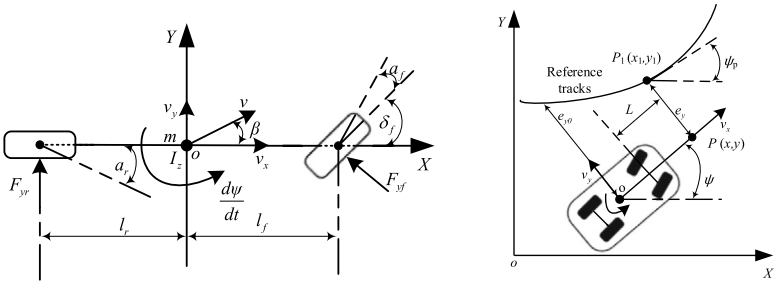
The two-degree-of-freedom model of the vehicle is depicted in Fig. 1b. The dynamic equations that govern this system are as follows [6]:

$$\begin{cases} \dot{v}_y = -\dfrac{K_f + K_r}{m v_x} v_y - \left(\dfrac{l_f K_f - l_r K_r}{m v_x} + v_x \right) \dot{\psi} + \dfrac{K_f}{m} \delta_f \\ \ddot{\psi} = -\dfrac{l_f K_f - l_r K_r}{I_z v_x} v_y - \dfrac{l_f^2 K_f + l_r^2 K_r}{I_z v_x} \dot{\psi} + \dfrac{l_f K_f}{I_z} \delta_f \end{cases} \tag{1}$$

In the formula, m represents the mass; v signifies the velocity; $\dot{\psi}$ indicates the transverse angular velocity; I_z represents the rotational moment of inertia; l denotes the distance from the center to the axis; K represents the lateral deflection stiffnesses; δ_f represents the front wheel angle.

The underlying concept of the single-point pre-scanning model is illustrated in Fig. 1b. The lateral displacement deviation and the heading angle deviation are defined as follows [7]:

$$\begin{cases} \dot{e}_y = v_y + v_x e_\psi + L\dot{\psi} \\ \dot{e}_\psi = \dot{\psi} - \dot{\psi}_p = \dot{\psi} - k_d v_x \end{cases} \tag{2}$$



(a) Two degree of freedom vehicle model (b) Single-point pre-scanning model

Fig. 1 Vehicle model

In the formula, k_d signifies the curvature of the point p_1; L denotes the pre-scanning distance.

2.2 MPC Controller Design

The state variables of the system are defined as $\boldsymbol{x} = (e_\mathrm{y}, v_\mathrm{y}, e_\psi, \dot{\psi})$, the control variable is designated as $u = \delta_\mathrm{f}$, and the system outputs are represented as $\boldsymbol{y} = (e_\mathrm{y}, e_\psi, \dot{\psi})$. The system state space equation, as follows:

$$\begin{cases} \dot{\boldsymbol{x}} = \boldsymbol{A}\boldsymbol{x} + \boldsymbol{B}u + \boldsymbol{w} \\ \boldsymbol{y} = \boldsymbol{C}\boldsymbol{x} \end{cases} \tag{3}$$

In the formula, $\boldsymbol{A} = \begin{bmatrix} 0 & 1 & v_x & L \\ 0 & -\frac{K_f+K_r}{mv_x} & 0 & \frac{l_rK_r-l_fK_f}{mv_x} - v_x \\ 0 & 0 & 0 & 1 \\ 0 & \frac{l_rK_r-l_fK_f}{I_zv_x} & 0 & -\frac{l_f^2K_f+l_r^2K_r}{I_zv_x} \end{bmatrix}$, $\boldsymbol{B} = \begin{bmatrix} 0 \\ \frac{K_f}{m} \\ 0 \\ \frac{l_fK_f}{I_z} \end{bmatrix}$, $\boldsymbol{C} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$, $\boldsymbol{w} = (0, 0, -k_d v_x, 0)^T$.

In order to smooth the corner curve, the state variables are redefined as $\boldsymbol{\xi}_\mathrm{k} = (\boldsymbol{x}_\mathrm{k}, u_{\mathrm{k}-1})^\mathrm{T}$. This allows the state space equations after dimension expansion to be obtained as follows:

$$\begin{cases} \boldsymbol{\xi}_{k+1} = \widetilde{\boldsymbol{A}}_d\boldsymbol{\xi}_k + \widetilde{\boldsymbol{B}}_d\Delta u_k + \widetilde{\boldsymbol{w}}_k \\ \boldsymbol{y}_k = \widetilde{\boldsymbol{C}}_d\boldsymbol{\xi}_k \end{cases} \tag{4}$$

In the formula, $\widetilde{\boldsymbol{A}}_d = T\begin{bmatrix} \boldsymbol{I}+\boldsymbol{A} & \boldsymbol{B} \\ 0_{1\times 4} & \boldsymbol{I} \end{bmatrix}; \widetilde{\boldsymbol{B}}_d = T\begin{bmatrix} \boldsymbol{B} \\ \boldsymbol{I} \end{bmatrix}; \widetilde{\boldsymbol{C}}_d = (\boldsymbol{C}, 0_{3\times 1}); \widetilde{\boldsymbol{w}}_k = T\begin{bmatrix} \boldsymbol{w} \\ \boldsymbol{I} \end{bmatrix}$.

The system state variables are iterated through Eq. 4 to obtain the predicted time-domain state variables. Upon substituting the predicted time-domain state variables into output Eq. 4 of the system, the following prediction model is obtained:

$$\boldsymbol{Y}_k = \widetilde{\boldsymbol{C}}_p \boldsymbol{X}_k + \widetilde{\boldsymbol{D}}_p \Delta \boldsymbol{U}_k + \widetilde{\boldsymbol{E}}_p \widetilde{\boldsymbol{W}}_k \tag{5}$$

In the formula, $\boldsymbol{Y}_{\mathrm{k}}$ represents the output of the system; $\boldsymbol{X}_{\mathrm{k}}$ denotes the state variable of the system; $\triangle \boldsymbol{U}_{\mathrm{k}}$ indicates the control increment of the system; $\tilde{\boldsymbol{W}}$p stands for the disturbance of the system; $\tilde{\boldsymbol{C}}$p, $\tilde{\boldsymbol{D}}$p, $\tilde{\boldsymbol{E}}$p indicate the coefficient matrix.

The objective function for the MPC lateral controller is defined by the following equation:

$$\boldsymbol{J} = \boldsymbol{Y}^T \boldsymbol{Q} \boldsymbol{Y} + \Delta \boldsymbol{U}^T \boldsymbol{R} \Delta \boldsymbol{U} + \rho \varepsilon^2 \tag{6}$$

In the formula, $\boldsymbol{Q}$ represents the output weight factor, $\boldsymbol{R}$ denotes the control increment weight factor, ε indicates the relaxation factor, and ρ stands for the relaxation factor weight factor.

The variables e_{y}, e_{ψ}, $\dot{\psi}$ and $\triangle u$ are constrained to a specified range. The objective function is to be solved using quadratic programming, as follows:

$$\begin{aligned} \min \quad & \boldsymbol{J} = \left(\Delta \boldsymbol{U}^T, \varepsilon\right)^T \boldsymbol{H} \left(\Delta \boldsymbol{U}^T, \varepsilon\right) + \boldsymbol{G}\left(\Delta \boldsymbol{U}^T, \varepsilon\right) \\ s.t. \quad & \boldsymbol{y}_{\min,t+k} - \varepsilon \le \boldsymbol{y}_{t+k} \le \boldsymbol{y}_{\max,t+k} + \varepsilon \\ & u_{\min,t+k} \le u_{t+k} \le u_{\max,t+k} \\ & \Delta u_{\min,t+k} \le \Delta u_{t+k} \le \Delta u_{\max,t+k} \end{aligned} \tag{7}$$

In the formula, $\boldsymbol{H} = diag\left(\widetilde{\boldsymbol{D}}_p^T \boldsymbol{Q} \widetilde{\boldsymbol{D}}_p + \boldsymbol{R}, \rho\right)$, $\boldsymbol{G} = \left(2\left(\widetilde{\boldsymbol{C}}_p \boldsymbol{\xi}_k\right)^T \boldsymbol{Q} \widetilde{\boldsymbol{D}}_p, 0\right)$.

3 Longitudinal Controller Design

3.1 PID Upper Controller

The PID controller offers a straightforward principle. This paper employs PID as the upper controller, defining the input variable e(t) as the discrepancy between the actual and desired speeds, and the output quantity u(t) as the desired acceleration. The relationship between input and output quantities is as follows:

$$u(t) = k_p e(t) + k_I \int e(t) d_t + k_D \dot{e}(t) \tag{8}$$

In the formula, k_{p} represents the proportionality coefficient; k_{I} denotes the integration coefficient; k_{D} signifies the differentiation coefficient.

3.2 Lower Layer Controller Build

Under the assumption that the vehicle is travelling on a level road surface, the vehicle driving equation is given by the following equation:

$$\frac{\eta_T T_{tq} i_g i_0}{r} = mgf + \frac{1}{2} C_D A \rho_w v_x^2 + ma_x + KP_{des} \tag{9}$$

In the formula, η_{T} represents the mechanical efficiency; T_{tq} denotes the engine torque; i_{g} signifies the transmission ratio; i_0 represents the main reduction gear ratio; r signifies the rolling radius; f stands for the rolling resistance coefficient; a_{x} represents the acceleration; C_{D} denotes the air resistance coefficient; A stands for the windward area; ρ_{w} signifies the air density; P_{des} signifies the braking pressure; and K denotes the fixed ratio of braking force to braking pressure.

The desired torque is as follows when the vehicle is in motion:

$$T_{des} = \frac{\left(mgf + \frac{1}{2} C_D A \rho_w v_x^2 + ma_{des}\right) r}{i_g \eta_T i_0} \tag{10}$$

Furthermore, the relationship between throttle opening and desired torque is as follows:

$$\alpha_{\mathrm{thr}} = F(T_{des}, n) \tag{11}$$

In the formula, a_{thr} represents the throttle opening and n denotes the engine speed.

When the vehicle is in a state of deceleration, the driving force is equal to zero, while the braking force is not equal to zero. the braking pressure is defined as follows:

$$P_{des} = \frac{\left| ma_{des} + mgf + \frac{1}{2} C_D A \rho_w v^2 \right|}{K} \tag{12}$$

4 GA-Based Global Offline Optimisation of MPC Weight Coefficients

In this paper, the fitness function of the GA is presented in the following equation:

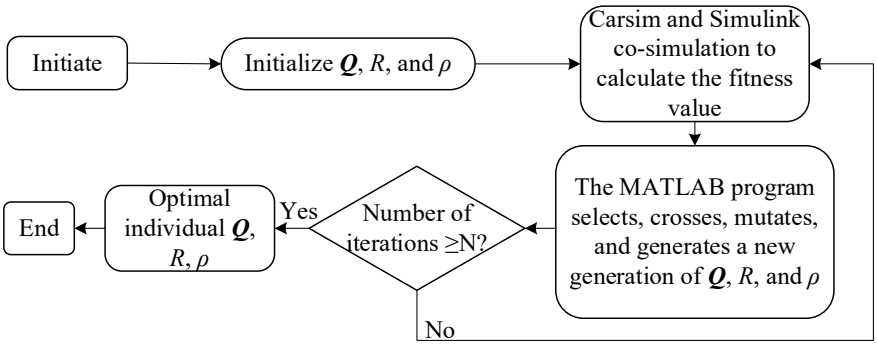


Fig. 2 GA optimisation flowchart

$$j = \int_0^t \sigma \left| e_{y0} \right| + (1 - \sigma) \left| \dot{\psi} \right| \tag{13}$$

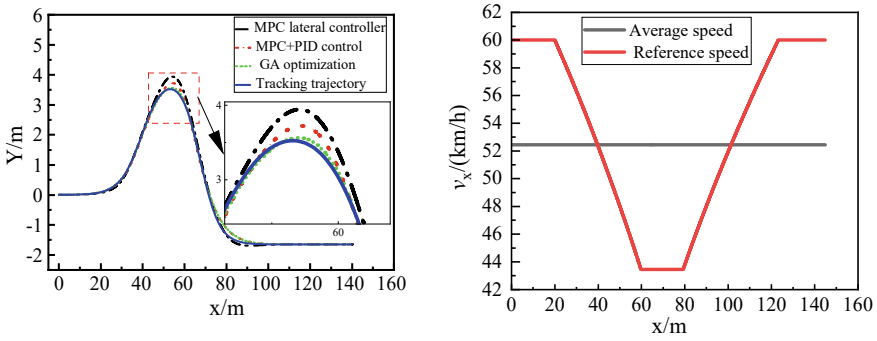
In the formula, σ signifies the weighting factor; e_{y0} denotes the center displacement deviation.

The weight coefficients to be optimised for the MPC controller include $\boldsymbol{Q} = \text{diag}(q_1, q_2, q_3)$, R, and ρ. Consequently, this paper sets the number of the optimisation parameters are set to five, the number of populations is 50, and the number of iterations is 12. The crossover probability is 0.8, the elite probability is 0.05. The optimisation process is illustrated in Fig. 2 below. Finally, the optimal individuals are $\boldsymbol{Q} = \text{diag}(65, 8, 48)$, $R = 9452$, and $\rho = 3975$, respectively.

5 Simulation Verification

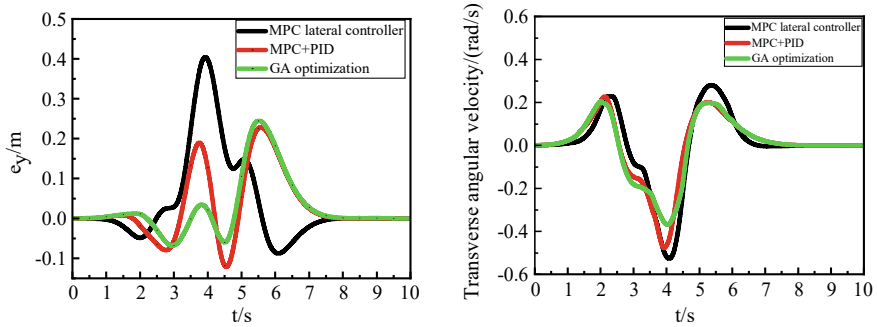
The tracking trajectory is illustrated in Fig. 3a, the reference speed [8] is depicted in Fig. 3b, and the road friction coefficient is 0.85. Prior to GA unoptimisation, the weight coefficients of MPC are $\boldsymbol{Q} = \text{diag}(100, 100, 100), \rho = 1000, R = 1 \times 10^5$. The average vehicle speed was employed as the driving speed under the MPC controller.

From Figs. 3a and 4a, it can be observed that the GA-optimised e_y is the smallest at the bends from 40 to 60m, followed by the e_y under MPC + PID controllers, and the largest e_y under the MPC controller. As evidenced in Table 1, the e_{y_avg} is diminished by 16.76% and 19.62%, respectively. The e_{y_max} is markedly reduced under the MPC+PID controller. As illustrated in Fig. 4b, the transversal angular velocities that have been globally optimised by the GA demonstrate the most favourable performance. The mean transverse angular velocity was found to be reduced by 6.79% and 1.59%, respectively. It was observed that the maximum transverse angular velocity was markedly diminished.



(a)Tracking trajectory (b) Reference speed

Fig. 3 Tracking trajectory and reference speed



(a) Lateral displacement deviation (b) Transverse angular velocity

Fig. 4 Tracking accuracy simulation results

Table 1 Comparison of control strategies

Index	MPC	MPC + PID	GA
e_{y_avg}/m	0.0637	0.053	0.0427
e_{y_max}/m	0.4041	0.2286	0.24488
$\dot{\psi}_{avg}$/(rad/s)	0.1012	0.0944	0.0928
$\dot{\psi}_{max}$/(rad/s)	0.5267	0.4774	0.3695

6 Conclusion

This paper is to propose an MPC+PID control strategy for vehicle trajectory tracking. The results demonstrate that, in comparison to the MPC transverse controller, the MPC+PID have been demonstrated to reduce the average lateral displacement deviation by 16.76%. The average transverse angular velocity is reduced by 6.79%. The MPC+PID controllers demonstrate superior control compared to the MPC

controller, resulting in a notable enhancement in the accuracy and stability of the vehicle's trajectory tracking. Following the global offline optimisation of the MPC controller parameters by GA, significant enhancements have been made to the performance of the controller. Future, we will continue to study the relationship between road curvature and vehicle speed in order to improve the controller's performance.

References

1. Li Y, Yang W, Zhang X, Kang X, Li M (2022) Research on automatic driving trajectory planning and tracking control based on improvement of the artificial potential field method. Sustainability 14(19):12131
2. Fan X, Wang J, Wang H, Yang L (2023) LQR trajectory tracking control of unmanned wheeled tractor based on improved quantum genetic algorithm. Machines 11(1):62
3. Ren Y, Zheng L, Khajepour A (2019) Integrated model predictive and torque vectoring control for path tracking of 4-wheel-driven autonomous vehicles. IET Intel Transp Syst 13(1):98–107
4. Gao HB, Li SB, Xie GT, Cheng B (2018) Discrete time model predictive control for lateral trajectory tracking of intelligent vehicles. J Command Control 4(04):297–305
5. Wu HD, Li ZH, Si ZL (2021) Trajectory tracking control for four-wheel independent drive intelligent vehicle based on model predictive control and sliding mode control. Adv Mech Eng 13(9):16878140211045142
6. Massey JC, Langella I, Swaminathan N (2019) A scaling law for the recirculation zone length behind a bluff body in reacting flows. J Fluid Mech 875:699–724
7. Liang Y, Li Y, Khajepour A, Zheng L (2020) Holistic adaptive multi-model predictive control for the path following of 4WID autonomous vehicles. IEEE Trans Veh Technol 70(1):69–81
8. Zhang P, Jiang SZ, Chen YF, Zhang B, Han Y (2023) Intelligent vehicle path tracking based on transverse-longitudinal integrated control. J Chongqing Jiaotong Univ (Nat Sci Ed) 42(04):153–160

Open Access This chapter is licensed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License (http://creativecommons.org/licenses/by-nc-nd/4.0/), which permits any noncommercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if you modified the licensed material. You do not have permission under this license to share adapted material derived from this chapter or parts of it.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.



Full-Wheel Active Steering Control Method for Heavy-Duty Vehicles Based on GA-LQR



Jining Fan, Jue Gong, Jian Zhao, Pengbo Liu, Jihong Zhou, and Xin Wang

Abstract The working environment of heavy-duty special vehicles is complex, so during operation and driving, high demands are placed on the vehicle's steering performance. To improve both the motility of low-speed steering and stability of high-speed steering of five-axis heavy-duty special vehicle, a full-wheel active steering control method is proposed according to the genetic algorithm optimized linear quadratic regulator (GA-LQR). The LQR is designed considering the cost of vehicle state error and the increment of wheel angle input. The vehicle monorail model is established and simulation is conducted at 15 and 60 km/h in Matlab/Simulink environment. The numerical results show that the designed control system enables the vehicle to have a minimum turning radius of 7.9 m comparing to that of 8.4 m from the uncontrolled model at low speeds and can also balance the handling stability and driver's driving experience at high speeds. Furthermore, compared to the LQR based controller, it can significantly reduce the maximum values of sideslip angle by 83% at low speeds and 97% at high speeds, as well as the time to reach steady state.

Keywords Heavy-duty vehicle · Full-wheel active steering · GA-LQR · Motility · Handling stability

1 Introduction

Heavy-duty special vehicles are commonly employed in special working environments like lifting, transportation, and rescue, which generally possess characteristics like long body length, large tonnage, and multiple axles. Nevertheless, research on

J. Fan · J. Gong · J. Zhao (✉) · P. Liu · X. Wang
Dalian University of Technology, Dalian, China
e-mail: jzhao@dlut.edu.cn

J. Zhou
Taiyuan Heavy Industry Co., Ltd., Taiyuan, China

© The Author(s) 2026

H. Saman et al. (eds.), *The 9th International Conference on Advances in Construction Machinery and Vehicle Engineering*, Lecture Notes in Mechanical Engineering,
https://doi.org/10.1007/978-981-96-5527-4_103

the selection of steering strategies and handling characteristics for multi-axle vehicles is scarce. Thus, multi-axle steering technology is an important research direction in the field of heavy-duty vehicles [1].

Bayar et al. [2] proposed extending the four-wheel steering concept to n-axle vehicles for determining the optimal steering strategy. Huh et al. [3] established an 18-degree-of-freedom system for a six-wheeled vehicle and studied the handling performance of six-wheeled special vehicles. Williams [4] investigated the influence of rear axle steering of a three-axle vehicle on steady-state handling effects and verified the ability of the three-axle steady-state handling model to predict rear axle steering performance.

Research on active steering of multi-axle vehicles mainly concentrates on PID control [5, 6], LQR method [7, 8], robust stability control, as well as sliding mode control and model predictive control, etc. [9]. Compared with other control methods, the LQR controller has the advantages of simple solution method and good stability. However, the weight matrix values in the LQR controller are usually determined empirically, which can sometimes result in suboptimal control effects.

This paper researches a new optimization strategy for vehicle dynamic control which presents an LQR controller with weight matrices optimized by the Genetic Algorithm (GA). It aims to improve the performance of a five-axle heavy-duty special vehicle when steering at different speeds by actively controlling the steering of the rear three axles. By optimizing the weight matrix $\boldsymbol{Q}$ through genetic algorithm, the steering performance of the vehicle in the target scenario is further improved, and we provide a new controller parameter optimization method for active steering control strategy of multi-axle heavy vehicles.

2 Establishment of Two-Degree-of-Freedom Linear Model

2.1 Vehicle Monorail Model

The vehicle steering conforms to Ackermann's principle, and the steering center is located on the center line of the 3rd and 4th axles. In order to reduce the complexity of the model, the car is simplified into a linear model on the longitudinal symmetric plane of the body, as follows. As shown in Fig. 1, the four wheels of the first and second axles are equivalent to one wheel on the symmetric plane and located in the center of the two axles. The other three axles are equivalent to three wheels on the symmetrical surface of the vehicle body, and a simplified monorail model containing four wheels is obtained.

The parameters involved in the model are as follows: m_t is vehicle mass (kg); I_z is yaw moment of inertia ($\mathrm{kg} \cdot \mathrm{m}^2$); v_x is longitudinal speed (m/s); ω is yaw rate (rad/s); β is sideslip angle (rad); l_i is the distance of the equivalent wheel from the center of mass; l_{01} is the distance between the first axle and the center of mass; l_{02} is the

x: The *x*-axis is parallel to the ground and points towards the front of the vehicle.
y: The *y*-axis points to the driver's left.

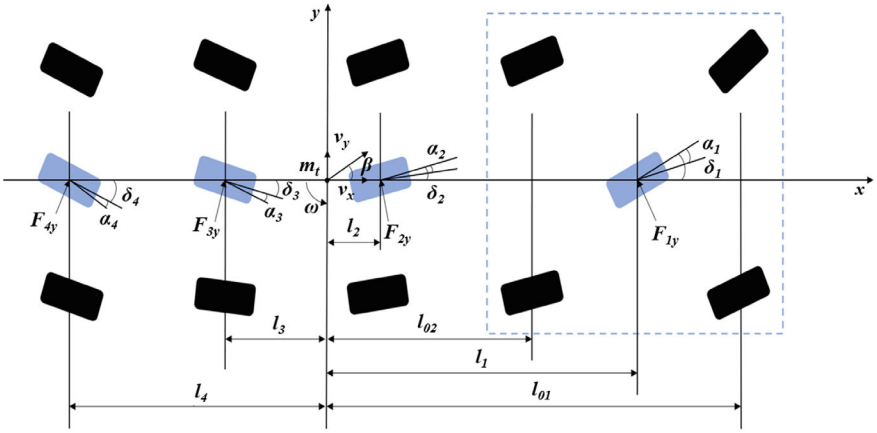


Fig. 1 Simplified vehicle model

distance between the second axle and the center of mass (m); δ_i is equivalent wheel angle (°); F_{iy} is lateral force (N).

The vehicle dynamic model is established by considering the factors affecting the wheel force and making small-angle assumptions, as follows.

$$m_t v_x(\dot{\beta} + \omega) = F_{1y} + F_{2y} + F_{3y} + F_{4y} \tag{1}$$

$$I_z\dot{\omega} = F_{1y}l_1 + F_{2y}l_2 - F_{3y}l_3 - F_{4y}l_4 \tag{2}$$

The automobile continuous system state space model and its coefficient matrix are sorted out.

$$\begin{cases} \dot{x} = A_1 x + B_1 u \\ y = C_1 x + D_1 u \end{cases} \tag{3}$$

$$A_1 = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} = \begin{bmatrix} \frac{\sum_{i=1}^{4} k_i}{m_t v_x} & \frac{k_1 l_1 + k_2 l_2 - k_3 l_3 - k_4 l_4}{m_t v_x^2} \\ \frac{k_1 l_1 + k_2 l_2 - k_3 l_3 - k_4 l_4}{I_z} & \frac{\sum_{i=1}^{4} k_i l_i^2}{I_z v_x} \end{bmatrix} \tag{4}$$

where, the state vector $\boldsymbol{x} = [\boldsymbol{\beta}, \boldsymbol{\omega}]^T$; input vector $\boldsymbol{u} = [\boldsymbol{\delta}_1, \boldsymbol{\delta}_2, \boldsymbol{\delta}_3, \boldsymbol{\delta}_4]^T$. The coefficient matrix $\boldsymbol{A}_1$ and $\boldsymbol{B}_1$ are derived from the mathematical model. a_{ij} and b_{ij} represent each element of matrix $\boldsymbol{A}_1$ and matrix $\boldsymbol{B}_1$ respectively, as shown in Eqs. (4) and (5). The coefficient matrix $\boldsymbol{C} = \boldsymbol{diag}[\mathbf{1}, \mathbf{1}]$ and the coefficient matrix $\boldsymbol{D} = 0$ of the output equation.

$$B_1 = \begin{bmatrix} b_{11} & b_{12} & b_{13} & b_{14} \\ b_{21} & b_{22} & b_{23} & b_{24} \end{bmatrix} = \begin{bmatrix} -\frac{k_1}{m_t v_x} & -\frac{k_2}{m_t v_x} & -\frac{k_3}{m_t v_x} & -\frac{k_4}{m_t v_x} \\ -\frac{k_1 l_1}{I_z} & -\frac{k_2 l_2}{I_z} & \frac{k_3 l_3}{I_z} & \frac{k_4 l_4}{I_z} \end{bmatrix} \tag{5}$$

Continuous state space model is not conducive to controller design, so the sampling period $T_S = 0.0005$ s is used to discretize Eq. (3), and the vehicle's discretized state space model is obtained as follows.

$$\begin{cases} x_{[k+1]} = Ax_{[k]} + Bu_{[k]} \\ y_{[k]} = Cx_{[k]} + Du_{[k]} \end{cases} \tag{6}$$

Among them, matrices $\boldsymbol{A}$ and $\boldsymbol{B}$ are derived from $\boldsymbol{A}_1$ and $\boldsymbol{B}_1$ variations according to Eqs. (7) and (8), while matrices $\boldsymbol{C}$ and $\boldsymbol{D}$ are the same as $\boldsymbol{C}_1$ and $\boldsymbol{D}_1$, respectively.

$$A = e^{A_1(t_{k+1}-t_k)} = e^{A_1 T\mathrm{s}} \tag{7}$$

$$B = \int_{t_k}^{t_{k+1}} e^{A_1(t_{k+1}-\tau)} \mathrm{d}\tau B_1 = A_1^{-1}(A - I)B_1 \tag{8}$$

2.2 Trucksim Model

In order to verify the accuracy of the reference model, it is necessary to establish the corresponding dynamic model in the simulation software and compare it, so the vehicle model is established in Trucksim. In order to achieve the purpose of all-wheel steering, wheel steering control channel is needed. So that the wheel rotation can be directly controlled by generating wheel angle in Matlab/Simulink. All wheels are steerable and the angle relation satisfies Ackermann's principle. Set the speed and road simulation environment, and the ground friction coefficient $u = 0.85$. Detailed parameters of vehicle model are shown in Table 1.

2.3 Validation of the Reference Model

The low-speed working condition is 15 km/h and the equivalent first axle wheel angle step input is 30°. The high-speed working condition is 60 km/h and the equivalent wheel angle step input of the first axle is 5°. The above two working conditions were selected to verify the accuracy of the reference model, and the deviation between the two-degree-of-freedom linear model and the output results of Trucksim software is compared.

Table 1 Parameters required for vehicle model

Symbol	Unit	Value
m_t	kg	86,000
I_z	kg · m^2	638,976
l_{01}	m	5.098
l_{02}	m	2.448
l_2	m	0.798
l_3	m	2.102
l_4	m	3.752
k_1	N/rad	− 1,916,870.88
k_2	N/rad	− 958,435.44
k_3	N/rad	− 958,435.44
k_4	N/rad	− 958,435.44

As shown in Fig. 2, the sideslip angle and yaw rate calculated by the reference model and the Trucksim model have the same variation trend regardless of the low speed or high speed working conditions, and the numerical deviation between the two is small and the fitting degree is high. This means that the reference model can accurately reflect the motion characteristics of the vehicle during the steering process, and the reference model can be applied to the design of steering controller.

3 Control Objective and Process

3.1 Control Objectives

In order to pursue the low-speed steering performance of the vehicle, considering the body attitude, turning radius and tire wear, the control objectives at low speed are determined as follows. (1) The sideslip angle is zero during the steering process. (2) The turning radius is the minimum value under the corresponding working condition.

In high-speed steering, the traditional front-wheel steering can obtain better handling stability, but it is still difficult to track the ideal heading. Therefore, the control objectives at high speed are determined as follows. (1) The sideslip angle is zero during the turning process. (2) Ensure that the steady-state value of yaw rate follows that of traditional front-wheel steering vehicles.

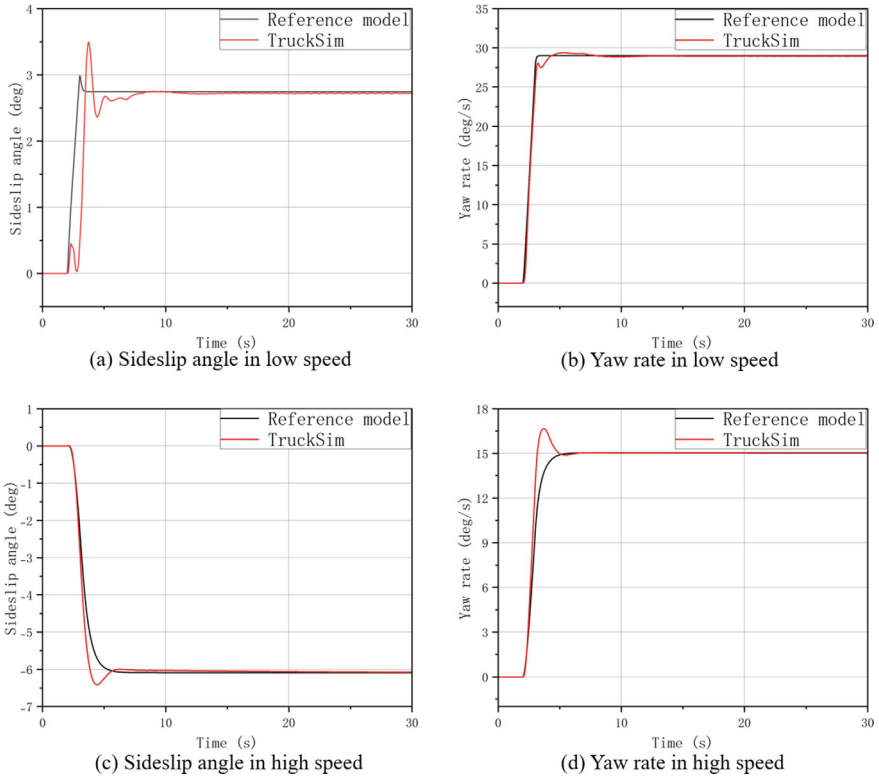


Fig. 2 Simplified vehicle model

3.2 Control Process

As shown in Fig. 3, the LQR control system is designed by using genetic algorithm to optimize the weight matrix $\boldsymbol{Q}$. The angular step input and vehicle speed are given as the initial input conditions. The vehicle motion state vector is calculated, compared with the target state vector of the current sampling point, and input to the GA-LQR controller together with the control vector of the previous sampling point. The GA-LQR controller obtains the control input increment of the current sampling point through the feedback matrix $\boldsymbol{K}$, and obtains the wheel Angle control output through the angle calculator, which is fed back to the reference model to realize the multi-axis active steering control.

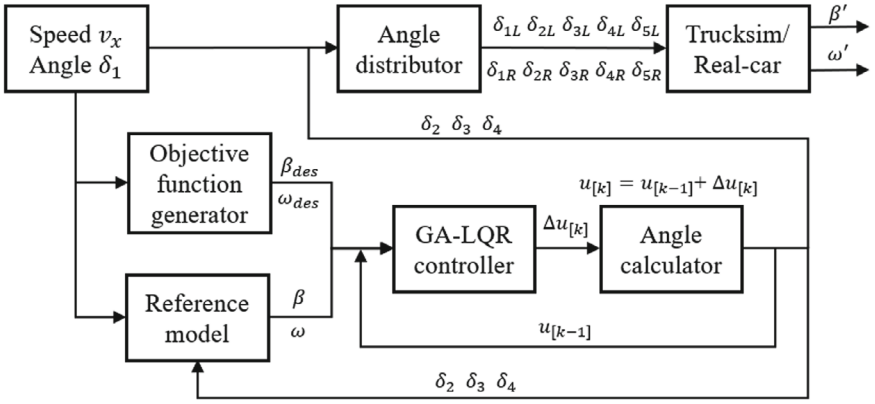


Fig. 3 GA-LQR steering controller structure

4 Controller Design

4.1 Ideal Vehicle Model

Ideal Model at Low Speed. The ideal model for low-speed steering needs to be established based on the control objective, so it is necessary to calculate the minimum steady-state steering radius under the current operating conditions. Since $\dot{\beta} = 0$ and $\dot{\omega} = 0$ during steady-state steering, by substituting Eqs. (1) and (2) and combining them, the turning radius R_w of the vehicle during steady-state can be obtained.

$$R_w = \frac{v_x(a_{11}a_{22} - a_{21}a_{12})}{\sum_{i=1}^{4}(a_{21}b_{1i} - a_{11}b_{2i})\delta_i} \tag{9}$$

When the wheel angle δ_1 and speed v_x are given, the turning radius can be minimized by optimizing the rotation angle of each wheel.

Optimization variable: δ_2, δ_3, δ_4.

Objective function: min $R_w = f(\delta_2, \delta_3, \delta_4)$.

Constraints: Considering the tire wear problem, the actual rotation angle of the rear three axles and the rotation angle error under the Ackermann principle are controlled within 2°.

$$\delta_{iA} - \frac{\pi}{90} \le \delta_i \le \delta_{iA} + \frac{\pi}{90} \tag{10}$$

where δ_{iA} is the rotation angle of the wheel under Ackermann's principle, which can be calculated as follows.

$$\delta_{iA} = \frac{s_i}{s_1}\delta_1 \tag{11}$$

where s_i is the distance between the wheel of the i axis and the steering center line.

When $\delta_2 = \delta_{2A} + \pi/90$, $\delta_3 = \delta_{3A} - \pi/90$, $\delta_4 = \delta_{4A} - \pi/90$, the turning radius is the minimum value $R_{\min}$, and then the corresponding ideal yaw rate can be obtained.

$$\omega_{desL} = v_x / R_{\min} \tag{12}$$

Ideal Model at High Speed. The control objective of high-speed steering is to ensure that the steady state of yaw rate follows the steady state of yaw velocity of the traditional front-wheel steering vehicle. When $\delta_2 = \delta_3 = \delta_4 = 0$, the steady state value of yaw velocity of high-speed front wheel steering vehicle can be obtained by substituting Eqs. (9) and (12).

$$\omega_{desH} = \frac{a_{21}b_{11} - a_{11}b_{21}}{a_{11}a_{22} - a_{21}a_{12}}\delta_1 \tag{13}$$

4.2 Steering LQR Controller

In order to achieve control objectives, the cost function of the LQR controller is constructed based on the end error, state error and control input increment, taking into account the smoothing of the control input and the tracking ability of the reference target. The vehicle state reference target vector is $\boldsymbol{x_{des[k]}}$, the error vector $\boldsymbol{e_{[k]}} = \boldsymbol{x_{[k]}} - \boldsymbol{x_{des[k]}}$, and the input increment vector is $\boldsymbol{\Delta u_{[k]}}$. The cost function of the control system is obtained.

$$J = \frac{1}{2}e_{[N]}^T S e_{[N]} + \frac{1}{2}\sum_{k=0}^{N-1}[e_{[k]}^T Q e_{[k]} + \Delta u_{[k]}^T R \Delta u_{[k]}] \tag{14}$$

where $\boldsymbol{S}$ is the end error weight matrix, $\boldsymbol{Q}$ is the state error weight matrix, and $\boldsymbol{R}$ is the control input increment weight matrix.

$$\Delta u_{[k]} = u_{[k]} - u_{[k-1]} \tag{15}$$

$$x_{d_{[k+1]}} = A_D x_{d_{[k]}} \tag{16}$$

Substituting Eq. (15) into Eq. (6) yields:

$$x_{[k+1]} = Ax_{[k]} + B \triangle u_{[k]} + Bu_{[k-1]} \tag{17}$$

Based on Eq. (6), we define the augmented matrix $\boldsymbol{x_{a[k]}} = [\boldsymbol{x_{[k]}}, \boldsymbol{x_{d[k]}}, \boldsymbol{u_{[k\text{-}1]}}]^T$, which includes reference values and input increments, and combine Eqs. (15)–(17)

to construct equations related to $\boldsymbol{x_{a[k]}}$ and $\boldsymbol{\Delta u_{[k]}}$.

$$x_{a_{[k+1]}} = A_a x_{a_{[k]}} + B_a \Delta u_{[k]} = \begin{bmatrix} A & 0 & B \\ 0 & A_D & 0 \\ 0 & 0 & I \end{bmatrix} x_{a_{[k]}} + \begin{bmatrix} B \\ 0 \\ I \end{bmatrix} \Delta u_{[k]} \tag{18}$$

Define the error vector:

$$e_{[k]} = x_{[k]} - x_{des_{[k]}} = [\, I \; -I \; 0 \,][\, x_{[k]} \; x_{des_{[k]}} \; u_{[k-1]} \,]^T = C_a x_{a_{[k]}} \tag{19}$$

Substitute Eq. (19) into Eq. (14) and rewrite the cost function as:

$$J = \frac{1}{2} x_{a_{[N]}}^T S_a x_{a_{[N]}} + \frac{1}{2} \sum_{k=0}^{N-1} [x_{a_{[k]}}^T Q_a x_{a_{[k]}} + \Delta u_{[k]}^T R \Delta u_{[k]}] \tag{20}$$

where $\boldsymbol{S_a} = \boldsymbol{C_a}^T \boldsymbol{S C_a}$, $\boldsymbol{Q_a} = \boldsymbol{C_a}^T \boldsymbol{Q C_a}$.

Based on $\boldsymbol{A_a}$, $\boldsymbol{B_a}$, $\boldsymbol{Q_a}$, $\boldsymbol{S_a}$, and $\boldsymbol{R}$ from Eqs. (18) and (20), construct a dedicated vehicle active steering LQR controller. Iterative solution of the discrete Riccati equation can yield the gain matrix $\boldsymbol{K}$. The ideal steering angle control input is:

$$u_{[k]} = u_{[k-1]} + \Delta u_{[k]} = u_{[k-1]} - K x_{a_{[k]}} \tag{21}$$

4.3 Parameter Optimization Based on Genetic Algorithm

The appropriate weight matrices $\boldsymbol{S}$, $\boldsymbol{Q}$, and $\boldsymbol{R}$ is selected to improve the control effect. Many studies have usually determined the weight matrix through empirical trial-and-error methods, which is less efficient and has poor results. In order to meet the control objectives, this paper selects the terminal error weight matrix $\boldsymbol{S} = \boldsymbol{diag}[\mathbf{1}, \mathbf{1}]$ and the control input increment weight matrix $\boldsymbol{R} = 1$, and optimizes the state error weight matrix $\boldsymbol{Q}$. The fitness function is constructed by taking the sideslip angle β and the value $\Delta\omega$ of the difference between the yaw rate ω and the ideal yaw rate ω_{des}. Since the numerical range of the two is very close, the fitness function is set to the square of the maximum value of the two within the simulation time range.

$$fitness = [max(|\beta|)]^2 + [max(|\Delta\omega|)]^2 \tag{22}$$

The fitness function can evaluate the control effect of the LQR controller in the sideslip angle β and yaw rate ω. The optimal state error weight matrix $\boldsymbol{Q_{GA}}$ is obtained through the genetic algorithm, which makes the controller adjust the cost function $\boldsymbol{J}$ in the desired direction to achieve the best control effect.

The genetic algorithm is used for optimization, with an initial population size of 50, a crossover rate of 0.8, and a maximum iteration limit of 100.

The optimal state error weight matrix $\boldsymbol{Q_{GAL}} = \boldsymbol{diag}[\mathbf{394.727,832.16}]$ at low speed and $\boldsymbol{Q_{GAH}} = \boldsymbol{diag}[\mathbf{937.88,779.015}]$ at high speed were obtained through iteration.

5 Simulation Experiment

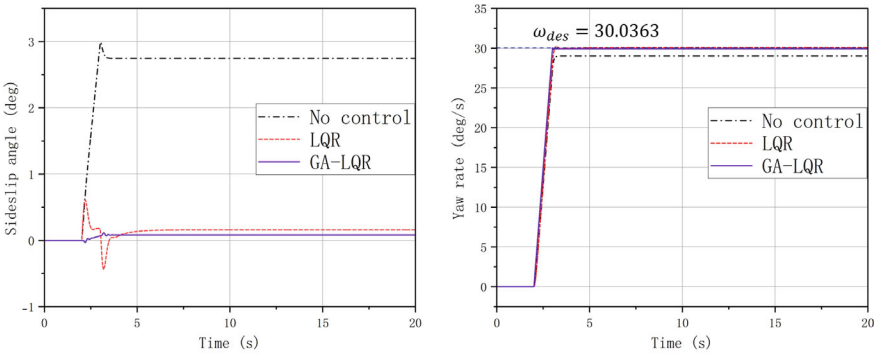
5.1 Simulation Preset

To verify the control effect of the GA-LQR control system on the steering performance of a certain five-axle heavy-duty special vehicle, simulation is conducted on uncontrolled vehicles, unoptimized LQR vehicles, and GA-LQR vehicles respectively. The simulated low-speed working condition is set as a vehicle speed of 15 km/h and a 30° angular step input of the equivalent first axle tire. The high-speed working condition is set as a vehicle speed of 60 km/h and a 5° angular step input of the equivalent first axle tire. The centroid sideslip angle and yaw rate are selected as evaluation indicators.

5.2 Simulation Experiment and Results

The simulation results of the angular step input at low speed are shown in Fig. 4.

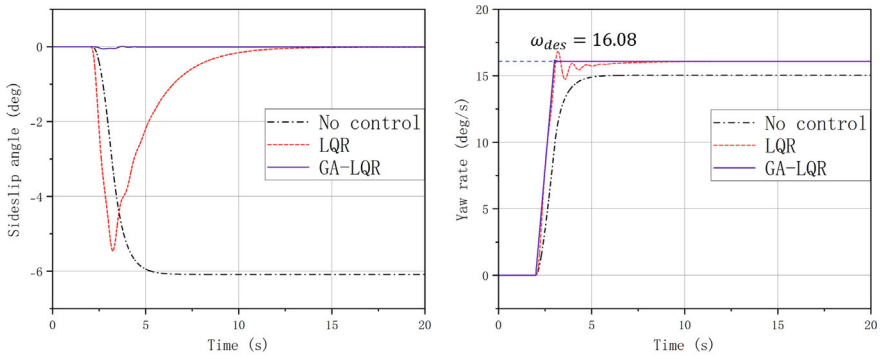
From Fig. 4, it can be seen that both the unoptimized LQR controller and GA-LQR controller can control the yaw rate around the ideal value of 30.04°/s during low-speed steering. GA-LQR controller reduces the steering radius by 5.2% compared to the uncontrolled yaw rate 28.57°/s. Regarding the sideslip angle, the GA-LQR



(a) Comparison of sideslip angle in low speed

(b) Comparison of yaw rate in low speed

Fig. 4 Low-speed angular step response curve



(a) Comparison of sideslip angle in high speed

(b) Comparison of yaw rate in high speed

Fig. 5 High-speed angular step response curve

controller reduces the maximum absolute deviation to 0.11°, which further improves the control effect compared to the 2.99° for the uncontrolled result and 0.67° for the LQR controlling result. In addition, compare with LQR controlling and uncontrolled result, the time to reach steady state for the GA-LQR controlling result is shorter, as well as the deviation from the ideal centroid sideslip angle after steady state is smaller.

The simulation results of the angular step input at high speed are shown in Fig. 5.

From Fig. 5, it can be seen that both the unoptimized LQR controller and GA-LQR controller can control the yaw rate around the ideal value of 16.08°/s during high-speed steering. However, the difference lies that GA-LQR significantly reduces the time to reach steady state and overshoot. While, the maximum absolute sideslip angle decreased from 6.09° for the uncontrolled state to 5.45° for the LQR controller and 0.14° for the GA-LQR controller, respectively. The LQR controller optimized by genetic algorithm further reduced the maximum absolute deviation value by 5.40° and adjusted the centroid sideslip angle to zero faster, significantly improving the control effect.

6 Conclusion

This paper aims at the problem of the motility of low-speed steering and handling stability of high-speed steering of heavy-duty special vehicles, establishes a vehicle dynamics model and designs a GA-LQR steering controller. Specific conclusions are as follows:

(1) A monorail vehicle dynamics model of a five-axis heavy-duty special vehicle is established, and the accuracy of the vehicle model is verified by contrast with the Trucksim model, which provides a foundation for the design of a multi-axis active steering vehicle train controller.

(2) Then we designed a LQR multi-axis steering controller which is at the cost of system state error and control input increment, so that the control-input varies more smoother and also it can track the target value function. We use genetic algorithms to optimize the state error weight matrix of the LQR controller and use simulation methods to verify it.
(3) The results show that the GA-LQR controller can accurately track the ideal yaw rate when steering at low or high speed, compared with the uncontrolled vehicle. Compared with LQR controller, the maximum value of sideslip angle and the time to steady state are significantly reduced. This indicates that the GA-LQR controller that based on state error and control input increment can effectively improve the lateral stability of the five-axis heavy duty special vehicle when steering and motility at low speeds and the handling stability at high speed which enables the driver to maintain the original steering feeling without increasing his operating burden. In this paper, a new idea of controller design and control parameter optimization is proposed. It can further improve the steering performance of multi-axle heavy-duty vehicle, so that it can have better performance in the face of complex working environment.

Acknowledgements The authors wish to express gratitude to Key Core Technology Research Project of Taiyuan City.

References

1. Jagirdar VV, Maskar VP, Trikande MW (2017) Steering strategy for a multi-axle wheeled vehicle. In: 2017 International conference on advances in mechanical, industrial, automation and management systems. IEEE, Yokohama, pp 164–171
2. Bayar K, Unlusoy YS (2008) Steering strategies for multi-axle vehicles. Int J Heavy Veh Syst 15(2–4):208–236
3. Huh K, Kim J, Hong J (2000) Handling and driving characteristics for six-wheeled vehicles. Proc Inst Mech Eng Part D J Automob Eng 214(2):159–170
4. Williams DE (2010) Handling benefits of steering a third axle. Proc Inst Mech Eng Part D J Automob Eng 224(12):1501–1511
5. Abroshan M, Taiebat M, Goodarzi A et al (2017) Automatic steering control in tractor semi-trailer vehicles for low-speed maneuverability enhancement. Proc Inst Mech Eng Part K J Multi-Body Dyn 231(1):83–102
6. Deng ZW, Zhao QX, Zhao YQ et al (2022) Active LQR multi-axle-steering method for improving maneuverability and stability of multi-trailer articulated heavy vehicles. Int J Automot Technol 23(4):939–955
7. Zhang Q, Zhang N, Wang Z et al (2022) Stability analysis of a closed–loop articulated vehicle system with trailer active steering based on LQR control. In: 6th CAA International conference on vehicular control and intelligence. IEEE, Nanjing, pp 1–5
8. Ahmed AA, Sami A, Alsharif A et al (2022) Evaluation of electric vehicle stability using fractional PID and fuzzy PID based controllers. In: 2nd International Maghreb meeting of the conference on sciences and techniques of automatic control and computer engineering. IEEE, Sabratha, pp 116–120

9. Metzler M, Tavernini D, Gruber P et al (2020) On Prediction model fidelity in explicit nonlinear model predictive vehicle stability control. IEEE Trans Control Syst Technol 29(5):1964–1980

Open Access This chapter is licensed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License (http://creativecommons.org/licenses/by-nc-nd/4.0/), which permits any noncommercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if you modified the licensed material. You do not have permission under this license to share adapted material derived from this chapter or parts of it.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.



Convolutional Neural Network-Based Autonomous Perception and Intelligent Decision System for Cabin Cleaning Machines



Xinguo Yu, Yan Huang, Chunyang Liu, and Siyuan Zhou

Abstract The existing various types of cabin cleaning machines are sufficient in terms of transportation capacity, endurance, and cleaning efficiency, but there are still some small problems. This article can focus on addressing the shortcomings of cabin cleaning machines in obstacle recognition and avoidance, target detection and positioning, enabling cabin cleaning machines to have certain intelligent decision-making capabilities and improving the efficiency of clearance work. This article uses the Residual Neural Network (ResNet) from Convolutional Neural Network (CNN) to build a system. After data collection, image denoising and adjustment, model training, and integration into autonomous perception and intelligent decision-making systems, the final model is integrated into the cabin cleaning machines. Finally, other algorithm models are applied for comparative experiments to verify the superiority and robustness of the proposed model. In the first set of experiments, the recognition accuracy of the model established by the algorithm in this article was between 0.9111 and 0.9428 during daytime, nighttime, sunny and rainy days, and the decision accuracy was between 0.9532 and 0.9797. The average recognition accuracy of the second group of experiments was between 0.9311 and 0.9502 during clear day, no rain at night, rainy day, and rainy night, and the average decision accuracy was between 0.9452 and 0.9841. The other four algorithms generally have accuracies below 0.9 or even 0.8 in both cases, and the fluctuation in accuracy is greater than that of the algorithm proposed in this article. The autonomous perception and intelligent decision-making system of the cabin cleaning machine based on convolutional neural network optimizes the cleaning process, reduces manual operation requirements, and

X. Yu (✉) · C. Liu · S. Zhou
Ores Terminal Branch Company of Yantai Port, Yantai, Shandong, China
e-mail: yxgguo@126.com

S. Zhou
e-mail: zhsy_1008@163.com

Y. Huang
Polytechnic Lei Kezhitu (Beijing) Technology Co., Ltd., Beijing, China
e-mail: huangyan@racobit.com

© The Author(s) 2026

H. Saman et al. (eds.), *The 9th International Conference on Advances in Construction Machinery and Vehicle Engineering*, Lecture Notes in Mechanical Engineering,
https://doi.org/10.1007/978-981-96-5527-4_104

successfully improves the perception and cleaning efficiency of the cabin cleaning machine towards the cabin environment, with strong robustness.

Keywords Autonomous perception · Intelligent decision making · Target detection · Recognition accuracy · Decision making accuracy

1 Introduction

In ship unloading operations, a cabin cleaning machine refers to a specialized mechanical equipment used to concentrate the remaining cargo at the bottom of the hold and dead corners of the hold towards the bottom of the hatch for further unloading operations [1–3]. These equipment include but are not limited to scrapers, bulldozers, etc. Bilge residue affects the operation of loading and unloading equipment, leading to equipment damage or operational accidents, and can cause pollution to the marine environment. The cleaning machine is created for these problems. However, unfortunately, due to the complexity of the cabin environment [4, 5], there are many types of residues at the bottom of the cabin and they are usually located in dead corners of the cabin, so there are certain blind spots in the cargo collection of the cabin cleaning machine at the bottom and hatch areas. Cabin cleaning machines often rely on pre-set path planning and simple mechanical operations, making it difficult to accurately identify and locate target objects in the bottom and hatch areas. It is unable to make timely and effective decisions based on environmental changes and job requirements, which can easily lead to external interference and cause job interruptions or errors. The existence of these problems seriously affects the smooth progress of ship unloading operations.

At present, research on ship clearance in the market mainly focuses on two aspects, namely improving the operation mode and applying intelligent management facilities [6, 7]. Zhou et al. [8] analyzed the current status of port clearance technology and proposed a new type of operation mode, which consists of one operation platform and four working devices. As a basic platform, the excavator loader can provide power, travel, and centralized control functions. The front end is equipped with hydraulic quick change, quick change buckets, enclosed sweepers, vibration cleaning shovels, etc., which can achieve the function of cleaning and collecting materials inside the cabin. The tail end is equipped with an enlarged bucket, which can collect and transfer concentrated materials on the slope of the cabin bottom to the middle of the cabin. This homework mode ensures the safety of personnel from all parties and balances safety and economic interests. The bilge of FPSO (Floating Production Storage and Offloading) cargo tanks and production water tanks is difficult to clean and remove due to long-term accumulation of contaminated oil and sludge. Wang [9] constructed a mobile automated integrated cleaning system for this. It is mainly divided into four subsystems: sludge flushing and disturbance subsystem, sludge suction and pumping subsystem, sludge transportation pipeline subsystem, and sludge reduction subsystem. This significantly improves work efficiency and reduces operational risks.

Although these two methods have achieved efficiency improvement in cabin cleaning operations, they require high technical skills from personnel in practical operations and are difficult to efficiently utilize in practical operations.

In terms of applying intelligent management facilities, Sui et al. [10] utilized key technologies such as artificial intelligence [11, 12], augmented reality, robot algorithms, as well as remote start/stop, collision warning [13], perspective technology, and third-party perspectives to achieve remote control and assisted driving of mobile machinery during clearance operations. Through this system, operators do not need to enter the cabin to complete the cleaning operation. In order to solve the problem of ship cleaning operations, improve the efficiency of ship cleaning, and reduce manual participation, Huo et al. [14] used various information technology methods to carry out remote control automation transformation of loaders and excavators, carry out automation transformation of gantry cranes, and develop specialized tools and electronic pole systems for ship cleaning. This can achieve unmanned operation of cabin cleaning operations, allowing all mechanical drivers, cabin cleaning workers, and hatch command personnel to evacuate. While ensuring clearance efficiency, this completely eliminates issues such as personal injury and epidemic infections. The above research methods have solved the high technical requirements for operators mentioned above, but only achieved automated processes, resulting in limited efficiency improvement in cabin cleaning operations.

For this purpose, this article applies Convolutional Neural Network (CNN) [15] to design autonomous perception and intelligent decision-making systems for cabin cleaning machines [16, 17]. The target recognition algorithm carried by the system can help the cleaning machine accurately identify target objects in the bottom and hatch areas, allowing the cleaning machine to quickly find and clean up dead corners in the bottom and hatch areas of residual cargo, including scattered cargo, oil stains, etc. Moreover, the system enables the cabin cleaning machine to have environmental awareness capabilities, enabling real-time recognition and localization of obstacles. This can optimize the operation path and obstacle avoidance strategy, avoid collision and jamming problems, and improve the safety and stability of the cleaning operation. The most important thing is that the intelligent decision-making system of the system can autonomously adjust the operation strategy, select the optimal cleaning path, adjust the operation speed, etc., based on the results of target recognition and environmental perception. This enables the cleaning machine to have the intelligent decision-making ability to adapt to the cabin, thereby improving the operational efficiency and accuracy of the cleaning machine. The system designed in this article solves the problem of high requirements for operators in the improvement plan of homework mode mentioned earlier. The introduction of intelligent management facilities in the previous text may seem similar to the method in this article, but in reality, it is different. The system in this article not only solves the automation of cabin cleaning, but also automatically adjusts the cabin cleaning strategy according to different complex environments.

2 Method

2.1 Data Collection and Preprocessing

Data Collection. The image data in this article was collected from Tianjin Port using various tools such as camera monitoring, industrial cameras, and drone cameras. Camera monitoring ensures that there is no missed opportunity for manual data collection, while also ensuring real-time data collection without affecting the progress of cabin cleaning work. Industrial cameras can provide high-resolution images, while drone cameras can capture dead corners of the cabin floor or body in complex and ever-changing environments. To ensure the diversity of image data, this article also adopts various methods to collect data. It not only collected data manually offline, but also collected some publicly available cabin datasets on online platforms, collecting a total of three types of image data. The image data of the cleaning machine operation is collected when the cleaning machine is operating in the actual marine environment, including the working status of the cleaning machine at different positions and angles. The target object image refers to the image of the target object that can be collected in the bottom and hatch areas, such as scattered goods, oil stains, etc. Obstacle images include images of obstacles in the bottom and hatch areas, such as seaweed, rocks, debris, etc. Finally, considering the changes in the working environment of the cabin cleaning machine, this article can collect data under four conditions: day, night, rainy, and sunny. Table 1 shows the number of types of data collection.

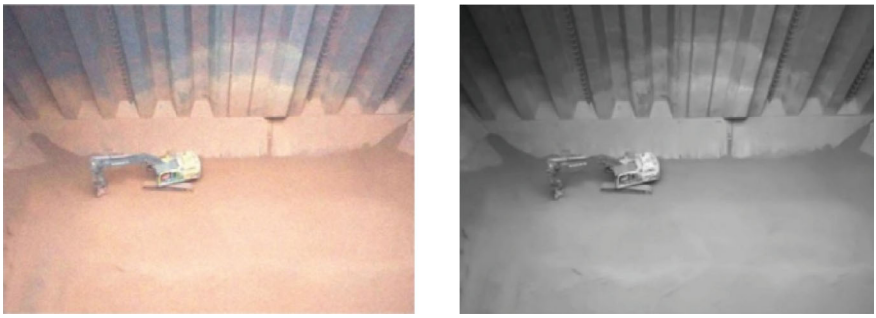
Data Collection

(1) Gaussian Filtering Denoising

The images of cabin cleaning machine operations generally have large dimensions and complex structures. Directly inputting the model may increase the computational complexity and training time of the model, such as sensor noise, vibrations caused by ocean waves, and sharp edges or fine textures in the captured image data. Therefore, in this article, Gaussian filtering [18, 19] is used in advance to denoise the image. This makes the image smoother, helping the model better learn the overall features and structure of the image. The size of the filter used in this article is 3×3, with a standard deviation of 0.5. Figures 1, 2, and 3 show the Gaussian filtering results of the cleaning machine operation, target objects, and obstacles, respectively. The grayscale image represents the Gaussian filtering denoising result.

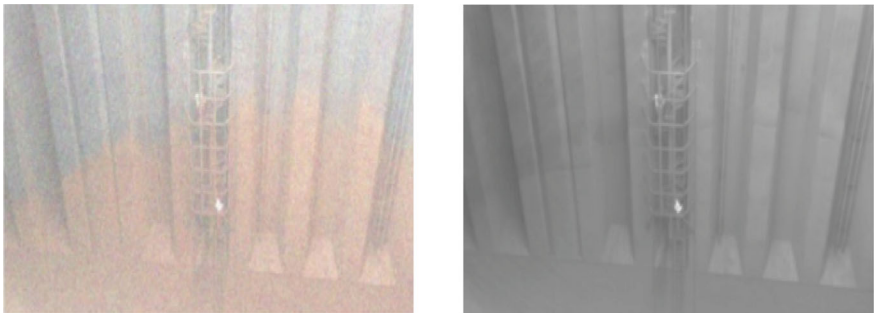
Table 1 Types and quantities of images

	Day	Night	Rainy	Sunny
Cabin cleaning machines operation	1636	2865	1624	1545
Target object	1978	1964	2721	1951
Obstacles	2741	1862	1752	1693



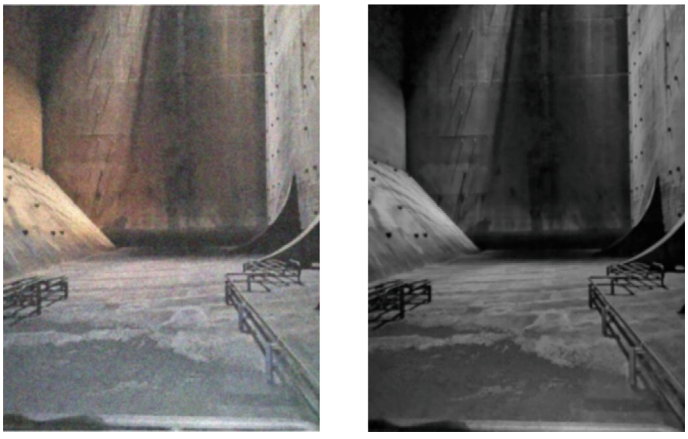
(a) lmage of the original clearer at work **(b)** Grayscale denoised image

Fig. 1 Cleaning operation diagram and its Gaussian filtering denoising diagram



(a) lmage of original rust **(b)** Grayscale denoised image

Fig. 2 Target object (rust) and its Gaussian filtering denoising image



(a) lmage of the original obstacle **(b)** Grayscale denoised image

Fig. 3 Obstacles and their Gaussian filtering denoising image



Fig. 4 Image adjustment results

(2) Image Adjustment

Firstly, as the data collection in this article is divided into different environments, the article can adjust the brightness, contrast, and sharpening of the image to enhance the features and details in the image, improve the quality and distribution of the image data, and thus enhance the model's generalization ability. Next, the processed image can be cropped to meet the requirements of the model input and unify the image size. This can reduce computational complexity and improve processing efficiency. Figure 4 shows the result of image processing.

2.2 Model Construction and Training

Residual Neural Network. Residual Neural Network (ResNet) [20] is a type of convolutional neural network, which is improved by applying residual connections to solve the gradient vanishing problem in deep neural network training. In practice, ResNet is often used for tasks such as image classification, object detection, and semantic segmentation, fully demonstrating the powerful capabilities of convolutional neural networks in the field of image processing. ResNet has achieved good performance on large-scale datasets such as ImageNet and performs excellently in various image recognition tasks. Through transfer learning, pre-trained ResNet models in other tasks can be used for reference in the process of building autonomous perception system models for cabin cleaning machines. In the autonomous perception system of the cabin cleaning machine in this article, a deeper network is required to extract complex image features. The residual connection of ResNet allows the network to cross multiple levels, allowing for the training of deeper networks and solving this problem. Therefore, this article can apply ResNet to establish a model.

(1) Residual Learning

In ResNet, the concept of residual learning [21] is proposed, which is different from direct learning mapping. If the output of a layer is H (x) and the mapping that should be learned is F (x), then the goal of residual learning is to learn H (x) − F (x), which is the residual. If the network can learn identity maps, then the residual is 0, making it easier for the network to learn the remaining parts.

(2) Residual Block

The basic building block of ResNet is the residual block. Each residual block contains two main parts: Main Path and Residual Connection. The main path consists of a series of convolutional layers used for learning feature representations. The calculation of residual blocks is shown in formula (1):

$$y = \mathcal{F}(x, \{W_i\}) + x \tag{1}$$

Among them, x is the input; y is the output; $\mathcal{F}(x, \{W_i\})$ is the part of residual learning; $\{W_i\}$ represents the weight parameters in the residual block. The ResNet model constructed in this article plays a major role in the system, which is used to perceive the surrounding environment, including lighting, water quality, cabin space conditions, etc. By perceiving and analyzing environmental data, the system can conduct real-time evaluation of the cleaning operation environment. Based on ResNet's environmental perception data, suitable job strategies and path planning are formulated through deep learning algorithms and decision models, providing key support for the system's autonomous perception and intelligent decision-making.

Construction of Autonomous Recognition Model for Cabin Cleaning Machines. The dataset prepared in this article can be divided into training, validation, and testing sets at a ratio of 70, 15, and 15%. The training set is used for model training, the validation set is used to adjust model parameters and prevent overfitting, and the test set is used to evaluate the performance of the model. Convolutional layers [22], activation layers [23], pooling layers [24], and fully connected layers can be selected as their basic components, and finally, cross entropy loss [25–27] can be selected as the loss function.

(1) Convolutional Layer

Convolutional layers are used for feature extraction, and each layer of the convolutional layer is used to extract pixel features by sliding one or more convolutional kernels (filters) at different positions in the image pixel matrix. The cabin data collected in this article needs to capture local spatial information in the image, such as texture, color, and other features of the cabin bottom and walls. This requires the abstract feature extraction function of convolutional layers to improve the perception system's understanding of the environment. Therefore, it is reasonable to choose the convolutional layer as the fundamental component of the perception system. The calculation formula for the convolutional layer is shown in Formula (2), where conv (X, K) represents the convolution operation between the input feature map X and the convolution kernel K.

$$Y = \mathrm{conv}(X, K) + b \tag{2}$$

(2) Activation Layer

The function of the activation layer is to perform nonlinear transformations on the output of convolutional or fully connected layers to generate nonlinear responses of neural networks. Autonomous perception and intelligent decision-making of cabin cleaning machines are complex classification problems in practical operations, such as dividing images of areas such as the bottom and walls into different categories. This type of task often involves complex nonlinear relationships, and the convolutional layer itself is linear, requiring the introduction of nonlinear characteristics through activation functions to enhance the model's expressive power. This article adopts the following two activation functions, as shown in Formulas (3) and (4):

Tanh Function

$$f(x) = \frac{e^{x} - e^{-x}}{e^{x} + e^{-x}} \tag{3}$$

ReLU Function

$$f(x) = \max(0, x) \tag{4}$$

Gradient vanishing is a common problem in deep neural networks. The output range of the Tanh function is between $[-1, 1]$, and compared to the Sigmoid function, its gradient is larger, which helps with gradient propagation. The ReLU function has a gradient of 1 when the input is greater than 0, which can avoid the problem of vanishing gradients. The calculation speed of Tanh function and ReLU function is relatively fast, especially ReLU function, because its calculation only involves simple comparison operations and does not require exponential operation. Therefore, it has advantages in training this large-scale dataset of cabin images in this article.

(3) Pooling Layer

The characteristics of areas such as the bottom and walls of the cabin often have certain translational invariance and scale invariance. Therefore, pooling operations can be used to reduce the dimensionality of feature maps, reduce computational complexity, and maintain important feature information, which helps to improve the efficiency and performance of perception systems. This article adopts the average pooling operation, and the pooling calculation is shown in Formula (5):

$$y = \frac{1}{n^2} \sum_{i=1}^{n} \sum_{j=1}^{n} x_{ij} \tag{5}$$

(4) Fully Connected Layer

The fully connected layer can map the abstract features extracted by convolutional neural networks to specific operations and decision spaces, such as determining whether the bilge needs to be cleaned and the type of cargo, providing input for intelligent decision-making systems. The calculation of the fully connected layer is shown in Formula (6), where x is the input vector, W is the weight matrix, and b is the bias term:

$$y = W \cdot x + b \tag{6}$$

(5) Cross Entropy Loss

The main task of the system designed in this article is to perform binary classification on the image data received by sensors. For example, when the transmitted image is bilge oil pollution, it is hoped that the model can accurately classify the bilge image as "requiring cleaning" or "not requiring cleaning". The cross entropy loss function is a commonly used loss function in the field of machine learning for classification tasks. It calculates the cross entropy between the predicted probability distribution of each category and the true label of the model, and then takes the average cross entropy of all samples as the loss value of the model. The calculation of cross entropy loss is shown in Formula (7):

$$L(y, \hat{y}) = -\frac{1}{N}\sum_{i=1}^{N}\sum_{c=1}^{C} y_{i,c}\log\left(\hat{y}_{i,c}\right) \tag{7}$$

In the formula, N is the number of samples; C is the number of categories; $y_{i,c}$ is the true label of sample i belonging to category c; $\hat{y}_{i,c}$ is the model's predicted probability that sample i belongs to category c. The optimization process of the cross entropy loss function aims to minimize the difference between the predicted probability distribution and the true labels, so that the model can make more accurate classification predictions. The cross entropy loss function can be selected as the loss function in the autonomous perception and intelligent decision-making system of the cabin cleaning machine, which can effectively measure the classification accuracy of the model and promote the model to better distinguish images of different categories during the optimization process.

Model Training and Evaluation. When evaluating the performance of this model, this article uses accuracy, precision, recall, and F1 score. In terms of accuracy, due to the limited number of categories, only the overall accuracy is calculated, without discussing classification. The above indicators can comprehensively evaluate the performance of the autonomous perception and intelligent decision-making system for cabin cleaning machines based on convolutional neural networks, reflecting the classification accuracy, recognition ability, and generalization ability of the model from different perspectives.

Firstly, the input data of the training set can be fed into the model, and the predicted output of the model can be calculated through the forward propagation of the model.

Forward propagation refers to the propagation process of data from the input layer through the hidden layer to the output layer. The calculation of each layer depends on the output of the previous layer and the weight parameters of the current layer. Next, the cross entropy loss function can be used to compare the predicted output of the model with the true labels of the training set, and calculate the loss value. Then, through the backpropagation algorithm, the loss value can be propagated layer by layer from the output layer to the input layer, and the gradient of each parameter on the loss value can be calculated. Then, the gradient descent algorithm can be used to update the model parameters based on their gradient values, gradually reducing the loss value. Finally, the above steps can be repeated continuously, and through multiple iterations, the model parameters can be continuously adjusted on the training set, gradually improving the performance of the model. The condition for stopping training is to reach the maximum number of iterations. In this article, the maximum number of iterations is set to 350. Figure 5 shows the number of iterations and the changes in the four evaluation indicators.

As shown in Fig. 5, the accuracy, precision, recall, and F1 score of the model in this article all reached high values and stabilized after 330 iterations. Among them, the highest value of accuracy reached 0.8780, and then stabilized between 0.8492 and 0.8780. The highest value of accuracy reached 0.8667, and then stabilized at 0.8134–0.8667. The highest recall rate reached 0.8935, and then stabilized at 0.8446–0.8935. The highest F1 score reached 0.9290, and then stabilized at 0.8734–0.9200.

Establishment of Intelligent Decision-making System. The main function of the autonomous recognition and intelligent decision-making system of the cabin cleaning machine is to make decisions after binary classification. Therefore, this article applies an intelligent decision modulation recognition network based on multi-criteria fusion. In a multi-criteria fusion network, the principle of local optimal solution can be

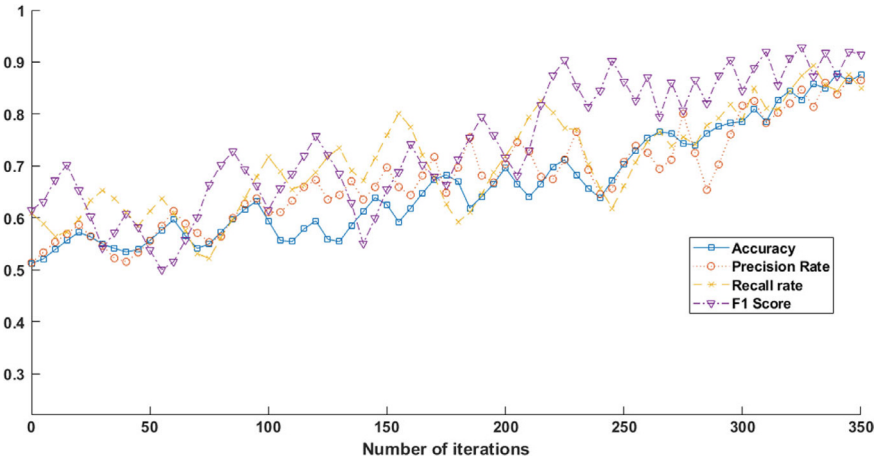


Fig. 5 Model iterative training

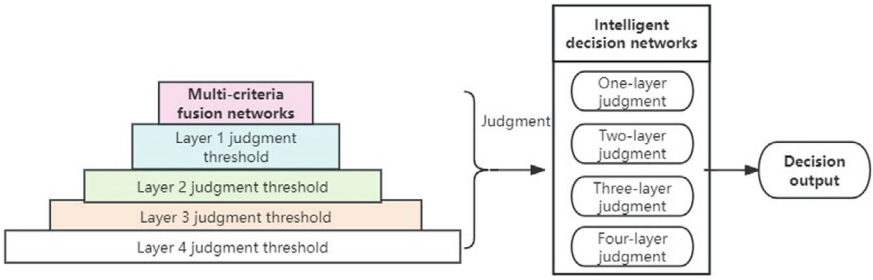


Fig. 6 Overall architecture of criterion recognition intelligent decision-making network

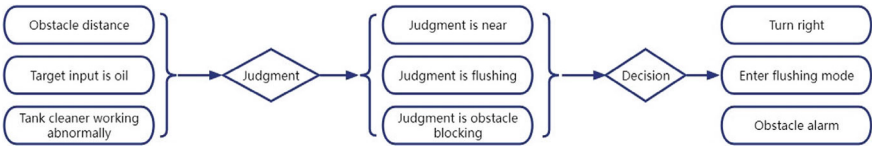


Fig. 7 Demonstration of intelligent decision-making for cabin cleaning machines

used, and a four layer decision threshold can be obtained by comparing two different threshold determination principles based on information gain and Gini coefficient [28]. The four layer decision threshold and corresponding modulation system can be input into the intelligent decision-making network, and the network can be constructed and iteratively optimized based on the support vector machine algorithm to form a modulation recognition network based on multi-criteria fusion and intelligent decision-making. Figure 6 shows the overall architecture of this intelligent decision-making network. In this article, the output of the residual neural network model is analyzed to obtain the characteristic information of the cleaning machine's operation and its surrounding environment. Then, a series of decision criteria are set up to evaluate the status and environment of the cleaning machine. The decision criteria are shown in Fig. 7:

3 Experimental Design

This section can integrate the proposed autonomous recognition model and intelligent decision-making network into an autonomous recognition and intelligent decision-making system, which can be integrated into the cabin cleaning machine. Due to the extensive use of image processing and computation in the experiment, this section adopts a cloud platform for the experiment. The p3.16xlarge instance type from Amazon AWS service provider can be used, which is equipped with 64 virtual CPU cores and 8 NVIDIA Tesla V100 GPUs. Each GPU has 16 GB of graphics memory, as well as 488 GB of memory and 2000 GB of SSD storage space. This configuration

can provide powerful computing power and storage resources, suitable for processing ultra large datasets and conducting large-scale model training. By utilizing multiple GPUs and fully utilizing their parallel computing capabilities, experiments can be accelerated to improve the training efficiency of deep learning models.

In addition, to verify the superiority of ResNet, this article can apply four algorithms for comparative experiments, namely Recurrent Neural Networks (RNN) [29], Decision Trees (DT) [30], Deep Reinforcement Learning (DRL) [31], and Support Vector Machines (SVM) [32]. These four algorithms, like convolutional neural networks, can be adjusted and learned according to environmental changes and input data, and can all handle nonlinear relationships, making them suitable for the complex data analysis and decision-making problems of the cabin cleaning machine in this article. Each algorithm has its unique advantages in the experiment of autonomous perception and intelligent decision-making of cabin cleaning machines. For example, DRL can learn the optimal strategy through interaction with the environment, and DT has good robustness to missing data. RNN can capture temporal dependencies between data, making it suitable for modeling and predicting historical data. SVM performs well in handling high-dimensional and small sample datasets. These commonalities and specificities have posed challenges for this comparative experiment.

The experiment in this section obtained 4000 images from the established training dataset for algorithm validation, with 1000 images taken from day, night, sunny, and rainy days respectively. Recognition accuracy and decision accuracy can be used as evaluation criteria for algorithms to evaluate the capabilities of these five algorithms in these four environments. The recognition accuracy and decision accuracy of the five algorithms are shown in Tables 2 and 3.

Table 2 Comparison of recognition accuracy of five algorithms

Algorithm	Day	Night	Sunny	Rainy
ResNet	0.9323	0.9223	0.9428	0.9111
RNN	0.8421	0.8245	0.8512	0.8002
DT	0.8322	0.8121	0.8412	0.7963
DRL	0.8341	0.8032	0.8541	0.7994
SVM	0.8361	0.8063	0.8661	0.8126

Table 3 Comparison of decision accuracy of five algorithms

Algorithm	Day	Night	Sunny	Rainy
ResNet	0.9632	0.9598	0.9797	0.9532
RNN	0.8632	0.8554	0.8824	0.8218
DT	0.8821	0.8531	0.9102	0.8463
DRL	0.8469	0.8142	0.8698	0.8014
SVM	0.8797	0.8563	0.9012	0.8369

From Table 2, it can be seen that the recognition accuracy of the autonomous perception and intelligent decision-making system based on these five algorithms in all four environments exceeds 0.7900, which is in a high state. The accuracy during the day is higher than that at night, and the accuracy on sunny days is higher than that on rainy days. The recognition accuracy of the algorithm in this article exceeded 0.9 in all four environments, and the recognition accuracy during the day reached 0.9323. In terms of stable recognition accuracy in different environments, the algorithm in this article has a recognition accuracy of only 0.01 lower at night than during the day, and 0.0317 lower on rainy days than on sunny days. The other four algorithms have the lowest recognition accuracy at night, which is 0.0176 lower than during the day, and on rainy days, it is 0.0449 lower than on sunny days. The decision accuracy shown in Table 3 is similar to Table 2, with higher accuracy during the day than at night and higher accuracy on sunny days than on rainy days. The decision accuracy of the algorithm in this article is above 0.9500 in all four environments, while the other four algorithms are below 0.9200. In terms of stable decision-making accuracy in environmental changes, the algorithm proposed in this article is also generally superior to the other four algorithms. This comparative experiment fully demonstrates that the autonomous perception and intelligent decision-making system based on ResNet can effectively improve recognition accuracy and decision accuracy, and ensure the stability of recognition accuracy and decision accuracy under environmental changes.

To further demonstrate the credibility of the comparative experimental results in this article, this section further arranges and combines these four environments, dividing the image set into four new situations: sunny day, no rain at night, rainy day, and rainy night. The recognition accuracy and decision accuracy of the five algorithms can be verified separately in these four scenarios, and 30 experiments can be conducted to ensure the reliability of the experiments. To more intuitively represent the difference between the recognition accuracy and decision accuracy of these five algorithms, the experimental results can be presented in the form of a bar chart, with the average of the experimental results taken, as shown in Figs. 8 and 9.

Based on the data in Figs. 8 and 9, the average recognition accuracy of ResNet (algorithm proposed in this article) is 0.9502, 0.9425, 0.9402, and 0.9311 for clear day, no rain at night, rainy day, and rainy night, respectively. The average decision accuracy is 0.9841, 0.9802, 0.9611, and 0.9452, respectively. Numerically, it can be seen that regardless of the environmental combination, the average recognition accuracy and average decision accuracy of the article's algorithm are higher than the other four algorithms, further proving that the article's algorithm is superior to the other four algorithms. The two types of accuracy of the algorithm in this article have decreased steadily, which reflects the strong robustness of the algorithm in the face of environmental changes. It is worth noting that regardless of which algorithm is used, the accuracy of recognition and decision-making shows a decreasing trend in these four environmental combinations. Among them, the algorithm performs the best during sunny days. Therefore, the system established in this article is most suitable for clearing the cabin during sunny days.

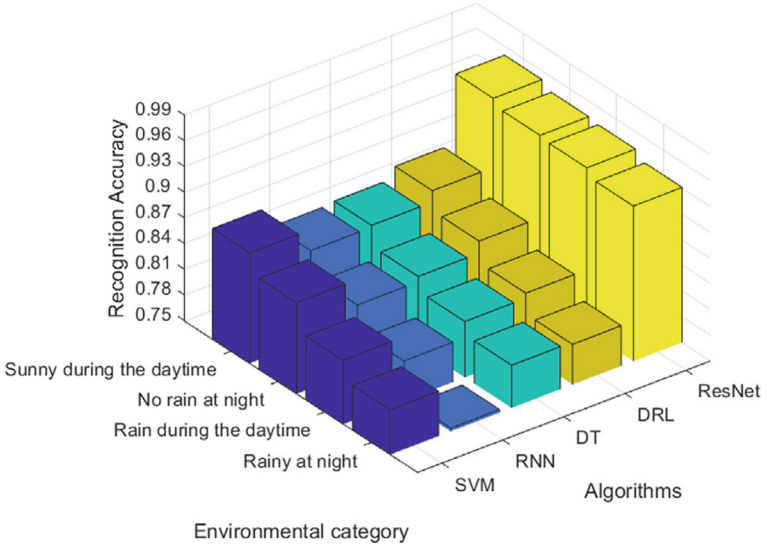


Fig. 8 Recognition accuracy of the algorithm in four new environmental combinations

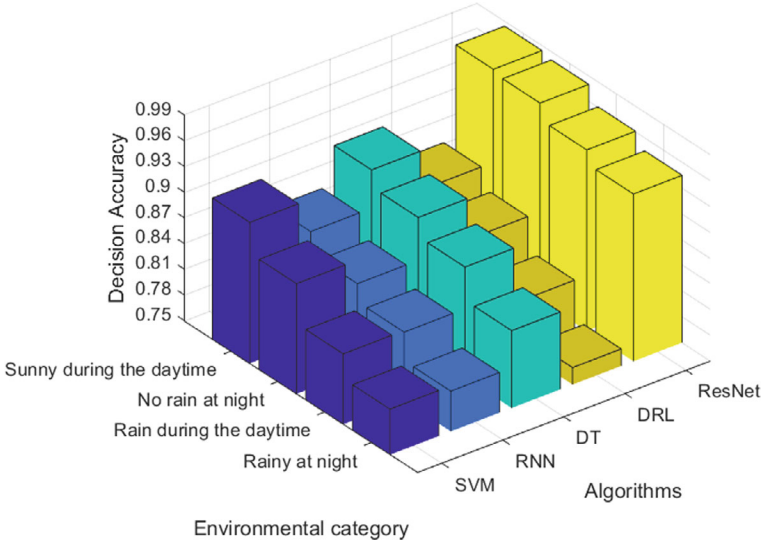


Fig. 9 Decision accuracy of the algorithm under four new environmental combinations

4 Conclusions

The system established in this article aims to solve the problems of low efficiency and high labor costs caused by the inability of the clearing machine to autonomously recognize and process obstacle recognition and avoidance, target detection and positioning during operation. This article designed two sets of algorithm comparison experiments in different environments through data collection and system integration. Experimental results have shown that the autonomous perception and intelligent decision-making system of the cabin cleaning machine established by the algorithm used in this article can better improve the accuracy of target recognition and event decision-making during the operation of the cabin cleaning machine, and has stronger robustness under the interference of environmental changes. However, there are still some shortcomings in this study, such as covering all possible cabin environments during the data collection and preparation stages, and not classifying and discussing different cabin environments, which may lead to insufficient generalization ability of the system established in this article. In future research, it is necessary to comprehensively discuss various cabin cleaning work. In terms of environmental changes, it is not only limited to weather, but also takes into account changes in the state of the marine environment, such as rising and falling tides. In the selection of convolutional neural networks, it can not be limited to ResNet, and other convolutional neural networks can be applied to build an autonomous perception and intelligent decision-making system for cabin cleaning machines. In this way, suitable CNN models can be selected for customized systems based on specific task requirements and resource constraints. Attempting to use new models can drive technological innovation and progress, bringing new possibilities and development directions for autonomous perception and intelligent decision-making systems for cabin cleaning machines.

References

1. Gao Z, Sun D, Zhao R et al (2021) Ship-unloading scheduling optimization for a steel plant. Inf Sci 544:214–226
2. Carlan V, Ceulemans D, van Hassel E et al (2023) Automation in cargo loading/unloading processes: do unmanned loading technologies bring benefits when both purchase and operational cost are considered? J Shipping Trade 8(1):20
3. Fontana D, Cardenia C, Pietrantonio M et al (2023) Applications of advanced oxidative processes for the recovery of water from bilge water. Int J Environ Sci Technol 20(8):8561–8570
4. Wang J, Cui X, Zhang R et al (2021) Study on the mass loss rate of liquid pool fire in a well-confined ship cabin. Int J Therm Sci 166:106984
5. Huang YK (2023) Design of a smart cabin lighting system based on internet of things. Cloud Comput Data Sci 112–121
6. Pan Y, Zhang L (2021) Roles of artificial intelligence in construction engineering and management: a critical review and future trends. Autom Constr 122:103517
7. El Khatib M, Almteiri M, Al Qasemi SA (2021) The correlation between emotional intelligence and project management success. iBusiness 13(1):18–29

8. Zhou Y, Yu Z, Han M et al (2023) Research on the application of multifunctional clearance equipment for bulk cargo ships in ports. Constr Mach Technol Manage 36(04):38–40. https://doi.org/10.13824/j.cnki.cmtm.2023.04.003
9. Wang D (2019) Construction of FPSO automated cabin clearance integrated system. Mar Eng 48(03):83–86
10. Sui S, Liu S, Zhang Y (2023) Remote intelligent control system for mobile machinery clearance operations. Port Technol 10:32–37
11. Ahmed I, Jeon G, Piccialli F (2022) From artificial intelligence to explainable artificial intelligence in industry 4.0: a survey on what, how, and where. IEEE Trans Industr Inf 18(8):5031–5042
12. Thombre S, Zhao Z, Ramm-Schmidt H et al (2020) Sensors and AI techniques for situational awareness in autonomous ships: a review. IEEE Trans Intell Transp Syst 23(1):64–83
13. Lu Z, Shu W, Li Y (2023) An anti-collision algorithm for self-organizing vehicular ad-hoc network using deep learning. IEEE Trans Intell Transp Syst
14. Huo G, Zhang J, Zhao W et al (2023) Construction of intelligent clearance system for ships in dry bulk terminals. China Ports 7:22–27
15. Tao J, Han F, Wang L et al (2024) Research on unmanned intelligent loading and unloading based on machine vision and convolutional neural networks. Autom Technol Appl 43(04):26–30. https://doi.org/10.20033/j.1003-7241.(2024)04-0026-05
16. Fayyad J, Jaradat MA, Gruyer D et al (2020) Deep learning sensor fusion for autonomous vehicle perception and localization: a review. Sensors 20(15):4220
17. Purwono P, Ma'arif A, Rahmaniar W et al (2022) Understanding of convolutional neural network (CNN): a review. Int J Robot Control Syst 2(4):739–748
18. Yun T, Jiang K, Li G et al (2021) Individual tree crown segmentation from airborne LiDAR data using a novel Gaussian filter and energy function minimization-based approach. Remote Sens Environ 256:112307
19. Poluri SM, Dey A (2023) Adaptive Gaussian filters for nonlinear state estimation with one-step randomly delayed measurements. Asian J Control 25(2):867–885
20. Wang S, Mo B, Zhao J (2021) Theory-based residual neural networks: a synergy of discrete choice models and deep neural networks. Transp Res Part B Methodological 146:333–358
21. Fang W, Yu Z, Chen Y et al (2021) Deep residual learning in spiking neural networks. Adv Neural Inf Process Syst 34:21056–21069
22. Hoshi I, Shimobaba T, Kakue T et al (2020) Single-pixel imaging using a recurrent neural network combined with convolutional layers. Opt Express 28(23):34069–34078
23. Szandala T (2021) Review and comparison of commonly used activation functions for deep neural networks. In: Bio-inspired neurocomputing, pp 203–224
24. Jie HJ, Wanda P (2020) RunPool: a dynamic pooling layer for convolution neural network. Int J Comput Intell Syst 13(1):66–76
25. Ruby U, Yendapalli V (2020) Binary cross entropy with deep learning technique for image classification. Int J Adv Trends Comput Sci Eng 9(10)
26. Boudiaf M, Rony J, Ziko IM et al (2020) Metric learning: cross-entropy vs. pairwise losses. arXiv preprint arXiv:2003.08983
27. Qu Z, Mei J, Liu L et al (2020) Crack detection of concrete pavement with cross-entropy loss function and improved VGG16 network model. IEEE Access 8:54564–54573
28. Thompson AE, Chase ASZ, Feinman GM (2023) Measuring inequality: the effect of units of analysis on the Gini coefficient. Anc Mesoam 34(3):e2
29. Hewamalage H, Bergmeir C, Bandara K (2021) Recurrent neural networks for time series forecasting: current status and future directions. Int J Forecast 37(1):388–427
30. Charbuty B, Abdulazeez A (2021) Classification based on decision tree algorithm for machine learning. J Appl Sci Technol Trends 2(01):20–28
31. Gronauer S, Diepold K (2022) Multi-agent deep reinforcement learning: a survey. Artif Intell Rev 55(2):895–943
32. Bansal M, Goyal A, Choudhary A (2022) A comparative analysis of K-nearest neighbor, genetic, support vector machine, decision tree, and long short term memory algorithms in machine learning. Decis Anal J 3:100071

Open Access This chapter is licensed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License (http://creativecommons.org/licenses/by-nc-nd/4.0/), which permits any noncommercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if you modified the licensed material. You do not have permission under this license to share adapted material derived from this chapter or parts of it.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.



Research on Mobile Grasping and Path Planning of Robotic Arms Based on Deep Learning



Zhanhui Wang, Zhifei Ji, Ning Chen, Xiaochuan Lin, Kangyi Wu, and Xinrui Ma

Abstract In view of the increasing aging of the Chinese population, this paper studies the mobile grasping robot for helping the elderly, because the environment in the bedroom is very complex, which requires the robot to have the ability to sense the indoor environment and pose estimation, so this paper studies the pose estimation algorithm, robotic arm grasping and mobile robot mapping and positioning in the environment, and the experiment proves that the size of our improved DOPE network model is reduced by 38.8% The inference speed is increased by 1.13 FPS, the planning time of the RRT algorithm is shortened by 9.5%, and the wheeled odometer fused with IMU information has a more robust positioning effect.

Keywords Pose estimation · Trajectory planning · Mapping and positioning

1 Introduction

With the aging of the Chinese population, the development and design of elderly products suitable for their needs has become a hot topic of social concern. The mobile grasping robot equipped with the mechanical arm is a good choice, and the bedroom is a very complex and disordered environment, and the indoor items are placed disorderly, so this requires the robot to have the ability to sense the indoor environment and pose estimation, so this paper studies the pose estimation algorithm, the robotic arm grasping and the mapping and positioning of the mobile robot in the environment. Object pose estimation and grasping has always been a difficult point in the work of robot grasping, and the current mainstream methods are 2D plane grasping and 6DOF grasping. In plane grabbing, Dong [1] proposes to divide data-driven methods into scraping of known and unknown objects. Then, the typical grasping detection methods and the related characteristics of each classification in

Z. Wang · Z. Ji (✉) · N. Chen · X. Lin · K. Wu · X. Ma
School of Marine Equipment and Mechanical Engineering, Jimei University, Xiamen, Fujian, China
e-mail: Zfji18@163.com

© The Author(s) 2026

H. Saman et al. (eds.), *The 9th International Conference on Advances in Construction Machinery and Vehicle Engineering*, Lecture Notes in Mechanical Engineering,
https://doi.org/10.1007/978-981-96-5527-4_105

the grasping of unknown objects are described in detail. Song [2] a single-stage multi-task robot grasping and detection method based on the Faster R-CNN framework was proposed. The designed network independently detects the class of objects and their possible gripping positions by using a loss function. A new crawl matching strategy was designed to determine the relationship between object categories and predicted crawl locations. Depierre [3] proposed a framework that combines crawl regression and score evaluation, and uses crawl quality scores to improve regression quality to generate crawl boxes. Zhao [4] generated the optimal grasping posture in a noisy environment, using a monoscopic point cloud as input, and the network consisted of three stages: a scoring network, a grasping area network, and a refining network, in which the scoring network selected positive points with high confidence, then made a crawling prediction on the selected positive points, and used the refining network to generate a more accurate grasping posture. Ni [5] used an end-to-end method to directly predict the attitude, score, and category of all grabbing, using a sparse point cloud as input, without the need for sampling and search process. Yang [6] proposed a grasping planning system to extract an accurate 6-degree-of-freedom grasping attitude from a single RGB-D image, using two neural networks to process the same image for 3D point cloud reconstruction and 6-degree-of-freedom, grasping suggestions, combining 3D reconstruction with grasping prediction to improve the flexibility of grasping prediction. The numerical algorithm [7] converges the inverse solution results through iteration. The numerical solution uses the Jacobian matrix to represent the mapping relationship between the input speed and the output speed. For the manipulator, in the joint space, it is the joint speed and the end execution Based on the idea of micro-elements, the inverse solution can be approximated using the Jacobian matrix inversion method or the Jacobian matrix transpose method. Li et al. [8], a new high-order extended unscented Kalman filter based on unscented Kalman filter (HUKF) was designed by taking advantage of the high-order statistical characteristics of approximation error. References [9, 10] discusses the key issues of map merging for multi-terrestrial robot SLAM, and reviews typical map merging methods for several important map types in SLAM applications: occupancy grid maps, feature-based maps, and topology maps. These map consolidation methods are categorized according to how they work or the types of features they work with.. In 2007, Grisetti et al. [11] proposed a GMapping scheme based on RBPF by improving the proposed distribution and adaptive resampling algorithm, which was a milestone for laser SLAM by reducing the amount of computation and improving the positioning accuracy.

2 Lightweight Research on Pose Estimation Algorithm

The 6D pose estimation algorithm based on deep convolutional network can well solve the problem that the pose estimation algorithm of traditional algorithms cannot deal with the occlusion of objects, so the research on the pose estimation algorithm

of deep learning is carried out, and the lightweight processing is carried out to meet the needs of more real-time detection.

2.1 DOPE Pose Estimation Algorithm

DOPE is a convolutional neural network algorithm framework based on key point detection, which is used to estimate the pose of objects with six degrees of freedom, proposed by Jonathon Tremblay et al. [12] of NVIDIA Labs. The RGB image is used as the input, and the 50 × 50 128-dimensional feature map is obtained through the feature extraction network, and the feature map will be sent into the six-stage confidence graph prediction and vector map prediction network respectively, and then the effective points are extracted by combining the confidence map and the vector map, and finally the PnP algorithm is used to restore the pose of the object in the real world and visualize it. DOPE's feature extraction network.

The first 10 layers of VGG19 were used for three pooling, and the image resolution was reduced by 8 times, and then two 3 × 3 convolutional layers were spliced to reduce the feature dimension from 512 to 256 and from 256 to 128. The large number of parameters of the task branch network model is due to the fact that five 7 × 7 convolution kernels are used in the second to sixth stages, and they run through the whole process.

2.2 Lightweight Task Branch Network

Inspired by the VGG network to replace the 11 × 11, 7 × 7 and 5 × 5 convolutional kernels in the AlexNet network with different numbers of 3 × 3 convolutional kernels, the task branch network is improved, and the improved refining stage network is shown in Fig. 1, in the second stage to the sixth stage, all 7 × 7 convolutional kernels are replaced by three 3 × 3 convolution kernels, and each 3 × 3 convolutional kernel is followed by the ReLU activation function. In this way, the same receptive field is ensured, which not only reduces the number of parameters, but also increases the depth of the network, introduces more nonlinear activation functions, and improves the abstraction ability of model features.

2.3 Dataset Production

Pose estimation has high requirements for datasets, a large number of real scene pictures and accurate annotation information will affect the pose estimation effect, because the pose estimation network is for specific objects to detect, and the detected objects still have occlusion problems in reality detection, so it is very difficult to

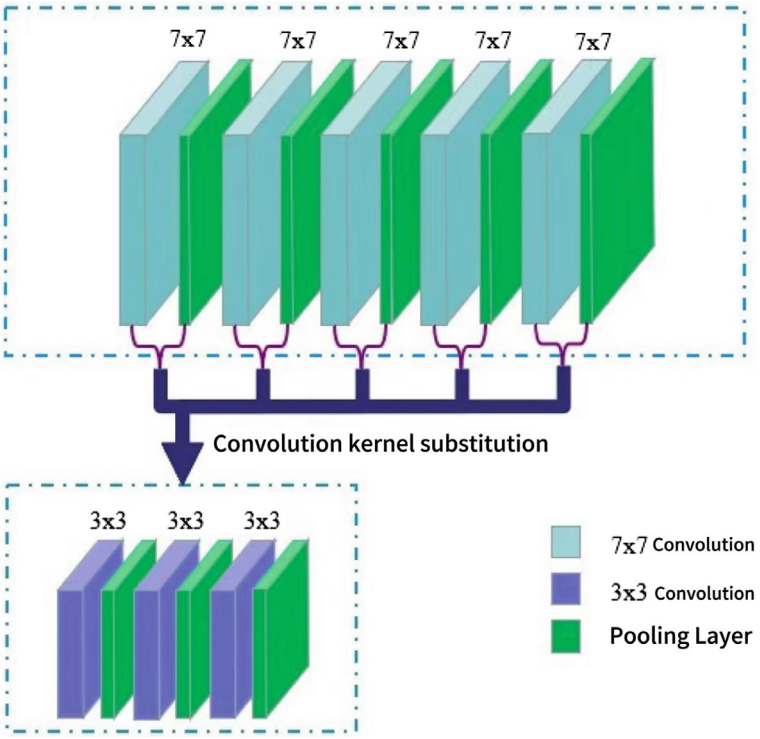


Fig. 1 Improved refining stage network

generate a large number of accurate datasets, this section uses the virtual dataset synthesized by Unreal Engine to optimize the occlusion problem of the dataset. To solve the occlusion problem, with the help of random components in NDDS, interfering objects are added to the scene, and the random movement of interfering objects causes different degrees of occlusion of the medicine box.

In order to make the trained model have better generalization ability, the images of the COCO dataset are added to the random background as a supplement, and random lighting intensity is added to the lighting to better simulate the changes of lighting in the environment. Finally, configure the camera module, set the size of the images captured by the camera to 640 × 480, and output the dataset as shown in Fig. 2.

2.4 Experimental Data Analysis

In this section, Unreal Engine is used to generate a drug box dataset for the DOPE network training before and after improvement, and the qualitative analysis of the experimental results is carried out by comparing the training data with the evaluation

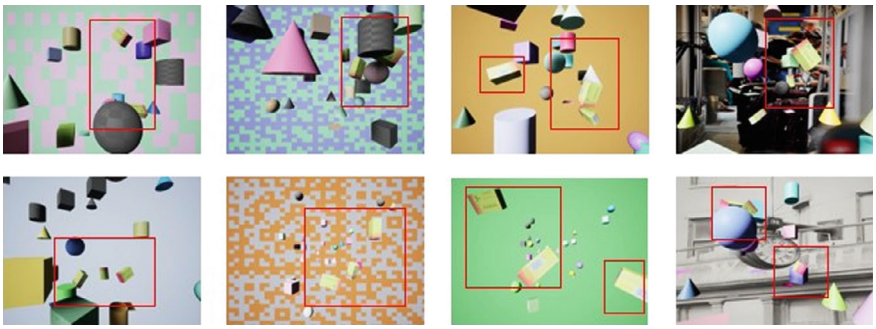


Fig. 2 Randomly generated dataset for domains

indicators. The evaluation index of pose estimation is generally measured by mean distance (ADD), and the main consideration is pose $[R|t]$. The points in the 3D model will generate corresponding pairs of points under the transformation of the pose estimate and the attitude truth, and the estimation effect will be evaluated by measuring the average distance between the corresponding points, and the average distance less than 10% of the model diameter will be considered accurate, as in Eq. 1:

$$ADD = \frac{1}{n}\sum_{i=1}^{n} \left\| (R_e P_i + T_e) - (R_{gt} P_i + T_{gt}) \right\|_2 \tag{1}$$

In the formula, R_e and T_e are the estimated rotation and translation matrices, R_{gt} and T_{gt} are the true rotation and translation matrices, and P_i is the key points in the 3D model. In this paper, the ADD threshold is set to 0.1, if the value of ADD is less than the given threshold, it means that the pose detection is successful, which is recorded as TP, and if it is greater than the given threshold, it means that the pose is misaligned, which is recorded as FN, and the ADD pass rate is calculated as follows: Eq. 2, using AUC as the ratio of the area under the curve to the total area.

$$ADD_{\text{pass rate}} = \frac{TP}{TP + FN} \tag{2}$$

The hardware running environment of the pose estimation experiment in this paper is: processor i7-7700K, graphics card GeFoece RTX 1080, video memory 8G, software running.

We trained the network according to the training parameters in Table 1.The size of the improved network model is reduced from the original 201 to 124M, the loss rate before the network improvement is shown in Fig. 3, the vertical axis is the loss rate, and the horizontal axis is the number of iterations, at the beginning of training, the loss rate decreases relatively quickly, with the increase of the number of iterations, the loss rate decreases slowly, and begins to converge after 15k iterations, and the loss rate finally remains at about 0.0026. Figure 4 shows that the loss function converges

Table 1 Network training parameters

Training parameters	Value
Batchsize	14
Epochs	120
Learning rate	0.0001
Input imagesize	400

faster than the improved network, which is close to convergence after 10K iterations, and the loss rate remains at about 0.0024.

The test data was evaluated using the ADD index, the threshold was set to 0.1, and the results were plotted as a curve. Figure 5 shows the AUC before network improvement, the AUC value of the threshold distance from 0 to 0.1 m is 0.461, and the ADD pass rate is 0.858. As shown in Fig. 6, the AUC value increased to 0.491 and the ADD pass rate was 0.893. It can be seen that the improved network model has a

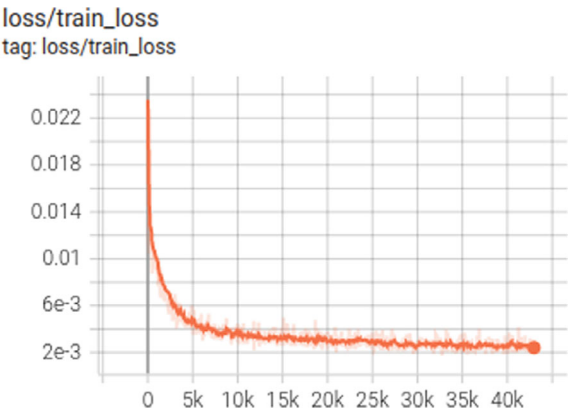


Fig. 3 Loss curve before improvement

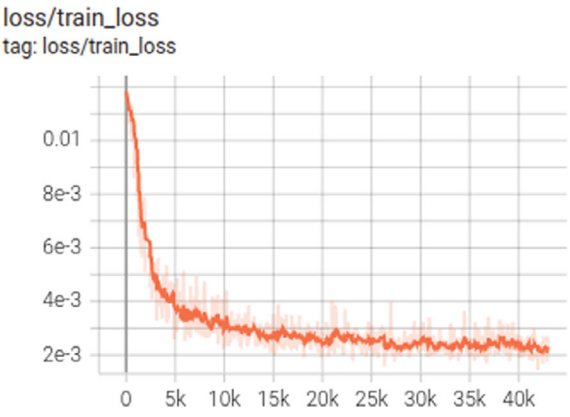


Fig. 4 Improved loss curve

certain improvement in prediction accuracy, and the inference speed is increased by 1.13FPS compared with the original network.

Next, the pose estimation effect is visualized, and the interference factors of strong light and occlusion are added to the experiment to continue to test the performance of the network model Fig. 7 shows the visualization page.

The visible slight occlusion texture feature of the hand is still well recognized, but in bright light, even if the pose can be estimated, the detected pose is wrong, and in the case of obvious occlusion, there is an inaccurate estimation. Compared with the original network, the improved network has a certain improvement in inference speed and detection accuracy, and the detection accuracy can meet the requirements of grabbing.

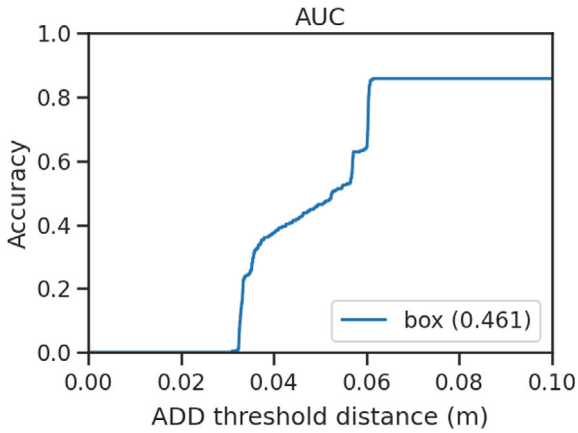


Fig. 5 AUC curve before improvement

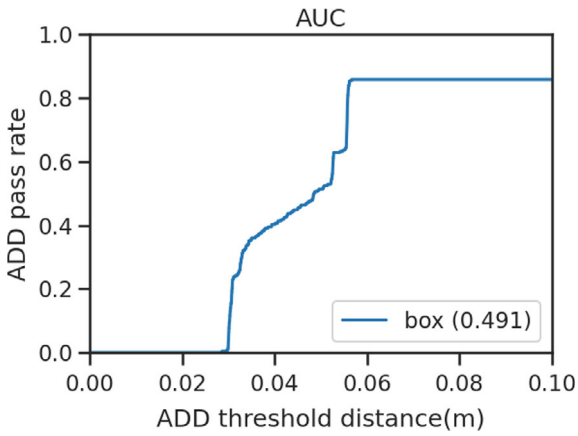


Fig. 6 Improved AUC curve

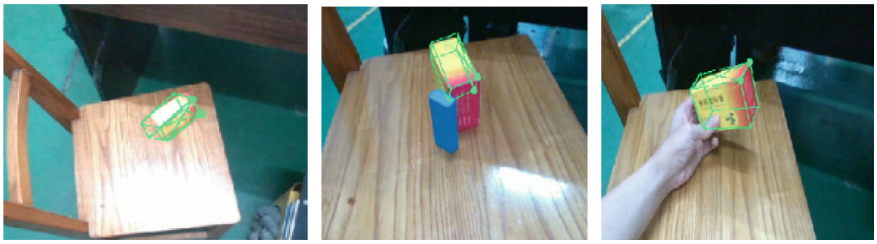


Fig. 7 Estimated detection of interfering factors

3 Improved RRT Algorithm

The core idea of the RRT algorithm is to generate a random tree in the configuration space, and continuously expand outward through random sampling iteration until the target point is found, but due to the randomness of the expansion, the planning time is long, so the expansion process of the random tree near the point is improved in this section. The random tree sampling process of the RRT algorithm is shown in Fig. 8.

Because the RRT algorithm selects random points by random sampling in space, therewill be more redundant bifurcated branches when generating random trees, which leads to the algorithm spending too much time on the expansion of useless branches and leaves, and the q_{rand} expansion of the adjacent nodes of random points is optimized by the vector angle operation to improve the operation efficiency of the algorithm. The expansion of the random tree datum node q_{nea} with increasing the direction bias is shown in Fig. 9, in the direction of the datum node q_{nea} to the random point q_{rand} taken, the vector end point q1 position is determined according to the step size parameter with the datum node q_{nea} as the vector starting point, and its vector is $\overrightarrow{a}$, and the datum node qnear is in the direction of the target point, and the end point q2 position is determined by the same method, its vector is $\overrightarrow{b}$, and the modulus length of the two vectors is the step size, and then the new node q_{new} can be obtained by the sum operation of these two vectors, and the new extension direction is obtained. The improved RRT algorithm is more biased towards the target point in the search direction, which reduces the extension on the useless path. Matlab was

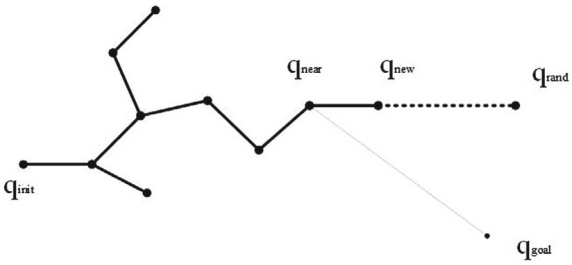


Fig. 8 Improved pre-RRT algorithm

used to generate the trajectory of the RRT algorithm before Fig. 10 and after the improvement Fig. 11, and the two-dimensional plan with geometric obstacles was used as the planning environment, and then the coordinates of the starting point and the target point were set up to compare the experimental results. Both the improved RRT algorithm and the improved RRT algorithm can plan trajectories that avoid geometric shapes Fig. 11, and the expanded branches and leaves of the improved algorithm are significantly reduced compared with the previous algorithm, which reduces the computational cost of planning.

The improved algorithm is planned 10 times, and the path length and time consumption of the two algorithms are shown in Table 2.

The improved RRT algorithm plans a shorter path length and runs faster, reducing the average time by 9.5% compared to the original RRT algorithm.

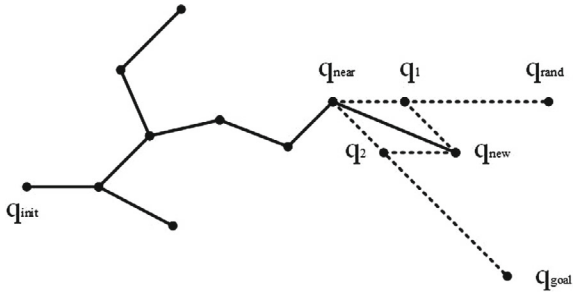


Fig. 9 Improved RRT algorithm

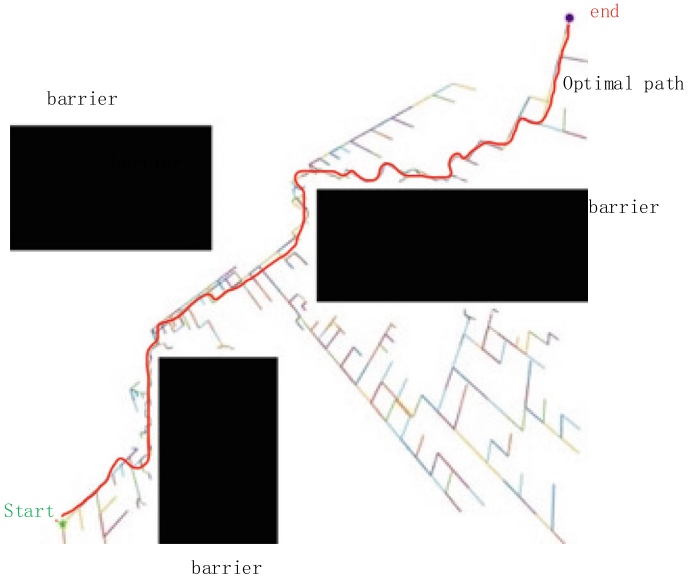


Fig. 10 Improved pre-RRT algorithm trajectory [13]

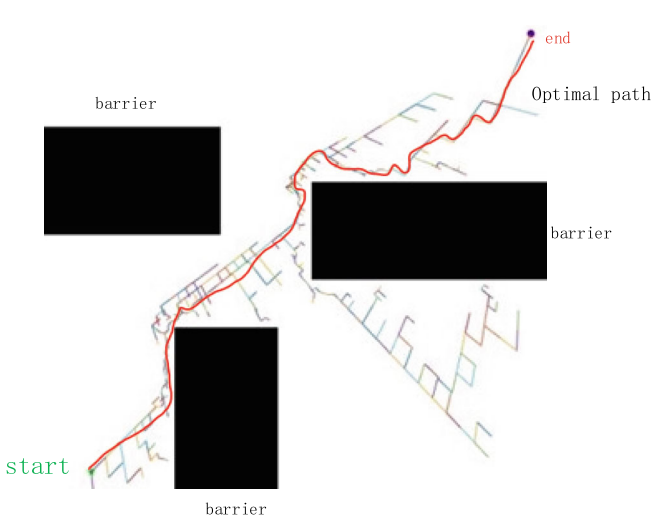


Fig. 11 Improved RRT algorithm trajectory

Table 2 Path length and time before and after RRT improvement

Algorithm	Average total path distance (pixels)	Average time elapsed (s)
RRT	793.0811	11.4639
Improving RRT	770.3242	10.3755

4 Research on Localization and Mapping Algorithms

The mobile robot is composed of a robotic arm and a mobile chassis, and the mobile chassis also needs to have the function of autonomous movement in complex environments, and positioning and mapping are the key to the autonomous movement function. Therefore, in view of the phenomenon that the wheeled odometer of the handling robot is easy to drift for a long time, the position and pose correction of the extended Kalman filter are carried out by combining the wheeled odometer and IMU of the handling robot, and the navigation framework of the handling robot is constructed, which paves the way for the subsequent mapping and navigation experiments.

4.1 Principles of the Gmapping Algorithm

GMapping is an SLAM based on RBPF filtering algorithm, which separates positioning and mapping, and realizes localization before mapping. In the SLAM algorithm of RBPF, $z_{1:t} = z_1, z_2, \cdots, z_t$ is measured by the observation of LiDAR-Mobile robot on wheelMeter $u_{1:t-1} = u_1, u_2, \cdots, u_{t-1}$ to estimate the trajectory $x_{1:t} = x_1, x_2, \cdots, x_t$ Posterior probability $p(x_{1:t}, m|z_{1:t}, u_{1:t-1})$ with the environment map mThe joint distribution function of the posterior probability can be further decomposed into the product of the posterior probability of the trajectory and the posterior probability of the map:

$$p(x_{1:t}, m|z_{1:t}, u_{1:t-1}) = p(x_{1:t}|z_{1:t}, u_{1:t-1})p(m|x_{1:t}, z_{1:t}) \tag{3}$$

In the formula, $p(x_{1:t}|z_{1:t}, u_{1:t-1})$ is the posterior probability density function of the mobile robot's trajectory, which is estimated by the particle filter method, the particles in the particle swarm represent the trajectory of the mobile robot's possible movement, and $p(m|x_{1:t}, z_{1:t})$ is the posterior probability density function of the map information, which is calculated by the known pose of the mobile robot. The above is the principle of the Gmapping algorithm. In this paper, the mobile grasping robot uses the GMapping algorithm to construct a map, which needs to obtain its odometer information and the depth information of the lidar, and at the same time add the IMU information, and improve the positioning accuracy of the robot through the fusion of the odometer information and the IMU information, so that the depth information provided by the lidar is more accurate, and finally the surrounding environment is established in the form of a two-dimensional raster map.

4.2 Research on Data Fusion Localization Based on Extended Kalman Filter

The optimal pose of the mobile robot will be estimated by fusing the data of the wheel odometer and the inertial measurement unit IMU through the extended Kalman filter [14] method. Kalman filtering is a method of estimating the optimal value of the state of a system through the observation of the system, and is often used to deal with linear systems. Both the wheeled odometer and the IMU as sensors are nonlinear systems. In order to deal with the nonlinear system, the extended Kalman filter (EKF) expands the Taylor series of the nonlinear function, discarding the higher-order term and linearly approximating the nonlinear function.

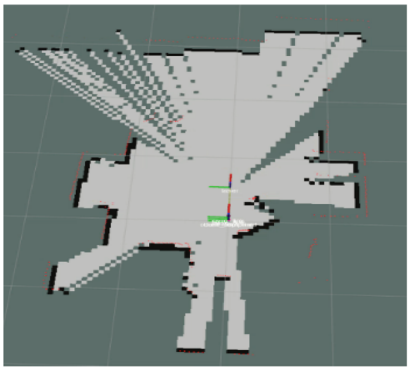
The mobile robot uses LiDAR to construct the mapping algorithm of GMapping algorithm, combined with the positioning information fused by the two sensors based on the extended Kalman filter method, and transmits the corrected attitude to the SLAM system to construct a raster map of the surrounding environment, which provides a prior map for the navigation work of the handling robot. First, for the

transformation of the coordinate system, the center of the wheel track on the ground is used as the motion center of the handling robot, and the coordinate systems such as IMU and lidar are transformed into the coordinate system. Immediately after that, the motor drive, LiDAR drive and IMU analysis module are started, and then the ROS node of the chassis and the extended Kalman filter node are opened, and the node of the chassis manages the release of the wheeled odometer and IMU information, and similarly, also manages the motion control after navigation planning. The extended Kalman filter receives the information from the odometer and the IMU, combines the information of the two to correct the motion state of the entire handling robot, makes its own positioning more accurate, and finally starts the GMapping node.

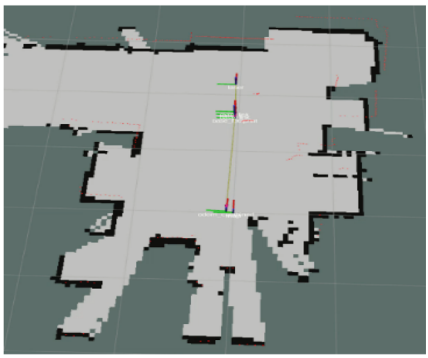
Start Rviz to view the status information of the mapping, as shown in Fig. 12, you can see that the occupied raster map has completely appeared in the area where the lidar is emitted to the area reflected back from the obstacles, the map presented in the initial stage is the map scanned by the handling robot at the starting position, in order to make the map of the surrounding environment be completely established, the handling robot is remotely controlled, and the handling robot is controlled to move to scan the surrounding environment until a complete indoor map is presented, in the intermediate process, There is no obvious edge drift on the occupied raster map, and the positioning effect of the robot itself is good. Finally, save the scanned indoor map, as shown in Fig. 12c, the established map is not much deformed, and can better restore the real environment.

5 Conclusion

In this paper, the convolutional kernel is replaced on the task branch network in the pose estimation network, and the experimental results show that the parameters of the improved pose estimation network are reduced, and the prediction accuracy is improved to a certain extent. The RRT algorithm is improved in the trajectory planning of the robotic arm, and the experimental results show that the average time of the improved model is reduced by 9.5% compared with the original model. However, the improved pose estimation network in this paper has a probability of estimation error in the case of large area occlusion, and the network can be optimized to improve the accuracy of pose estimation under large area occlusion and strong light.



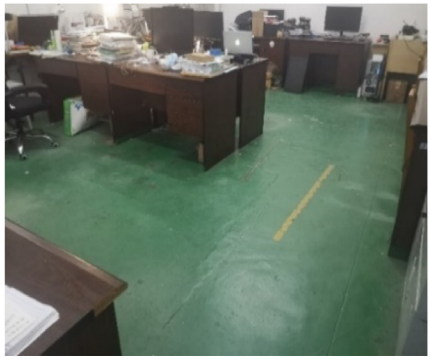
(a) Initial stage of mapping



(b) Intermediate mapping process



(c) Raster maps



(d) Mapping environment

Fig. 12 Mapping effect of mobile robots

Acknowledgements The authors gratefully acknowledge the support of the Xiamen Production and Research Post-subsidy Project (No. 2023CXY0316), Fujian Provincial Department of Science and Technology—University-industry cooperation project (No. 2024H6012).

References

1. Dong M, Zhang J (2023) A review of robotic grasp detection technology. Robotica 41(12):3846–3885. https://doi.org/10.1017/S0263574723001285
2. Song Y, Gao L, Li X (2022) A novel vision-based multi-task robotic grasp detection method for multi-object scenes. Science 65:1–13. https://doi.org/10.1007/s11432-021-3558-y
3. Depierre A, Dellandréa E, Chen L (2021) Scoring graspability based on grasp regression for better grasp prediction. In: International conference on robotics and automation (ICRA). Xian, pp. 4370–4376

4. Zhao B, Zhang H, Lan X et al (2021) REGNet: region-based grasp network for end-to-end grasp detection in point clouds. In: International conference on robotics and automation (ICRA). Xian, pp 13474–13480
5. Ni P, Zhang W, Zhu X et al (2020) Pointnet++ grasping: learning an end-to-end spatial grasp generation algorithm from sparse point clouds. In: International conference on robotics and automation (ICRA). Chengdu, pp 3619–3625
6. Yang D, Tosun T, Eisner B et al (2021) Robotic grasping through combined image-based grasp proposal and 3D reconstruction. In: International conference on robotics and automation (ICRA). Xian, pp 6350–6356
7. Sekiguchi M, Takesue N (2020) Fast and robust numerical method for inverse kinematics with prioritized multiple targets for redundant robots. Adv Robot 34(16):1068–1078. https://doi.org/10.1080/01691864.2020.1780151
8. Li C, Wen C (2024) A novel extended unscented Kalman filter is designed using the higher-order statistical property of the approximate error of the system model. Actuators 13:169. https://doi.org/10.3390/act13050169
9. Yu S, Fu C, Gostar AK, Hu M (2020) A review on map-merging methods for typical map types in multiple-ground-robot SLAM solutions. Sensors 20:6988. https://doi.org/10.3390/s20236988
10. Montemerlo M, Thrun S (2023) Simultaneous localization and mapping with unknown data association using FastSLAM. In: International conference on robotics and automation. Taipei, pp 1985–1991
11. Grisetti G, Stachniss C, Burgard W (2007) Improved techniques for grid mapping with Rao-Blackwellized particle filters. IEEE Trans Robot 23:34–46
12. Tremblay J, To T, Sundaralingam B et al (2018) Deep object pose estimation for semantic robotic grasping of household objects. https://doi.org/10.48550/arXiv.1809.10790
13. Steven M. LaValle. Rapidly-exploring random trees:a new tool for path planning. The Annual Research Report. https://api.semanticscholar.org/CorpusID:14744621
14. Van der Merwe R, Wan EA (2011) The square-root unscented Kalman filter for state and parameter-estimation. In: International conference on acoustics, speech, and signal processing. Proceedings, Salt Lake City, UT, USA, pp 3461–3464. https://doi.org/10.1109/ICASSP.2001.940586

Open Access This chapter is licensed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License (http://creativecommons.org/licenses/by-nc-nd/4.0/), which permits any noncommercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if you modified the licensed material. You do not have permission under this license to share adapted material derived from this chapter or parts of it.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.



Research on the Key Technology of Multi-sensor Fusion Environment Perception Based on Unmanned Patrol Vehicle

Min Lin, Dongsheng Lin, Pengxin Yue, Wenyang Chi, Xinrui Ma, and Yingbing Huang

Abstract The rapid development of China's automotive industry has driven advancements in inspection vehicle technology, with an increasing number of intelligent inspection vehicles being applied in urban management and infrastructure monitoring. They not only enhance inspection efficiency but also collect data in real time, aiding decision-makers in making more informed judgments. As electrification and autonomous driving technologies advance, the functions and application scenarios of inspection vehicles will further expand, becoming the "eyes" and "hands" of urban management. However, due to complex backgrounds, lighting conditions, and limitations of individual sensors, the detection accuracy of single-sensor-based inspection technologies still needs improvement. To address this, this paper proposes constructing a multi-sensor system for unmanned inspection vehicles, conducting systematic research on issues such as target detection, data processing, and information fusion for intelligent inspection perception systems, and designing a complete experimental plan to validate the constructed perception system. Experimental results indicate that this approach effectively resolves the issue.

Keywords Inspection · Multi-sensor fusion · Image segmentation · Object detection · 3D point cloud

1 Introduction

As an extension and application of artificial intelligence in the automotive industry, intelligent inspection technology is making rapid progress. As shown in Fig. 1, this technology holds significant market potential in various scenarios, such as private car travel, campus inspections, logistics delivery, and sanitation cleaning. In recent

M. Lin (✉) · D. Lin · P. Yue · W. Chi · X. Ma · Y. Huang
School of Marine Equipment and Mechanical Engineering, Jimei University, Xiamen, Fujian, China
e-mail: 201761000009@jmu.edu.cn

© The Author(s) 2026

H. Saman et al. (eds.), *The 9th International Conference on Advances in Construction Machinery and Vehicle Engineering*, Lecture Notes in Mechanical Engineering,
https://doi.org/10.1007/978-981-96-5527-4_106

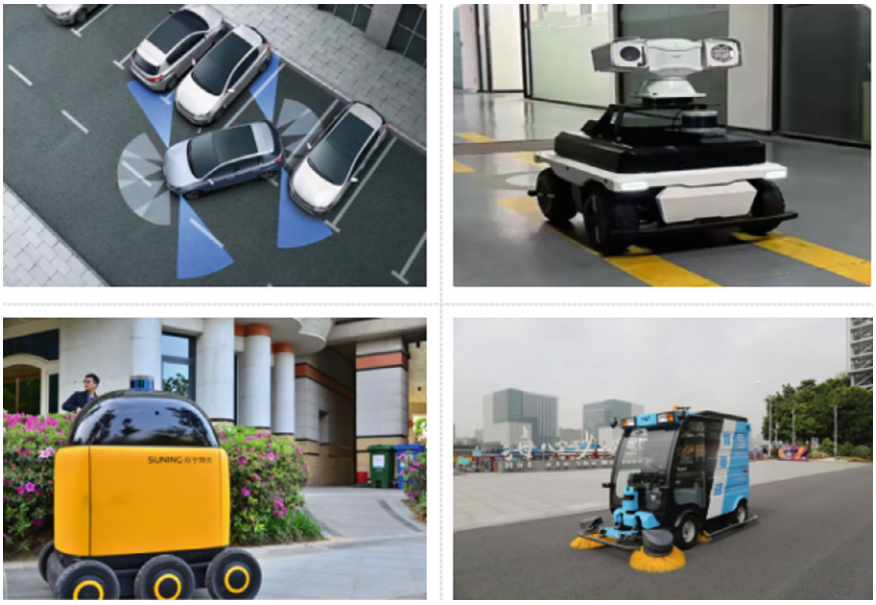


Fig. 1 Application scenarios of intelligent driving

years, intelligent inspection technology has not only garnered positive responses from academia and industry but has also received close attention at the national level. Therefore, vigorously developing intelligent inspection technology not only offers a more convenient lifestyle for people but also plays a crucial role in the transformation and upgrading of China's manufacturing industry.

The framework of smart inspection technology mainly consists of three core modules: perception, decision planning, and control execution. In the task of perceiving the external environment, the core work is object detection. Among these, three-dimensional object detection is a key technology in the field of smart vehicle inspection and has become a focal point for researchers both domestically and internationally, leading to the proposal of many solutions. Depending on the input data sources, these solutions can be broadly categorized into three types: vision-based three-dimensional object detection, lidar-based three-dimensional object detection, and multi-sensor fusion-based three-dimensional object detection. However, while using a single visual sensor for three-dimensional object detection saves on economic costs, the inherent limitations of cameras in capturing depth information significantly compromise both accuracy and data processing convenience [1]. Additionally, their reliability is often questioned due to the substantial impact of environmental conditions.

In the study of lidar and vision fusion strategies, Cui et al. [2] proposed a novel classification method, categorizing fusion models into two major types: strong fusion and weak fusion, thereby enhancing the reliability of lidar and vision fusion strategies. However, due to factors such as significant noise in the raw data collected by this

strategy, the limited size of multimodal datasets, information loss, the substantial differences between point cloud and RGB data, and the challenges of multi-sensor calibration and data alignment, it has been difficult to achieve satisfactory practical results. This area remains open for further research.

Therefore, this paper focuses on 3D object detection technology based on the fusion of lidar and vision in the context of outdoor environment perception for intelligent inspection vehicles. By utilizing lidar and cameras mounted on the vehicle to collect environmental information, the data from both sensors are processed and fused, allowing for mutual assistance and validation. This enables the accurate detection of surrounding vehicles, pedestrians, riders, and other objects. Such advancements provide crucial information support for subsequent planning, decision-making, and control execution, significantly enhancing the overall intelligence of the vehicle system.

2 Related Work

The lidar and camera on unmanned inspection vehicles are heterogeneous sensors from multiple sources. To fuse their information, preliminary data processing and matching are required, with the primary tasks being time synchronization and spatial alignment. Therefore, it is essential to first conduct research on the calibration of lidar and camera models, as well as joint calibration techniques. Following this, the datasets and evaluation metrics necessary for subsequent object detection must be analyzed and selected to ensure the effectiveness of the experimental tools.

For sensors such as lidar and cameras, each hardware component operates based on inherent principles and theoretical models. This leads to inevitable errors during data collection, either due to the working environment or the stability of the hardware itself, which cannot be overlooked [3]. To enhance the accuracy and consistency of the perception data, streamline subsequent data processing, and improve overall efficiency, it is essential to perform calibration on each sensor.

Mechanical lidar determines angles through horizontal rotational scanning and measures distances by calculating the time and phase differences of the laser signals. However, because the coordinate system for each point measured by the lidar differs within each frame during motion, this results in motion distortion in the point cloud. The principle for removing motion distortion involves first estimating the lidar's movement during data collection, then compensating for this movement in each frame by applying rotational and translational transformations based on the relative timestamp of each point. A comparison of the effects before and after distortion correction is shown in Fig. 2. The black rectangular box represents the true boundary of the object, while the blue rectangular box in Fig. 2a shows the scan results before lidar correction, and the red rectangular box in Fig. 2b displays the scan results after correction.

The vast majority of cameras collect data based on the pinhole imaging principle [4]. This principle uses an ideal mathematical model to describe the mapping between

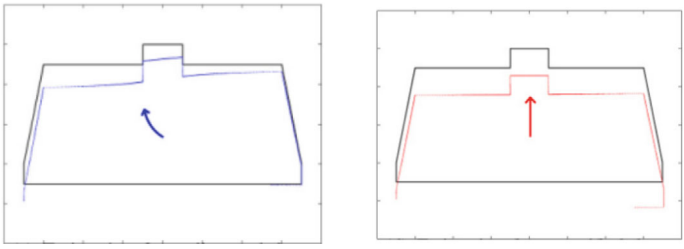


Fig. 2 Comparison of distortion correction effects

3D points in the real world and pixel coordinates on an image plane. However, in practical applications, errors often occur. The primary source of these errors is the distortion of light propagation, which arises from the camera-lens combination [5]. These distortions are typically classified into two categories: (1) radial distortion, caused by the irregular shape of the optical lens itself, and (2) tangential distortion, which occurs when the optical lens is not perfectly parallel to the camera's image plane during assembly. Both radial and tangential distortions can be corrected by establishing relevant distortion correction coefficients in a calibration model.

In the construction of perception systems for inspection vehicles, the combination of lidar and cameras is one of the ideal approaches for achieving comprehensive sensing. The foundation of data fusion between lidar and cameras lies in the synchronization of the temporal and spatial relationships of the information collected by both sensors. Currently, the mainstream methods for time synchronization include hardware synchronization and software synchronization. Software synchronization involves providing both the camera and lidar with a common time source and assigning timestamps to the data recorded by each. This creates virtual frames based on the sensors' inherent acquisition frequencies, enabling data synchronization across different timestamps to capture information at the same moment.

In this paper, a software synchronization method is employed. The synchronization process is illustrated in Fig. 3. First, the time sources of the lidar and camera are unified. Then, the sampled data from both sensors are buffered into a unified time sequence. Using the camera's longer sampling interval as the reference, the lidar data closest in time to the camera's sampling moment is selected as the synchronized data. Finally, the information from the lidar and camera is fused.

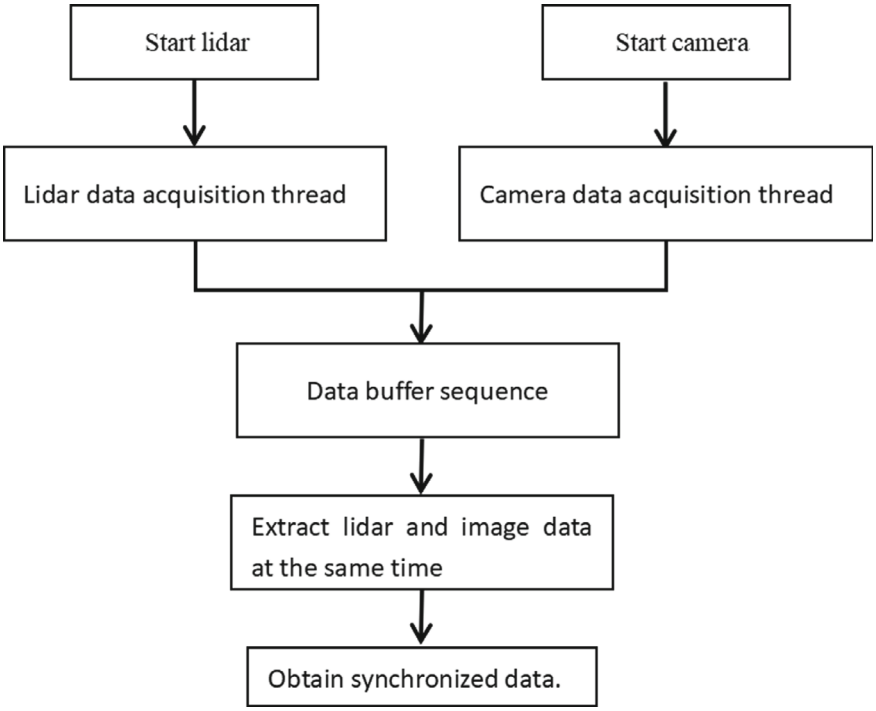


Fig. 3 Synchronization process

3 Detection Algorithm Optimization

3.1 Point Cloud Filtering

Commonly used point cloud filtering methods include radius filtering, statistical filtering, voxel filtering, and geometry-based methods. Voxel filtering, in particular, involves partitioning a cluster of point clouds into three-dimensional grid cells of equal volume, with each cell representing a voxel. Subsequently, operations are performed on the points within each voxel. This study focuses on the application of voxel filtering for noise reduction.

Voxel filtering includes two methods: voxel centroid downsampling and voxel center downsampling [6]. When comparing these two sampling methods, voxel center downsampling, while faster in computational speed, tends to exhibit lower accuracy and is more susceptible to distortion. Therefore, this study selects the voxel centroid downsampling method for processing the point cloud. The specific computation process is as follows.

Create a cubic bounding box that completely encapsulates the target point cloud cluster with dimensions l, w, h. Select a suitable voxel edge length L to partition

the target point cloud cluster into smaller grid cells, where each cell represents a voxel. Each voxel is then assigned a unique identifier and calculated. Let us assume a non-empty voxel V containing m points. The method for calculating the centroid coordinates $P_c(x_c, y_c, z_c)$ of this voxel is as follows:

$$x_c = \frac{\sum_{i=1}^{m} x_i}{m} x_c = \frac{\sum_{i=1}^{m} x_i}{m} z_c = \frac{\sum_{i=1}^{m} z_i}{m} \tag{1}$$

3.2 Ground Point Cloud Segmentation Algorithm

During object detection, the primary criterion for clustering discrete 3D data points into the same object class is the distance between points. In most environments, a significant portion of the collected lidar point cloud data consists of interference data related to ground points. This type of data often appears textured, which can easily disrupt subsequent point cloud recognition and classification processes. Therefore, the removal of ground point clouds is a critical step in the preprocessing of point clouds.

This paper proposes an adaptive ground point cloud segmentation algorithm that performs ground point cloud segmentation based on the inherent horizontal angular resolution of the lidar. First, laser data from n lines at the same horizontal angular resolution is processed, followed by calculations across various angles to determine the overall plane. By evaluating the distances and local slopes of the laser point clouds between consecutive scanning circles at the same plane angle, parameters such as adaptive horizontal distance, local height, and global height thresholds are utilized to ascertain whether a point is classified as ground, thereby mitigating interference from ground point clouds.

3.3 Improvement of the PointPillars Network Architecture

In the task of 3D object detection, this study initiates its exploration from point cloud-based 3D object detection methods, applying and enhancing the concepts of the PointPillars model for detection. The architecture of the PointPillars network is illustrated in Fig. 4 and consists of three main components: (1) the Pillar Feature Net (feature encoder network), (2) the Backbone (2D CNN), and (3) the detection head (SSD).

Although the original PointPillars algorithm demonstrates a certain level of detection performance, it still has limitations. Due to the constraints of the local receptive fields in its backbone network, the algorithm struggles to capture the complete feature information of the entire object, resulting in decreased detection performance

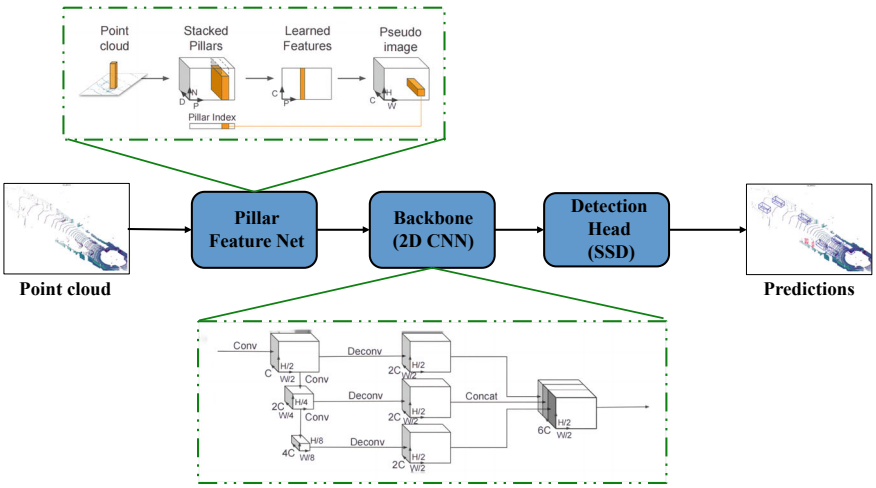


Fig. 4 PointPillars model architecture

for some targets. This study improves upon the original PointPillars network architecture, as depicted in Fig. 5. In the feature encoding phase, the original point cloud data is processed into pillar format. However, the original backbone feature extraction component is replaced with a ResNet-18 network, which utilizes a series of residual blocks to downsample the pseudo-image and extract high-level features. Subsequently, the ASPP module is introduced, employing convolution operations with varying dilation rates to expand the receptive field and enhance the understanding of contextual information and the integration of multi-scale feature maps. Finally, the concatenated feature information is fed into the SSD detection head for bounding box regression.

4 Fusion Model Construction Experiment

Although the mainstream 3D object detection algorithms currently available on the market are primarily based on single-sensor detection, the increasing complexity of operational scenarios for unmanned inspection vehicles and the rising safety requirements suggest that multi-sensor fusion for object detection will be a key trend in the future [7]. Therefore, incorporating semantic information such as visual color to complement pure point cloud data can significantly enhance the informational basis for detection, thereby greatly improving accuracy. Building upon the related work discussed in Sect. 2, this study investigates a pre-fusion approach to maximize the utilization of sensor information.

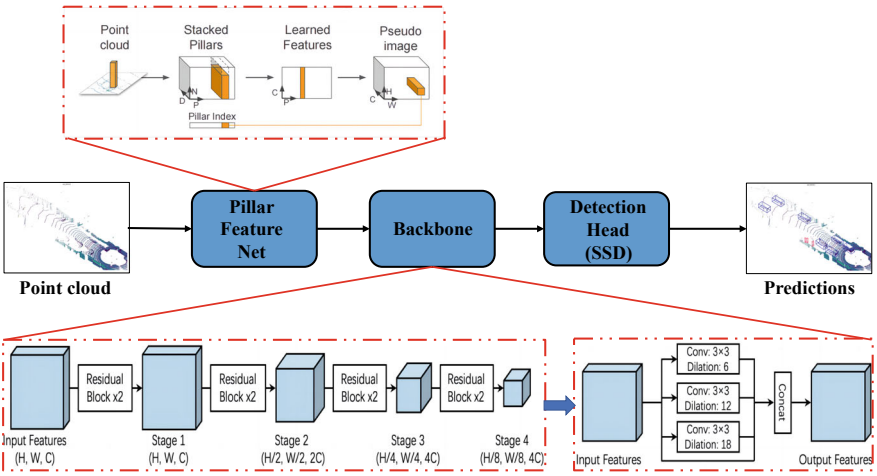


Fig. 5 Improved algorithm architecture

4.1 Analysis of Semantic Segmentation Network Selection

Common semantic segmentation networks include Fully Convolutional Networks (FCN), SegNet, U-Net, and the DeepLab series. Among these, the DeepLab network introduces atrous convolution and employs Conditional Random Fields to enhance the model's ability to capture intricate details. DeepLabV3+ utilizes an encoder-decoder architecture, with V3 serving as the encoder structure and a separately designed decoder structure, which improves the accuracy and precision of image recognition [8]. Considering the performance characteristics of these networks, this study selects the DeepLabV3+ network to filter the experimental targets from images. The overall architecture of the DeepLabV3+ network is illustrated in Fig. 6.

4.2 Point Cloud and Image Fusion Module

Once the semantic segmentation of the image portion is completed, it is necessary to match and fuse the lidar point clouds with the results of the semantic segmentation. The foundation for integrating point clouds with image semantic information is the establishment of a mapping between the point cloud coordinate system and the camera imaging plane coordinates. Based on the camera imaging model and the principles of joint calibration between the lidar and the camera, it can be determined that a series of matrix transformations can project the point clouds onto the camera's imaging plane. In the KITTI dataset, since the coordinate relationships of the original sensors used for data acquisition are well-defined, the mapping parameters from the lidar coordinate system to the camera coordinate system can be directly utilized for processing, as expressed in the following equation:

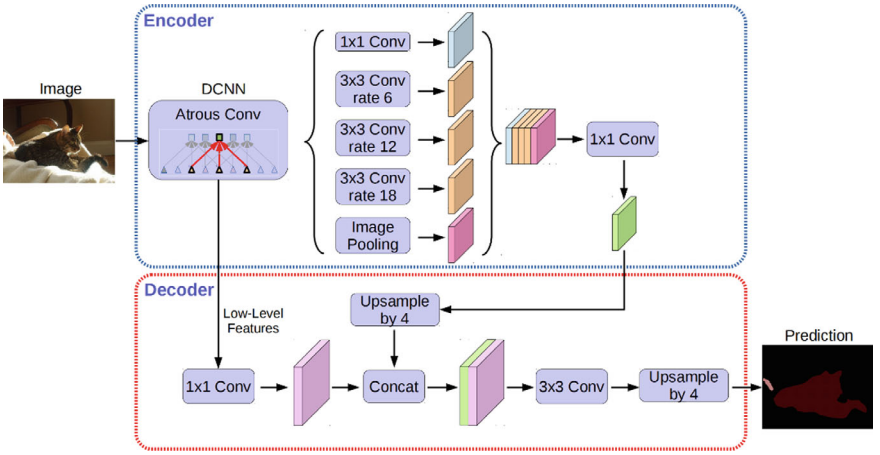


Fig. 6 DeepLabV3+ network architecture

$$T_{ki} = T_{(cam_to_image)} T_{(rect_to_cam)} T_{(lidar_to_cam)} \tag{2}$$

After the aforementioned transformations, the lidar point cloud data is projected onto the camera's imaging plane, establishing a specific coordinate correspondence between the point clouds and the image pixels. Before performing point cloud and image registration, it is essential to eliminate point clouds that are outside the camera's field of view to enhance overall operational efficiency. The area of the point clouds on the RGB image can be calculated using Eq. (3), which employs the point cloud coordinates in the pixel coordinate system. In this equation, u_p represents the pixel coordinates in the horizontal direction, v_p denotes the pixel coordinates in the vertical direction, I_u refers to the width of the aligned image dimensions, and I_v indicates the height of the aligned image dimensions.

$$\left\{(u_p, v_p) | (0.5 <= u_p < I_u - 0.5), (0.5 <= v_p < I_v - 0.5)\right\} \tag{3}$$

4.3 Point Cloud and Image Fusion Module

In this study, the semantic segmentation experiment employs DeepLabV3+ as the processing model for image information. Initially, the model is pretrained on the Cityscapes dataset to obtain a training model, which is subsequently tested for performance on the same dataset. The experimental results are illustrated in Fig. 7, where each pixel has been accurately assigned fixed label information. The primary detection target of this research, pedestrians, is rendered in red, vehicles in dark blue, and bicycles in brown; other background elements, such as the sky, are rendered



Fig. 7 Image segmentation results

in light blue, and the ground is depicted in pink. The experimental results demonstrate that even in scenes with dense pedestrian presence (Scene 2) and high vehicle density (Scene 3), satisfactory segmentation performance is achieved, meeting the requirements for subsequent information fusion.

5 Experimental Validation

5.1 Configuration Parameters of the Experimental Platform

The research group has established an experimental platform based on the Autolabor Pro vehicle, as shown in Table 1. This platform is equipped with the Hesai Pandar64 64-line lidar, the Intel RealSense D455 camera, and the Chenyang Nuvo-6180GC industrial computer.

5.2 Camera Calibration Experiment

In this study, the camera calibration experiment employs the Zhang Zhengyou calibration method based on a reference object. During calibration, the camera's resolution is set to 1280 × 720. The calibration tool used is a 9 × 7 checkerboard calibration board, printed on A4 paper, with each square measuring 24.5 mm on each side. Throughout the experiment, the calibration board is moved to various positions within the camera's field of view, allowing the camera to capture images of the checkerboard from multiple angles, thereby facilitating the extraction of feature points, as illustrated in Fig. 8.

In the camera calibration experiment of this study, multiple sets of images were collected, with over 30 sets selected for the final calibration. The spatial relationships of these 30 images with respect to the camera are depicted in Fig. 9a, while the average re-projection error of the calibration results is shown in Fig. 9b. The calibration data obtained in this study meets the experimental requirements.

Table 1 Selection and parameters of key components of the experimental platform

Module	Selection	Name	Parameters
Small vehicle		Autolabor	CAN protocol communication, steer-by-wire, hub motor drive
Computer		Nuvo-6180GC	CPU i7-6700, 32GB memory, 500GB storage space, GPU RTX-3060, and 480W power supply
Laser sensor		Pandar64	A measurement range of up to 150m, vertical field of view of 40°: + 15° to − 25°, vertical angular resolution of 0.1°, 360° field of view, rotation rate of 10–20 Hz, and dual return
Visual sensor with RGB camera		Intel Realsense D455	Resolution of 1280 × 720, maximum frame rate of 90 FPS, and RGB sensor dimensions (H × V): 86 × 57 (±3)

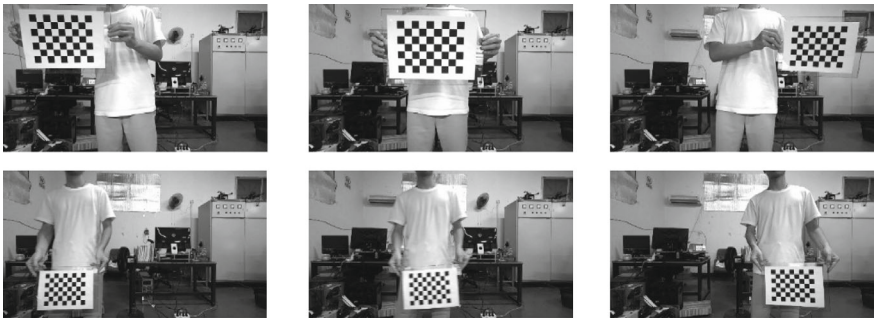
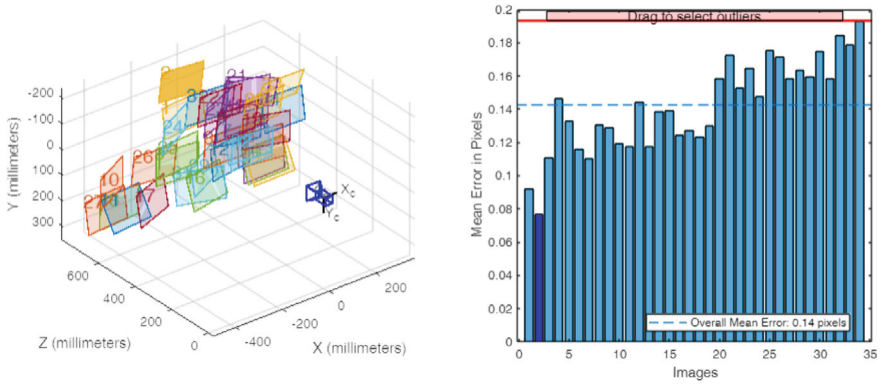


Fig. 8 Camera calibration diagram

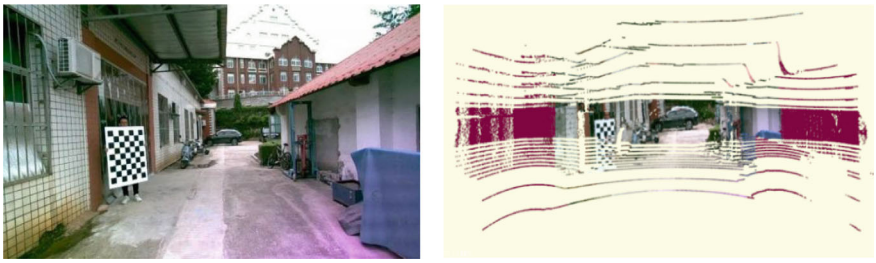
5.3 Sensor Joint Calibration Experiment

Building on the specific parameters obtained for the camera, this experiment utilizes the Calibration Toolkit from Autoware for the joint calibration of the lidar and camera. After acquiring the calibration parameters for both the camera and lidar, these parameters are employed to fuse the RGB information from the camera with the point cloud



(a).Camera Calibration Diagram (b).Average Reprojection Error

Fig. 9 Camera calibration experiment diagram



(a).Camera Calibration Diagram (b). Point Cloud Information Fusion Diagram

Fig. 10 Fused point cloud map

data from the lidar. The results are illustrated in Fig. 10, where the left side displays the camera image and the right side shows the corresponding fused point cloud. A comparison of the two images reveals that both the nearby calibration board and the distant electric bike and car are clearly visible, with the point clouds within the camera's field of view accurately assigned color information.

5.4 Three-Dimensional Object Detection Experiment

To validate the compensation effect achieved by incorporating visual semantic information, this study compares the three-dimensional object detection results after fusion with those obtained from pure point cloud data.

The experimental results are presented in Fig. 11, where the upper image of each frame depicts the original image, the middle image displays the detection results based on pure point cloud data, and the lower image shows the results after fusion.

The detection results for Frame 21 indicate that after integrating visual information, the perception system can detect objects at greater distances with improved accuracy. Frame 115 illustrates the elimination of false detection issues during the detection process. In this scenario, two trees located on the right side of the road were initially misidentified as pedestrians due to the similarity in point cloud shapes; however, after improvements, these false detections were successfully removed. The fused detection model enhances the richness of the overall information and increases the accuracy of the detections.

This study also conducted a quantitative analysis of the experimental results based on the data. Table 2 summarizes the accuracy of the 3D detection boxes. In terms of 3D detection box performance, there was no significant improvement in vehicle detection compared to the previously improved PointPillars algorithm; however, enhancements were observed in the detection accuracy for pedestrians and cyclists.



(a) The 21st frame. (b) The 115st frame.

Fig. 11 Detection result comparison diagram

Table 2 Average precision table for 3D object detection on the KITTI test set

Model	Vehicle			Pedestrian			Cyclist		
	Easy	Moderate	Hard	Easy	Moderate	Hard	Easy	Moderate	Hard
VoxelNet	81.97	65.46	62.85	57.86	53.42	48.87	67.17	47.65	45.11
SECOND	83.13	73.66	66.20	51.07	42.56	37.29	70.51	53.85	53.84
PointRCNN	85.94	75.76	68.32	49.43	41.78	38.63	73.93	59.60	53.59
PointPillars	86.59	76.97	75.30	54.32	49.23	46.01	81.27	64.75	60.66
PointPillars	87.11	77.61	76.17	54.85	50.70	46.46	82.12	64.88	61.42
After Optimization	87.63	77.83	76.28	59.23	54.37	49.30	84.19	67.18	63.65
After Fusionn	90.01	87.65	85.10	66.10	59.67	54.96	90.29	70.05	66.23

6 Conclusion

This study focuses on the perception system of an intelligent unmanned inspection vehicle, specifically addressing the issue of three-dimensional object detection for surrounding obstacles. A three-dimensional object detection scheme that integrates LiDAR and visual information is proposed, optimizing the input for subsequent three-dimensional object detection by incorporating visual label information into the pure point cloud.

Initially, semantic segmentation is employed to extract pixel information from images and classify it. Subsequently, coordinate transformation is used to project the point cloud into the image coordinate system, thereby adding the extracted semantic information to the feature dimensions of the point cloud, facilitating the fusion of both information sources. Finally, an experimental platform was independently established to validate the proposed point cloud preprocessing methods, pure point cloud detection algorithms, and the fusion detection algorithms. Experiments were conducted using the laboratory's unmanned inspection vehicle, and the results demonstrated the effectiveness of the proposed research approach.

However, several challenges encountered in this study warrant further exploration and optimization in future work. Due to model computational limitations, the research primarily focused on detecting pedestrians, vehicles, and cyclists. In real operational environments for intelligent vehicles, a variety of obstacles may be present. Thus, future research could consider including additional types of obstacles in the detection scope.

Acknowledgements The authors gratefully acknowledge the support of the Natural Science Foundation of Fujian Province, China (No. 2023J01792), Major Project of Science and Technology of Quanzhou (No. 2022GZ8), Xiamen Science and Technology Subsidy Project (No. 2023CXY0316), and the Young and Middle-aged Teacher Education Scientific Program of Fujian Province (Grant No. JAT220180), Fujian Provincial Department of Science and Technology-University-industry cooperation project (No. 2024H6012).

References

1. Weng X, Kitani K (2020) Monocular 3D object detection with pseudo-LiDAR point cloud
2. Cui Y, Ren C, Wenbo C (2022) Deep learning for image and point cloud fusion in autonomous driving. IEEE Trans Intell Transp Syst 23(2):722–739
3. Sindagi V-A, Yin Z, Oncel T (2019) Multimodal VoxelNet for 3D object detection. Int Conf Robot Autom 8794:7276–7282
4. Juarez-Salazar R, Juan Z, Diaz-Ramirez VH (2020) Distorted pinhole camera modeling and calibration. Appl Opt 59(36):11310–11318
5. C Ricolfe-Viala, Sanchez-Salmeron A-J (2010) Lens distortion models evaluation. Appl Opt 49(30):5914-5928
6. Benhabiles H, Aubreton O, Barki H (2013) Fast simplification with sharp feature preserving for 3D point clouds. IEEE, pp 47–52
7. Zheng W, Tang W, Jiang L (2021) Self-ensembling single-stage object detector from point cloud. CVPR, pp 14489–14498
8. Shelhamer E, Jonathan L, Trevor D (2016) Fully convolutional networks for semantic segmentation. IEEE—Trans Pattern Recognit Mach Intell 50(19):474–478

Open Access This chapter is licensed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License (http://creativecommons.org/licenses/by-nc-nd/4.0/), which permits any noncommercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if you modified the licensed material. You do not have permission under this license to share adapted material derived from this chapter or parts of it.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.



Research of Ergonomic Test on Human Whole-Body Vibration of Armored Vehicle Cabin



Tushun Huang, Meng Cai, Wei Chang, and Weiwei Zhang

Abstract In order to measure and evaluate the impact of whole-body vibration in the armored vehicle on human work efficiency, a certain type of off-road vehicle is taken as the research object, and the vibration of seat from the front and rear seat in the vehicle cabin is measured during the process of driving on undulating soil roads. Four methods from standards are used to analyze the relations among the vibration and the fatigue-reduction efficiency limits and exposure limits of human beings, also they are compared with each other. Then the shortcomings of existing standards are pointed out, which may go some way towards optimizing the ergonomics test methods and upgrading the related national standards to a certain extent.

Keyword Whole-body vibration · Armored vehicle cabin · Fatigue-reduction efficiency limits · Exposure limits

1 Introduction

The vibration generated by armored vehicles during operation is one of the main environmental factors affecting human work efficiency, vibration can be classified into hand transmitting vibration, equipment vibration, and whole-body vibration by its transmission mode [1, 2]. Among them, whole-body vibration has the most serious impact on human health and work efficiency, the whole-body vibration refers to the vibration transmitted to the entire body through the body supporting surface. From the perspective of the mechanical structure of vibration transmission, the human body is a multi-degree of freedom vibration system, and each local system has its own resonance frequency. When vibration is transmitted in the human body, high-frequency components quickly decay, while low-frequency components resonate in certain areas, especially sensitive to vibrations among 1–100 Hz [3, 4]. For example, vibration below 10 Hz has a greater impact on the chest and abdomen; vibration

T. Huang · M. Cai (✉) · W. Chang · W. Zhang
Huayin Ordnance Test Center of China, Huayin, China
e-mail: 645670777@qq.com

© The Author(s) 2026

H. Saman et al. (eds.), *The 9th International Conference on Advances in Construction Machinery and Vehicle Engineering*, Lecture Notes in Mechanical Engineering,
https://doi.org/10.1007/978-981-96-5527-4_107

below 20 Hz has a greater impact on the lower limbs; vibration from 20 to 50 Hz has a greater impact on the spine, especially the lumbar spine, which can lead to lumbar muscle pain or even herniated intervertebral discs; and vibration from 20 to 90 Hz has a greater impact on the eyeballs. In terms of the impact on work efficiency, vibration of 1–20 Hz affects language, vibration of 1–2 Hz affects the ability of the eye to track the target movement, vibration of 1–25 Hz affects reading, vibration of 3–8 Hz affects manual tracking, vibration of 6–8 Hz affects the head movement, and it is difficult for human body to maintain a stable operating posture at 100 Hz, especially below 30 Hz [5]. Therefore, it is necessary to conduct ergonomics test on the whole-body vibration.

Carrying out the human whole-body vibration test of armored vehicle compartment in the testing range has two purposes: firstly, by analyzing the characteristics of whole-body vibration, it can make the research of vibration-damping measures more targeted; and secondly, it can provide a theoretical basis for the testing range to carry out training of human and weaponry scientifically. Taking a certain type of off-road vehicle as the test object, the human whole-body vibration in the process of driving is measured practically based on the relevant test standards, and various evaluation methods are compared and analyzed to derive the fatigue-reduction efficiency limits and exposure limits of a person when he is driving this type of vehicle. It provides a reference for carrying out relevant tests at the testing range.

2 Testing

2.1 Conditions and Elements of Test

The test is based on the national military standard GJB59.15 “Armored Vehicle Test Procedure Field Vibration Test”, GJB965 “Specification for Measurement of Human Exposure to Whole-body Vibration Environment”, GJB966 “Comfort Reduction Limits and Evaluation Criteria for Human Body Whole-body Vibration Exposure”, GB/T13441 “Specification for Measurement of Human Exposure to Whole-body Vibration Environment” and GB/T13442 “Reduced Comfort Boundary and Evaluation Criteria for Human Exposure to Whole-body Vibration” are also referred.

The test object is an off-road vehicle simulating an armored vehicle, the driving road surface is a naturally occurring undulating dirt road in rural areas, which length is about 2 km. One test person is in the sitting posture during the vehicle’s driving process, and the whole-body vibration is mainly transmitted to the human body through the seat supporting surface. Therefore, the measurement points are placed at the center of the seat supporting surfaces on the right side of the front seat and the left side of the rear seat, at speeds of 15 km/h and 25 km/h, respectively. the test equipment is a PCB seat cushion 3D acceleration sensor, and the frequency response range is 1–100 Hz. The 3D vibration direction is referenced to the anatomical coordinates of the human body as shown in Fig. 1, X-axis is from back to chest, Y-axis is from

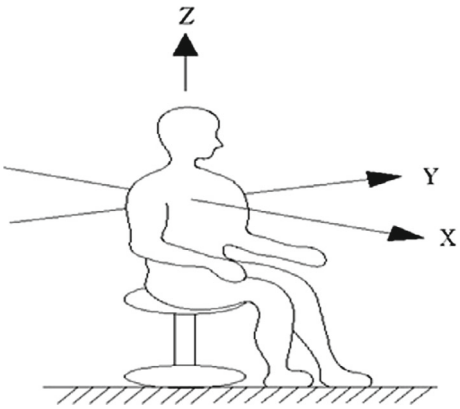


Fig. 1 Schematic diagram of human anatomy

right to left, and Z-axis is from buttocks to head [6]. The height of the subject is 170 ± 0.5 cm and the weight is 70 ± 5 kg. Figure 1 is the Schematic diagram of human anatomy.

2.2 Test Data

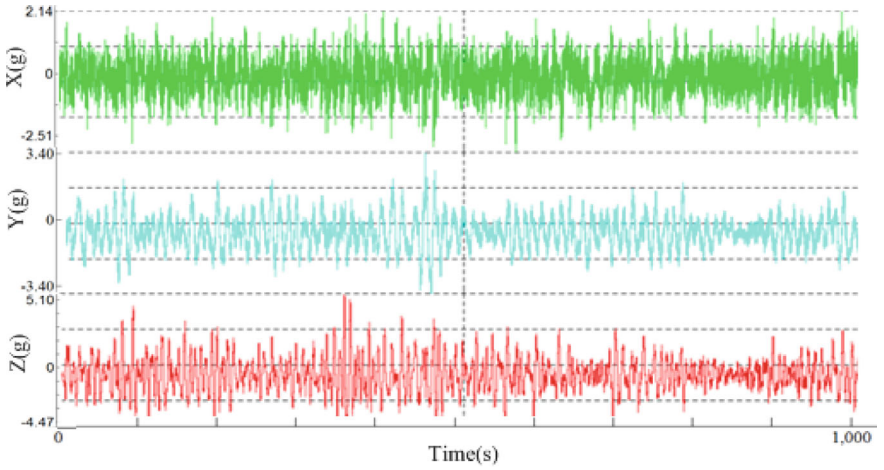
The time-domain data of acceleration can be obtained directly through the test, and the RMS (Root Mean Square) of 1/3-octave acceleration spectrum can be obtained after data process. Figures 2a, b and 3a, b show the time-domain and frequency-domain curves of one test point at 25 km/h, respectively, and this is a typical driving status for an off-road vehicle on a bumpy road.

The test results show that there is a non-uniform energy distribution in the most sensitive frequency range from 1 to 80 Hz, and the low-frequency component is relatively rich, mainly distributed in the range from 1 to 50 Hz. The vibration is a broadband random vibration, and its basic evaluation parameters are the RMS values of acceleration for each 1/3-octave bandwidth in the frequency range, and the specific data are shown in Table1.

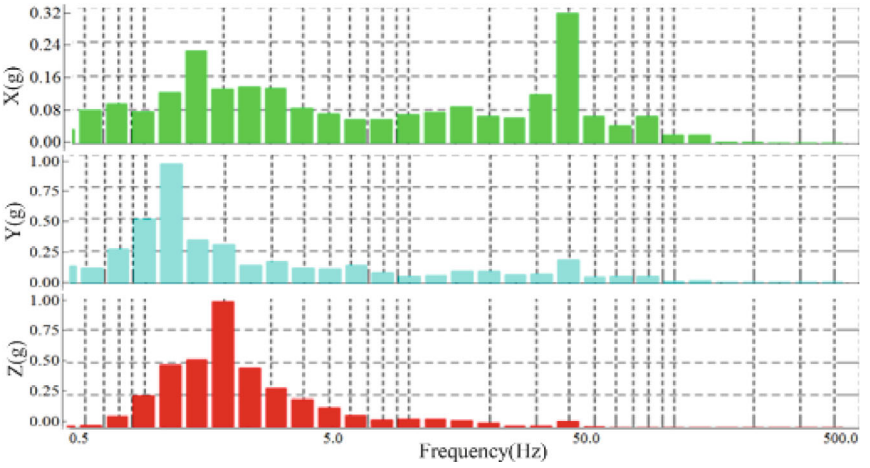
2.3 Data Validation

In order to prevent underestimation of the effect of peak acceleration on the human body, the test data need to be verified with a peak coefficient as shown in Eq. (1) below [7].

$$CF = \frac{a_{(t)\max}}{a_{RMS}} \tag{1}$$



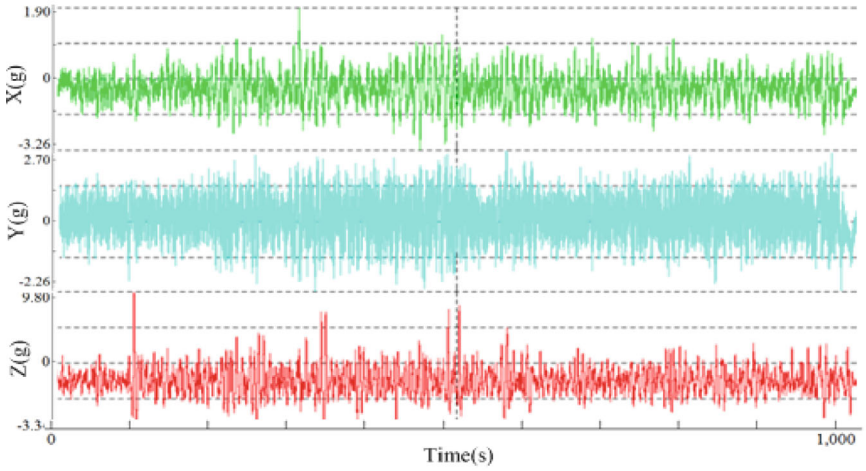
(a)



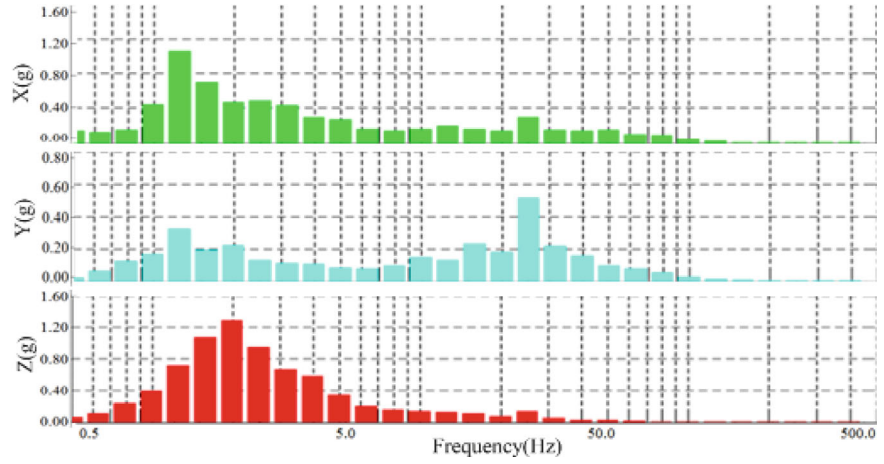
(b)

Fig. 2 **a** Time-domain curves on the right of the front seat at 25km/h; **b** 1/3 octave-frequency curves on the right of the front seat at 25 km/h

By calculating and comparing, the maximum peak coefficient of the nine groups of data in Table 1 is 5.06, and it is generally believed that when the peak coefficient is not greater than 6, the RMS values can be used as the basic parameters for evaluating the vibration, and all the collected data are valid [8].



(a)



(b)

Fig. 3 **a** Time-domain curves on the left of the front seat at 25 km/h; **b** 1/3 octave-frequency curves on the left of the front seat at 25 km/h

Table 1 RMS of acceleration values on the right of front seat and the left of rear seat

Sampling number	RMS of acceleration on the right of front seat (m/s^2)						RMS of acceleration on the left of rear seat (m/s^2)					
	15 km/h			25 km/h			15 km/h			25 km/h		
	X(g)	Y(g)	Z(g)	X(g)	Y(g)	Z(g)	X(g)	Y(g)	Z(g)	X(g)	Y(g)	Z(g)
1	0.0671	0.282	0.136	0.076	0.492	0.244	0.435	0.169	0.402	0.326	0.11	0.313
2	0.218	0.296	0.435	0.120	0.928	0.485	1.050	0.325	0.715	0.774	0.194	0.820
3	0.153	0.265	0.540	0.218	0.334	0.524	0.693	0.193	1.070	1.000	0.320	1.780
4	0.166	0.317	1.190	0.128	0.303	0.967	0.463	0.223	1.280	0.994	0.243	1.520
5	0.152	0.251	0.996	0.133	0.138	0.457	0.478	0.128	0.948	0.985	0.411	1.750
6	0.144	0.219	0.539	0.129	0.166	0.301	0.424	0.109	0.667	0.833	0.322	2.220
7	0.128	0.200	0.537	0.084	0.118	0.219	0.288	0.105	0.582	0.769	0.222	1.400
8	0.162	0.255	0.587	0.071	0.111	0.156	0.256	0.082	0.345	0.486	0.222	0.757
9	0.143	0.324	0.355	0.058	0.141	0.099	0.148	0.078	0.200	0.380	0.201	0.681
10	0.233	0.298	0.331	0.057	0.083	0.065	0.130	0.098	0.157	0.440	0.278	0.510
11	0.221	0.276	0.379	0.068	0.052	0.071	0.146	0.148	0.137	0.476	0.411	0.405
12	0.333	0.394	0.373	0.074	0.059	0.068	0.184	0.130	0.127	0.478	0.387	0.296
13	0.510	0.371	0.236	0.087	0.092	0.057	0.148	0.230	0.108	0.381	0.584	0.222
14	0.353	0.231	0.136	0.064	0.092	0.042	0.132	0.181	0.075	0.421	0.510	0.140
15	0.277	0.213	0.069	0.061	0.069	0.019	0.288	0.512	0.138	0.593	0.903	0.191
16	0.391	0.247	0.060	0.114	0.070	0.018	0.141	0.217	0.054	0.300	0.371	0.091
17	0.369	0.236	0.049	0.306	0.178	0.052	0.132	0.157	0.029	0.366	0.312	0.076
18	0.263	0.186	0.039	0.065	0.044	0.010	0.141	0.099	0.029	0.214	0.163	0.042
19	0.171	0.214	0.029	0.043	0.050	0.007	0.085	0.078	0.018	0.161	0.119	0.032
20	0.107	0.0649	0.012	0.064	0.054	0.007	0.073	0.056	0.011	0.125	0.065	0.025

3 Assessment and Analysis of Data

3.1 Assessment Criteria

For random whole-body vibration generated by an armored vehicle during running, the comfort reduction limits, fatigue-reduction efficiency limits, and exposure limits can be used as the final evaluation results [9, 10]. There is one correlation among these three indicators, with the fatigue-reduction efficiency limits being 3.15 times the comfort reduction limits, and the exposure limits being 6.3 times the comfort reduction limits. The data for the comfort reduction limits are given in GJB965, and Fig. 4 shows a curve representation of vertical (Z-axis) acceleration versus the comfort reduction limits.

Since military standards only provide acceleration limits and exposure time for individual points in tabular form, the data need to be fitted. In this paper, the following fitting Eq. (2) is used for the comfort reduction limits and fatigue-reduction efficiency limits.

$$a_t = \begin{cases} a_1, & t \le 10 \\ a_1\sqrt{t_0/t_1}, & 10 \le t \le 15 \end{cases} \tag{2}$$

where, a_1—boundary value for 1 min vibration, a_t—boundary value for t mins vibration, t_0—10 min, t_1—exposure time. And the following exponential fitting Eq. (3) is used for the exposure limits:

$$t_1 = ke^{\lambda a_t} \tag{3}$$

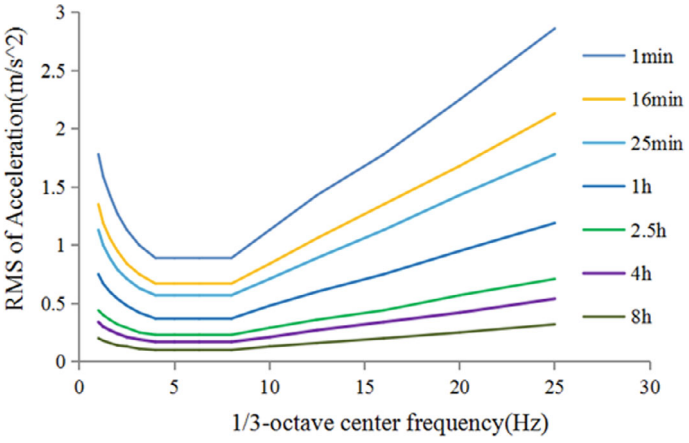


Fig. 4 Relation between vertical acceleration limits and comfort reduction limits

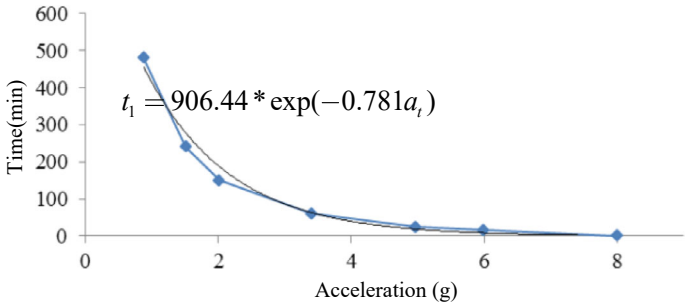


Fig. 5 Fitted curve of the exposure limits for 2 Hz vibration of the vertical axis at 25km/h

where, k, λ are the fitting coefficients. Figure 5 shows the fitted curve of the exposure limits for 2 Hz vibration of the vertical axis at 25km/h.

The relationship among the comfort reduction limits, the fatigue-reduction efficiency limits and the exposure limits is established by data fitting, so the fatigue-reduction efficiency limits and the exposure limits can be analyzed with data of the comfort reduction limits [11].

3.2 Evaluation Methods

Method 1—direct evaluation of RMS of unidirectional acceleration. The RMS values of acceleration in 1/3 octave frequency band for each direction vector in each of the three axial directions are rated according to the boundaries corresponding to the center frequency of the band. The corresponding allowable exposure time is calculated according to Eqs. (2) and (3). Because the comfort reduction limits in the horizontal direction (X and Y axes) is basically the same, no further distinction is made between the X and Y axes in the data table. The evaluation results utilizing method 1 are shown in Table 2.

It's proposed from method 1 that when there are multiple frequency bands of vibrational energy acting on the human body, there is no obvious connection during the frequency bands, and the impacts on the human body are mainly caused by the single most prominent frequency band.

Method 2—Unidirectional Weighted Acceleration RMS Direct Separate Evaluation Method. The basic idea of method 2 is that the acceleration components for different individual octave band do not truly reflect the human body's perception of vibration intensity. It is necessary to convert other frequencies of vibration to the most sensitive frequency range of the human body, with the most sensitive frequency range of the allowable acceleration RMS values for comparison, to determine the allowable exposure time. So the weighting coefficients indicating the degree of sensitivity to vibration at different frequencies are introduced, in which the human body's most

Table 2 Permitted exposure time (method 1)

Measuring position	Velocity (km/h)	Direction	Comfort reduction limits (min)	Fatigue-reduction efficiency limits (min)	Exposure limits (min)
Right of front seat	15	X, Y	39.5	391	358
		Z	11.4	113	
	25	X, Y	< 1	45	202
		Z	17.2	171	
Left of rear seat	15	X, Y	< 1	35	168
		Z	< 1	97	
	25	X, Y	< 1	39	102
		Z	< 1	20	

sensitive frequency range is 4–8 Hz on the vertical axis and 1–2 Hz on the horizontal axis. The RMS values a_{wi} of weighted acceleration are obtained according to Eq. (4).

$$a_{wi} = a_i \cdot w_j \tag{4}$$

where, a_i—RMS spectral value of the *ith* 1/3-octave bandwidth, w_j—weighting coefficient (from the table). The RMS values of the weighted acceleration are denoted as a_{zwi}, a_{xwi} and a_{ywi}, and the maximum values $a_{zw\,\max}$, $a_{xw\,\max}$ and $a_{yw\,\max}$ are obtained, the exposure limits in the 4–8 Hz band (with a weighting coefficient of 1) in a_Z axis and in the 1–2 Hz band (with a weighting coefficient of 1) in a_X and a_Y axes are evaluated, respectively. The evaluation results from method 2 are shown in Table 3.

Method 3—Unidirectional Full-domain Weighted Acceleration Vector Sums. It's considered from method 3 that both uniaxial and multiaxial vibrations should be evaluated according to the full-domain weighted acceleration vector sum in order

Table 3 Permitted exposure time (method 2)

Measuring position	Velocity (km/h)	Direction	Comfort reduction limits (min)	Fatigue-reduction efficiency limits (min)	Exposure limits (min)
Right of front seat	15	X, Y	39.5	391	347
		Z	11.1	110	
	25	X, Y	< 1	45	202
		Z	16.8	166.7	
Left of rear seat	15	X, Y	< 1	35	168
		Z	< 1	95.2	
	25	X, Y	< 1	39	98
		Z	< 1	20	

to reflect the total effect of vibration. The unidirectional fully weighted acceleration vector sums are calculated as Eq. (5).

$$a_w = \sqrt{\sum_{i=1}^{20} (a_i \cdot w_i)^2} \tag{5}$$

The exposure limits in the 4–8 Hz band in a_z and the 1–2 Hz band in a_x and a_y are evaluated, respectively. The results of the evaluation are shown in Table 4.

Method 4—Three-direction full-domain weighted acceleration vector sums. The full-domain weighted acceleration vector sums in the three axial directions can be obtained by converting the unidirectional full-domain vector sums to the Z-axis, which is calculated according to Eq. (6),

$$a = \sqrt{a_{zw}^2 + (1.4a_{xw})^2 + \left(1.4a_{yw}\right)^2} \tag{6}$$

The evaluation results are shown in Table 5.

Table 4 Permitted exposure time (method 3)

Measuring position	Velocity (km/h)	Direction	Comfort reduction limits (min)	Fatigue-reduction efficiency limits (min)	Exposure limits (min)
Right of front seat	15	X, Y	< 1	88.7	127
		Z	< 1	28.0	
	25	X, Y	< 1	29.4	128
		Z	< 1	80.9	
Left of rear seat	15	X, Y	< 1	17.6	85
		Z	< 1	26.4	
	25	XY	< 1	10.3	19
		Z	< 1	< 1	

Table 5 Permitted exposure time (method 4)

Measuring position	Velocity (km/h)	Direction	Comfort reduction limits (min)	Fatigue-reduction efficiency limits(min)	Exposure limits (min)
Right of front seat	15	X, Y, Z	< 1	< 1	37
	25	X, Y, Z	< 1	17.5	86
Left of rear seat	15	X, Y, Z	< 1	< 1	7
	25	X, Y, Z	< 1	< 1	< 1

3.3 Analysis of Results

(1) The calculation results from method 1 and method 2 are basically the same, because they are consistent in principle. Although method 2 compares the acceleration after weighting and then converting it to the most sensitive frequency range of the human body, the evaluation results of these two methods are necessarily the same because the limits value of each frequency is equal to the limits value of the most sensitive frequency range after weighting.
(2) It is obvious that the evaluation method of full-domain vector sums is a bit more rigorous, and as far as method 3 is concerned, it agrees with the results of the first two evaluation methods in the most favorable case, i.e., when all vibration energy is concentrated in a single 1/3 octave frequency band.
(3) In methods 3 and 4, the fatigue- reduction efficiency limits for the front seats are the same at 15 km/h and 25 km/h and the exposure limits are opposite, this is because the vertical vibration plays a decisive role at 15 km/h but its energy is concentrated at 2–4 Hz, which is not in the vertical human sensitivity range, whereas the transversal vibration plays a decisive role at 25 km/h, and its energy is exactly concentrated in the transversal human. The vibration in the transverse direction plays a decisive role at 25 km/h, and its energy is concentrated in the transverse human sensitivity range of 1–2 Hz.

4 Conclusions

The following conclusions can be drawn from the analysis and evaluation of test data above:

(1) It is desirable to use the fatigue-ergonomics reduction limits and the exposure limits as the rubrics for evaluating the effects of whole-body vibration on health and operational ergonomics in the compartment of an armored vehicle, because they are quantitative values that are easy to compare and understand, and they are also useful as a guide for the development of a training plan at the testing range.
(2) The four methods from this paper for evaluating the effects of whole-body vibration on the human body are reasonable, mainly depending on the application of the strict degree of vibration environmental requirements, but also according to the needs of a certain range of limit values within a certain range of appropriate amendments to the standard, so further optimization of the evaluation method is needed.
(3) Method 2 is the calculation method recommended by GJB59.15 “Armored Vehicle Test Procedure Field Vibration Test”, but whether the method truly reflects the true feelings and operation efficiency of occupants in armored vehicle is still a question to be explored, because its evaluation method is too loose, and

the exposure limits of the experiment does not seem to be in line with the subjective feelings of the participants. In this case, if the fatigue-ergonomics reduction limit and exposure limit are insisted on as the evaluation results. Then the fundamental research on the impact of the vibration environment of armored vehicles on the ergonomics of personnel operations should be strengthened. The current standards should be constantly revised and improved to effectively control and protect the vibration design of armored vehicle cabin and reduce the human work efficiency impact of armored vehicle cabin.

References

1. Li SG, Zou Y, Xiao N, Pang ZB, He LH (2016) Study on development of human factors and ergonomics in military equipment. Chin Med Equip J 37:145–148
2. Bai SH, Zhang CL, Li XL (2004) Fuzzy evaluation of the tank crew discomfort in the vibration of the armored tracklayer. Mach Des Manuf 2:8–10
3. Li SG, Zhang P, Xiao N, Zhang ZG, Fan LD, Yin J, Li Y, Chen H, Zhou L (2013) Res Ergon Dev Military Med Equip 9:83–85
4. Wang HP, Pan HX, Huang JY, Xue H (2009) Research of the crew discomfort in the vibration of vehicle. Equip Manufactring Technol 2:28–30
5. Wu MQ, Chen HY, Yao ZZ, Cui B, She XJ (2006) Meas Eval Tank Vib 24:81–84
6. Gao C, Yang H, Su Y, Fu Q (2009) Acquisition and analysis of vibration signal in armored car. Transducer Microsyst Technol 28:37–39
7. An GQ, Liang YF, Jiang ZY, Bai JC (2021) Non-intrusive load identification based on CF-MF-SE joint feature. J Hebei Univ Sci Technol 42:462–469
8. Zhang ZF, Xu ZM, He YS, Wang XG, Xie DY (2009) Evaluation method of human body response to vibration. Noise Vib Control 2:149–155
9. Wang H, Kao XB, Yan JH, Kong W, Li Y, Chen YN, Lu Q (2013) Comfort analysis of occupant seats in crawler special vehicles. Occup Health 29:1894–1896
10. Chen YW, Chen Y, Wang LY, Wang ZX, Jia N, Chen QS (2023) Whole-body vibration exposure characterization in bus drivers of Haikou city: a comparison between two health risk assessment methods. J Environ Occup Med 40(11):12690013.0–12690013.0
11. Zhang LY, Zhu JJ, Li M, Tu JJ, Zhou M (2018) Study on ergonomics layout method of off-road vehicle mounted weapon station. Trans Beijing Inst Technol 38:49–54

Open Access This chapter is licensed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License (http://creativecommons.org/licenses/by-nc-nd/4.0/), which permits any noncommercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if you modified the licensed material. You do not have permission under this license to share adapted material derived from this chapter or parts of it.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.



Workspace of Parallel-Legged Hexapod Walking Robot



Xinxing Tang, Shuwei Bai, Huiqiu Mu, Dachuan Fan, and Fangyuan Han

Abstract A parallel-legged hexapod walking robot is designed based on the 6UPUR branched chain. The inverse velocity kinematics model is derived. The displacement equation of each branched chain is obtained through the leg velocity. Then, the limit working space of the parallel leg is obtained respectively by the boundary search method and the branched chain parameter combination method. Finally, the accuracy of the solution results of the limit working space is verified by the trajectory method.

Keywords Parallel-legged hexapod walking robot · Path planning · Inverse velocity kinematics · Inverse kinematics solution · Limit workspace

1 Intorduction

Compared with crawler and wheeled robots, legged robots have stronger adaptability to complex environments [1–4]. Multi-legged support can ensure the stability of the robot's walking. The multi-degree-of-freedom structure can also make the robot's operation more flexible, enhancing the robot's ability to adapt to complex terrain, so that it can complete walking tasks in most complex terrains. Precisely because the parallel-legged hexapod walking robot needs to face complex and changeable working environments, if only simple position control is carried out, when encountering a complex uneven ground environment, the impact of the robot's feet colliding with ground obstacles will generate extremely large impact forces, which can easily cause damage to the robot, accelerate the wear of parts, and greatly reduce the service life of the robot. When the robot is loaded with supplies or patients, the robot body will tilt or sway due to uneven ground, and both the stability of supplies and the comfort of patients will be greatly affected. Therefore, research of the parallel-legged hexapod

X. Tang · S. Bai · D. Fan · F. Han (✉)
School of Mechatronic Engineering, Changchun University of Technology, Changchun, China
e-mail: Hanfangyuan@sina.com.cn

H. Mu
Automotive Chassis Department, BYD Auto Industry Company Limited, Shenzhen, China

© The Author(s) 2026

H. Saman et al. (eds.), *The 9th International Conference on Advances in Construction Machinery and Vehicle Engineering*, Lecture Notes in Mechanical Engineering,
https://doi.org/10.1007/978-981-96-5527-4_108

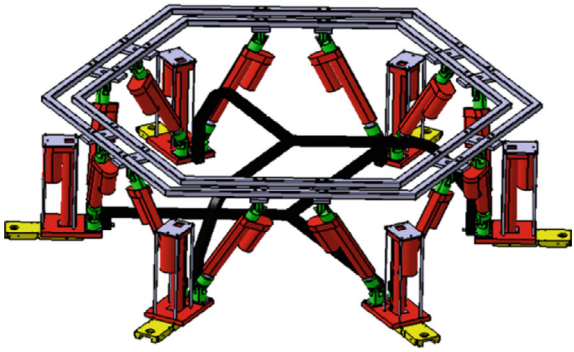


Fig. 1 3D Model of parallel-legged hexapod walking robot

walking robot is conducive to the promotion and application of robots and can also provide a reference for other new types of robots to carry out related research work.

2 The Structure Design and Leg Workspace Analysis

2.1 Structure Design

The main structure of the robot comprises an upper platform, two Y-shaped foot arches, and twelve UPUR-structured branches. The upper platform, which is a regular hexagon, is connected to the upper ends of the twelve branches. Every two adjacent branches constitute a pair, and their lower ends are joined to a single foot, thus creating a total of six feet. Three spaced feet are interconnected via Y-shaped foot arches to form the lower platform. The upper platform, the lower platform, and the six branches linking the two platforms constitute a 6-UPUR parallel mechanism, representing one leg of the robot. The robot has two legs, sharing one upper platform, and each leg has three feet. To enhance the robot's capability to overcome obstacles and adapt to rough terrain, each foot is furnished with an auxiliary leg. These extensible auxiliary legs can modify the working space of the robot's leg parallel mechanism, thereby increasing the flexibility of the robot's legs. The 3D model of the parallel-legged hexapod walking robot is presented in Fig. 1, and the advance direction is indicated in Fig. 1.

2.2 Inverse Velocity Kinematics

In order for a multi-legged walking robot to successfully walk in a rugged and complex environment, detailed planning and control of the movement of the multi-legged walking robot are still needed. In a rugged and complex terrain environment,

the multi-legged walking robot needs to continuously adjust the position and posture of its body center of gravity to ensure the stability of the robot and have an effective movement space at the same time, so that the robot can successfully pass through rugged terrain [5]. Therefore, precise control of the position and posture of the multi-legged walking robot is particularly important. The inverse kinematics of single-leg velocity of the robot is studied based on the parallel-leg hexapod walking robot platform. On this basis, the inverse velocity kinematics model of velocity of the parallel-leg hexapod walking robot is used to establish and realize the closed-loop control of the position and posture of the hexapod walking robot [6].

Define the generalized coordinate vector $q = [x, y, z, \varphi, \theta, \emptyset]^T$, where x, y, z represents the position of the origin of coordinate system relative to the world coordinate system, respectively, and $\varphi, \theta, \emptyset$ represents the spatial attitude of the origin of the moving coordinate system in the world coordinate system, respectively. According to the planned path, given the expected pose $q_d = [x_d, y_d, z_d, \varphi_d, \theta_d, \emptyset_d]^T$ and expected leg movement velocity $\dot{q}_d = [v_x, v_y, v_z, \omega_\varphi, \omega_\theta, \omega_\emptyset]^T$, and then obtaining the current pose $q_c = [x_c, y_c, z_c, \varphi_c, \theta_c, \emptyset_c]^T$, The linear velocities of the branch chains of each electric cylinder can be calculated, and then the elongations of the six branch chains of the six electric cylinders of the platform can be obtained. Thus, there is a kinematic relationship between the velocity or displacement of each electric cylinder and the linear velocity of the moving platform. This relationship can be described by Jacobian matrix as shown by Eq. (1).

$$l_i = \int \dot{l}_i = J\dot{q} \tag{1}$$

where, $\dot{l}_i$ $(i = 1, 2, 3, 4, 5, 6)$ is the velocity vector of each electric cylinder, J is the Jacobian matrix, and $\dot{q}$ represents the velocity of the moving platform for the parallel parallel-legged hexapod walking robot.

2.3 Parallel-Legged Hexapod Walking Robot Workspace

The size of the workspace of a parallel mechanism is an important indicator for measuring its working performance [7]. The analysis and solution of the workspace mainly refer to determining the activity range of the output component in space under the conditions of knowing the size parameters of the parallel mechanism, the interference and constraints of the branched chains. For robots, the larger the workspace, the wider the reachable activity space of the output component, and the stronger its working ability. Therefore, under certain overall size constraints of the robot, people always hope that its workspace can be as large as possible. Compared with serial mechanisms, due to the parallel mechanism itself being a closed-loop structure and the mutual interference between components, its workspace is relatively small and even singular phenomena may occur. Therefore, in the process of developing parallel-legged walking robots, it is necessary to analyze and solve its

workspace to avoid mechanism damage caused by direct experimentation or put into use without research. The research on the workspace of parallel mechanisms aims to clarify the size and boundary of its workspace, so as to better make a precise and reasonable evaluation of the performance of the robot, which is of great significance for the development of parallel-legged walking robots.

Limit Workspace Analysis. When solving the limit workspace of a parallel mechanism, it is usually necessary to first find all combinations of possible values of the branch chain parameters of the mechanism, and then use the algorithm of the forward position solution of the parallel mechanism to obtain the position set of the reference point of the moving platform, that is, the limit workspace.

The calculation process of the branch chain parameter combination method is as follows:

(1) Supported that the initial value $l_1 = l_2 = l_3 = l_4 = l_5 = l_6 = l_{\min}$.
(2) Determine the search step size Δl, $\Delta l = (l_{\max} - l_{\min})/n$.
(3) Starting from $l_1 = l_2 = l_3 = l_4 = l_5 = l_6 = l_{\min}$, gradually increase the value of l_i ($i = 1, 2, 3, 4, 5, 6$) until all combinations of values that can be taken by the six branched chains are found.
(4) Input all combinations of values that can be taken by the six branched chains as input values into the forward solution algorithm program of the parallel mechanism in turn. All the output values obtained are the set of positions that the reference point of the moving platform can reach.

According to the structural parameters of the parallel-legged hexapod walking robot and combined with the solution method of the forward kinematics position of the parallel mechanism, the above process is programmed and solved by using MATLAB software. The obtained limit workspace of the parallel-legged hexapod walking robot is shown in Fig. 2.

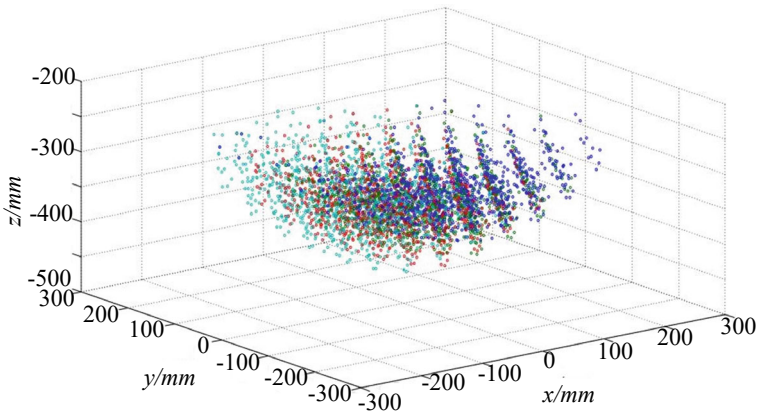
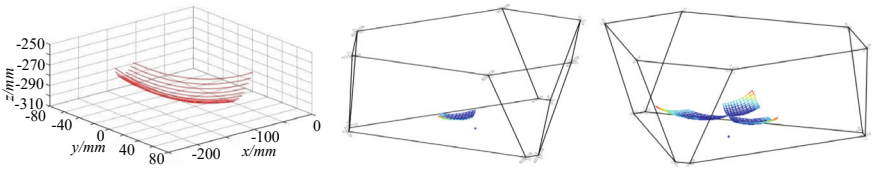


Fig. 2 3D limit workspace of the leg for the parallel-legged hexapod walking

Verification of limit workspace by trajectory method. The trajectory method is a simple and intuitive method for solving the workspace of parallel mechanisms with few degrees of freedom. After constraining some branched chains of the parallel mechanism (taking the maximum or minimum value), the lengths of the remaining branched chains are increased or decreased in turn. The reference point of the moving platform of the mechanism will move according to the specified path. Taking the length parameters of these branched chains as input values and inputting them into the forward kinematics position solution MATLAB program, the movement path of the reference point of the moving platform will be displayed as a series of smooth curves in its three-dimensional movement space. By constraining the branched chains of the parallel mechanism in different ways according to certain rules, multiple sets of path trajectory curves can be obtained. Fitting these sets of curves into multiple surfaces, and by observation, taking the outermost surface in space, the complete closed surface of the limit workspace can be obtained. The advantages of this method are high visualization and good intuitiveness. The disadvantage is that once the degree of freedom of the calculated parallel mechanism exceeds 5, finding the values of the branched chain lengths of the parallel mechanism at all limit positions is a very complex process. Moreover, every time an additional degree of freedom is added, the number of generated surfaces will greatly increase. It is not easy to screen out the outermost surface in space from the mutually enveloping and intersecting surfaces by observation. This is also the difficulty of this method. Therefore, when the trajectory method is selected to solve the 6-UPUR parallel mechanism with 6 degrees of freedom, only some surfaces located at the edge of the limit workspace of the parallel mechanism are calculated to verify the accuracy of the limit workspace solved in the previous text [8, 9]. The specific steps for finding the workspace of the 6-UPUR parallel mechanism by the trajectory method are as follows:

Supported that $l_{min} = 325$, $l_{max} = 475$, and $\Delta l = (l_{max} - l_{min})/n$. Let $l_1 = l_3 = l_4 = l_6 = l_{min}$, $l_2 = l_{min} + i\Delta l$ ($i = 0, 1,\ldots, n$), and $l_5 = l_{min} + j\Delta l$ ($j = 0, 1,\ldots, n$). When j is equal to 0, increase i from 0 to n. The obtained $n + 1$ sets of branched chain parameters are input into the forward kinematics position positive solution equation for solving. The reference point of the moving platform will generate $n + 1$ coordinates. Record the coordinate data in the form of smooth curves in the three-dimensional space where the parallel mechanism moves. Then let $j = 1$. Similarly, increase i from 0 to n, and another smooth curve can be obtained…. When $j = n$, $n + 1$ working curves can be obtained. Fit these curves into a working surface, as shown in Fig. 3. Among them, Fig. 3a is the trajectory curve of the reference point when the moving platform moves, and Fig. 3b is the surface obtained by fitting these trajectory curves. According to the structural characteristics of the 6-UPUR parallel mechanism, the obtained surface is rotated by 120° around the Z-axis for arraying, and three surfaces as shown in Fig. 3c can be obtained.

Similarly, let $l_1 = l_3 = l_4 = l_6 = l_{\max}$, $l_2 = l_{min} + i\Delta l$ ($i = 0, 1,\ldots, n$), and $l_5 = l_{min} + j\Delta l$ ($j = 0, 1,\ldots, n$). When j is equal to 0, increase i from 0 to n. The obtained $n + 1$sets of branched chain parameters are input into the forward kinematics position positive solution equation for solving. The reference point of the moving platform will generate n + 1 coordinates. Record the coordinate data in the form of smooth



(a)Trajectory curve of the reference point of the moving platform

(b)Surface fitted by trajectory curves

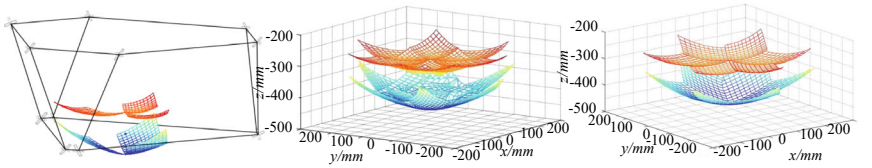
(c)Surfaces after arraying

Fig. 3 Workspace determination of 6-UPUR parallel mechanism using trajectory method

curves in the three-dimensional space where the parallel mechanism moves. Then let $j = 1$. Similarly, increase i from 0 to n, and another smooth curve can be obtained…. When $j = n$, n working curves can be obtained. Fit these curves into a working surface, and rotate the surface by 120° around the Z-axis for arraying. The newly obtained three surfaces are shown in Fig. 4a.

Similarly, by altering the constraints on different branched chains, multiple sets of trajectory curves of the reference points of the moving platform of the branched chains of the parallel mechanism at extreme positions are obtained. These trajectory curves are fitted into several curved surfaces, as depicted in Fig. 4b. By using the observation method, the outermost curved surfaces in space are selected, as shown in Fig. 4c. These curved surfaces constitute a part of the outer surface of the limit working space of the 6-UPUR parallel mechanism.

As shown in Fig. 5, by importing the outermost surface obtained by the trajectory method and the limit workspace obtained by the input parameter combination method into the same three-dimensional coordinate system, it can be seen that the position of the outermost surface obtained by the trajectory method and the edge of the limit workspace solved by the input parameter combination method are completely consistent in the three-dimensional coordinate system.



(a)Portion of the upper and lower surfaces of the extreme workspace

(b) set of surfaces fitted from trajectory curves

(c) The outermost curved surface located on the outermost side

Fig. 4 Verification of maximum workspace via trajectory method

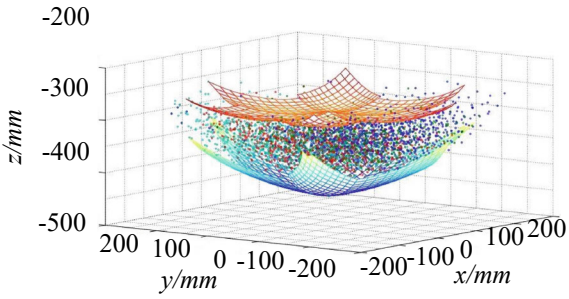


Fig. 5 The limit workspaces obtained by the two methods

3 Conclusions

The structure of a parallel-legged hexapod walking robot is designed. Based on the elongation obtained based on the inverse velocity kinematics model, the limit working space of the parallel legs is solved by the branched chain parameter combination method. Finally, part of the outer surface of the limit working space of the parallel mechanism is fitted by the trajectory method, and the obtained limit working space results are verified, which proves the accuracy of the limit working space solution results. In the future, based on the existing robot structure and workspace, in the future, closed-loop control of pose will be realized by combining path planning and inverse velocity kinematics to obtain compliant control of parallel-legged walking robots in unstructured environments.

Acknowledgements This research work was supported by the Science and Technology Development Project of Jilin Province (No. 20230101332JC).

References

1. Sugahara Y (2007) Experimental stiffness measurement of WL-16RII biped walking vehicle during walking operation. J Robot Mechatron 19(3):272–280
2. Plasberg C, Roennau A, Dillmann R (2022) Adaptation of walking-patterns to damage for the six-legged walking robot LAURON V. IEEE Trans Med Robot Bionics 4(3):550–557
3. Xu X, Qi C, Han L et al (2024) Multiple terrain traversal capabilities based mechanism dimension design for a six-legged robot using performance charts from analytical conditions. Mech Mach Theory 201:105731
4. Zhuang H, Wang N, Gao H et al (2023) Power consumption characteristics research on mobile system of electrically driven large-load-ratio six-legged robot. Chin J Mech Eng 36:26
5. Liu D, Wang J, Lei T, Wang S (2022) Active suspension control with consensus strategy for dynamic posture tracking of wheel-legged robotic systems on uneven surfaces. ISA Trans 131:628–638
6. Chen ZH, Li JH, Wang SK, Wang JZ, Ma LL (2022) Flexible gait transition for six wheel-legged robot with unstructured terrains. Robot Auton Syst 150:103989
7. Shivakumar MR, Srikrishna R, Sriram R, SreeSailesh et al (2022) Workspace analysis and optimal design of 3-PRR planar parallel manipulators. J Chin Soc Mech Eng 43(6):579–591

8. Fan D (2019) Kinematics analysis and gait design of a parallel leg hexapod walking robot. Changchun University of Technology, Changchun, Jilin Province
9. Han F (2012) A connection method for a parallel walking robot and its parallel walking robot. CN101973027B

Open Access This chapter is licensed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License (http://creativecommons.org/licenses/by-nc-nd/4.0/), which permits any noncommercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if you modified the licensed material. You do not have permission under this license to share adapted material derived from this chapter or parts of it.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.



Trajectory Tracking Control of Wind Turbine Blade Mountain Transporter Based on MPC with Disturbance Observer



Yahui Zhang, Yuhao Jiang, Xiaohong Jiao, Ting Zhang, Chengye Wu, Mingyang Wang, Yiying Wang, and Yuchen Wang

Abstract The trajectory tracking control problem for Wind Turbine Blade Mountain Transporter (WTBMT) is crucial for enhancing transportation efficiency and ensuring safety. These vehicles often navigate unstructured mountain roads, where they are significantly influenced by varying road conditions and external factors. Such influences can lead to perturbations in model parameters, adversely impacting the tracking performance of the controller. To address these challenges, this study proposes a Model Predictive Control (MPC) trajectory tracking method enhanced by a disturbance observer. The proposed controller consists of two main components: disturbance compensation input and optimal MPC control input. A disturbance observer is designed to detect and compensate for disturbances at the control input, while the MPC controller manages the trajectory tracking. The effectiveness of this method has been validated through simulations on the TruckSim and MATLAB/Simulink platforms.

Keywords Wind turbine blade mountain transporter · Semi-trailer tractor · Trajectory tracking control · MPC · Disturbance Observer

1 Introduction

The WTBMT plays a vital role in the deployment of wind power generation facilities in rugged terrains. These vehicles are specifically designed to transport wind turbine blades from production workshops to installation sites in mountainous areas. Due to the large size and significant weight of the blades, the transport chassis is

Y. Zhang · Y. Jiang · X. Jiao (✉) · C. Wu · M. Wang · Y. Wang · Y. Wang
Yanshan University, Qinhuangdao, Hebei Province, China
e-mail: jiaoxh@ysu.edu.cn

T. Zhang
Tsinghua University, Beijing, China

© The Author(s) 2026

H. Saman et al. (eds.), *The 9th International Conference on Advances in Construction Machinery and Vehicle Engineering*, Lecture Notes in Mechanical Engineering,
https://doi.org/10.1007/978-981-96-5527-4_109

highly susceptible to external factors such as uneven road surfaces, crosswinds, and changes in slope. These conditions can lead to vibrations and even the risk of rollover. Therefore, studying the trajectory tracking problem for these transport vehicles in mountainous operations is essential for ensuring safe and efficient transportation [1]. The transport chassis of the WTBMT is a semi-trailer tractor, which provides valuable insights into the trajectory tracking control problem associated with such vehicles. As a widely utilized controller in vehicle control applications [2–4], Model Predictive Control (MPC) has been extensively applied to the trajectory tracking control of semi-trailer tractors. Felez et al. [5] proposed a conceptual model for a fully electrical vehicle with autonomous driving, showing promising results in simulations. Li et al. [6] developed a reversing trajectory tracking algorithm for semi-trailer trucks with a two-layer path following controller based on a kinematics model. Liu et al. [7] presented a regional tractor semi-trailer path tracking strategy based on MPC, optimizing the front wheel angle of the tractor. Kumar et al. [8] proposed a lateral MPC algorithm for truck-trailers, linearizing system dynamics around a nominal trajectory. Chen et al. [9] introduced a parametric self-learning MPC method for lateral control of autonomous semi-trailer trucks. Dirckx et al. [10] presented ASAP-MPC, an asynchronous update scheme for online motion planning with MPC, ensuring real-time execution and smooth trajectory merging. While the aforementioned research demonstrates the effectiveness of the MPC controller in semi-trailer tractor trajectory tracking control, it does not adequately address the impact of external environmental interference on the controller's performance. Thus, in response to the unique operational scenario of the WTBMT, this paper proposes an MPC trajectory tracking control algorithm incorporating a disturbance observer, aiming to achieve high-precision trajectory tracking control of the vehicle under unstructured road conditions.

2 Vehicle Dynamics Model

The trajectory tracking problem of the WTBMT can be simplified to the trajectory tracking problem of its transport chassis—semi-trailer tractor. It consists of two parts: a tractor and a trailer, which are connected by an articulated mechanism. The model of a common off-axle semi-trailer tractor is shown in Fig. 1.

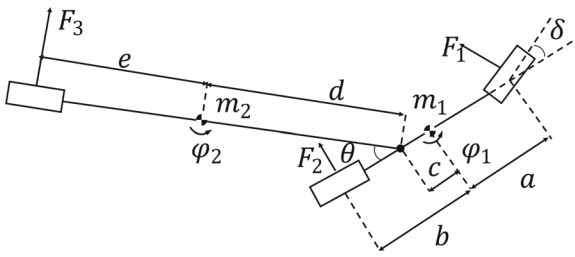


Fig. 1 Vehicle dynamics model

According to the dynamic model of the semi-trailer tractor and Newton's second law of motion, the dynamic equation of the semi-trailer tractor is derived, which is:

$$\begin{cases} m_1 v_{x1}\left(\dot{\varphi}_1 + \dot{\beta}_1\right) = F_1 \cos\delta + F_2 + F_4 \\ I_{ZZ1}\ddot{\varphi}_1 = F_1 a \cos\delta - F_2 b - F_4 c \\ m_2 v_{x2}\left(\dot{\varphi}_2 + \dot{\beta}_2\right) = F_3 - F_4 \cos\theta \\ I_{ZZ2}\ddot{\varphi}_2 = -F_4 d \cos\theta - F_3 e \\ \dot{x}_1 = v_{x1}\cos\beta_1 - v_{x1}\beta_1 \sin\varphi_1 \\ \dot{y}_1 = v_{x1}\sin\beta_1 + v_{x1}\beta_1 \cos\varphi_1 \\ \dot{x}_2 = v_{x2}\cos\beta_2 - v_{x2}\beta_2 \sin\varphi_2 \\ \dot{y}_2 = v_{x2}\sin\beta_2 + v_{x2}\beta_2 \cos\varphi_2 \end{cases} \tag{1}$$

The coupling relationship between the tractor and the trailer is:

$$v_{x2} = v_{x1}\cos\theta - (\beta_1 v_{x1} - \dot{\varphi}_1 b)\sin\theta \tag{2}$$

When the hinge angle θ is small, we have:

$$\dot{\beta}_1 - \dot{\beta}_2 + \varphi_1 - \varphi_2 - \tfrac{c}{v_{x1}}\dot{\varphi}_1 - \tfrac{d}{v_{x2}}\dot{\varphi}_2 = 0 \tag{3}$$

Assuming that the tires always work in the linear range, the lateral force F_1 of the front axle of the tractor, the lateral force F_2 of the rear axle, and the lateral force F_3 of the rear axle of the trailer can be given by the following formula:

$$\begin{cases} F_1 = k_1\left(\beta_1 + \dfrac{a\dot{\varphi}_1}{v_{x1}} - \delta\right) \\ F_2 = k_2\left(\beta_1 - \dfrac{b\dot{\varphi}_1}{v_{x1}}\right) \\ F_3 = k_3\left(\beta_2 - \dfrac{e\dot{\varphi}_2}{v_{x2}}\right) \end{cases} \tag{4}$$

According to formulas (1) to (4), the state space equation of the semi-trailer tractor can be expressed as:

$$\begin{cases} \dot{X} = AX + BU \\ \dot{Y} = CX \end{cases} \tag{5}$$

where $X = [\,\beta_1\ \dot{\varphi}_1\ \beta_2\ \dot{\varphi}_2\ x_1\ y_1\ \varphi_1\ \varphi_2\,]^T$, $A = \begin{bmatrix} A11_{4\times4} & 0_{4\times4} \\ A21_{4\times4} & A22_{4\times4} \end{bmatrix}$, $B = \begin{bmatrix} B1_{4\times1} \\ 0_{4\times1} \end{bmatrix}$, $C = \begin{bmatrix} 0\,0\,0\,0\,0\,1\,0\,0 \end{bmatrix}$, $U = \delta$. For the meaning of parameters A_{11}, A_{21}, A_{22}, B_1, O and other symbols, please refer to the Appendix.

3 MPC Controller Design Based on Disturbance Observer

This paper employs an MPC controller to address the trajectory tracking problem of a WTBMT. To mitigate the impact of external factors on the prediction model parameters in unstructured road conditions, we designed a disturbance observer at the input of the MPC controller. This observer estimates and compensates for disturbances in real time, enabling effective anti-disturbance control for the vehicle. Figure 2 illustrates the structure of the designed controller.

3.1 MPC Trajectory Tracking Controller Design

We consider the discrete state space equation of the controlled system as:

$$\begin{cases} \xi(k+1|t) = \tilde{A}_{k,t}\xi(k|t) + \tilde{B}_{k,t}\Delta U(k|t) \\ \eta(k|t) = \tilde{C}_{k,t}\xi(k|t) \end{cases} \tag{6}$$

where $\xi(k|t) = \begin{bmatrix} \tilde{X}(k|t) \\ \tilde{U}(k|t) \end{bmatrix}$, $\tilde{A}_{k,t} = \begin{bmatrix} A_k & B_k \\ 0_{m\times n} & I_m \end{bmatrix}$, $\tilde{B}_{k,t} = \begin{bmatrix} B_k \\ I_m \end{bmatrix}$, $\tilde{C}_{k,t} = \begin{bmatrix} I_m & 0_{m\times n} \end{bmatrix}$, $A_k = A\tau + I_m$, $B_k = B\tau$, $\tilde{X}(k|t)$ and $\tilde{U}(k|t)$ represents the state and control input of the system at the current time t after discretization, I_m is the unit matrix, m and n is the dimension of the control quantity.

Considering the complex working conditions, when the control quantity and control increment are constrained, the MPC optimization problem may not have a feasible solution. The weight coefficient ρ and relaxation factor ε are introduced to set the objective function:

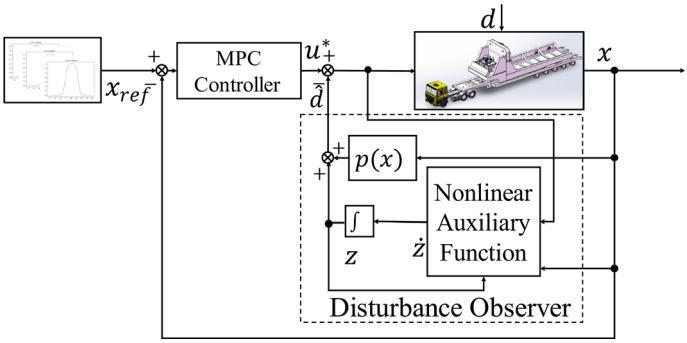


Fig. 2 Trajectory tracking controller design block diagram

$$J(k) = \sum_{i=1}^{N_p} \|\eta(k+i|t) - \eta_{ref}(k+i|t)\|_Q^2 + \sum_{i=1}^{N_{c-1}} \|\Delta U(k+i|t)\|_R^2 + \rho\varepsilon^2 \tag{7a}$$

$$s.t. U_{min}(t+k) \le U(t+k) \le U_{max}(t+k) \tag{7b}$$

$$\Delta U_{min}(t+k) \le \Delta U(t+k) \le \Delta U_{max}(t+k) \tag{7c}$$

where Q and R are the weight matrices of the system state and control variables, respectively.

3.2 Disturbance Observer Design

We redefine the controlled system as:

$$\begin{cases} \dot{X} = AX + B(U+d) \\ \dot{Y} = CX \end{cases} \tag{8}$$

where d represents the disturbance. Then we can design a disturbance observer for the above system as:

$$\begin{cases} \dot{Z} = -LB \cdot Z - L(B \cdot P + AX + BU) \\ \hat{d} = Z + P \end{cases} \tag{9}$$

where Z represents the internal state of the observer, P is the function we need to design, and the gain L of the observer is:

$$L = \frac{\partial P}{\partial X} \tag{10}$$

The estimated error is defined as: $e_d = d - \hat{d}$. This paper assumes that the disturbance is slow-varying compared to the dynamics of the observer, that is, $\left\|\dot{d}\right\| < \epsilon$, $\epsilon \approx 0$ is a given constant. Consider the dynamics of the estimated error:

$$\dot{e}_d = \dot{d} - \dot{\hat{d}} = -LB \cdot e_d \tag{11}$$

By designing a suitable gain L so that − LB satisfies the Hurwitz matrix, the convergence of the disturbance observer error can be guaranteed, and the designed disturbance observer is stable.

4 Simulation and Results

To verify the effectiveness of the designed controller, a vehicle model was established in TruckSim, and a simulated unstructured road scene was created. A joint simulation test was conducted using MATLAB/Simulink. The parameters of the semi-trailer tractor used in the simulation test are detailed in Table 1.

During the simulation, the vehicle speed was set to 10 m/s, and a reference trajectory of about 1.5 km was constructed based on the real road data of the TruckSim example. In order to better simulate the operation of unstructured roads, different elevation profiles were set on the left and right sides of the road centerline to produce uneven road surfaces. The road parameter information is shown in Fig. 3.

The prediction step of the MPC controller is selected as 15, the control step is 5, the relaxation factor is 10, the observer gain L is a constant that changes with time, the simulation step is 0.1, and the simulation time is 150 s. At the same time, in order to verify the effectiveness of the algorithm, a comparative experiment with MPC is carried out. The experimental results are shown in Figs. 4 and 5.

The green line in the figure represents the reference value, the blue line represents the control algorithm proposed in this article, and the orange line represents the MPC control algorithm. We observe that compared with MPC, the tracking error of the

Table 1 Vehicle parameters

Parameter	Value	Unit	Parameter	Value	Unit	Parameter	Value	Unit
Tractor wheelbase	3900	mm	I_{zz1}	34,802.6	Kg m^2	c	1890	mm
Trailer wheelbase	8500	mm	I_{zz2}	179,992	Kg m^2	d	5220	mm
a	1110	mm	m_1	4455	Kg	e	3280	mm
b	2790	mm	m_2	6000	Kg	$k_1k_2k_3$	300,000	N/rad

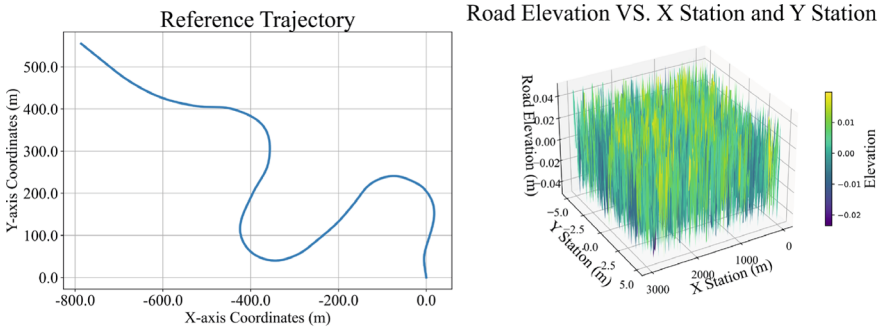


Fig. 3 Road parameter information

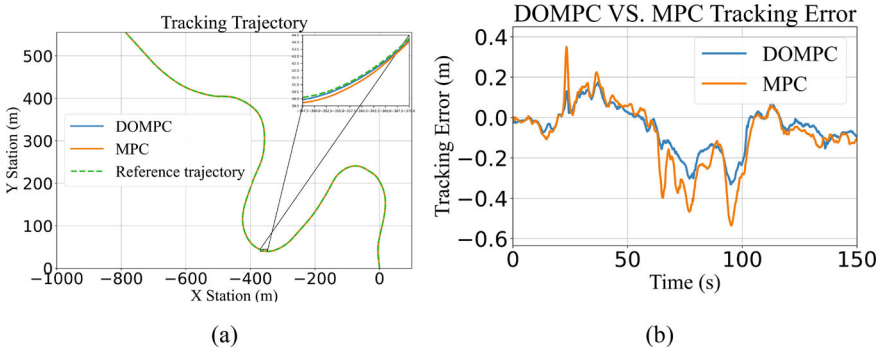


Fig. 4 Schematic diagram of trajectory tracking effect

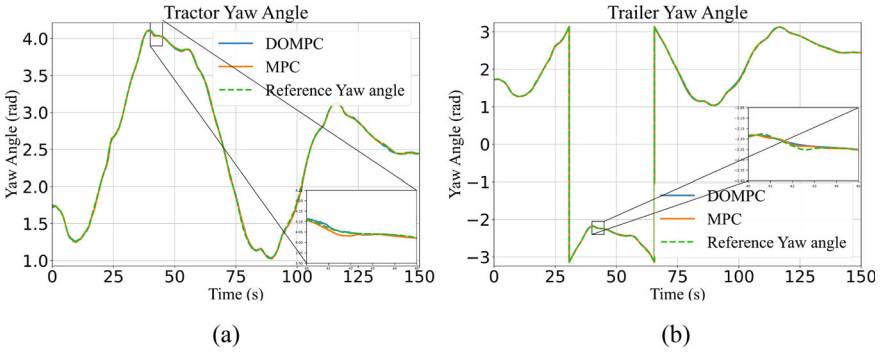


Fig. 5 Yaw angle tracking effect

control algorithm proposed in this article is significantly reduced. The maximum tracking error of MPC is 0.53 m and the mean square error is 0.031. The maximum tracking error proposed in this article is 0.33 m, and the mean square error is 0.014. In summary, in the unstructured road operation scenario, the control algorithm proposed in this article can effectively suppress the impact of external interference on the system and improve the control accuracy of trajectory tracking.

5 Conclusion

This paper presents a trajectory tracking control method for WTBMT. We begin by developing a vehicle dynamics model, which serves as the foundation for formulating the associated optimization problem. The controller design is structured into two primary components. First, we address the trajectory tracking control issue

using nominal Model Predictive Control (MPC). To account for disturbances encountered during vehicle operation, we design a disturbance observer that monitors and compensates for these disturbances at the MPC control input. Simulation results demonstrate a significant improvement in trajectory tracking accuracy when external disturbances are bounded. Future research will focus on investigating the coupling mechanism between the upper loading mechanism and the transport chassis, as well as its effects on the performance of the trajectory tracking controller.

Appendix

For some parameters mentioned in the article but without clear values, please refer to the Baidu network disk link:
https://pan.baidu.com/s/1T51J_GfriyRDuyu9N0FlfQ?pwd=nbyt
Extraction code: nbyt.

References

1. He Y, Yu R (2023) Current status and prospects of road transportation technology for large cargo. Prospect Sci Technol 2(3):106–117
2. Mata S, Zubizarreta A, Pinto C (2019) Robust tube-based model predictive control for lateral path tracking. IEEE Trans Intell Veh 4(4):569–577
3. Cheng S, Li L, Chen X, Wu J, Wang H (2021) Model-predictive-control-based path tracking controller of autonomous vehicle considering parametric uncertainties and velocity-varying. IEEE Trans Industr Electron 68(9):8698–8707
4. Meng Q, Qian C, Chen K, Sun Z-Y, Liu R, Kang Z (2024) Variable step MPC trajectory tracking control method for intelligent vehicle. Nonlinear Dyn
5. Felez J, GarcíA-Sanchez C, Lozano JA (2018) Control design for an articulated truck with autonomous driving in an electrified highway. IEEE Access 6:60171–60186
6. Li Z, Cheng H, Ma J, Zhou H (2020) Research on parking control of semi-trailer truck. In: 2020 4th CAA international conference on vehicular control and intelligence (CVCI), pp 424–429
7. Liu J, Zhao X, Guo H, Chen H (2020) Tractor semi-trailer modeling and regional path tracking control strategy. In: 2020 Chinese automation congress (CAC), pp 1961–1966
8. Kumar M, Haas A, Strauss P, Kraus S, Taş ÖŞ, Stiller C (2022) Conception and experimental validation of a model predictive control (MPC) for lateral control of a truck-trailer. In: 2022 IEEE intelligent vehicles symposium (IV), pp 1550–1557
9. Chen J, Jiang S, Zhou Z, Zhang M, Ming X, Guo N (2024) Lateral semi-trailer truck control using a parameter self-learning MPC method in urban environment. Proc Inst Mech Eng Part D: J Automob Eng 238(5):964–976
10. Dirckx D, Bos M, Vandewal B, Vanroye L, Decré W, Swevers J (2024) ASAP-MPC: an asynchronous update scheme for online motion planning with nonlinear model predictive control

Open Access This chapter is licensed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License (http://creativecommons.org/licenses/by-nc-nd/4.0/), which permits any noncommercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if you modified the licensed material. You do not have permission under this license to share adapted material derived from this chapter or parts of it.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.



Adaptive Control for Vehicle Active Suspension with Limited Working Space Based on Tire Stress Estimation



Zhenjun Lin, Shuang Liu, Mingxuan Li, Dingxuan Zhao, Nayu Su, and Cong Zhang

Abstract In this paper, a backstepping adaptive neural network (BANNC) active suspension control strategy is proposed to improve vehicle ride comfort. The concept of differential isomorphism is introduced, the state space equations containing body displacement and suspension dynamic travel are constructed, and the obstacle Lyapunov function is utilized to limit the range of suspension dynamic travel. Aiming at the problem that the tire stress is complex and difficult to calculate, the neural network function is used to estimate the nonlinear stress of the tire, and combined with the adaptive control algorithm, so that the control strategy of the system can still maintain good control effect when the tire stress changes with the system state. Finally, the effectiveness of the algorithm in this paper is proved by joint simulation. The BANNC algorithm in this paper shows an improvement of approximately 14% in spring deflection and acceleration compared to the passive suspension, and an improvement of around 22% in the dynamic/static load ratio.

Keywords Vehicle active suspension · Barrier Lyapunov function · Radial basis function neural network · Adaptive control strategy

1 Introduction

Excessive vibration of a vehicle during operation often brings many adverse effects, such as cargo damage and passenger discomfort. Active suspension, on the other hand, can maximize its efficiency by adjusting the output force to change the vehicle attitude. Therefore, active suspensions have attracted extensive academic attention [1, 2].

Z. Lin · S. Liu (✉) · M. Li · N. Su · C. Zhang
School of Electrical Engineering, Yanshan University, Qinhuangdao, China
e-mail: liush@ysu.edu.cn

D. Zhao
School of Mechanical Engineering, Yanshan University, Qinhuangdao, China

© The Author(s) 2026

H. Saman et al. (eds.), *The 9th International Conference on Advances in Construction Machinery and Vehicle Engineering*, Lecture Notes in Mechanical Engineering,
https://doi.org/10.1007/978-981-96-5527-4_110

At present, the research on active suspension system is becoming more and more mature, and a variety of control strategies have emerged. Reference [3] effectively addresses the issues present in non-fragile H∞ static output feedback within active suspension control by employing finite frequency constraints. This approach significantly enhances the ride comfort of vehicles, improving the smoothness of vehicle motion. In [4], an active seat suspension system that is robust to uncertainties and unknown road profiles is designed using a state and disturbance observer. Although the above literature have a significant effect on improving vehicle ride comfort, the optimization effect of suspension dynamic deflection are not obvious.

Considering the dynamic deflection of suspension is limited in practice, the solution in structural is usually to add limit block; In terms of system, methods such as setting threshold and introducing filter can be adopted. In reference [5], the adaptive backstepping method is employed to simultaneously improve the ride comfort of vehicles and reduce the suspension's dynamic deflection. On the other hand, reference [6] proposes an adaptive output feed-back control strategy with backstepping command filtering to address full-state constraints in nonlinear systems. When the input delay is less than the upper bound of the delay, the suspension dynamic deflection will be limited to the pre-set range. For the problem, a state equation containing both sprung mass displacement and suspension dynamic deflection is designed, and the barrier Lyapunov function is used to set the threshold for suspension dynamic deflection, so as to optimize the vehicle ride comfort and ensure that the suspension dynamic deflection is stable in a safe range.

Tire is also an important factor affecting vehicle ride comfort. During driving, the stress of tire is affected by many factors, which makes it difficult to be calculated, so that the expected effect of active suspension is often not achieved in practice. Reference [7] establishes a non-steady-state, nonlinear tire model, which provides a favorable platform for simulation and control under complex wheel motion inputs. In [8], an observer-based traction force control scheme that is robust with respect to variations of road conditions and tire model uncertainties is proposed. In this paper, the tire stress is regarded as the nonlinear force related to the system states and road inputs, and the radial basis function neural network is used to estimate it. By using the adaptive algorithm, the actuator control force can be adjusted according to the system states and road profile, which ensures that the suspension can maintain the best condition. The contributions of this paper can be summarized as follows:

(1) By utilizing barrier Lyapunov function and backstepping control algorithm, it improves the vehicle's ride comfort while ensuring limited suspension dynamic deflection.
(2) This paper treats tire forces as nonlinear forces that depend on system states and road excitation signals. A radial basis neural network function is employed to estimate the tire forces.

2 Establishment of Active Suspension System

2.1 Establishment of Active Suspension Model

By the Newton's second law, the dynamical model of the suspension system is given as follows:

$$\begin{cases} m_s\ddot{x}_s = k_s(x_s - x_u) + c_s(\dot{x}_s - \dot{x}_u) - u \\ m_u\ddot{x}_u = -k_s(x_s - x_u) - c_s(\dot{x}_s - \dot{x}_u) + f + u \end{cases} \tag{1}$$

where, m_s and m_u represent the sprung mass and unsprung mass, x_s represents the sprung mass displacement, x_u represents the unsprung mass displacement, x_r represents the road input, k_s and c_s are the spring stiffness and the sprung mass damping coefficient, f represents the nonlinear stress of tire and u is the actuator control force.

2.2 Establishment of State Space Expression

Let $x_1 = x_s, x_2 = \dot{x}_s, x_3 = x_u - x_s, x_4 = \dot{x}_u$ be the state variables of the system and $y = x_s$ be the output.

From the suspension system (1) and state variables, it is easy to derive the following state space model:

$$\begin{cases} x_1 = x_2 \\ x_2 = \dfrac{k_s}{m_s}x_3 + \dfrac{c_s}{m_s}(x_4 - x_2) + \dfrac{u}{m_s} \\ x_3 = x_4 - x_2 \\ x_4 = -\dfrac{k_s}{m_u}x_3 - \dfrac{c_s}{m_u}(x_4 - x_2) + \dfrac{f}{m_u} - \dfrac{u}{m_u} \end{cases} \tag{2}$$

However, due to the control quantity appears in the iterative process, backstepping cannot be directly used, so that the diffeomorphism is proposed to transform the expression, the specific form is as follows:

$$\begin{cases} z_1 = \dfrac{m_s + m_u}{m_u}x_1 + x_3 \\ z_2 = \dfrac{m_s}{m_u}x_2 + x_4 \\ z_3 = x_3 \\ z_4 = -x_2 + x_4 \end{cases} \tag{3}$$

Take the derivative of (3) and substitute (2) into it, we can obtain:

$$\begin{cases} \dot{z}_1 = z_2 \\ \dot{z}_2 = \dfrac{f}{m_u} \\ \dot{z}_3 = z_4 \\ \dot{z}_4 = -\left(\dfrac{k_s}{m_s} + \dfrac{k_s}{m_u}\right) z_3 - \left(\dfrac{c_s}{m_s} + \dfrac{c_s}{m_u}\right) z_4 + \dfrac{f}{m_u} - \left(\dfrac{1}{m_u} + \dfrac{1}{m_s}\right) u \end{cases} \tag{4}$$

It is obvious that the stress of the tire f is a function of z_3, so the following equation can be obtained by Lagrange transformation:

$$f = f|_{x_3=0} + g x_3 = f|_{x_3=0} + g z_3 \tag{5}$$

where, $g = \frac{\partial f}{\partial x_3}|_{x_3=x_3^0}$, and x_3^0 is a constant between 0 and x_3. Next, the system controller will be designed to improve the ride comfort of the vehicle.

3 Controller Design of Active Suspension System

3.1 Design of First-Order Filter

In order to avoid the problem of "differential explosion" in the process of backstepping, a first-order filter is introduced, the form is as follows:

$$\eta_i \dot{\alpha}_{if} + \alpha_{if} = \alpha_i \tag{6}$$

where, η_i is a positive constant, α_{if} is the output of the first-order filter and α_{i} is the virtual control law, and its specific form will be given below. Let $y_i = \alpha_{if} - \alpha_i$ be the error between the output of the first-order filter and the virtual control law.

3.2 Estimation of Tire Stress

Considering that the tire stresses during driving are nonlinear and difficult to calculate, a radial basis function neural network function is used as a substitute in the model to reduce the adverse effects of linear substitution. The specific form of the function is as follows:

$$f|_{x_3=0} = \vartheta_1^{\mathrm{T}} \phi_1(X_1) + \varepsilon_1(X_1) \tag{7}$$

where, $\phi_i(X_i)$ is the basis function of neural network, $\varepsilon_i(X_i)$ is the estimation error between neural network function and the tire stress, which meets $|\varepsilon_i(X_i)| \leq \overline{\varepsilon}_i$, $\overline{\varepsilon}_i$ is an unknown positive constant, $X_1 = (x_1, x_2, x_4, x_r, \dot{x}_r)$.

3.3 Design of System Controller

In this section, the backstepping and the adaptive algorithm is used to design the system controller. Firstly, for the active suspension system, the following coordinate transformation is introduced: $\xi_1 = z_1 - z_d$, $\xi_2 = z_2 - \alpha_{1f}$, $\xi_3 = z_3 - \alpha_{2f}$, $\xi_4 = z_4 - \alpha_{3f}$. Where, z_d is the expected output of the system. For the first-order subsystem, the Lyapunov function in the following form is constructed:

$$V_1 = \frac{1}{2}\xi_1^2 + \frac{1}{2}y_1^2 \tag{8}$$

The derivative of the above Lyapunov function is obtained:

$$\dot{V}_1 = \xi_1(\xi_2 + y_1 + \alpha_1 - z_d) + y_1\dot{y}_1 \tag{9}$$

According to Young's inequality, the following inequality holds:

$$\begin{aligned} &\xi_1\xi_2 \leq \frac{1}{2}\xi_1^2 + \frac{1}{2}\xi_2^2, \xi_1 y_1 \leq \frac{1}{2}\xi_1^2 + \frac{1}{2}y_1^2, \xi_1\dot{z}_d \leq \frac{1}{2\sigma_1}\xi_1^2\dot{z}_d^2 + \frac{\sigma_1}{2} \\ &y_1\dot{y}_1 = -\frac{y_1^2}{\eta_1} + y_1\tau_1 \leq \left(\frac{1}{4} - \frac{1}{\eta_1}\right)y_1^2 + \tau_1^2 \end{aligned} \tag{10}$$

where, $\tau_i = -\dot{\alpha}_i$.

The virtual controller is designed as follows:

$$\alpha_1 = -\rho_1\xi_1 - \xi_1 - \frac{\xi_1\dot{z}_d^2}{2\sigma_1} \tag{11}$$

where, ρ_i is a positive constant.

Substitute (10), (11) into (9), we can obtain:

$$\dot{V}_1 \leq -\rho_1\xi_1 + \frac{1}{2}\xi_2^2 + \left(\frac{3}{4} - \frac{1}{\eta_1}\right)y_1^2 + \tau_1^2 + \frac{\sigma_1}{2} \tag{12}$$

For the second-order subsystem, the Lyapunov function is constructed as follows:

$$V_2 = V_1 + \frac{1}{2}\xi_2^2 + \frac{1}{2}\tilde{\vartheta}_1^{\mathrm{T}}\tilde{\vartheta}_1 + \frac{1}{2}y_2^2 \tag{13}$$

where, $\tilde{\vartheta}_i = \hat{\vartheta}_i - \vartheta_i$, $\hat{\vartheta}_i > 0$ is the estimate of ϑ_i, ϑ_i is the unknown parameter of NN.

The derivative of the above Lyapunov function is obtained:

$$\dot{V}_2 = \dot{V}_1 + \xi_2\left(\frac{f|_{x_3=0}}{m_u} + \frac{g}{m_u}(\xi_3 + y_2 + \alpha_2) - \dot{\alpha}_{1f}\right) + y_2\dot{y}_2 + \tilde{\vartheta}_1^{\mathrm{T}}\dot{\hat{\vartheta}}_1 \tag{14}$$

Substitute (7) into (14), we can obtain:

$$\dot{V}_2 = \dot{V}_1 + \xi_2\left(\frac{\vartheta_1^{\mathrm{T}}\phi_1 + \varepsilon_1}{m_u} + \frac{g}{m_u}(\xi_3 + y_2 + \alpha_2) - \dot{\alpha}_{1f}\right) + y_2\dot{y}_2 + \tilde{\vartheta}_1^{\mathrm{T}}\dot{\hat{\vartheta}}_1 \tag{15}$$

According to Young's inequality, the following inequality holds:

$$\begin{aligned}
&\xi_2\varepsilon_1 \le \frac{[g]}{2}\xi_2^2 + \frac{1}{2\underline{[g]}}\bar{\varepsilon}_1^2 && \tfrac{1}{2}\xi_2^2 \le \frac{[g]}{2}\xi_2^4 + \frac{1}{2\underline{[g]}} \\
&\xi_2 g\xi_3 \le \frac{\mu_1[g]}{2}\xi_2^2 + \frac{\bar{g}^2}{8\mu_1\underline{[g]}} + \tfrac{1}{2}\xi_3^2 && \xi_2\dot{\alpha}_{1f} \le \frac{[g]}{2}\xi_2^2\dot{\alpha}_{1f}^2 + \frac{1}{2\underline{[g]}} \\
&\xi_2 g y_2 \le \frac{\gamma_1[g]}{2}\xi_2^2 + \frac{\bar{g}^2}{8\gamma_1\underline{[g]}} + \frac{y_2^2}{2} && \xi_2\hat{\vartheta}_1^{\mathrm{T}}\phi_1 \le \frac{\lambda_1[g]}{2}\xi_2^2\hat{\theta}_1\psi_1 + \frac{1}{2\lambda_1\underline{[g]}}
\end{aligned} \tag{16}$$

where, $\psi_i(x) = \phi_i(X_i)^2$, λ_1, γ_1, μ_1 are positive constants, $\bar{g}$ is the upper bound of g, let $[g] = -|g|$ and it meets $[g] \le \underline{[g]}$.

The virtual controller and the law of adaptive are designed as follows:

$$\alpha_2 = -\rho_2\xi_2 - \frac{1}{2}\xi_2 - \frac{1}{2}\xi_2^3 - \frac{\lambda_1}{2}\xi_2\hat{\theta}_1\psi_1 - \frac{\mu_1}{2}\xi_2 - \frac{\gamma_1}{2}\xi_2 - \frac{1}{2}\xi_2\dot{\alpha}_{1f}^2 \tag{17}$$

$$\dot{\hat{\vartheta}}_1 = \frac{\xi_2\phi_1}{m_u} - \beta_1\hat{\vartheta}_1 \tag{18}$$

where β_i, is a positive constant.

Substitute (16), (17) and (18) into (15), we can obtain:

$$\dot{V}_2 \le -\rho_1\xi_1^2 - \frac{\rho_2[g]\xi_2^2}{m_u} + \frac{1}{2m_u}\xi_3^2 + \frac{\beta_1}{2}\tilde{\vartheta}_1^{\mathrm{T}}\tilde{\vartheta}_1 + \sum_{i=1}^{2}\left(\frac{3}{4} - \frac{1}{\eta_i}\right)y_i^2 + d_1 \tag{19}$$

where, $d_1 = \frac{\bar{\varepsilon}_1^2}{2\underline{[g]}} + \frac{\bar{g}^2}{8\mu_1\underline{[g]}} + \frac{\bar{g}^2}{8\gamma_1\underline{[g]}} + \frac{1}{\underline{[g]}} + \frac{1}{2\lambda_1\underline{[g]}} + \frac{\sigma_1}{2} + \frac{\beta_1}{2}\vartheta_1^{\mathrm{T}}\vartheta_1 + \sum_{i=1}^{2}\tau_i^2$.

For the third-order subsystem, the Lyapunov function is constructed as follows:

$$V_3 = V_2 + \frac{1}{2}\log\frac{k^2}{k^2 - \xi_3^2} + \frac{1}{2}y_3^2 \tag{20}$$

where, k is a positive constant.

The derivative of the above Lyapunov function is obtained:

$$\dot{V}_3 = \dot{V}_2 + \upsilon_1\left(\xi_4 + y_3 + \alpha_3 - \dot{\alpha}_{2f}\right) + y_3\dot{y}_3 \tag{21}$$

where, $\upsilon_1 = \frac{\xi_3}{k^2-\xi_3^2}$.

According to Young's inequality, the following inequality holds:

$$\begin{aligned} \upsilon_1\xi_4 &\le \frac{1}{2}\upsilon_1^2 + \frac{1}{2}\xi_4^2 \\ \upsilon_1 y_3 &\le \frac{1}{2}\upsilon_1^2 + \frac{1}{2}y_3^2 \end{aligned} \tag{22}$$

The virtual controller is designed as follows:

$$\alpha_3 = -\rho_3\xi_3 - \upsilon_1 - \frac{1}{2m_u}\xi_3\left(k^2 - \xi_3^2\right) + \dot{\alpha}_{2f} \tag{23}$$

Substitute (22) and (23) into (21), we can obtain:

$$\dot{V}_3 \le -\rho_1\xi_1^2 - \frac{\rho_2\underline{[g]}\xi_2^2}{m_u} - \rho_3\upsilon_1\xi_3 + \frac{1}{2}\xi_4^2 + \frac{\beta_1}{2}\tilde{\vartheta}_1^{\mathrm{T}}\tilde{\vartheta}_1 + \sum_{i=1}^{3}\left(\frac{3}{4} - \frac{1}{\eta_i}\right)y_i^2 + d_2 \tag{24}$$

where, $d_2 = \frac{\bar{\varepsilon}_1^2}{2\underline{[g]}} + \frac{\bar{\sigma}^2}{8\mu_1\underline{[g]}} + \frac{\bar{\sigma}^2}{8\gamma_1\underline{[g]}} + \frac{1}{\underline{[g]}} + \frac{1}{2\lambda_1\underline{[g]}} + \frac{\sigma_1}{2} + \frac{\beta_1}{2}\vartheta_1^{\mathrm{T}}\vartheta_1 + \sum_{i=1}^{3}\tau_i^2$.

For the forth-order subsystem, the Lyapunov function is constructed as follows:

$$V_4 = V_3 + \frac{1}{2}\xi_4^2 + \frac{1}{2}\tilde{\vartheta}_2^{\mathrm{T}}\tilde{\vartheta}_2 \tag{25}$$

The derivative of the above Lyapunov function is obtained:

$$\begin{aligned} \dot{V}_4 = \dot{V}_3 + \xi_4\Bigg(&-\left(\frac{1}{m_s} + \frac{1}{m_u}\right)u - \left(\frac{k_s}{m_s} + \frac{k_s}{m_u}\right)z_3 \\ &-\left(\frac{c_s}{m_s} + \frac{c_s}{m_u}\right)z_4 + \frac{f}{m_u} - \dot{\alpha}_{3f}\Bigg) - \tilde{\vartheta}_2^{\mathrm{T}}\dot{\hat{\vartheta}}_2 \end{aligned} \tag{26}$$

The RBF neural network function is used to approximate the unknown nonlinear function as follows:

$$-\left(\frac{k_s}{m_s} + \frac{k_s}{m_u}\right)z_3 - \left(\frac{c_s}{m_s} + \frac{c_s}{m_u}\right)z_4 + \frac{f}{m_u} = \vartheta_2^T\varphi_2(X_2) + \varepsilon_2(X_2) \tag{27}$$

where, $X_2 = (x_1, x_2, x_3, x_4, x_r, \dot{x}_r)$.

According to Young's inequality, the following inequality holds:

$$\xi_4\varepsilon_2 \le \frac{1}{2}\xi_4^2 + \frac{1}{2}\overline{\varepsilon}_2^2 \tag{28}$$

The controller and the law of adaptive are designed as follows:

$$u = \frac{\left(\rho_4\xi_4 + \hat{\vartheta}_2^{\mathrm{T}}\phi_2 + \xi_4 - \dot{\alpha}_{3f}\right)m_u m_s}{m_s + m_u} \tag{29}$$

$$\dot{\hat{\vartheta}}_2 = \xi_4\phi_2 - \beta_2\hat{\vartheta}_2 \tag{30}$$

Substitute (28), (29) and (30) into (27), we can obtain:

$$\dot{V}_4 \le -\rho_1\xi_1^2 - \frac{\rho_2\underline{[g]}\xi_2^2}{m_u} - \rho_3\upsilon_1\xi_3 - \rho_4\xi_4^2 + \sum_{j=1}^{2}\frac{\beta_j}{2}\tilde{\vartheta}_j^{\mathrm{T}}\tilde{\vartheta}_j + \sum_{i=1}^{3}\left(\frac{3}{4} - \frac{1}{\eta_i}\right)y_i^2 + d_3 \tag{31}$$

where,$d_3 = \sum_{i=1}^{2}\frac{\beta_i}{2}\vartheta_i^{\mathrm{T}}\vartheta_i + \sum_{j=1}^{3}\tau_j^2 + \frac{1}{2}\overline{\varepsilon}_2^2 + \frac{1}{2\underline{[g]}}\overline{\varepsilon}_1^2 + \frac{\overline{g}^2}{8\mu_1\underline{[g]}} + \frac{\overline{g}^2}{8\gamma_1\underline{[g]}} + \frac{1}{\underline{[g]}} + \frac{1}{2\lambda_1\underline{[g]}} + \frac{\sigma_1}{2}$.

Up to now, the design of the controller has been completed.

3.4 *Proof of System Stability*

From the above, we can obtain:

$$V_4 = \frac{1}{2}\xi_1^2 + \frac{1}{2}\xi_2^2 + \frac{1}{2}\log\frac{k^2}{k^2 - \xi_3^2} + \frac{1}{2}\xi_4^2 + \frac{1}{2}\sum_{i=1}^{3}y_i^2 + \frac{1}{2}\sum_{j=1}^{2}\tilde{\vartheta}_j^{\mathrm{T}}\tilde{\vartheta}_j \tag{32}$$

Obviously, if $|\xi_3| < k$, the following inequality holds:

$$\log\frac{k^2}{k^2 - \xi_3^2} < \frac{\xi_3^2}{k^2 - \xi_3^2} \tag{33}$$

Substitute (33) into (32), we can obtain:

$$V_4 < \frac{1}{2}\xi_1^2 + \frac{1}{2}\xi_2^2 + \frac{1}{2}\upsilon_1\xi_3 + \frac{1}{2}\xi_4^2 + \frac{1}{2}\sum_{i=1}^{3}y_i^2 + \frac{1}{2}\sum_{j=1}^{2}\tilde{\vartheta}_j^{\mathrm{T}}\tilde{\vartheta}_j \tag{34}$$

According to (31) and (34), we can obtain:

$$\dot{V}_4 \le -\rho V_4 + d_3 \tag{35}$$

where, $\rho = \min\{2\rho_1, \frac{2[g]\rho_2}{m_u}, 2\rho_3, 2\rho_4, \beta_i, \left(\frac{3}{2} - \frac{2}{\eta_j}\right), \quad i = 1, 2; \quad j = 2, 3, 4\}$.

According to Lyapunov bounded stability theorem, the stability of the system can be proved.

Due to the system stability, the dynamic controllers as (10), (17), (23), (29) and adaptive control laws as (18), (30) designed above are bounded, which shows the effectiveness of the designed controllers. The final controller not only limits the suspension dynamic deflection to the preset range, but realizes the effective suppression of vehicle system vibration.

4 Simulation and Result Analysis

In this section, the rationality of the proposed control strategy will be validated through co-simulation using Amesim 2021 and Matlab 2021b. A simulation object of a quarter-vehicle active suspension system was built in Amesim 2021 to simulate the actual suspension system. A simulation program for the control algorithm presented in this paper was developed in Matlab 2021b. The control signals and suspension states are transmitted in real-time between the two software tools, The schematic diagram of the simulator is shown in Fig. 1. Reference [9], The suspension system model parameters and controller parameters in the joint simulation experiment are shown in Table 1.

The frequency of road surface variations in actual road conditions is complex and diverse. Simulation scenario 1: To validate the effectiveness of the proposed algorithm for suspension control under different frequency road excitations, a sinusoidal road input with frequencies ranging from 1 to 2.5 Hz was set up.

In this paper, the estimated tire force using a radial basis neural network function based on the system state is compared with the actual tire force output by Amesim, as

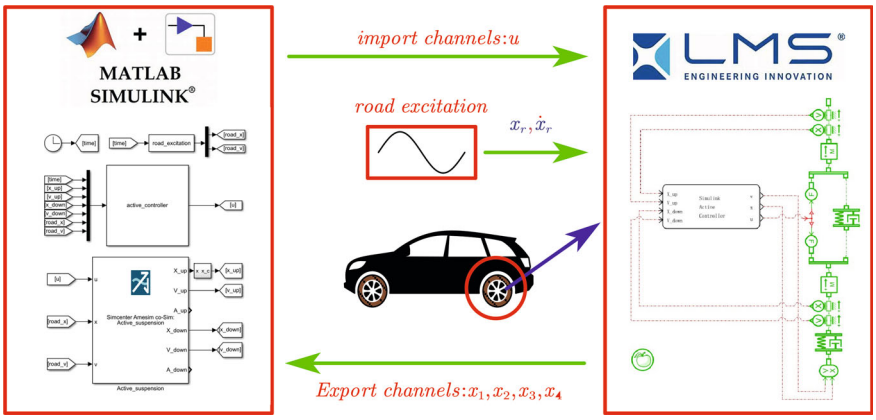


Fig. 1 Schematic diagram of Amesim and Matlab joint simulation

Table 1 Suspension system model parameters and controller parameters table

Parameter	Number	Parameter	Number	Parameter	Number
m_s	340 kg	m_t	40kg	k_s	8500 N/m
k_t	22000 N/m	c_s	2000 N s/m	c_t	300 N s/m
ρ_1	5.5	ρ_2	−2	ρ_3	2
ρ_4	−30	η_1	100000	η_2	1000000
η_3	1000	λ_1	0.0001	γ_1	0.01
μ_1	0.01	σ_1	1	β_1	0.1
β_2	0.1	k	0.15		

shown in Fig. 2. It can be observed that the algorithm accurately predicts the changes in actual force at different frequencies.

The comparison of the four indicators mentioned above (spring deflection, spring acceleration, dynamic travel, and dynamic/static load ratio) under different frequency sinusoidal road inputs between the BANNC algorithm proposed in this paper and the PID controller and passive suspension are shown in Fig. 3.

The BANNC control strategy has smaller amplitudes in spring deflection and spring acceleration. The suspension's dynamic travel also undergoes limited expansion and contraction within a specified range, with a smaller dynamic/static load ratio. Therefore, the BANNC algorithm proposed in this paper exhibits better driving smoothness than PID control and passive suspension under different frequency road excitations.

Random road surfaces are better able to reflect the actual conditions of vehicle motion. Scenario 2: A random road surface was designed. The comparison of spring

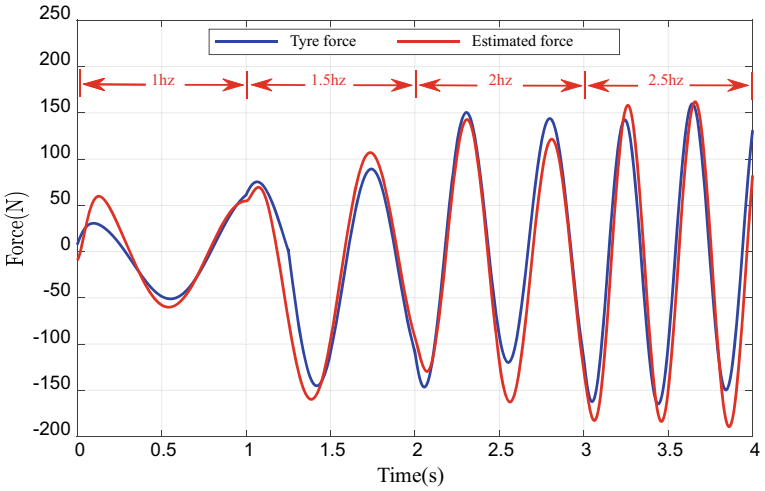


Fig. 2 Estimated force versus actual force diagram

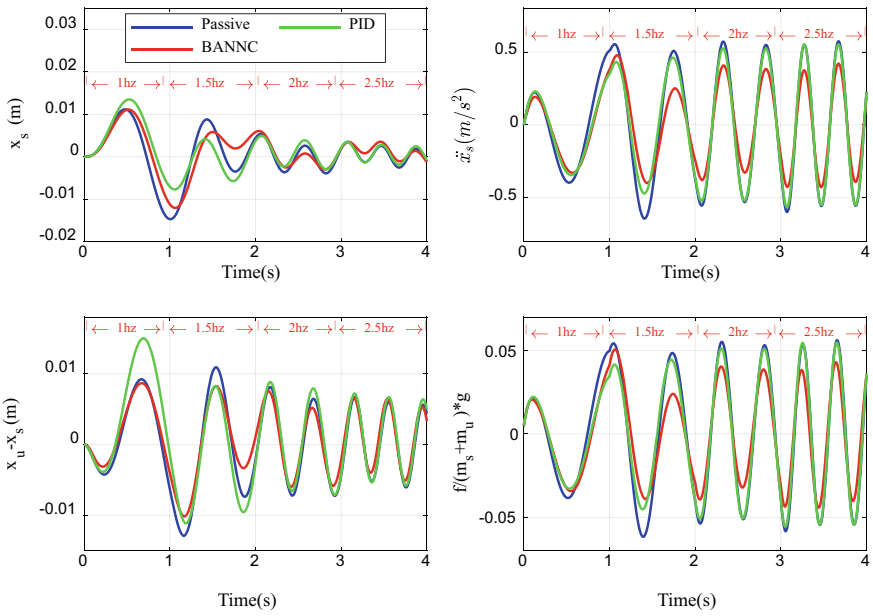


Fig. 3 Performance of suspension under variable frequency sinusoidal road excitation

deflection, spring acceleration, dynamic travel, and dynamic/static load ratio of the suspension through joint simulation under the random road input is shown in Fig. 4. Figure 5 shows the performance comparison.

It can be seen that under the random road input, the BANNC algorithm proposed in this paper exhibits smoother spring deflection and spring acceleration, and better dynamic travel and dynamic/static load ratio compared to the PID controller and passive suspension. The BANNC algorithm in this paper shows an improvement of approximately 14% in spring deflection and acceleration compared to the passive suspension, and an improvement of around 22% in the dynamic/static load ratio.

Through simulations in two scenarios, namely variable frequency sinusoidal road input and random road input, and by comparing with PID control and passive suspension, it can be observed that the proposed backstepping adaptive neural network controller in this paper has significant improvement in spring deflection and spring acceleration. It has to some extent enhanced the vehicle's driving smoothness and ride comfort. The range of suspension travel is within reasonable limits and it also exhibits a smaller dynamic/static load ratio, thereby improving the safety of vehicle driving.

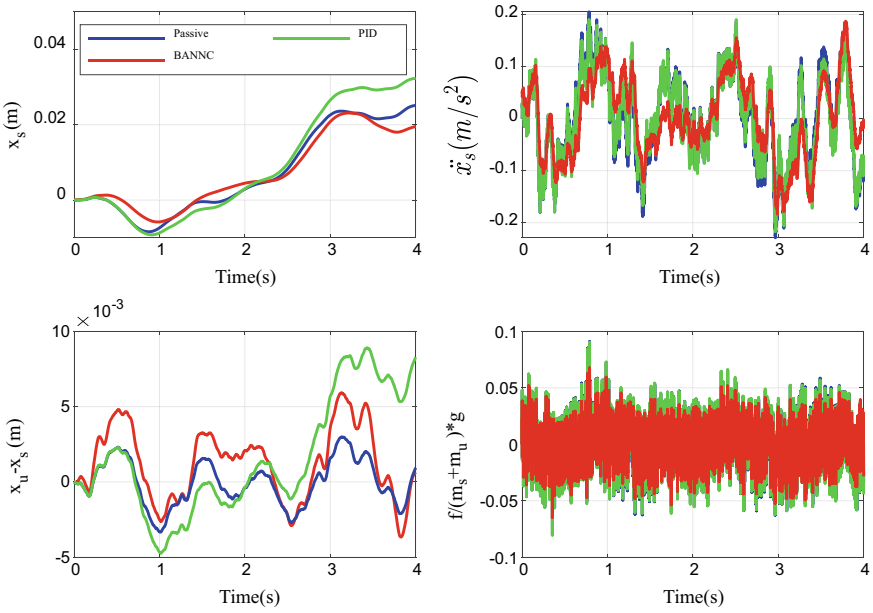


Fig. 4 Suspension performance under random road excitation

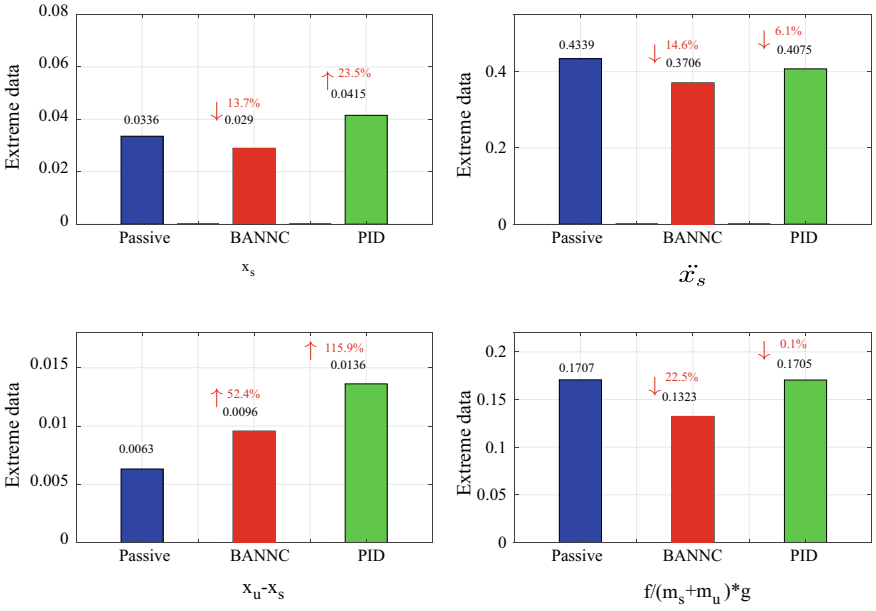


Fig. 5 Comparison of extreme values of different methods on random pavement

5 Conclusion

This paper addresses the issues of vehicle driving smoothness and ride comfort by establishing a two-degree-of-freedom quarter vehicle suspension system model. It investigates effective control strategies to improve vehicle performance, taking into account the complexities of tire forces and the limitations of suspension travel in real-world scenarios. By utilizing barrier Lyapunov functions and backstepping techniques, an adaptive control strategy for the system is designed. The stability of the system is verified based on Lyapunov stability criteria. Ultimately, the proposed approach achieves improvements in vehicle driving smoothness and ride comfort, while ensuring limited suspension travel. The effectiveness of the algorithm is validated through simulations.

References

1. Ni L, Wu L, Zhan H (2022) Parameter uncertainty analysis of posture control of a four-wheel-legged robot with series slow active suspension system. Mech Mach Theory 175:104699
2. Ahmad I, Ge X, Han Q-L (2021) Decentralized dynamic event-triggered communication and active suspension control of in-wheel motor driven electric vehicles with dynamic damping. IEEE/CAA J Autom Sin 8(5):971–986. https://doi.org/10.1109/JAS.2021.1003967
3. Wang G, Chen C, Yu S (2017) Robust non-fragile finite-frequency H_∞ static output-feedback control for active suspension systems. IEEE/CAA J Autom Sin0 (91):41–56
4. Lathkar MS, Shendge PD, Phadke SB (2020) Active control of uncertain seat suspension system based on a state and disturbance observer. IEEE Trans Syst Man Cybernetics: Syst 50(3):840–850. https://doi.org/10.1109/TSMC.2017.2734121
5. Sun W, Gao H, Kaynak O (2013) Adaptive backstepping control for active suspension systems with hard constraints. IEEE/ASME Trans Mechatron (18):1072–1079
6. Yu J, Zhao L, Yu H, Lin C (2019) Barrier Lyapunov functions-based command filtered output feedback control for full-state constrained nonlinear systems. Automatica 105:71–79
7. Konghui GUO, Dang LU (2007) UniTire: unified tire model for vehicle dynamic simulation. Veh Syst Dyn 45(S1):79–99
8. Hsiao T (2013) Robust estimation and control of tire traction forces. IEEE Trans Veh Technol 62(3):1378–1383. https://doi.org/10.1109/TVT.2012.2230656
9. Na J, Huang Y, Wu X et al (2021) Active suspension control of quarter-car system with experimental validation. IEEE Trans Syst Man Cybernetics: Syst 52(8):4714–4726

Open Access This chapter is licensed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License (http://creativecommons.org/licenses/by-nc-nd/4.0/), which permits any noncommercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if you modified the licensed material. You do not have permission under this license to share adapted material derived from this chapter or parts of it.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.



A Study on Automobile Wind Resistance Prediction of E-vehicle Based on AOA-CatBoost Hybrid Modeling



Chuanbo Liu, Yuhan Liu, Guangyuan Yang, and Zhu Bin

Abstract In this paper, an automobile wind resistance prediction method is proposed to address the problems of high cost and long time-consuming of traditional automobile aerodynamic shape design methods as well as the low accuracy of the current wind resistance prediction model. First, an MPV simulation model is constructed for computational fluid dynamics (CFD) simulation, from which the relevant input feature parameters are selected. Then, the AOA-CatBoost prediction algorithm is modeled. Finally, the model is evaluated by relevant indicators. The experimental results show that the AOA-CatBoost wind resistance prediction model has high consistency with the results of real-vehicle tests, and the model can effectively predict the wind resistance of the vehicle.

Keywords E-vehicles wind resistance prediction · Arithmetic optimization algorithm · CatBoost · Prediction model

1 Introduction

The energy consumed to overcome vehicle air resistance is one of the main energy consumption of vehicles, so reducing air resistance can significantly improve the energy utilization efficiency of electric vehicles and extend their range, which is crucial for the new energy vehicle industry to improve energy efficiency and meet environmental standards.

With the rapid development of artificial intelligence technology and intelligent algorithms, the data-driven automobile wind resistance prediction method has gradually become a hotspot for scholars at home and abroad. Urquhart et al. proposed

C. Liu · Y. Liu (✉) · G. Yang · Z. Bin
School of Mechanical and Electronic Engineering, Wuhan University of Technology, Wuhan, China
e-mail: YuhanL@whut.edu.cn

C. Liu
e-mail: Lchb72@whut.edu.cn

© The Author(s) 2026

H. Saman et al. (eds.), *The 9th International Conference on Advances in Construction Machinery and Vehicle Engineering*, Lecture Notes in Mechanical Engineering,
https://doi.org/10.1007/978-981-96-5527-4_111

an aerodynamic shape design method based on Proper Orthogonal Decomposition, and found that its prediction ability was more advantageous at low sample numbers [1]. Bakshi et al. studied Ahmed model using FFD technique. And an inverted rear wing was installed at the rear of the vehicle to increase the pressure and after using Simple algorithm in mesh deformation technique the drag was reduced by 0.0777% [2].

Although existing methods in the literature have achieved some success, wind resistance prediction is still complex and challenging due to the diversity and complexity of automobiles, which vary significantly from model to model. Reducing wind resistance is not only one of the main methods to improve the energy efficiency of electric vehicles, but also a key factor to enhance market competitiveness and user satisfaction [3]. Therefore, it is particularly important to be able to predict wind resistance quickly and accurately at an early stage in the design process to meet the fast-paced development cycle and immediate needs of the new energy electric vehicle industry.

In response to the above problems and existing research, this paper proposes a wind resistance prediction method for new energy electric vehicles based on an arithmetic optimization algorithm optimizing CatBoost. Considering the influence of different feature parameters on the wind resistance of a vehicle, after the wind resistance data are obtained by Computational Fluid Dynamics (CFD) simulation, the feature parameters with higher correlation are selected as inputs to the CatBoost prediction model. An arithmetic optimization algorithm (AOA) is introduced to train the hyperparameters in the CatBoost prediction model to find the optimal values of the wind resistance prediction results. Finally, relevant evaluation indexes are introduced to analyze and evaluate the performance of the prediction model to verify the feasibility of the method proposed in this paper.

2 Establishment of Prediction Algorithm

2.1 CatBoost for Whole Vehicle Wind Resistance Prediction

CatBoost is a member of the Gradient Boosted Decision Tree (GBDT) family, which consists of a combination of Gradient Boosting Gradient Boosting and Categorical Features. The algorithm has the ability to handle categorical features and missing data while exhibiting strong generalization capabilities. In addition, CatBoost has the ability to automatically perform feature selection and parameter tuning, providing good support for model performance. Its use of a fully symmetric tree as the base model effectively avoids overfitting and improves the reliability of prediction [4]. The formula is:

$$x_{\sigma_p,k} = \frac{\sum_{j=1}^{p-1} \left[x_{\sigma_j,k} = x_{\sigma_p,k}\right] Y_{\sigma_j} + \alpha p}{\sum_{j=1}^{p-1} \left[x_{\sigma_j,k} = x_{\sigma_p,k}\right] Y_{\sigma_j} + \alpha} \tag{1}$$

On the other hand, CatBoost adopts a method based on the statistical value of the prediction target in the processing of the category features, which transforms the category features into numerical features, reflecting the idea of mean value coding. The algorithm also employs a more efficient strategy to transform the values of the category features taken for each sample by random ordering, taking the mean value based on the category label values of the previous samples and introducing a weighting coefficient for the priority (a priori value). In terms of feature crossover, CatBoost uses a greedy strategy that does not use any crossover features for the first split, and uses all original features and crossover features used in generating the tree for subsequent splits, dynamically combining different types of features to form new features to obtain higher-order dependencies. These innovative designs and optimizations have made CatBoost a highly sought-after advanced tool in the machine learning field.

2.2 Arithmetic Optimization Algorithm

Arithmetic Optimization Algorithm (AOA) is a novel meta-heuristic optimization algorithm, which divides the optimization process into three key phases, namely initialization, exploration and exploitation, by selecting the optimization strategies through mathematical function gas pedals. In the exploration phase, multiplication and division operations realize global exploration to update the global optimal position to avoid local solutions; in the exploitation phase, addition and subtraction operations realize local exploitation to improve the solution accuracy [5].

In the initialization phase, AOA generates a uniformly distributed random population for the initial solution X. The initial population can be found by the following equation:

$$X_{(i,j)} = Rand \times (U_b - L_b) + L_b \tag{2}$$

In Eq. (2), U_b is the upper bound, L_b is the lower bound, Rand is the random number before [0,1], $X_{(i,j)}$ is the position of the ith solution in the jth dimension space.

The exploration phase or the development phase is selected before the start of the algorithm iteration by means of a mathematical optimizer acceleration function (MOA):

$$MOA(t) = Min + t \times \frac{(Max - Min)}{T} \tag{3}$$

In Eq. (3), r_1 is the random number between [0,1], MOA(t) is the current value of the acceleration function, Min is the minimum value of the acceleration function, Max is the maximum value of the acceleration function, T is the maximum number of iterations, t is the current number of iterations; when $r_1 <$ MOA(t), the function will

enter the global exploration phase, and vice versa, it will enter the local development phase [6].

In the exploration phase, a multiplication or division strategy is decided based on the random number. The random number $r_2 \in [0,1]$ is compared with 0.5 and the outermost division strategy is chosen if $r_2 > 0.5$ and vice versa for multiplication. Both strategies are highly discrete, which facilitates particle exploration in the algorithmic space.

$$X(t+1) = \begin{cases} X_{b(t)} \div (MOP + \varepsilon) \times ((U_b - L_b) \times \mu + L_b), r_2 < 0.5, \\ X_{b(t)} \times MOP \times ((U_b - L_b) \times \mu + L_b), otherwise \end{cases} \quad (4)$$

In Eq. (4), r_2 is the random number of [0,1], X(t + 1) is the position of the next-generation particle, $X_{b(t)}$ is the position of the particle with the best current fitness, μ is the control coefficient of the searching process (the value of 0.499), ε is the minima, and MOP is the Mathematical Optimizer Probability, which is calculated by the following formula:

$$MOP(t) = 1 - (t/T)^{(1/\alpha)} \quad (5)$$

In Eq. (5), α is the iteration sensitivity coefficient, the higher the value of α the more the number of iterations affects the MOP(t).

The development phase utilizes the additive and subtractive mechanisms as search strategies for local optimization, and this phase has a small search scope and high accuracy [7]. Their calculation formulas are as follows:

$$X(t+1) = \begin{cases} X_{b(t)} - MOP \times ((U_b - L_b) \times \mu + L_b), r_3 < 0.5, \\ X_{b(t)} + MOP \times ((U_b - L_b) \times \mu + L_b), otherwise \end{cases} \quad (6)$$

In Eq. (6), r_3 is a random number between 0 and 1. It is compared with 0.5 and if $r_3 > 0.5$, the subtraction strategy is chosen and vice versa. Both strategies are highly discrete, which facilitates particle exploration in the algorithmic space.

2.3 AOA-*CatBoost Model Combination*

The core idea of AOA-optimized CatBoost is to cleverly construct a stochastic population through clever arithmetic operator thinking combined with the hyperparameter space of CatBoost. Through the operations of initialization, exploration and exploitation, the individuals are iteratively evolved to find the optimal hyperparameter combination [8]. The method has the ability to perform global optimization search in the hyperparameter search space, and can flexibly adjust the hyperparameters to adapt to the characteristics of different datasets and problem domains, so as to effectively

optimize the performance of the CatBoost algorithm When using the AOA algorithm to optimize CatBoost, the specific operation steps are as follows:

(1) Randomly construct a population whose initial position represents the initial value of the relevant parameters, set the root mean square error as the fitness function, set the maximum number of iterations, select the optimization parameters and so on.
(2) Set the hyperparameter search range of Catboost algorithm, and perform model prediction by initializing the optimization parameters.
(3) If the current iteration value is less than the maximum value of the iteration (C_Iter < M_Iter), the fitness value will be calculated.
(4) Update the MOA, MOP, and select the development phase to update the parameters.
(5) If the number of evolutionary iterations reaches the maximum value, terminate the evolution and assign the global optimal individual hyperparameter values to the CatBoost model, otherwise go to step (3).
(6) Calculate the training error and update the hyperparameters.
(7) Judge whether the number of training times or training error meets the requirements, if not, go to step (6).
(8) Output the training results.

3 Acquisition of Experimental Data

3.1 CFD Model Construction

Before the simulation can be roughly divided into three processes. Pre-processing process: vehicle internal geometry cleanup, flow field calculation domain setting, vehicle mesh generation; Solution setting: turbulence model selection, discrete format determination, boundary condition setting; Post-processing process: experimental pressure, velocity vector, turbulence kinetic energy, and streamline analysis for the wind tunnel simulation of the vehicle model [9].

After making a large number of sketches through UG, a representative styling scheme for the tram was initially determined. After carrying out the processing of the model, it is introduced into STAR-CCM + for parameter setting, and finally the CFD simulation strategy is obtained as shown in Table 1 [10].

3.2 Selection of Characteristic Parameters

After determining the simulation strategy, based on the simulation post-processing results, 15 structural parameters are identified and their deformation ranges are

Table 1 CFD simulation strategy

Type Name	Parameters
Speed inlet	35 m/s
Velocity exit	0 Pa
Bottom surface of simulation domain	*Moving Wall*
Tire speed	109.375 rad/s
Turbulence model	*Realizable k-ε*
Calculation domain Iteration steps Number of grids	9L * 11W * 6H 4000 steps 42.671 million

Table 2 Deformation ranges of each parameter

Parameter name	Range of change
Front end width	1715–1895 mm
Rear gear tilt angle	53°–63°
Front gear tilt angle	23.3°–33.3°
Front overhang	855–1055 mm
Rear overhang	975–1175 mm
Wheelbase	2630–2880 mm
Hair cowl tilt angle	13°–23°
Approach angle	16.5°–26.5°
Departure angle	18.5°–25.5°
Top cover rear end down amplitude	40–140 mm
Ground clearance	95–195 mm
Rear end width	1715–1875 mm
Upper grille tilt angle	88.8°–106.8°
Vehicle height	1635–1835 mm
Vehicle width	1715–1915 mm

limited to their respective reasonable ranges, and the deformation ranges of each parameter are shown in Table 2.

4 Experimental Results and Discussion

In order to verify the feasibility of the AOA-CatGBoost wind resistance coefficient prediction model, this paper uses common machine learning algorithm models (ELM, CatBoost, XGBoost, GWO-CatBoost, SSA-CatBoost, PSO-CatBoost) to model and analyze the same set of samples, and compares the performance with the AOA-CatBoost model established in this paper. CatBoost for performance comparison. In the experimental setup, 90 sets of samples are randomly selected as the training set and then the remaining 30 sets of samples are randomly selected as the test set. Each set of samples is validated several times and then the average value is taken as the most output result.

In order to verify the accuracy of the prediction model, the coefficient of determination (R^2), root mean square error (RMSE) and mean absolute error (MAE) are used to evaluate the prediction performance of the model of automobile wind resistance in this paper. R^2 reflects the degree of fit between the predicted values of the model and the observed values, and the closer the value of R^2 is to 1 means the better the model is fitted; the RMSE represents the square root of the average difference between the predicted values of the model and the actual observations, and the smaller the value is, the higher the accuracy of the model is. RMSE indicates the square root of the average difference between the model predicted values and the actual observed values, the smaller the value, the higher the prediction accuracy of the model [11]; MAE indicates the average absolute difference between the model predicted values and the actual observed values, which is insensitive to outliers, and the smaller the value, the smaller the average error of the model on the prediction target.

Figure 1 shows that AOA-CatGBoost exhibits the best performance in terms of MAE, MSE and R^2 compared to other prediction methods. It achieves a maximum value of 0.976 for R^2, which is up to 0.028 higher than other prediction models and has a better fit with the simulation values; at the same time, the RMSE of this model is only 0.0101, which is up to 42% lower than that of other prediction models; and the MAE of this model of 0.0075 is reduced by 19.20–40.48% compared to that of other prediction models.

Figure 2 illustrates the C_d prediction results of 30 test sets under different prediction models, and there are some differences between the predicted C_d of each model and the simulated C_d. In terms of overall performance, the AOA-CatBoost model and

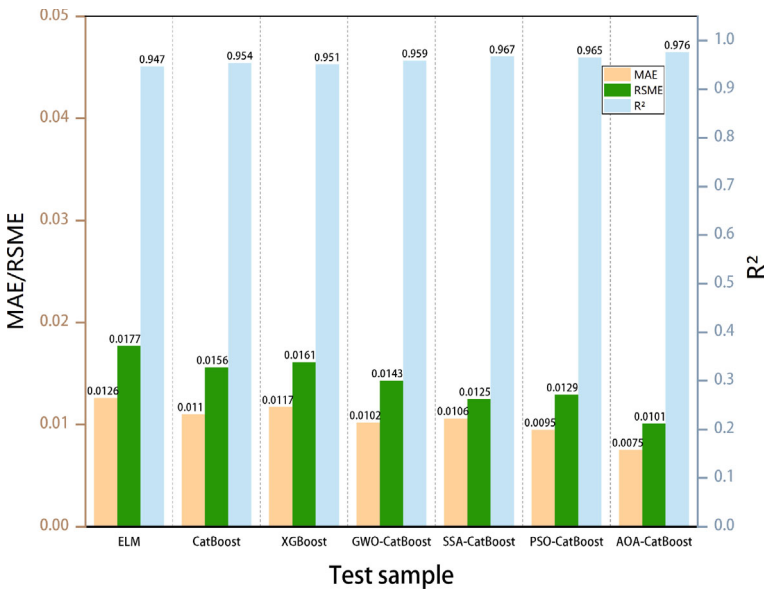


Fig. 1 Comparison of evaluation indicators under different models

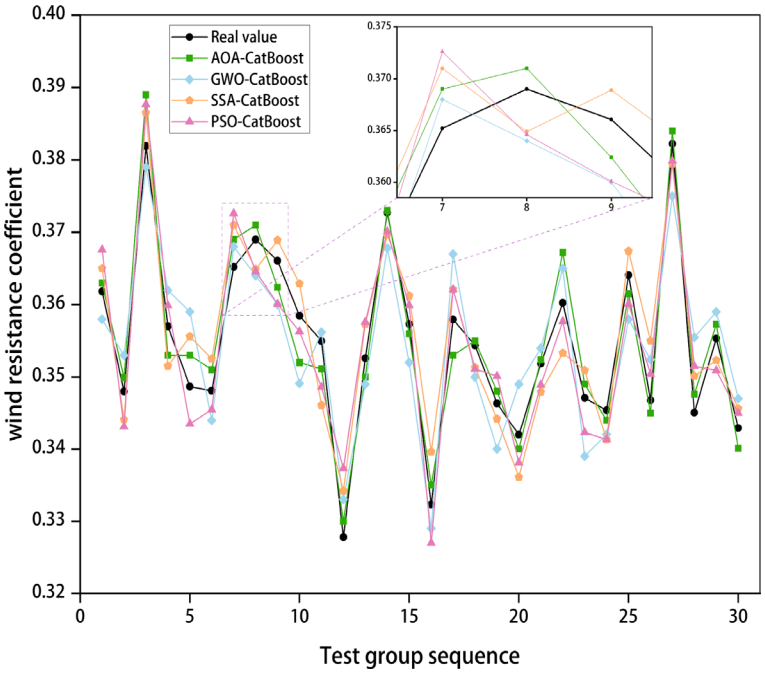


Fig. 2 Comparison of predictions of the four models in the test set

simulated C_d vary consistently in the predicted distribution of the test set samples, especially in the specific shape combinations showing high accuracy. Compared with the other three prediction models, the C_d prediction method proposed in this paper has less bias and better fit. This indicates that the method proposed in this paper has significant advantages in predicting C_d for cars with different aerodynamic profiles.

5 Conclusions

In this paper, we have conducted an in-depth discussion around the problem of automobile wind resistance prediction, and utilized advanced techniques and methods to accurately and exhaustively present the significant laws between the key characteristic parameters and wind resistance coefficients of a certain MPV passenger sedan. By quickly, accurately and reliably predicting the Cd of this tram under various aerodynamic profiles, we have comprehensively examined the method proposed in the paper, which fully demonstrates its excellent advantages and practical value in engineering practice.

Firstly, the MPV wind resistance simulation model is constructed in the article, and after establishing the computational domain and setting the solution conditions, the CFD simulation is completed to determine the characteristic parameters. The CFD simulation is carried out to obtain sample data under several combinations of different characteristic parameters. In the in-depth study, the article focuses on the prediction ability of the AOA-CatBoost prediction model on this MPV. By comparing the results of all hybrid prediction models on the test set, we find that AOA-CatBoost achieves the optimal values in R^2, RMSE, and MAE, and also has a higher goodness-of-fit. Combining all the studies, we conclude that the MPV wind resistance prediction model constructed by AOA fused with CatBoost performs best in terms of prediction accuracy and stability. This research result is expected to better assist in predicting the Cd value of unknown aerodynamic shape combinations, thus improving the limitations of the current tram shape design.

Acknowledgements This work is partially supported by Key Research and development project of Guangxi Province (No. AA23062066).

References

1. Urquhart M, Ljungskog E, Sebben S (2022) Surrogate-based optimisation using adaptively scaled radial basis functions. Appl Soft Comput J 88:106050
2. Bakshi S, Jawad B, Arslan S, Liu L, Yee K (2016) Optimization of modified car body using mesh morphing techniques in CFD. SAE Technical Paper 2016-01-0009
3. Shen SR, Han T, Pang JC (2024) Car drag coefficient prediction using long–short term memory neural network and LASSO. Measurement 225:113982
4. Prokhorenkova L, Gusev G, Vorobev A, Dorogush AV, Gulin A (2018) CatBoost: unbiased boosting with categorical features. In: 32nd Conference on neural information processing systems (NIPS) 2018, Montreal, CANADA, pp 6639–6649
5. Abualigah L, Diabat A, Mirjalili S, Abd EM, Gandomi AH (2021) The arithmetic optimization algorithm. Comput Methods Appl Mech Eng 376:113609
6. Ruqayah DA, Almunem, Dhiaa HM, Haider TH, Tamer K (2024) A novel method for modeling of photovoltaic modules based on arithmetic optimization algorithm and cuckoo search. Optik 298:171591
7. Murat A, Mehmet B (2022) Realization of Turkey's energy demand forecast with the improved arithmetic optimization algorithm. Energy Rep 8(9):18–32
8. Rastgoo A, Khajavi H (2023) A novel study on forecasting the airfoil self-noise, using a hybrid model based on the combination of CatBoost and arithmetic optimization algorithm. Expert Syst Appl 229(A):120576
9. de Souza F, Raeesi A, Belzile M, Caffrey C, Schmitt A (2019) Investigation of drag reduction technologies for light-duty vehicles using surface, wake and underbody pressure measurements to complement aerodynamic drag measurements. SAE Int J Adv Curr Pract Mobil 1(3):1233–1250
10. Kounenis C, Bonitz S, Ljungskog E, Sims-Williams D, Lofdahl L, Broniewicz A, Larsson L, Sebben S (2016) Investigations of the rear-end flow structures on a Sedan car. SAE Technical Paper 2016-01-1606
11. Pan TY, Zhang S, Li FD, Chen JL, Li AM (2023) A meta network pruning framework for remaining useful life prediction of rocket engine bearings with temporal distribution discrepancy. Mech Syst Signal Process 195:110271

Open Access This chapter is licensed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License (http://creativecommons.org/licenses/by-nc-nd/4.0/), which permits any noncommercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if you modified the licensed material. You do not have permission under this license to share adapted material derived from this chapter or parts of it.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.



Yaw Rate Synchronisation Steering Control of Multi-axle Heavy Vehicle



Mingyuan Han, Zhong Jin, Mingde Gong, Jingming Chang, Wei Zhang, and Dingxuan Zhao

Abstract Steering control is an important research direction of multi-axle heavy vehicles. A yaw rate synchronisation control strategy is proposed based on the steering dynamic model of multi-axle heavy vehicle. This strategy obeys Ackerman's law and the yaw rate is controlled by adjusting the steering angle of each rear wheel synchronously to maintain its consistency with the rotation angular speed of the vehicle to help the operator obtain a stable operation experience. This control strategy also solves the limitation of the yaw rate by the traditional zero side-slip angle control strategy and further improves the steering manoeuvrability of multi-axle heavy vehicles. The simulation results show that compared with zero side-slip angle control, the proposed control strategy reduces the backward shift of the vehicle turning centre under the same input conditions, improves the vehicle yaw rate, reduces the turning radius and improves the vehicle turning mobility.

Keywords Lateral dynamics · Yaw rate synchronisation control · Multi-axle steering · Zero side-slip angle control

1 Introduction

Vehicle steering has two basic theories: tire and yaw moment steering. Tire steering is divided into front wheel and multi-wheel steering. Different from front wheel steering, multi-wheel steering increases the steering combination with each rear wheel. Therefore, the number of corresponding variables is increased in the dynamic model, which provides a variety of feasible control strategies for the steering model. Most of the early strategies are to study the control strategy that the other steering wheels have a certain functional relationship with the front wheel angle [1, 2]. These studies meet the problem that the subjective evaluation of control mode is difficult to unify with the objective experience. With the increase in research on multi-axle

M. Han · Z. Jin · M. Gong · J. Chang · W. Zhang · D. Zhao (✉)
School of Mechanical Engineering, Yanshan University, Qinhuangdao, China
e-mail: zdx-yw@ysu.edu.cn

© The Author(s) 2026

H. Saman et al. (eds.), *The 9th International Conference on Advances in Construction Machinery and Vehicle Engineering*, Lecture Notes in Mechanical Engineering,
https://doi.org/10.1007/978-981-96-5527-4_112

heavy-duty vehicles, the control strategies usually have three schemes: side-slip angle control, ideal yaw rate tracking control and turning centre distance adjustment control. The main purpose of sideslip angle control is to control the side-slip angle to zero. Zhang [3] analyses the two degrees of freedom (2-DOF) model of multi-axle vehicles by using the zero side-slip angle control strategy and calculated the yaw rate, lateral acceleration and other parameters by using the specific vehicle data. Then Wang [4] analyses the motion characteristics of vehicles again and indicated that the lateral acceleration attenuation is large when driving at a high speed. Liu [5] combined sliding mode control theory with zero side-slip angle control to improve steering trajectory. These studies are based on the side-slip angle to investigate various aspects and have achieved promising results but the direct measurement of side-slip angle is an expensive and difficult problem [6]. Chen [7] proposes a switching strategy for advantage determination based on kinematic and dynamic schemes, utilizing a VSA soft measurement method for the switching strategy, and conducts hardware-in-the-loop experiments. Song [8] integrated satellite GPS signals with vehicle Inertial Navigation Systems (INS) to enhance signal quality in the presence of model uncertainties. By employing an adaptive square-root Cubature Kalman Filter (CKF), they estimated the vehicle's centroid slip angle. Experimental validation under real-world conditions verified the superiority of this approach. Obtaining the yaw rate compared with the side-slip angle is easy. Accordingly, implementing the yaw rate control method on vehicles is substantially easy. Du [9] designed a super-twisting sliding mode method for a three-axis vehicle under extreme operating conditions, enabling it to track the desired reference yaw rate effectively. Zhang [10] designed a trajectory tracking control method for all-wheel steering vehicles through Model Predictive Control (MPC) theory and developed a yaw rate tracking controller using sliding mode control theory. Finally, coordination between centroid slip angle control and yaw rate control was achieved through weight adjustment. Yu [11] proposed a chassis coordinated control strategy that integrates torque coordination with active rear steering. This strategy aims to effectively reduce body roll motion while tracking the target values of yaw rate and sideslip angle. Deng [12] addressed the issues of parameter uncertainties and disturbances in the model by proposing a nonlinear integral sliding mode control method based on a Kalman filter extended state observer. This method controls the steering angles of the front and rear wheels to track the desired yaw rate and has been validated through hardware-in-the-loop experiments. Xia [13] proposed a control strategy based on differential flatness and linear active disturbance rejection control to address issues such as underactuation, nonlinearity, strong coupling, disturbances, and measurement noise in the tracking control of four-wheel steering (4WS) vehicles. The turning centre control strategy takes the turning centre as output to ensure the preset turning radius. Li [14] set the target rotation centre and used the parallel control algorithm of steering angle and actuator PID to realise the given rotation path. Zhang [15] proposed six steering modes of multi-axis robot car according to the different positions of rotation centre. Chen [16] used the offset of the turning centre as the control output to derive the minimum turning radius of the vehicle. SEO [17] fixed the turning centre on an axis

extension line, added driver operation module in the model and conducted model predictive steering control.

Zero side-slip angle control is the most widely used strategy. The steady driving direction of the vehicle can be consistent with the longitudinal axis by using this strategy. This strategy has two shortcomings. Firstly, the yaw rate is limited, and the manoeuvrability steering operation is reduced. Secondly, obtaining the side-slip angle directly and economically is difficult; thus, the accuracy and real-time performance are unsatisfied. The advantage of vehicle turning centre control is that it can control the turning centre of the vehicle to determine a clear turning path. The disadvantage is that the turning centre is no longer related to the vehicle speed; thus, this strategy is only suitable for low speed steering. Only E. Esmailzadeh proposed a model for the yaw rate control strategy [18]. However, this model does not provide an optimal index for the model coefficients in combination with the target state of vehicle operation. The DYC strategy requires each wheel to conduct relatively independent drive and brake control, which no longer abides by Ackerman's law. This strategy also improves vehicle mobility and increases tire wear simultaneously. This paper is an extension of the yaw rate control strategy. On the basis of vehicle dynamics theory, this paper proposes the yaw rate synchronisation control (YRSC) strategy for multi-axle heavy-duty vehicles and completes the analysis of vehicle steering characteristics, such as vehicle cornering angle, lateral acceleration and rotation centre response. This paper also compares the aforementioned steering characteristics with those of the zero side-slip angle control strategy to show the uniqueness of the control strategy.

2 Steering Model of Multi-axle Heavy Vehicles and YRSC Strategy

2.1 2-DOF Steering Dynamics Model

Referring to the general vehicle body coordinate system, the 2-DOF steering model is established as shown in Fig. 1.

Descriptions of the parameters in Fig. 1 are shown in the nomenclature part.

The vehicle maintains the speed V. The side-slip angle β is usually small. Therefore, the projection of velocity V on the x-axis can be directly equal to itself. Newton's second law indicates that the turning motion can be calculated as follows:

$$\begin{cases} \sum_{i=1}^{n} F_i = ma_y = mV(\dot{\beta} + \omega_z) \\ M_z = \sum_{i=1}^{n} F_i L_i = I_z \dot{\omega} \end{cases} \tag{1}$$

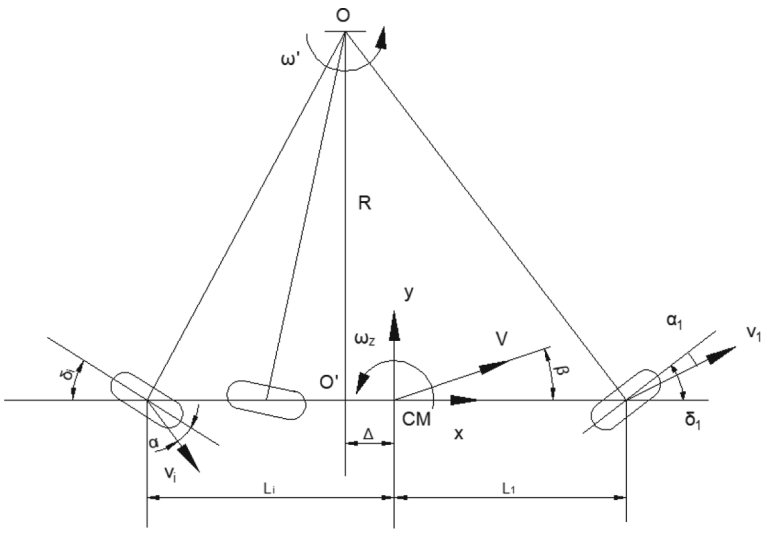


Fig. 1 Steering model of 2-DOF

where F_i is the i-Axis vehicle side force in Y axis, and a_y is the lateral acceleration. Considering that the steering angle of the tire is usually small, its value is consistent with the tire cornering force. Where m is the centroid mass. I_z represents the moment of inertia around the z axis.

For any tire, the cornering force is as follows:

$$F_i = -C_i \alpha_i \tag{2}$$

where C_i is the comprehensive cornering stiffness of the corresponding tire, and α_i is the cornering angle of the i-Axis tire. According to the definition, α_i can be expressed as:

$$\alpha_i = \beta + \frac{\omega_z L_i}{V} - \delta_i \tag{3}$$

where δ_i represents the steering angle of the i-th tire, β represents the side-slip angle of the vehicle's mass centre, ω_z represents the yaw rate and L_i represents the distance between i-Axis and mass centre.

The steering angle follows the Ackerman constraint to reduce tire wear:

$$\frac{\delta_i}{\delta_1} = \frac{L_i - \Delta}{L_1 - \Delta} \tag{4}$$

where Δ represents the distance between the projection point of the turning centre on the longitudinal axis of the vehicle body and the centre of mass. The geometric relationship in the figure indicates the following:

$$\Delta = R * \sin\beta \cong R * \beta \tag{5}$$

The following results can be obtained by substituting Eqs. (2)–(4) into Eq. (1).

$$\begin{cases} mV\left(\dot{\beta} + \omega\right) = \sum_{i=1}^{n} C_i \left(\frac{L_i - \Delta}{L_1 - \Delta}\delta_1 - \beta - \frac{\omega L_i}{V}\right) \\ I\,\dot{\omega} = \sum_{i=1}^{n} C_i L_i \left(\frac{L_i - \Delta}{L_1 - \Delta}\delta_1 - \beta - \frac{\omega L_i}{V}\right) \end{cases} \tag{6}$$

The transfer functions of the side-slip angle gain and yaw rate gain can be obtained by Laplace transform as follows:

$$\frac{\beta}{\delta_1} = \frac{b_1 s + b_0}{p_2 s^2 + p_1 s + p_0} \tag{7}$$

$$\frac{\omega}{\delta_1} = \frac{o_1 s + o_0}{p_2 s^2 + p_1 s + p_0} \tag{8}$$

The expressions of the coefficients are as follows:

$$\begin{aligned} b_1 &= I_z V\left(\sum C_i L_i - \Delta \sum C_i\right) \\ b_0 &= \Delta\left(\left(\sum_{i=1}^{n} C_i L_i\right)^2 - \sum_{i=1}^{n} C_i \sum_{i=1}^{n} C_i L_i^2\right) + mV^2\left(\Delta \sum_{i=1}^{n} C_i L_i - \sum_{i=1}^{n} C_i L_i^2\right) \\ p_2 &= mV^2 I_z (L_1 - \Delta) \\ p_1 &= \left(I_z V \sum_{i=1}^{n} C_i + mV \sum_{i=1}^{n} C_i L_i^2\right) * (L_1 - \Delta) \\ p_0 &= \left(\sum_{i=1}^{n} C_i \sum_{i=1}^{n} C_i L_i^2 - \left(\sum_{i=1}^{n} C_i L_i\right)^2 - mV^2 \sum_{i=1}^{n} C_i L_i\right) * (L_1 - \Delta) \\ o_1 &= mV^2\left(\sum_{i=1}^{n} C_i L_i^2 - \Delta \sum_{i=1}^{n} C_i L_i\right) \\ o_0 &= V\left(\sum_{i=1}^{n} C_i \sum_{i=1}^{n} C_i L_i^2 - \left(\sum_{i=1}^{n} C_i L_i\right)^2\right) \end{aligned} \tag{9}$$

2.2 YRSC Strategy

The definition of the YRSC strategy is as follows. In the stable steering state, the yaw rate ω_z is consistent with the angular velocity ω' around the turning centre by adjusting the δ_i.

$$\omega_z = \omega' = \frac{V}{R} = \frac{V\beta}{\Delta} \tag{10}$$

The flow chart of the control strategy is shown in Fig. 2.

In Fig. 3, the diagram of the YRSC strategy is presented as.

Under the YRSC strategy, the rotation angle b_1 and b_2 along the z-axis of its coordinate system are equal when the vehicle rotates the angle b_2 along the rotation centre. The control strategy can ensure that the position and attitude of the vehicle relative to the turning centre are constant when the vehicle is steering to provide a stable sense of control for the driver.

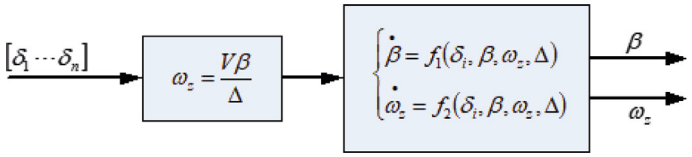


Fig. 2 Schematic diagram of control process

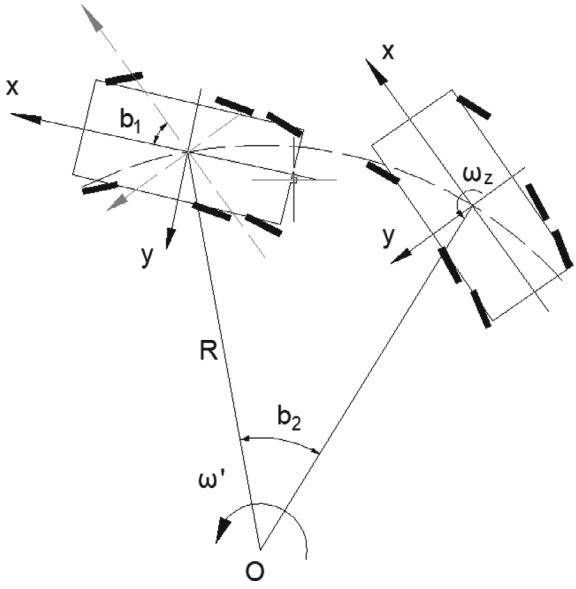


Fig. 3 Schematic of YRSC strategy

Under the YRSC strategy, the following conditions will be met:

$$\begin{cases} \dot{\beta} = 0 \\ \dot{\omega}_z = 0 \\ \omega_z = \dfrac{V}{R} = \dfrac{V\beta}{\Delta} \end{cases} \tag{11}$$

By substituting the above conditions into the system of differential Eqs. (6), the distance Δ from the projection point of the vehicle body to the centre of mass under the steady-state condition can be expressed:

$$\Delta = \frac{mV^2 \sum_{i=1}^{n} C_i L_i^2}{mV^2 \sum_{i=1}^{n} C_i L_i + 2 \sum_{i=1}^{n} \sum_{j=1}^{i} C_i C_j \left(L_i - L_j\right)^2} \tag{12}$$

The body yaw rate gain is as follows:

$$\frac{\omega}{\delta_1} = \frac{V}{\Delta} * \frac{\beta}{\delta_1} \tag{13}$$

The vertical distance R from the turning centre to the longitudinal axis body is calculated as follows:

$$R = \frac{L_1 - \Delta}{\delta_1} \tag{14}$$

3 Simulation Experiments

In Table 1, a five-axle engineering vehicle launched by a company is taken as an example, and its physical parameters are shown.

Table 1 Parameters of a five-axle heavy vehicle

Parameters	Value	Parameters	Value
m (kg)	54,800	I_z (kg * m^2)	195,300
L_1 (m)	4.45	C_1 (kN/rad)	440
L_2 (m)	1.50	C_2 (kN/rad)	440
L_3 (m)	− 0.37	C_3 (kN/rad)	474
L_4 (m)	− 1.72	C_4 (kN/rad)	474
L_5 (m)	− 3.40	C_5 (kN/rad)	472

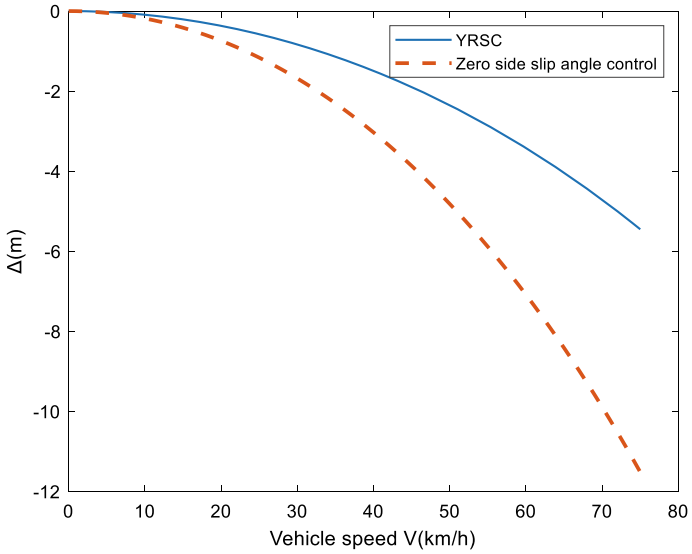


Fig. 4 Comparison of Δ under different control strategies

The speed V range from 0 to 75 km/h.

In the following analysis, the YRSC strategy is compared with the most commonly used zero side-slip angle controler [3] to show the characteristics of YRSC.

The displacement change Δ can be obtained by substituting the parameters in Table 1 into formula (10). The corresponding Δ under zero side-slip angle control strategy can also be obtained by substituting the same parameters into the model. The comparison is shown in Fig. 4. The figure shows that the YRSC strategy has an inhibitory effect on the backward shift of the steering centre.

The gain of the side-slip angle of the centre of mass can be obtained under the steady-state condition by substituting the above parameters and Δ into formula (12), as shown in Fig. 5.

The figure shows that under the YRSC strategy, the gain of the side-slip angle will increase along the negative direction with the increase in vehicle speed. Thus, the side-slip angle under this control strategy will slightly lag behind the longitudinal axis.

The relationship between the steering angle ratio coefficient of each axle and the vehicle speed is shown in Fig. 6.

The figure reveals that when the vehicle speed is low, the steering of the second tire is the same as that of the front wheel, whilst that of the rear three tires is opposite to that of the front wheel. The tire steering angles of the two control strategies are consistent with the front wheels with the increase in vehicle speed. The YRSC strategy changes more smoothly than the zero side-slip angle control strategy.

Figure 7 further shows the ratio of the steering angle of each axle to the front wheel under the YRSC strategy. The vehicle yaw rate gain curve is shown in Fig. 8.

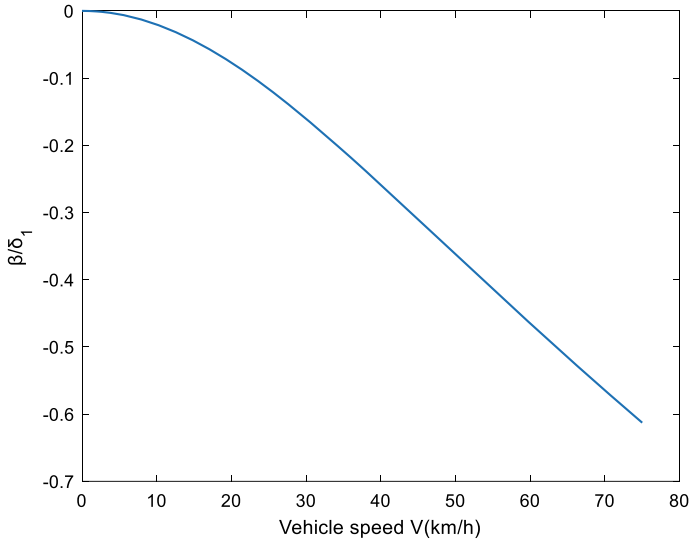


Fig. 5 Gain curve of the side-slip angle of the mass centre

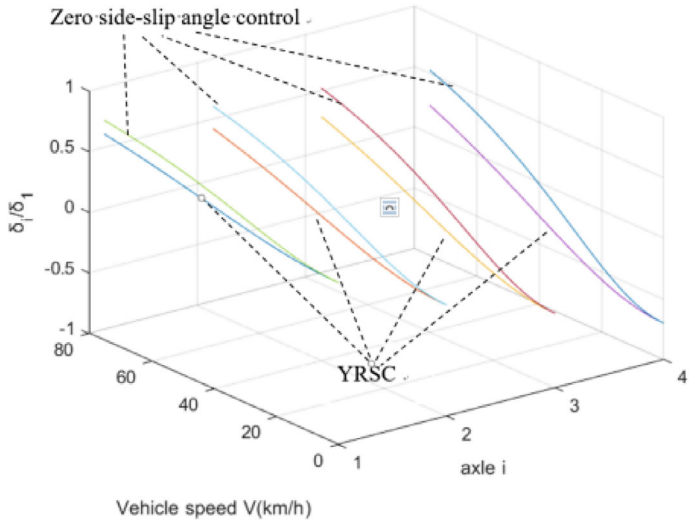


Fig. 6 Comparison of the tire steering angle

It can be seen that the yaw rate gain is slightly larger than that of the zero side-slip angle control strategy by using the YRSC strategy. The yaw rate gain gradually slows down with the increase in speed, and the vehicle begins to show the under-steer trend, which ensures the comfort and smoothness of the vehicle at high speed.

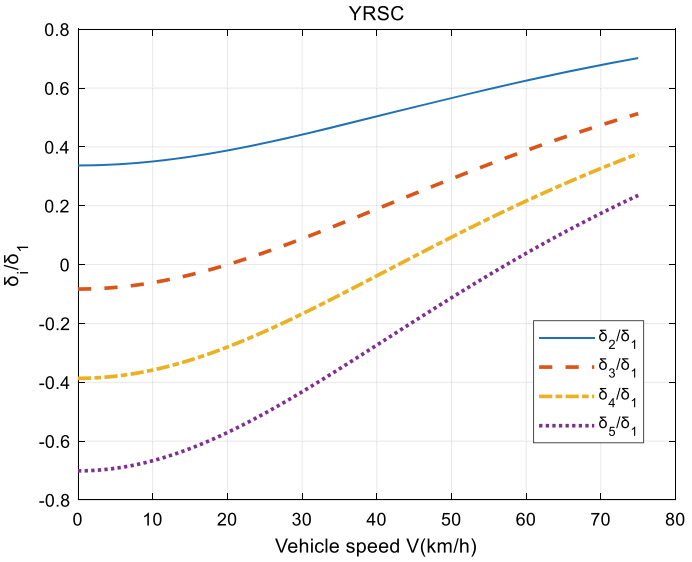


Fig. 7 Relationship of the steering angle of each axle

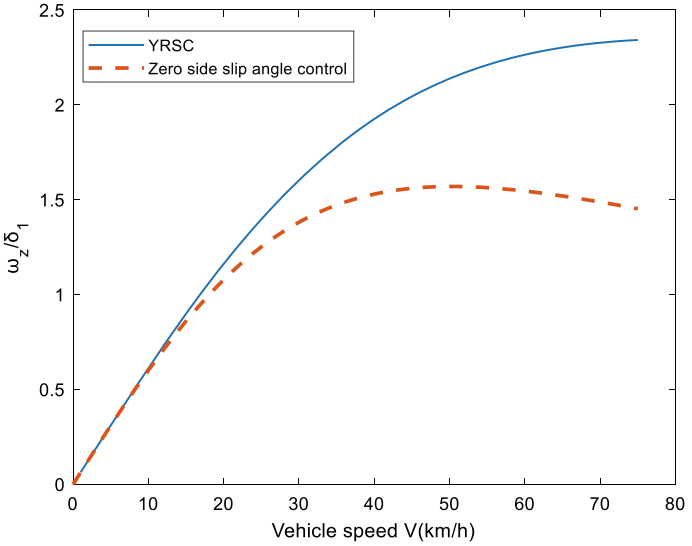


Fig. 8 Comparison of yaw rate gain

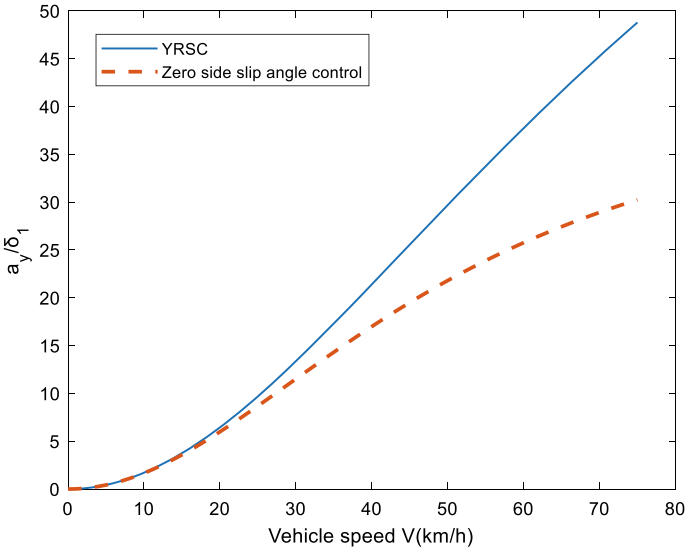


Fig. 9 Comparison of lateral acceleration gain

The lateral acceleration gain curve is shown in Fig. 9. The lateral acceleration of the YRSC strategy is higher than that of the zero side-slip control strategy due to the influence of yaw rate gain and increases with the vehicle speed.

Taking the front wheel rotation angle δ_1 as 0.1–0.4 rad, the reference turning radius R curve can be calculated as shown in Fig. 10.

The figure shows that compared with zero side-slip angle control strategy, the YRSC strategy can obtain small reference radii of gyration at different front wheel steering angles, especially at high vehicle speed and small δ1. This phenomenon shows that the YRSC strategy has high manoeuvrability. The response curve of the front wheel step input under different control strategies for vehicle yaw rate, side-slip angle and lateral acceleration is shown in Fig. 11.

It shows that the step response value of the yaw rate is stable under the YRSC strategy when the vehicle speed reaches 60 km/h, and the overall response value is slightly higher than that of the zero side-slip angle control strategy.

Figure 12 shows that the step response of the side-slip angle of the mass centre decreases with the increase in vehicle speed under the YRSC strategy and finally stabilizes to a negative value. This finding indicates that the body x-axis is slightly ahead of the motion direction of the mass centre and maintains a stable difference value during the turning process.

Figure 13 shows that the unit step response of the lateral acceleration to the front wheel input increases with the speed by using the YRSC strategy. Overall, the stable value of the lateral acceleration of the YRSC strategy is higher than that of the zero side-slip angle control strategy. This also explains the enhanced manoeuvrability of the YRSC strategy in steering.

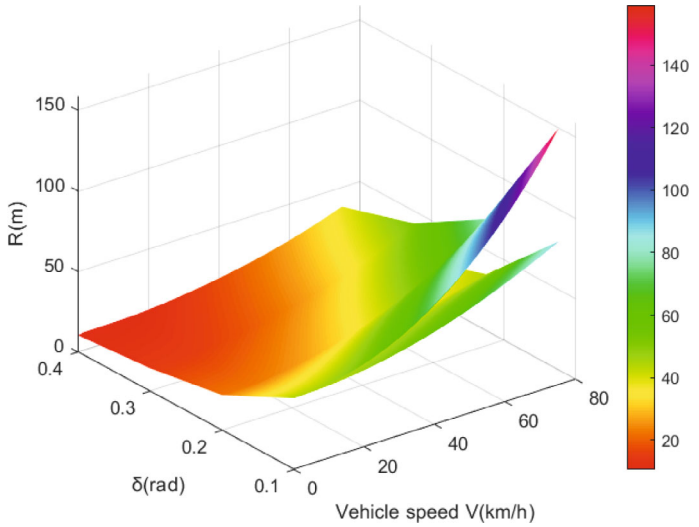


Fig. 10 Comparison of turning radius R

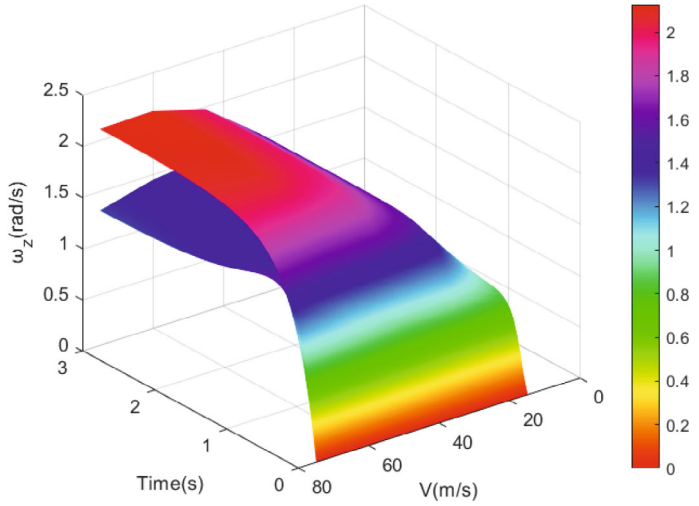


Fig. 11 Comparison of yaw rate response

4 Conclusions

The YRSC strategy of multi-axle heavy-duty vehicles is designed in this paper based on the vehicle dynamics, and the steady-state and transient models under the control strategy are deduced. The steady and transient results of yaw rate, side-slip angle

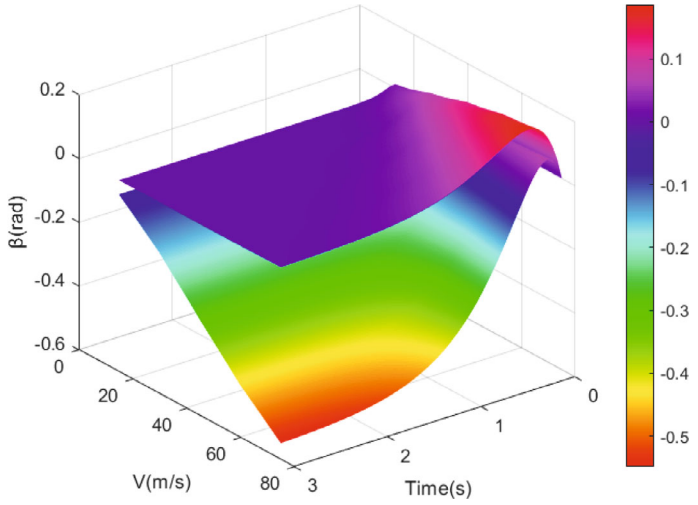


Fig. 12 Comparison of side-slip angle of mass centre response

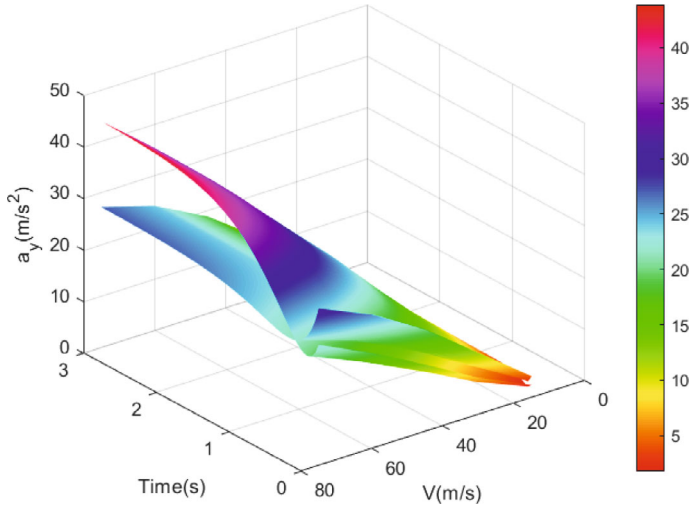


Fig. 13 Comparison of lateral acceleration response

and lateral acceleration are calculated on the basis of an actual vehicle and compared with the zero side-slip angle control strategy.

The YRSC strategy aims to maintain the synchronicity of the vehicle yaw rate with the turning angular speed of the vehicle. This strategy maintains a smooth steering angle ratio of wheels, reduces the backward shift of the turning centre and radius, increases the yaw rate and lateral acceleration and further improves the steering

mobility. However, the control strategy, which has high requirements for the overall stability of the vehicle, increases the yaw rate and lateral acceleration response.

Nomenclature

C_M	Potion of the vehicle mass centre
V	Velocity at the centre of mass of the vehicle
R	Distance from the vehicle turning centre to the body x-axis
O'	Projection point of the vehicle turning centre on the vehicle body
Δ	Distance between the mass centre and the projection point of the rotation centre
O	Vehicle turning centre
α_i	Side-slip angle of the i-axis tire
L_i	Distance from the i-th axis to the centre of mass
ω_z	Vehicle yaw rate
δ_i	Steering angle of the i-axis tire
β	Side-slip angle of the mass centre
ω'	Angular velocity of the vehicle turning.

References

1. Abe M (2015) Vehicle handling dynamics. Elsevier Butterworth-Heinemann, Amsterdam. ISBN 978-0-08-100390-9
2. Guo K, Ba H (1998) Control method of vehicle four wheels steering system. J Jilin Univ Technol 3–5
3. Zhang X, Gao X, Zhao J (2008) The zero sideslip angle proportional control research on multi-axle steering vehicle. Agric Equip Veh Eng
4. Wang Y, Chen N, Gao X (2009) Research on lateral dynamics of multi-axle steering vehicle. J Chongqing Inst Technol (Nat Sci)
5. Liu Q (2014) Research on trajectory of multi-axle steering control system for ALL-terrain cranes. Master, Dalian University of Technology
6. Pan S, Li H, Li J, Xia X, Ma K (2020) Estimation of side-slip angle of vehicle centroid based on unscented Kalman filter. J Milit Transp Univ 40–45
7. Chen J, Wu Q, Ge X, Zhao J (2024) High-performance acquisition for vehicle sideslip angle based on switch strategy. Qiche Gongcheng/Autom Eng 46:346–355. https://doi.org/10.19562/j.chinasae.qcgc.2024.02.017
8. Song R, Fang Y (2021) Vehicle state estimation for INS/GPS aided by sensors fusion and SCKF-based algorithm. In: Mechanical systems and signal processing, vol 150. https://doi.org/10.1016/j.ymssp.2020.107315
9. Du M (2021) Control strategy research on active suspension system of multi-axle emergency rescue vehicles. Doctor, Jilin University, Jilin
10. Zhang Z, Zheng L, Li Y, Li S, Liang Y (2023) Cooperative strategy of trajectory tracking and stability control for 4WID autonomous vehicles under extreme conditions. IEEE Trans Veh Technol 72:3105–3118. https://doi.org/10.1109/TVT.2022.3216486

11. Yu Z, Luo R, Wu P, Hou Z (2024) Chassis coordinated control for lateral stability of four-in-wheel-motor-drive vehicles. Automot Eng 46:1733–1743
12. Deng B, Li W, Wu D, Zhang B, Zhao H (2024) Integral sliding mode control for four-wheel active steering based on Kalman filter extended state observer. Qiche Gongcheng/Autom Eng 46:100–108. https://doi.org/10.19562/j.chinasae.qcgc.2024.01.011
13. Xia Y, Lin M, Zhang J, Fu M, Li C, Li S, Yang Y (2021) Trajectory planning and tracking for four-wheel steering vehicle based on differential flatness and active disturbance rejection controller. Int J Adapt Control Signal Process 35:2214–2244. https://doi.org/10.1002/acs.3315
14. Li Y, Yang L, Yang G (2007) Network-based coordinated motion control of large-scale transportation vehicles. IEEE/ASME Trans Mechatron 12:208–215. https://doi.org/10.1109/TMECH.2007.892830
15. Zhang P, Gao L, Zhu Y (2016) Study on control schemes of flexible steering system of a multi-axle all-wheel-steering robot. Adv Mech Eng 8:1687814016651556. https://doi.org/10.1177/1687814016651556
16. Chen X (2018) Analysis of steering characteristics and multi-objective angle control for multi-axle steering vehicle. Doctor, Jilin University
17. Seo J-H, Oh K-S, Noh H-J (2019) Model predictive control-based steering control algorithm for steering efficiency of a human driver in all-terrain cranes. Adv Mech Eng 11:1687814019859783. https://doi.org/10.1177/1687814019859783
18. Ckermann J (1997) A robust control prevents car skidding. IEEE Control Syst Mag 17:23–31. https://doi.org/10.1109/37.588073

Open Access This chapter is licensed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License (http://creativecommons.org/licenses/by-nc-nd/4.0/), which permits any noncommercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if you modified the licensed material. You do not have permission under this license to share adapted material derived from this chapter or parts of it.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.



A Self-adaptive PID Control Algorithm for Electromagnetic Bearings Based on RBF Adam



Haiwu Zheng, Yinying Ren, Hao Xiong, Zeguang Hu, Dingxuan Zhao, and Ziqi Huang

Abstract Aiming at the control problem of electromagnetic bearing system, in order to improve the dynamic characteristics and anti-interference ability of the system, a single degree of freedom mathematical model is established based on electromagnetic bearings, and an adaptive PID controller is designed based on an improved RBF neural network and Adam algorithm to adjust the system hyperparameters in real time. The experimental results show that under the given parameters of electromagnetic bearings, the adaptive PID controller reduces the maximum overshoot of step response by 26.68% compared to conventional PID controllers; After being subjected to the same 0.1 mm disturbance, the maximum offset of the system decreased by 21.78%, verifying that the controller has good control effect and robustness. The control effect of the electromagnetic bearing system is superior to the conventional PID control algorithm.

Keywords Electromagnetic bearings · RBF neural network · Adam algorithm · Adaptive control

1 Introduction

Traditional mechanical bearings have short lifespan, low rotational speed, and difficult motion sealing, making it difficult to meet special environmental requirements such as high speed and vacuum. As a result, electromagnetic bearings have gradually gained attention in the industrial field [1]. Due to its frictionless and lubrication free properties, it plays an irreplaceable role in some high-precision mechanical equipment. For electromagnetic bearing systems, the ideal control effect can include two aspects: dynamic characteristics and anti-interference ability. On the one hand, the system can quickly respond to changes in input signals or reference positions, achieve lower response times and smaller overshoot, and achieve accurate position

H. Zheng (✉) · Y. Ren · H. Xiong · Z. Hu · D. Zhao · Z. Huang
Yanshan University, Qinhuangdao, Hebei, China
e-mail: zhw@ysu.edu.cn

© The Author(s) 2026

H. Saman et al. (eds.), *The 9th International Conference on Advances in Construction Machinery and Vehicle Engineering*, Lecture Notes in Mechanical Engineering,
https://doi.org/10.1007/978-981-96-5527-4_113

control, so that the rotor can be accurately suspended in the desired position and achieve the required position accuracy; And at the same time, it can automatically adjust the control strategy according to the dynamic characteristics and parameter changes of the system to adapt to different working conditions and requirements. On the other hand, the system can resist the influence of external interference and parameter changes, maintain control stability and accuracy, and be able to remain stable near a given working point or expected position.

Due to the open-loop instability of electromagnetic bearings, closed-loop control must be adopted to ensure their normal operation. The choice of control method has a decisive impact on the working performance of electromagnetic bearings [2]. The traditional PID control method has been widely applied in the control of electromagnetic bearings, where the stiffness of the magnetic bearing can be adjusted by changing the proportional coefficient, the damping characteristics can be adjusted by dividing the process, and integral feedback can improve the static stiffness of the system [3]. Due to the particularity of the controlled object, the mathematical models established by classical control theory cannot reflect the reality of the system well. Although some improved algorithms for PID algorithm have emerged successively, such as fuzzy PID control [4, 5], improved incomplete differential control method [6], expert PID method [7], etc., the control effect is still limited by some limitations:

(1) Nonlinear systems are difficult to model: Electromagnetic bearing systems exhibit nonlinear characteristics at different operating points and conditions, making it challenging to establish accurate mathematical models. Traditional PID algorithms are usually designed based on linear models, which limits their control effectiveness for nonlinear systems.
(2) Poor robustness: Traditional PID control algorithms are highly sensitive to changes and disturbances in system parameters. During the operation of electromagnetic bearings, they may be affected by external environmental factors such as eddy currents and temperature changes, which can cause changes in system parameters. Therefore, the parameters of the PID controller need to be adjusted accordingly to adapt to the new operating conditions.
(3) High order control is difficult to achieve: The dynamic characteristics of electromagnetic bearing systems may have high-order complexity, and traditional PID algorithms are often unable to meet the control requirements of such high-order systems. When faced with complex dynamic characteristics, PID controllers may not provide sufficient control performance.

In recent years, artificial intelligence algorithms represented by deep learning have developed rapidly due to their powerful ability in high-dimensional information processing. Neural networks can establish nonlinear models by learning sample data, thus better adapting to the nonlinear characteristics of electromagnetic bearing systems. In addition, neural network control algorithms can compensate for changes and disturbances in system parameters through online adaptive learning, thereby improving the robustness of the system. In addition, neural networks can achieve higher-order controller structures, providing stronger control performance for the system.

Aiming at the problem of insufficient robustness of traditional PID algorithms, an innovative adaptive PID control strategy is proposed by combining Adam algorithm and RBF neural network to improve the robustness of the control system:

(1) From a mathematical perspective, an accurate model of electromagnetic bearings has been constructed, providing a solid theoretical foundation for the design of subsequent control strategies. At the same time, this article also designs an adaptive PID controller and explores in depth how to construct an effective control framework in the controller structure, especially the optimization design for the characteristics of electromagnetic bearings.
(2) By introducing the Adam algorithm for adaptive adjustment of PID parameters, the excellent performance in the field of deep learning is incorporated into the control strategy of electromagnetic bearings. This method not only enables the controller to converge quickly, but also greatly improves the robustness of the system.
(3) Introducing RBF neural network for system identification to obtain more accurate model information. Compared to traditional model identification methods, RBF neural networks can better capture the nonlinear characteristics of electromagnetic bearings and provide more accurate information for control strategies. This method, combined with RBF neural network, further enhances the adaptability and accuracy of the control strategy.

In the Results and Discussion section, the effectiveness of the proposed control strategy was verified in detail through experimental mechanical devices. The experimental simulation results clearly demonstrate the significant advantages of this strategy in increasing speed, reducing vibration, and enhancing stability.

2 Mathematical Model of Electromagnetic Bearings

A typical eight pole electromagnetic bearing can be simplified into the model shown in Fig. 1 [8]. Under the magnetization effect of the stator winding, two opposing magnetic controllable electromagnets appear inside the bearing. By adjusting the magnitude of the four electromagnetic forces, the rotor is suspended in the air gap and kept stable.

Neglecting factors such as air gap leakage and iron core reluctance, the electromagnetic force generated by an electromagnet in an eight pole electromagnetic bearing can be derived by calculating the partial derivative of the air gap based on Ampere's loop theorem and the energy stored in the magnetic field:

$$F = \frac{1}{4}\mu_0 N^2 A \frac{I^2}{d^2}\cos\alpha \tag{1}$$

In the formula:

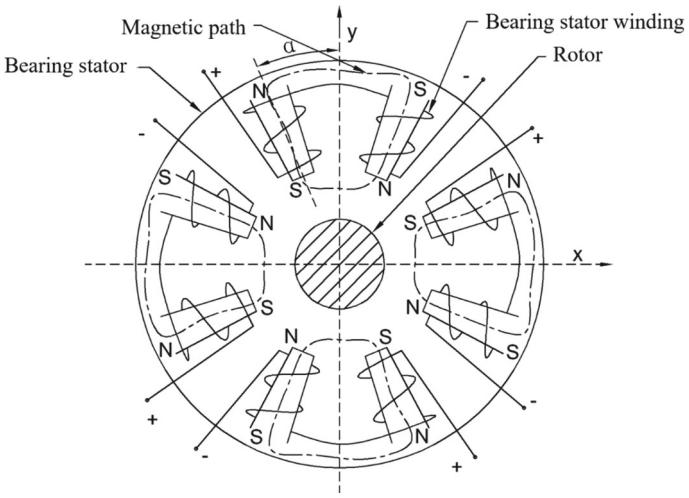


Fig. 1 Octupole electromagnetic bearing structure

μ_0 is the air gap permeability, $\mu_0 = 4\pi \times 10^{-7}$ H/m;

N is the total number of turns of the coils on the two magnetic poles that make up the electromagnet;

A is the effective projection area of one magnetic pole of the electromagnet on the rotor;

I is the current magnitude on the electromagnetic coil;

d is the size of the air gap between the electromagnet and the rotor;

The magnetic pole angle with α being 1/2;

The control of electromagnetic bearings usually adopts a differential winding drive strategy that uses bias current to set the bias magnetic field [9]. Two electromagnets on a single degree of freedom of the electromagnetic bearing are placed in a control system for overall differential control. Taking the two electromagnets in the x-direction in Fig. 1 as an example, the control principle is shown in Fig. 2 [10].

Assuming that the air gap between the rotor and the electromagnet at the balance position is δ_0, when the actual position of the rotor has an offset of δ, the air gap between the rotor and the electromagnet at one side changes from the rated air gap δ_0 to $\delta_0 - \delta$, and the air gap between the rotor and the electromagnet at the other side changes to $\delta_0 + \delta$. At this time, the sensor can detect the offset and convert it into an electrical signal and transmit it to the controller. The controller calculates the control current i_δ internally, and generates two output currents according to the control current i_δ and the preset offset current i_0, one of which is $i_0 - i_\delta$, the other is $i_0 + i_\delta$, and then sends them to two electromagnet coils through power amplifiers

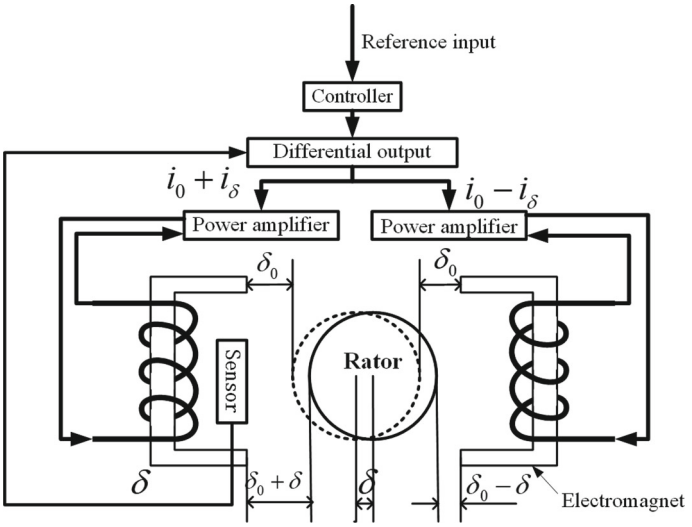


Fig. 2 Control principle of electromagnetic bearing in x direction

to adjust the electromagnetic force generated by the electromagnet, so that the rotor position offset decreases and maintains stable suspension.

By substituting the above parameters into the electromagnetic force formula (1), the electromagnetic forces generated by the two electromagnets in the x-direction can be obtained. Ignoring the coupling between the degrees of freedom and the influence of gyroscopic effects, the difference between the two forces can be used to obtain the resultant force formula of the two electromagnetic attraction forces acting on the rotor in the x-direction at this time:

$$F_x = \frac{1}{4}\mu_0 N^2 A\left[\frac{(i_0 - i_\delta)^2}{(\delta_0 - \delta)^2} - \frac{(i_0 + i_\delta)^2}{(\delta_0 + \delta)^2}\right]\cos\alpha \tag{2}$$

In practical applications, during normal suspension, the rotor is always in the vicinity of the equilibrium position. To simplify control, Eq. (2) is Taylor expanded at the equilibrium position ($\delta = 0, i_\delta = 0$) and the infinitesimal quantities above second order are omitted

$$F_x = k_i i_\delta + k_\delta \delta \tag{3}$$

In the formula (3), k_i and k_δ are the current stiffness coefficient and displacement stiffness coefficient of the electromagnetic bearing in the x-direction, respectively. Their specific expressions are:

$$k_i = \frac{\mu_0 N^2 A i_0 \cos\alpha}{{\delta_0}^2} \tag{4}$$

$$k_\delta = \frac{\mu_0 N^2 A {i_0}^2 \cos\alpha}{{\delta_0}^3} \tag{5}$$

Ignoring the external interference on the electromagnetic bearing rotor, according to Newton's second law, the motion equation of the rotor is:

$$m\ddot{\delta} = k_i i_\delta + k_\delta \delta \tag{6}$$

In the formula, m is the mass of the rotor. By performing Laplace transform on Eq. (6), the transfer function formula for controlling current $i_\delta(t)$ in the coil to rotor offset $\delta(t)$ in the x-direction can be obtained:

$$G(s) = \frac{k_i}{ms^2 - k_\delta} \tag{7}$$

For the convenience of control system simulation, it is necessary to discretize the continuous system. Using the zero pole matching method to perform z-transform discretization on Eq. (7), the discrete transfer function equation can be obtained:

$$G(z) = \frac{\frac{k_i}{2k_\delta}(m-2)\left(z^{-1} + z^{-2}\right)}{1 - mz^{-1} + z^{-2}} \tag{8}$$

Among them:

$$m = e^{\sqrt{\frac{k_\delta}{m}}T} + e^{-\sqrt{\frac{k_\delta}{m}}T}$$

In the formula, T is the sampling period.

The difference equation of the system can be obtained from the discrete transfer function Eq. (8) as follows:

$$\begin{aligned}\delta(kT) - m\delta[(k-1)T] + \delta[(k-2)T] &= \frac{k_i}{2k_\delta}(m-2)i_\delta[(k-1)T] \\ &+ \frac{k_i}{2k_\delta}(m-2)i_\delta[(k-2)T]\end{aligned} \tag{9}$$

In the formula: $\delta(kT)$ is the offset of the rotor at time kT; $i_\delta[kT]$ controls the current magnitude at kT time.

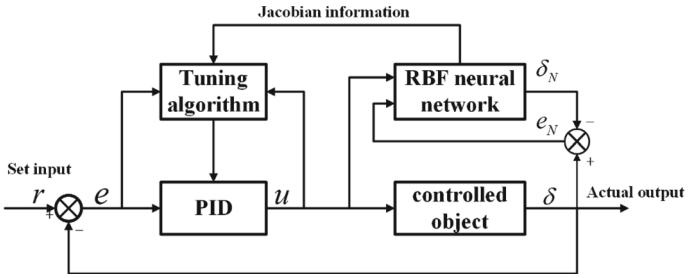


Fig. 3 Schematic diagram of adaptive PID controller structure

3 Adaptive PID Controller

3.1 Controller Structure

The structural principle of the adaptive PID controller designed in this article is shown in Fig. 3, where r is the reference input; δ is the actual output of the controlled object, obtained through sampling in industrial applications; e is the control error, and $e = r-\delta$; u is the control signal output by the controller; δ_N is the output of RBF neural network; e_N is the identification error of RBF neural network, and $e_N = \delta - \delta_N$. Due to the nonlinear force displacement and force current characteristics of electromagnetic bearings, as well as factors such as magnetic saturation and hysteresis of iron core materials, the electromagnetic bearing system has nonlinearity and parameter uncertainty, making it difficult to obtain an accurate mathematical model. Therefore, the adaptive PID controller designed in this paper uses RBF neural network to identify the controlled object online, providing the controller with approximate Jacobian information of the controlled object (i.e. the sensitivity information of the output of the controlled object to the input control information), so as to adjust the PID control parameters in real time through the tuning algorithm.

3.2 PID Controller

This article uses incremental PID control, and its control algorithm is:

$$\begin{aligned} u(k) = u(k-1) + K_P[e(k) - e(k-1)] + K_I e(k) \\ + K_D[e(k) - 2e(k-1) + e(k-2)] \end{aligned} \tag{10}$$

In the formula, K_P is the proportional coefficient of PID control;

K_I is the integral coefficient of PID control;

K_D is the differential coefficient of PID control;

$u(k)$ is the output of PID control at time k;

$e(k)$ is the control error of the controlled object at time k, and $e(k) = r(k) - \delta(k)$.

3.3 Adam Algorithm Tuning PID

The adjustment of PID parameters is a nonlinear optimization process, and the simplest and most commonly used nonlinear optimization algorithm is the gradient method. The idea of gradient method is to use the negative gradient direction of the current position as the optimization direction, because this direction is the fastest descent direction of the current position. Traditional gradient descent maintains a single learning rate to update all parameters, and the learning rate does not change throughout the entire adjustment process. However, the Adam algorithm designs independent adaptive learning rates for different parameters by calculating the first-order and second-order moment estimates of the gradient. Therefore, this paper adopts the Adam algorithm to tune the PID control parameters for higher convergence efficiency.

The performance indicators for PID controller parameter tuning are defined as:

$$E(k) = \frac{1}{2}[r(k) - \delta(k)]^2 = \frac{1}{2}[e(k)]^2 \tag{11}$$

The method for adjusting the parameter K_P using the Adam algorithm is as follows:

$$K_P(k+1) = K_P(k) - \gamma \frac{\widehat{m}_p(k)}{\sqrt{\widehat{v}_p(k)} + \varepsilon} \tag{12}$$

In the formula, γ is the learning step size; ε is a very small constant used for numerical stability to prevent division by zero errors during algorithm operation; $\widehat{m}_p$ and $\widehat{v}_p$ are the first and second moment estimates after correcting the bias of parameter K_P gradient, and the correction formula is:

$$\begin{cases} \widehat{m}_p(k) = \frac{m_p(k)}{1-\beta_1{}^k} \\ \widehat{v}_p(k) = \frac{v_p(k)}{1-\beta_2{}^k} \end{cases} \tag{13}$$

In the equation, m_p and v_p are biased first-order and second-order moment estimates of the gradient of parameter K_P; β_1 and β_2 represent the exponential decay rates of the first-order and second-order moment estimates of the gradients of each parameter. The update method for the biased first-order moment estimation and biased second-order moment estimation of parameter K_P gradient is as follows:

$$\begin{cases} m_p(k) = \beta_1 m_p(k-1) + (1-\beta_1)\frac{\partial E(k-1)}{\partial K_P} \\ v_p(k) = \beta_2 v_p(k-1) + (1-\beta_2)\frac{\partial E(k-1)}{\partial K_P} \end{cases} \tag{14}$$

In the formula, according to the structure principle of the adaptive PID controller, there are:

$$\frac{\partial E(k-1)}{\partial K_p} = -e(k-1)\cdot\frac{\partial \delta(k-1)}{\partial u(k-1)}\cdot[e(k-1)-e(k-2)] \tag{15}$$

In the formula, $\frac{\partial \delta(k-1)}{\partial u(k-1)}$ represents the sensitivity information of the output of the controlled object to the input.

Combining Eqs. (12), (13), (14), and (15), it can be concluded that the Adam algorithm is used to tune the PID control parameter K_P as follows:

$$K_I(k+1) = K_I(k) - \gamma\frac{\frac{\beta_1 m_i(k-1)-(1-\beta_1)\cdot e(k-1)\cdot\frac{\partial\delta(k-1)}{\partial u(k-1)}\cdot[e(k-1)]}{1-{\beta_1}^k}}{\sqrt{\frac{\beta_2 v_i(k-1)-(1-\beta_2)\cdot e(k-1)\cdot\frac{\partial\delta(k-1)}{\partial u(k-1)}\cdot[e(k-1)]}{1-{\beta_2}^k}}+\varepsilon} \tag{16}$$

Similarly, the methods for tuning the parameters K_I and K_D of PID control using the Adam algorithm are as follows:

$$K_I(k+1) = K_I(k) - \gamma\frac{\frac{\beta_1 m_i(k-1)-(1-\beta_1)\cdot e(k-1)\cdot\frac{\partial\delta(k-1)}{\partial u(k-1)}\cdot[e(k-1)]}{1-{\beta_1}^k}}{\sqrt{\frac{\beta_2 v_i(k-1)-(1-\beta_2)\cdot e(k-1)\cdot\frac{\partial\delta(k-1)}{\partial u(k-1)}\cdot[e(k-1)]}{1-{\beta_2}^k}}+\varepsilon} \tag{17}$$

$$K_D(k+1) = K_D(k) - \frac{\gamma\frac{\beta_1 m_d(k-1)-(1-\beta_1)\cdot e(k-1)\cdot\frac{\partial\delta(k-1)}{\partial u(k-1)}\cdot[e(k-1)-2e(k-2)+e(k-3)]}{1-{\beta_1}^k}}{\sqrt{\frac{\beta_2 v_d(k-1)-(1-\beta_2)\cdot e(k-1)\cdot\frac{\partial\delta(k-1)}{\partial u(k-1)}\cdot[e(k-1)-2e(k-2)+e(k-3)]}{1-{\beta_2}^k}}+\varepsilon} \tag{18}$$

It can be seen that in the PID control parameter tuning process, only the precise value of the sensitivity information $\frac{\partial \delta(k)}{\partial u(k)}$ of the output of the controlled object to the input cannot be obtained. In industrial applications, the sign function sgn $[\frac{\partial \delta(k)}{\partial u(k)}]$ is often used to approximate it. In this paper, the RBF neural network is used to identify the controlled object and obtain the approximate Jacobian information of the controlled object. By improving the accuracy of the value $\frac{\partial \delta(k)}{\partial u(k)}$, the control accuracy is further improved.

3.4 RBF Neural Network Identification

In recent years, neural networks have been widely studied and applied in the field of control due to their powerful nonlinear mapping and self-learning abilities. The BP (backpropagation) neural network is the most commonly used neural network, but BP networks have the disadvantages of being prone to local minima and low convergence efficiency. RBF networks have been widely used in the field of control because they have overcome the shortcomings of BP networks to some extent [11]. RBF neural network is a three-layer feedforward fully connected network consisting of an input layer, a single hidden layer, and an output layer [12, 13].

According to the differential Eq. (9) of the controlled object, the discrete description of the controlled object can be written as:

$$\delta(k) = f[\delta(k-1), \delta(k-1), u(k-1), u(k-2)] \tag{19}$$

In the equation, $\delta(k)$ and $u(k)$ represent the output and input of the controlled object at the kth sampling time.

The input of RBF network should be selected to have a significant impact on the network output. In this article, the output of the network is an approximate value of the output of the controlled object. Therefore, based on Eq. (19), the input vector of the RBF network is selected as:

$$X = [x_1, x_2, x_3, x_4]^T = [\delta(k-1), \delta(k-1), u(k-1), u(k-2)]^T \tag{20}$$

Assuming the number of neurons in the hidden layer of the RBF network is n and the excitation function is a Gaussian function:

$$\mathrm{h}_p = e^{-\frac{\|X - C_p\|^2}{2b_p{}^2}}, p = 1, 2, \ldots, n \tag{21}$$

Among them, the center vector of the p-th neuron in the hidden layer of the RBF network is:

$$C_p = \left[c_{p1}, c_{p2}, c_{p3}, c_{p4}\right]^T \tag{22}$$

The basis width vector of RBF network is:

$$B = \left[b_1, b_2, \cdots, b_p, \cdots, b_n\right]^T \tag{23}$$

The weight vector of RBF network is:

$$W = \left[w_1, w_2, \cdots, w_p, \cdots, w_n\right]^T \tag{24}$$

The output of RBF network is:

$$\delta_N(k) = w_1 h_1 + w_2 h_2 + \cdots + w_n h_n \tag{25}$$

The performance metric function of RBF network is defined as:

$$E_N(k) = \frac{1}{2}[\delta(k) - \delta_N(k)]^2 = \frac{1}{2}[e_N(k)]^2 \tag{26}$$

The iterative learning algorithm for the output weights, node base width parameters, and node center values of the RBF network is as follows:

$$\begin{cases} w_p(k+1) = w_p(k) - \eta \frac{\partial E_N(k)}{\partial w_p} + \lambda\left[w_p(k) - w_p(k-1)\right] \\ b_p(k+1) = b_p(k) - \eta \frac{\partial E_N(k)}{\partial b_p} + \lambda\left[b_p(k) - b_p(k-1)\right] \\ c_{pq}(k+1) = c_{pq}(k) - \eta \frac{\partial E_N(k)}{\partial c_{pq}} + \lambda\left[c_{pq}(k) - c_{pq}(k-1)\right] \\ p = 1, 2, \ldots, n; q = 1, 2, 3, 4 \end{cases} \tag{27}$$

In the formula, η is the learning rate of RBF network parameters; λ is the inertia coefficient; And according to Eqs. (20), (21), (22), and (23):

$$\begin{cases} \frac{\partial E_N(k)}{\partial w_p} = -e_N(k) \cdot h_p \\ \frac{\partial E_N(k)}{\partial b_p} = -e_N(k) \cdot w_p \cdot h_p \frac{\|X - C_p\|^2}{{b_p}^3} \\ \frac{\partial E_N(k)}{\partial c_{pq}} = -e_N(k) \cdot w_p \cdot \frac{x_q - c_{pq}}{{b_p}^2} \\ p = 1, 2, \ldots, n; q = 1, 2, 3, 4 \end{cases} \tag{28}$$

The purpose of constructing an RBF neural network is to obtain the Jacobian information $\frac{\partial \delta(k)}{\partial u(k)}$ of the controlled object. According to Eqs. (20), (21), and (25), it can be inferred that:

$$\frac{\partial \delta(k)}{\partial u(k)} \approx \frac{\partial \delta_N(k)}{\partial u(k)} \approx \frac{\partial \delta_N(k)}{\partial u(k-1)} = \sum_{p=1}^{n} w_p h_p \frac{c_{p3} - x_3}{{b_p}^2} \tag{29}$$

4 Experimental Results and Discussion

4.1 Experimental Mechanical Equipment

The experimental setup is shown in Fig. 4.

Figure 4 shows the mechanical structure diagram of the combined support test device, which includes the following components. Rotor: represents the rotating part and plays a key role in rotation, represented as a fine circular design in Fig. 4. The rotor is supported by two radial combination bearings and one axial magnetic levitation

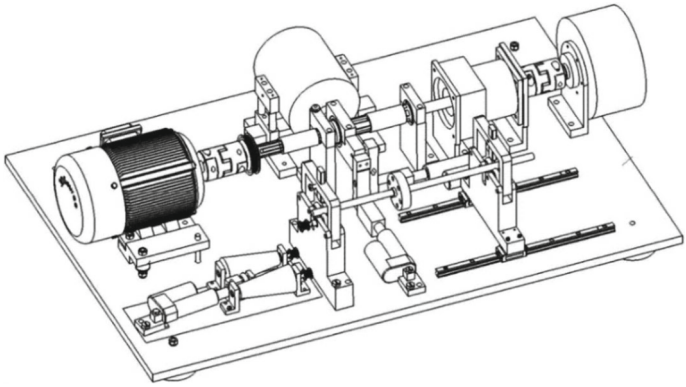


Fig. 4 Schematic illustration of mechanical structure

bearing. Radial combination bearing: Refers to two circular objects connected to the rotor, used to support the rotor in the radial direction. Combination bearings are usually composed of multiple sub bearings to provide better support performance. Axial magnetic levitation bearing: Refers to a linear object connected to the rotor, used to support the rotor in the axial direction. Magnetic levitation bearings use the principle of magnetic force to support the rotor, providing non-contact and frictionless support force. High frequency motor: Refers to a motor built into the rotor, used to drive the rotor to rotate at high speed. High frequency motor is the power source of the rotor, providing mechanical energy through the rotation of the rotor.

The mechanical structure of the device mainly includes a rotor, radial composite bearings, axial magnetic levitation bearings, and high-frequency motors. The rotor is supported by two radial combination bearings and axially supported by an axial magnetic levitation bearing, while also incorporating a high-frequency motor to achieve high-speed rotation of the rotor.

4.2 Experimental Simulation

The relevant parameters of the given electromagnetic bearing are shown in Table 1.

By substituting these parameters into Eq. (7), the approximate transfer function of the electromagnetic bearing in one degree of freedom can be obtained as follows:

$$G(s) = \frac{4.75}{s^2 - 8.55 \times 10^4} \tag{30}$$

Based on the transfer function Eq. (30) and the structural principle of the adaptive PID controller described above, programming simulation was conducted in MATLAB [12], and a comparison was made between conventional PID control and

Table 1 Schematic diagram of adaptive PID controller structure

Parameter name (variable)	Numerical value	Unit
Rotor quality (m)	12	kg
Rated air gap (δ_0)	0.5	mm
Magnetic circuit cross-sectional area (A)	340	mm^2
Number of turns of magnetic pole coil (N)	190	Number of turns
Magnetic pole inclination angle (α)	22.5	°
Bias current (i_0)	3	A

the adaptive PID control designed in this paper. Select sampling period $T = 0.1$ ms during simulation; The number of hidden layer neurons in RBF neural network is $n = 6$ the learning rate of network parameters is $\eta = 0.9$, and the inertia coefficient is $\lambda = 0.2$; The learning step size of Adam algorithm for tuning PID control parameters is $\gamma = 0.9$, the exponential decay rate of first-order moment estimation is $\beta_1 = 0.9$, and the exponential decay rate of second-order moment estimation is $\beta_2 = 0.999$, which is used to prevent the small constant $\varepsilon = 10^{-8}$ from zero division. For the initial biased first-order moment estimation and initial biased second-order moment estimation of the Adam algorithm, both are set to zero. For the center vector matrix C, base width vector B, and weight vector M of the RBF neural network, random numbers within [0,1] are selected. The parameters of traditional PID control in simulation were tuned using MATLAB's PID Tuner toolbox. Adjustments were made for the system and simulation experiments, with values of $K_P = 399{,}178$, K_I=8856, and $K_D = 4{,}391{,}118$, respectively. The obtained PID controller parameters are the optimal values under limited conditions, but they may not necessarily be globally optimal.

The simulation experiment of step response for the control system is shown in Fig. 5. From Fig. 5, it can be seen that the maximum overshoot of the conventional PID control step response is 0.4173 mm, the rise time is 0.5 ms, and the adjustment time is 4.9 ms; The maximum overshoot of the adaptive PID control step response designed in this article is 0.3033 mm, the rise time is 0.5 ms, and the adjustment time is 5 ms. It can be seen that the maximum overshoot of the step response of the adaptive PID controller designed in this article is smaller than that of traditional PID control. The rise time and adjustment time of the step response of the adaptive PID controller designed in this article are similar to those of traditional PID control. Although traditional PID control may be able to adjust for better performance in certain situations, the method proposed in this paper is expected to achieve better control effects in a wider range of systems and conditions.

Simulate the starting process of the electromagnetic bearing, that is, the process of the offset from the initial value of − 0.5 mm to the equilibrium position. At the same time, at t = 25 ms after reaching the equilibrium position, a disturbance with an amplitude of 0.1 mm is applied to the electromagnetic bearing. The simulation experiment output is shown in Fig. 6. From Fig. 6, it can be seen that the conventional PID control takes 11.3 ms to reach the equilibrium position after starting, and the

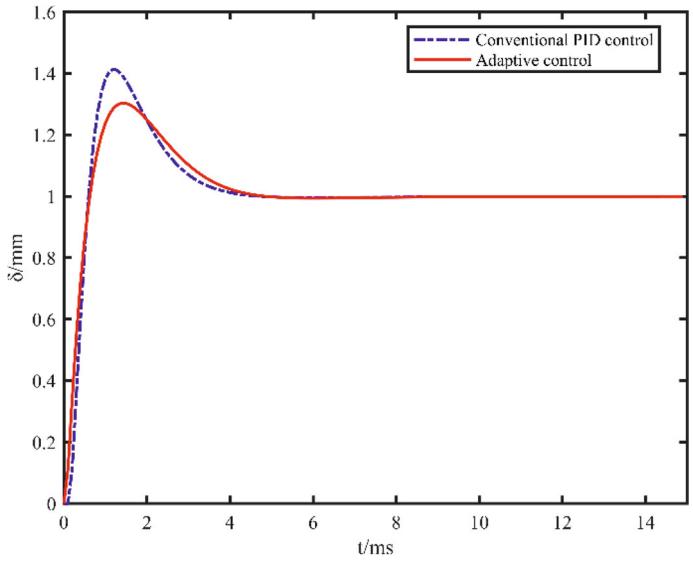


Fig. 5 Step response of control system

maximum deviation after being disturbed by 0.1 is 0.4147 mm. It takes 11.7 ms to return to the equilibrium position after being disturbed by 0.1 mm; The adaptive PID control designed in this article takes 11.8 ms to reach the equilibrium position after starting, with a maximum offset of 0.3447 mm after being disturbed by 0.1 mm, and 11.8 ms to return to the equilibrium position after being disturbed by 0.1 mm. It can be seen that the maximum offset of the adaptive PID controller designed in this article is smaller after being disturbed by 0.1 mm compared to traditional PID control. The time required for the adaptive PID controller designed in this article to reach the equilibrium position after starting and returning to the equilibrium position after being disturbed by 0.1 mm is similar to the time required by traditional PID control.

PID controllers may perform well under specific conditions, but their performance is limited when facing complex systems, nonlinear systems, or strong external disturbances. To overcome these limitations, more advanced control algorithms and methods are needed to improve the robustness and stability of the system.

For other forms of external disturbances, such as direct force acting on the rotor, the system's response may experience greater fluctuations and may lead to edge rubbing and amplifier saturation instability issues. This is because traditional PID controllers may not be able to adjust their control strategies in a timely and effective manner when facing strong external disturbances, resulting in the system being unable to maintain stability. Through experimental simulation, compare the control effect when force is directly applied to the rotor. As shown in Fig. 7.

Under rotor disturbance interference, both conventional PID control algorithm and adaptive PID control algorithm can achieve the goal of stabilizing the speed

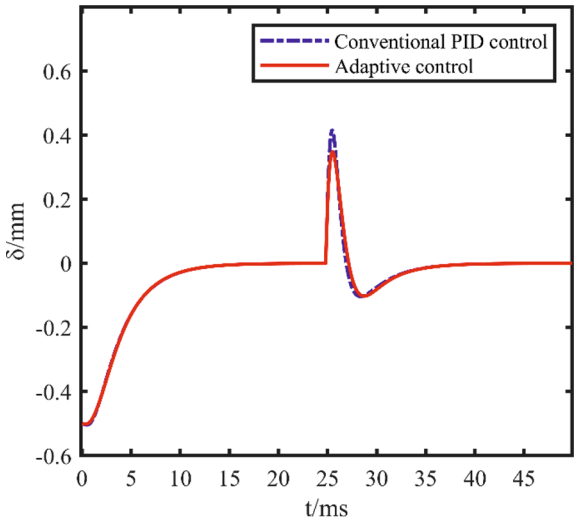


Fig. 6 Step response under interference

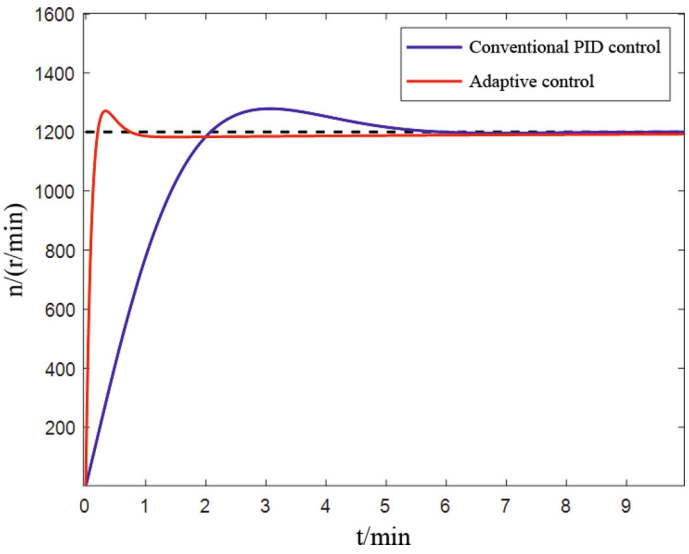


Fig. 7 Control effect under rotor disturbance

at 1200 n (r/min). Compared with conventional PID control algorithms, adaptive PID control algorithms have faster output response speed, smaller overshoot, shorter adjustment time, and better dynamic and steady-state performance of the system.

5 Conclusion

On the basis of in-depth analysis of the electromagnetic bearing control system, this paper constructs an approximate mathematical model of the single degree of freedom of the electromagnetic bearing, and designs an adaptive PID controller based on neural network to identify the controlled object, using Adam algorithm to adjust the PID control parameters in real time. The results show that the adaptive PID controller designed in this paper reduces the maximum overshoot of the step response of the electromagnetic bearing with given parameters by 26.68% compared to traditional PID control, and also reduces the maximum offset after being disturbed by 0.1 mm by 21.78%. Overall, the adaptive PID controller proposed in this article surpasses traditional PID control in terms of dynamic characteristics and anti-interference ability, significantly improving the control effect of electromagnetic bearings. In addition, this study has not yet compared the adaptive PID controller with other advanced PID controllers, which is an important direction for our future research.

References

1. Yuanyan L, Chao S, Bin Z et al (2023) Optimization design of electromagnetic bearing for long-life centrifugal pump. J Phys Conf Ser 2658(1)
2. Guorong Z, Xikui W, Hansen Z et al (2023) Research on time delay vibration reduction of nonlinear systems of rotors electromagnetic bearings. J Dyn Control 21(08):99–109
3. Bin L (2018) Research on active disturbance rejection control of electromagnetic bearing systems. Harbin Engineering University
4. Ye Y, Hu J (2020) Design of magnetic suspension rotor system controller based on fuzzy PID dual mode control. Electr Technol 12:15–16+19
5. Wang B, Geng H, Li H et al (2021) Particle swarm optimization-based fuzzy PID controller for stable control of active magnetic bearing system. In: Journal of Physics: conference series, the 5th international conference on mechanical, aeronautical and automotive engineering (ICMAA), Singapore, vol 1888, no 1
6. Yao L (2023) Research on static suspension control and dynamics of active magnetic bearings. Chongqing University of Technology
7. Lei Y, Li Q, Duan J (2007) Research on expert PID control of a new type of magnetic levitation rotor system. J Eng Des 1:22–26
8. Zheng Z (2014) Research on the application of single neuron adaptive PID controller in active magnetic levitation system. Mechan Des Manuf 4:268–271
9. Su Y, Ma Y, Shi Q et al (2018) Optimization of PID parameters for magnetic levitation water pumps based on BP neural network model. Fluid Mach 46(1):20–24+33
10. Kingma DP, Ba J (2014) Adam: a method for stochastic optimization. Computer Science
11. Yining Z, Zhongkui S (2023) Control of stability in semi-passive robot based on RBF neural network. Chaos, Solit Fract Interdiscipl J Nonl Sci Nonequilibrium Complex Phenom 173:113623
12. Xin Z, Ying Q (2023) Fuzzy sliding mode control of manipulator based on disturbance observer and RBF neural network. Autom Control Comput Sci 57(2):123–134
13. Jialin L ,Tao P, Shen Z et al (2023) Improved PID controller based on BP neural network. J Phys Conf Ser 2479(1)

Open Access This chapter is licensed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License (http://creativecommons.org/licenses/by-nc-nd/4.0/), which permits any noncommercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if you modified the licensed material. You do not have permission under this license to share adapted material derived from this chapter or parts of it.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.



Research on Active Suspension Preview Control of Vehicles Based on Real-Time Road Elevation Information

Dingxuan Zhao, Liqiang Wang, Lirong Yang, Jinming Chang, Haoyu Jiang, Yang Li, and Rui Cheng

Abstract This article addresses the issues of comfort and safety caused by excessive changes in vehicle posture while driving on rugged terrain. Based on active suspension control and combined with the collection of road surface elevation information in front of the vehicle and predictive control, it ensures that the vehicle's pitch and roll are minimized during driving, enhancing smoothness and handling. First, a mathematical model of a typical three-axle engineering vehicle is established, and a linear optimal controller is designed. A laser altimeter is installed in front of the vehicle to collect real-time height information of the road surface under the left and right wheels, and an optical speedometer is used to measure the longitudinal and lateral speeds of the vehicle. Finally, driving tests are conducted on trapezoidal obstacles, and experimental data show that compared to passive oil–gas suspension, the active hydraulic suspension combined with predictive control ensures smaller changes in vehicle posture, significantly improving driving comfort.

Keywords Active suspension · Road elevation information · Preview control

1 Introduction

With the increasing demands for vehicle comfort and handling in recent years, the field of vehicle control has transitioned from planar control technologies dominated by Anti-lock Braking System (ABS) and Electronic Stability Control (ESC) to vertical control primarily based on semi-active and active suspension control (ASC). With the development of intelligent vehicles, the recognition of road and surrounding

D. Zhao · L. Wang (✉) · J. Chang · H. Jiang
Hebei Key Laboratory of Special Carrier Equipment, Yanshan University, Qinhuangdao, Hebei, China
e-mail: wangliqiang@ysu.edu.cn

L. Yang · Y. Li · R. Cheng
Tai'an Aerospace Special Vehicle Co., Ltd., Taian, Shandong, China

© The Author(s) 2026

H. Saman et al. (eds.), *The 9th International Conference on Advances in Construction Machinery and Vehicle Engineering*, Lecture Notes in Mechanical Engineering,
https://doi.org/10.1007/978-981-96-5527-4_114

information through visual and LiDAR sensors has become a new development in intelligent vehicle control.

Using a complete vehicle model as the research object and utilizing wheelbase predictive information, which relates the input of the front and rear wheels, combined with a Kalman filter as a state estimator, an optimal control strategy for vehicle suspension has been designed. The feasibility of this model and algorithm has been verified [1, 2]. through simulation and emulation results, demonstrating significant improvements in comfort based on the Electric Controlled Air Suspension System (ECAS) [3]. Additionally, a fuzzy control strategy targeting the vertical acceleration of the vehicle body is applied to the front semi-active suspension, while a wheelbase predictive fuzzy control is used for the rear semi-active suspension. Simulation results indicate that compared to traditional passive systems, the vehicle's driving smoothness is significantly improved [4, 5]. Active suspension systems are greatly assisted by road preview schemes, consisting of special sensors usually based on laser scanners (e.g. LiDAR sensors, Binocular vision), which detect road irregularities ahead of the vehicle and feed this information to a control system, designed to manipulate the active suspension accordingly. Comparisons to a passive suspension system, as well as a standard MPC controller with a fully linear plant model, demonstrate the performance potential of using RBF prediction models in a road preview MPC context [6–8]. The new output-feedback H∞ control problem for active half-vehicle suspension systems is proposed controller in terms of linear matrix inequality (LMI) [9, 10].

The advantages semi-active and active suspensions is to automatically adjust important parameters such as stiffness, damping, and deflection based on road conditions and driver settings. Predictive control refers to a control technology that adjusts dampers in advance by obtaining road height information. This includes control methods based on road grade, inter-axle predictive control methods, and front vehicle predictive control methods.

This article employs road height measurement sensors installed at the front of the vehicle to scan the road elevation of the driving trajectories of both sides of the wheels in real time, inputting this data to the controller. The hydraulic system controls the servo valve opening to achieve the lifting and lowering of the hydraulic cylinder, complemented by feedback control from displacement sensors, minimizing changes in vehicle driving posture.

2 System Model

To describe vehicle performance, body deformation and damping are ignored, simplifying the vehicle body to a rigid body. Tire damping is also neglected, and the wheels are simplified to spring-mass blocks. The hydraulic suspension is simplified to a variable-stroke force generator, establishing a three-axis active suspension vehicle model as shown in Fig. 1.

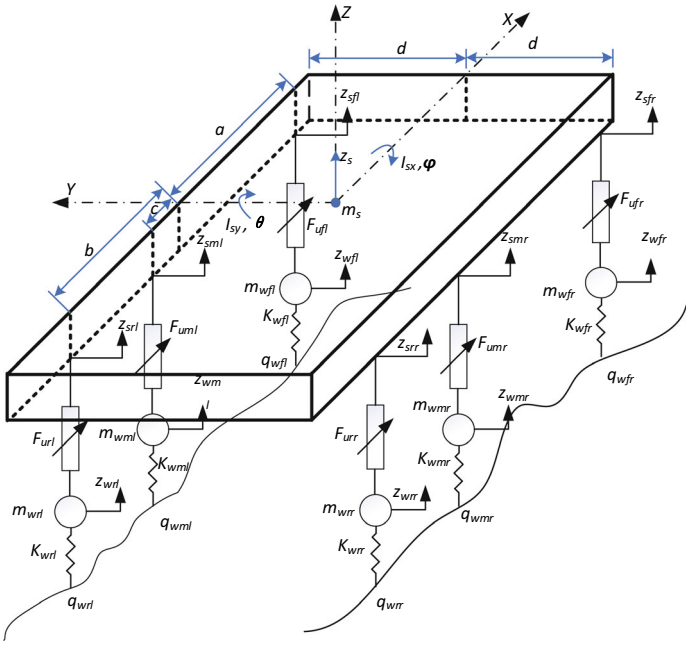


Fig. 1 Active hydraulic suspension model

When considering the vehicle as a rigid body in three-dimensional space, it has six degrees of freedom: three translational movements (*x, y, z*) and three rotational movements (the yaw angle (ψ), pitch angle (θ), and roll angle (φ)). A vehicle coordinate system is established with the Earth as the inertial reference frame. According to ISO 8855:2011 standards, the reference point is the center of mass, the forward direction of the vehicle is the positive X-axis, the left side of the vehicle is the positive Y-axis, and the direction perpendicular to the center of mass going upwards is the positive Z-axis.

In Fig. 1, q_{wij} represents the road surface elevation information at each wheel's hub position, the unsprung mass is denoted by m_{wij}, the tire stiffness is represented by K_{wij}, the hydraulic cylinder force is indicated by F_{uij}, the displacement at the suspension support point connected to the vehicle body is denoted by z_{sij}, the bouncing displacement of the unsprung mass centroid is represented by z_{wij}, the rising and falling displacement of the sprung mass centroid is denoted by z_s, the vehicle body pitch angle is represented by θ, the vehicle body roll angle is denoted by φ, the distance from the centroid to the front axle is a, to the mid axle is b, to the rear axle is c, and the distance from the centroid to the two side wheels is d, where $i = f, m, r$ indicates the front, middle, and rear axles. $j = r, l$ indicates the left and right sides.

When the pitch angle and roll angle are small, the vertical displacements at the six endpoints of the vehicle body have the following relationship.

$$\begin{aligned} z_{sfl} &= z_s - a\theta + d\varphi \\ z_{sfr} &= z_s - a\theta - d\varphi \\ z_{sml} &= z_s + c\theta + d\varphi \\ z_{smr} &= z_s + c\theta - d\varphi \\ z_{srl} &= z_s + c\theta + d\varphi \\ z_{srr} &= z_s + c\theta - d\varphi \end{aligned} \tag{1}$$

According to Newton's laws of motion, the vertical motion equation at the center of mass of the vehicle body is:

$$m_s\ddot{z}_s = F_{ufl} + F_{ufr} + F_{uml} + F_{umr} + F_{url} + F_{urr} \tag{2}$$

The pitch motion equation of the vehicle body is:

$$I_{sy}\ddot{\theta} = -a(F_{ufl} + F_{ufr}) + c(F_{uml} + F_{umr}) + b(F_{url} + F_{urr}) \tag{3}$$

The roll motion equation of the vehicle body is:

$$I_{sx}\ddot{\varphi} = d\left(F_{ufl} - F_{ufr} + F_{uml} - F_{umr} + F_{url} - F_{urr}\right) \tag{4}$$

The vertical motion equations for the four unsprung masses are:

$$\begin{aligned} m_{wfl}\ddot{z}_{wfl} &= K_{wfl}\left(q_{wfl} - z_{wfl}\right) + F_{ufl} \\ m_{wfr}\ddot{z}_{wfr} &= K_{wfr}\left(q_{wfr} - z_{wfr}\right) + F_{ufr} \\ m_{wml}\ddot{z}_{wml} &= K_{wml}(q_{wml} - z_{wml}) + F_{uml} \\ m_{wmr}\ddot{z}_{wmr} &= K_{wmr}(q_{wmr} - z_{wmr}) + F_{umr} \\ m_{wrl}\ddot{z}_{wrl} &= K_{wrl}(q_{wrl} - z_{wrl}) + F_{url} \\ m_{wrr}\ddot{z}_{wrr} &= K_{wrr}(q_{wrr} - z_{wrr}) + F_{urr} \end{aligned} \tag{5}$$

The above nine differential equations represent the nine degrees of freedom vehicle dynamics model.

State variables Equation

$$x = \left[\dot{z}_s\dot{\theta}\dot{\varphi}z_{wfl}z_{wfr}z_{wml}z_{wmr}z_{wrl}z_{wrr}\right]^T \tag{6}$$

Control variables Equation

$$u = \left[F_{ufl}F_{ufr}F_{uml}F_{umr}F_{url}F_{urr}\right]^T \tag{7}$$

Disturbance variables Equation

$$w = \left[q_{wfl}q_{wfr}q_{wml}q_{wmr}q_{wrl}q_{wrr}\right]^T \tag{8}$$

Select controlled output,

$$z = \left[\ddot{z}_s \ddot{\theta} \ddot{\varphi} L_{fl} L_{fr} L_{ml} L_{mr} L_{rl} L_{rr} F_{ufl} F_{ufr} F_{uml} F_{umr} F_{url} F_{urr}\right]^T \tag{9}$$

Select measurement variables

$$L_{ij} = z_{wij} - z_{sij} (i = f, m, r; j = l, r) \tag{10}$$

$$y = \left[\dot{\theta} \dot{\varphi} \ddot{z}_s L_{fl} L_{fr} L_{ml} L_{mr} L_{rl} L_{rr}\right]^T \tag{11}$$

Then the state-space model of the active suspension system is:

$$\begin{cases} \dot{x} = Ax + B_1 w + B_2 u \\ z = C_1 x + D_{11} w + D_{12} u \\ y = C_2 x + D_{21} w + D_{22} u \end{cases} \tag{12}$$

Then the transfer function of the system is

$$G_0 = C(sI - A)^{-1} B + D \tag{13}$$

where, $B = [B_1 B_2]^T$, $C = [C_1 C_2]^T$, $D = [D_{11} D_{12}; D_{21} D_{22}]^T$

3 Sensor Preview Model

To obtain the elevation information of the contact points between the wheels and the ground, two laser height gauges are installed at a fixed position on the vehicle body relative to the front wheels. These gauges scan the elevation information of the road surface along the driving trajectories of the left and right wheels, serving as control inputs. The installation position is related to the vehicle's horizontal projection, as shown in Fig. 2.

In Fig. 2, SensorL and SensorR represent the left and right height sensors, respectively, and V_x represents the forward speed. The distance from the left height sensor to the center of the left front wheel is X_{SL}, and the distance from the right height sensor to the center of the right front wheel is X_{SR}. D_{ij} represents each wheel, where $i = f, m, r$ stands for front, middle, and rear; $j = l, r$ indicates the left and right sides. From the formula (8) in the first part, it can be seen that the elevation information of each wheel is a delayed input of the elevation measured by the sensors, derived from the relationships in the figure.

$$q_{wfl} = SL + \tau_{1L}$$
$$q_{wml} = SL + \tau_{1L} + \tau_2$$

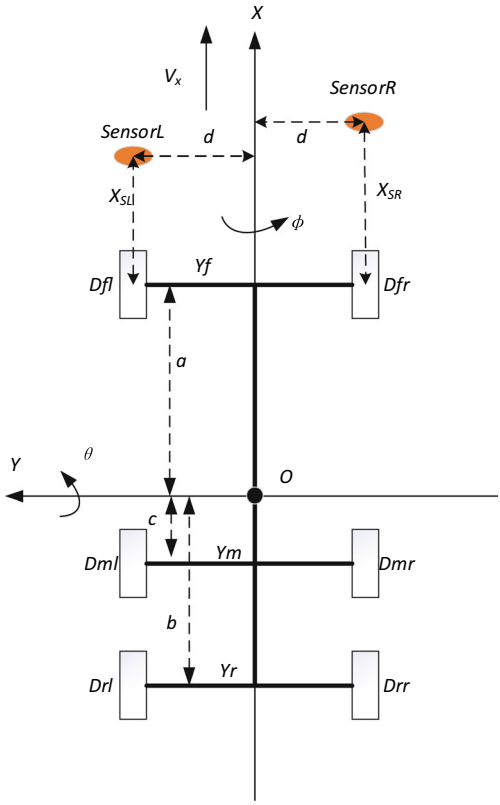


Fig. 2 Preview sensor schemetic diagram

$$\begin{aligned} q_{wrl} &= SL + \tau_{1L} + \tau_2 + \tau_3 \\ q_{wfr} &= SR + \tau_{1R} \\ q_{wmr} &= SR + \tau_{1R} + \tau_2 \\ q_{wrr} &= SR + \tau_{1R} + \tau_2 + \tau_3 \end{aligned} \tag{14}$$

Among them, SL and SR represent the displacement measurements from the left and right laser height gauges, respectively. And $\tau_{1L} = \frac{X_{SL}}{v}$, $\tau_{1R} = \frac{X_{SR}}{v}$, $\tau_2 = \frac{a+c}{v}$, $\tau_3 = \frac{b-c}{v}$.

4 Controller Design

For the system design of the H∞output feedback controller, the block diagram of the output feedback closed-loop control system is as follows (see Fig. 3).

The purpose of H_∞ control is to design a controller K(s) that ensures the closed-loop system is internally stable and minimizes $\|T_{zw}(s)\|_\infty$, where $T_{zw}(s)$ represents

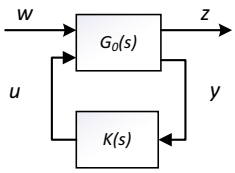


Fig. 3 Control system closed-loop feedback block diagram

the closed-loop transfer function from the disturbance input to the controlled output z. Since,

$$\begin{bmatrix} z \\ y \end{bmatrix} = G_0(s) \begin{bmatrix} w \\ u \end{bmatrix} = \begin{bmatrix} G_{11}(s)G_{12}(s) \\ G_{21}(s)G_{22}(s) \end{bmatrix} \begin{bmatrix} w \\ u \end{bmatrix} \tag{15}$$

$$T_{zw}(s) = G_{11} + G_{12}K(I - G_{22}K)^{-1}G_{21} \tag{16}$$

According to H_∞ control theory, to effectively suppress the influence of disturbance input, the design of the controller is transformed into a standard design problem of H_∞ systems, which can be solved using the LMI toolbox.

5 Road Testing

5.1 Obstacle Model

The trapezoidal top surface of the obstacle is 160 mm above the ground, with a top surface length of 500 mm. The inclined surfaces on both sides are symmetrical, with a projection length of 800 mm on the ground. The angle between the inclined surface and the ground is 11.3°, and the thickness is a 10 mm iron plate (see Fig. 4).



Fig. 4 Road surface obstacle profile

Table 1 The parameters of the test vehicle

Parameter	Unit	Value
Sprung mass, M_s	kg	21,000
Unsprung mass, M_w	kg	860
Moment of inertia, I_x	Kg m^2	63,020
Moment of inertia, I_y	Kg m^2	120,700
Tire stiffness, K	kN/m	360
Wheelbase, a + c	m	4.9
Wheelbase, b − c	m	1.5
Axle track, 2d	m	2.5

5.2 Vehicle Parameters

Table 1 gives the parameters of the test vehicle which is provided by Tai'an Aerospace Special Vehicle Co., Ltd.

5.3 System Composition

To achieve preview control, it is necessary to install a laser altimeter and a speedometer on the vehicle to measure the elevation information of the road surface located at the set positions in front of the left and right wheels. The sensor installation positions are shown in Fig. 5.

Figure 5a shows Kistler's S-Motion DTI sensors, (b) and (c) show Kistler's Optical Laser Height Sensors. The measurement range of Kistler S-Motion DTI sensors is − 250 to 250 km/h, and the measurement accuracy is 0.2 km/h. The measurement range of Height Sensors is 150–900 mm, the resolution is 0.3 mm, and the sampling rate is up to 8 kHz.

The vehicle suspension is an oil–gas spring independent suspension. An inertial measurement unit is installed near the vehicle's center of gravity. Displacement sensors are installed in parallel with the hydraulic cylinders. The hydraulic cylinders are controlled by the drive current of servo valves. The hydraulic pump station provides high-pressure hydraulic oil, which can be directly driven by the engine or powered by a high-pressure battery, with the motor controller adjusting the pump's speed and torque. The inertial measurement unit measures the vehicle's angular velocity and acceleration in three directions in real-time, and the electronic control unit integrates algorithms to collect altitude and attitude data, calculates the corresponding output based on the model, and sends it to the servo-valve (see Fig. 6).



(a) Position of S-Motion DTI sensors



(b) Position of Optical Laser Height Sensors

Fig. 5 Preview sensor installation position

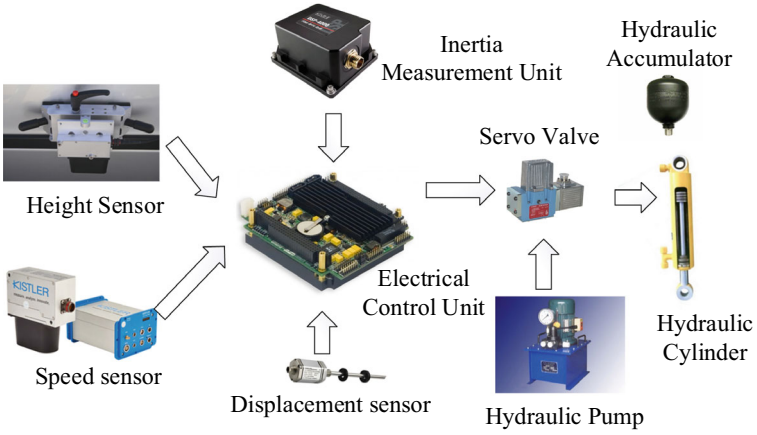
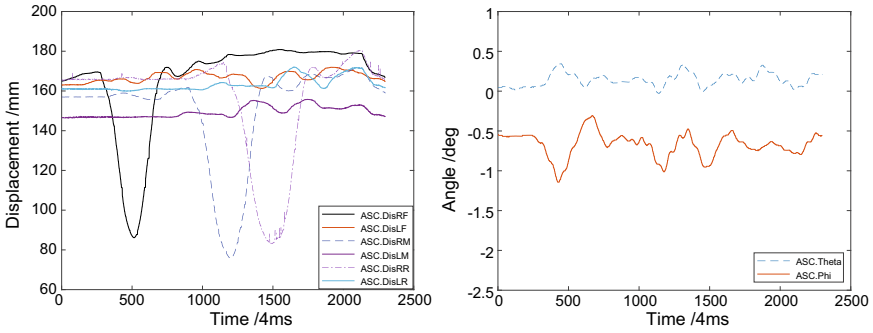


Fig. 6 Control system main component relationship diagram

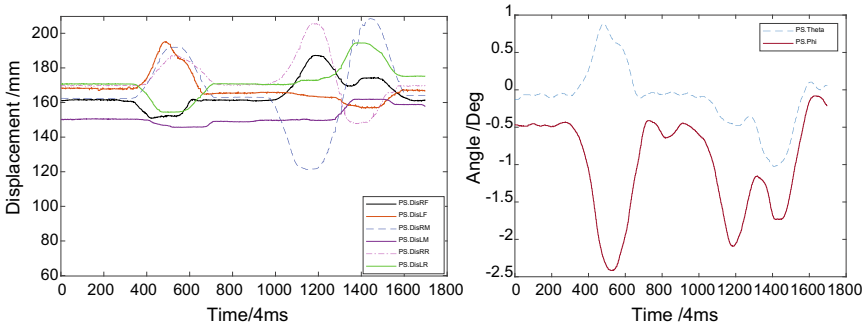
5.4 *Experiment and Data Analysis*

The vehicle operating condition is on a level surface, where during testing, the obstacle in Fig. 4 is placed approximately 2 m in front of the right front wheel. The right side wheels successively cross the obstacle at a speed range of 10–30 km/h. An algorithm is developed in the controller to adjust the extension of each hydraulic cylinder of the right side suspension based on height measurement values, to minimize the vehicle's pitch and roll. The test curves are shown in Figs. 7 and 8. Active suspension is referred to as Active Suspension Control (ASC), while passive suspension is referred to as Passive Suspension (PS). In the figure, the curves *DisRF, DisLF, DisRM, DisLM, DisRR,* and *DisLF* represent the right front displacement, left front displacement, right middle displacement, left middle displacement, right rear displacement, and left rear displacement, respectively; *Theta* represents the vehicle's pitch angle, and *Phi* represents the vehicle's roll angle.



(a) Hydraulic Cylinder Displacement Curve (b) Body Pitch and Roll Angle Curve

Fig. 7 Active suspension testing results



(a) Hydraulic Cylinder Displacement Curve (b) Body Pitch and Roll Angle Curve

Fig. 8 Passive suspension testing results

Figure 7a shows the stroke of the hydraulic cylinders during the active suspension's traversal over an obstacle. The three hydraulic cylinders on the left side retract by approximately 80 mm in sequence, while about 5–10 mm for the left side. Figure 7b presents the pitch and roll angles of the vehicle body, the pitch angle ranges from 0.1 to 0.4°, while the roll angle ranges from − 1.2 to − 0.4°. Figure 8a illustrates the passive suspension vehicle's hydraulic cylinders on the right side retracting by approximately 10, 40, and 20 mm, respectively. Figure 8b shows pitch and roll angles of the vehicle body during the passive suspension's traversal over an obstacle, the pitch angle ranges from − 1 to 1°, and the roll angle ranges from − 2.4 to − 0.2°.

From the comparative data, it can be observed that when crossing a trapezoidal obstacle, the compression amount of the passive suspension hydraulic cylinders is relatively small, resulting in excessive fluctuations in the vehicle body roll angle, while the pitch angle is also similarly affected, leading to an increased fluctuation range.

6 Conclusion

The preview control based on the front height gauge, compared to the preview control method based on the front axle, not only improves the vehicle's body posture when the rear wheels encounter obstacles, but also effectively filters out the fluctuations in the pitch and roll angles caused by obstacles when the front wheels go over them. This greatly enhances comfort, and the range of tire contact performance fluctuations is reduced compared to passive suspension, thereby also increasing maneuvering stability.

Acknowledgements This work was supported by the National Natural Science Foundation of China (No. U20A20332).

References

1. Yu F, Guo K (1999) A study on wheelbase preview control for active vehicle suspension system with a Kalman filter. Autom Eng 2:72–80
2. Bo D (2002) A study on optimal control for vehicle active suspension system using a whole vehicle model. Autom Eng 5:422–425
3. Zhimin M, Qingze Y, Zuozhu W, Huafang L (2019) Research on preview algorithm for stiffness adjustment of electronically controlled air suspension system. Mechan Sci Technol 38(03):398–403
4. Gang S, Changcheng X (2013) Stochastic optimal preview control of active vehicle suspension with time-delay consideration. Trans Chin Soc Agric Machin 44(06):1–7
5. Shaoyi B, Chuanyi Y, Long C et al (2010) Fuzzy neural network control of semi-active suspension based on wheelbase preview. Autom Eng 32(12):1067–1070+1082
6. Qianman Z, Zhi Feng W, Jie X (2024) Research on semi-active suspension preview control based on visual perception. Command Control Simul 46

7. Papadimitrakis M, Alexandridis A (2022) Active vehicle suspension control using road preview model predictive control and radial basis function networks. Appl Soft Comput 120:108646
8. Kim J, Lee T, Kim CJ, Yi K (2023) Model predictive control of a semi-active suspension with a shift delay compensation using preview road information. Control Eng Pract 137:105584
9. Viadero-Monasterio F, Boada BL, Boada MJL, Díaz V (2022) H∞ dynamic output feedback control for a networked control active suspension system under actuator faults. Mech Syst Signal Process 162:108050
10. Liu Z (2024) Ride comfort oriented integrated design of preview active suspension control and longitudinal velocity planning. Mech Syst Signal Process 208:110992

Open Access This chapter is licensed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License (http://creativecommons.org/licenses/by-nc-nd/4.0/), which permits any noncommercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if you modified the licensed material. You do not have permission under this license to share adapted material derived from this chapter or parts of it.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.



Defending Against Adversarial Examples with Adaptive Quantization Table-Guided Image Compression



Zhuang Zhang, Lijun Zhang, Dejian Meng, Wei Tian, Ye Han, and Kaikun Pei

Abstract Intelligent vehicles rely on deep learning-based image recognition models for accurate environmental perception, which is critical for driving safety. However, recent research has identified the significant threat from adversarial examples, which can mislead models by adding subtle perturbations to input images, thereby endangering intelligent vehicles. Existing defenses typically use image preprocessing to remove perturbations, however, most works overlook the diversity in adversarial perturbation amplitudes, leading to performance degradation in varying conditions. To address this limitation, this paper investigates the relationship between changes in image spatial frequency and varying adversarial perturbation amplitudes, revealing that increasing perturbation amplitudes lead to heightened high-frequency variations in the spectrum. Based on this finding, we propose an adaptive quantization table-guided image compression method that modifies quantization tables based on the image's frequency spectrum, which enhances the defense performance compared to existing methods that rely on fixed quantization tables. Experimental results demonstrate that our adaptive method effectively responds to the dynamic nature of adversarial examples, significantly outperforming state-of-the-art defenses against common adversarial attacks, such as FGSM, PGD, Deepfool, CW, AutoAttack, and BPDA. The proposed method shows potential for application in intelligent vehicles, as it offers robust defense capabilities across a range of perturbation amplitudes.

Keywords Image classification · Adversarial defense · Adversarial examples · JPEG compression · Quantization table

Z. Zhang · L. Zhang · D. Meng · W. Tian (✉) · Y. Han · K. Pei
School of Automotive Studies, Tongji University, Shanghai, China
e-mail: tian_wei@tongji.edu.cn

© The Author(s) 2026

H. Saman et al. (eds.), *The 9th International Conference on Advances in Construction Machinery and Vehicle Engineering*, Lecture Notes in Mechanical Engineering,
https://doi.org/10.1007/978-981-96-5527-4_115

1 Introduction

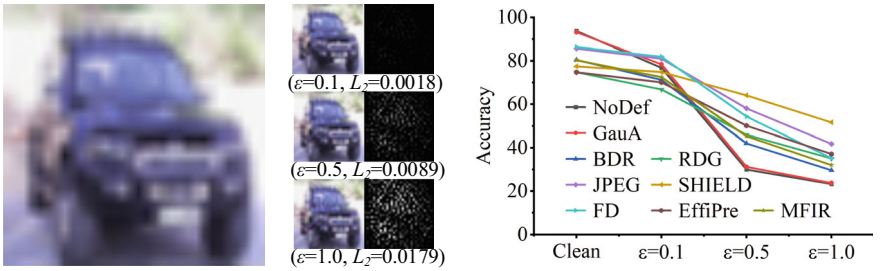
Deep Neural Networks (DNNs)-based image classification has become crucial for many safety–critical systems such as autonomous driving, etc. [1, 2]. Meanwhile, DNNs are also found with vulnerabilities [3–5]. In particular, by adding carefully crafted and imperceptible perturbations to images, which are called adversarial examples (AEs), one can mislead DNNs to false inferences [6, 7]. For example, by digitally or physically adding visual perturbations, road signs can be misclassified by DNNs [8, 9], which may cause traffic accidents for vision-based autonomous vehicles. Therefore, the defense of AEs has drawn increasing attention.

To defend against adversarial examples (AEs), various image processing-based techniques have been proposed. GauA [10] mitigates adversarial perturbations through Gaussian filtering. BDR [11] compresses pixel values into fewer bits to reduce the attack surface. RDG [12] introduces random distortions to the image mesh, and MFIR [13] enhances the limitations of fixed filters by integrating multiple image processing operations, including varying filtering parameters, random rotations, and averaging classification probabilities.

Among these approaches, JPEG compression-based defenses are particularly noteworthy. JPEG [14, 15], a widely used image file format, compresses high-frequency components that are often imperceptible to human vision. Early research demonstrated the efficacy of JPEG compression in defending against AEs, achieving notable results [15]. However, conventional JPEG compression employs quantization tables designed primarily based on human visual characteristics rather than the specific properties of AEs. This gap has led to the development of new quantization strategies aimed at enhancing the effectiveness of JPEG-based defenses. SHIELD [16] generates quantization tables with random factors for each grid. FD [17] analyzes the image spectrum to preserve frequency components critical for deep neural network (DNN) predictions. EffiPre [18] summarizes the spectral differences between AEs and clean samples, designing quantization tables based on the components exhibiting maximal absolute spectrum differences.

Despite the effectiveness of image preprocessing as a defense mechanism, existing research [15–18] has not adequately considered the variations in adversarial perturbation amplitudes, which leads to a significant decrease in the defense performance when the perturbation amplitude increases. A more detailed experiment is presented in Fig. 1, where the perturbation amplitude ε increases with greater attacking parameters in the Fast Gradient Sign Method (FGSM) [5]. Although the perturbations added to the AEs remain imperceptible, the performance of state-of-the-art (SOTA) defenders [15–18] shows a significant decrease as the perturbation amplitude increases, as illustrated in Fig. 1c. This performance drop is particularly concerning, as AEs can exhibit a wide range of amplitudes in real-world scenarios, thus limiting the practical applicability of these defense mechanisms.

To address the issue of performance degradation when perturbation amplitudes vary, this section investigates how these variations affect the image frequency spectrum, revealing that as the perturbation amplitude increases, the high-frequency



(a) Clean Image (b) AEs and perturbations (c) Accuracy of SOTA defenders

Fig. 1 AEs and performance of SOTA defenders. **a** Example of a clean image. **b** AEs and perturbations generated by FGSM attackers with different ε. Perturbation amplitude is indicated by the L_2 norm. For a better representation, the amplitude is multiplied by 50 in the perturbation image. **c** Performance decrease of SOTA defenders

variations in the image spectrum also increase. Based on this characteristic, this paper proposes an image compression method that utilizes adaptive quantization tables. Unlike existing methods [15–18] that rely on fixed quantization tables, this approach learns the image spectrum via a deep learning module, enabling the adaptive adjustment of compression quantization tables according to changes in the spectrum. Experiments conducted under various attack methods and perturbation amplitudes demonstrate that the proposed approach outperforms existing methods in terms of defensive effectiveness.

The key contributions of this paper are as follows:

- This work explores the relationship between perturbation amplitude variations and changes in the image frequency spectrum, revealing that increasing perturbation amplitudes lead to heightened high-frequency variations in the spectrum.
- We propose an adaptive quantization table-guided image compression, in contrast to existing methods that utilize fixed quantization tables [15–18]. This method learns the image spectrum with a deep learning module, facilitating the adaptive adjustment of quantization tables in response to spectral changes. This adaptation enhances the defense mechanism's effectiveness against adversarial examples, particularly as perturbation amplitudes vary.

2 Spectral Characteristics of Adversarial Examples and SOTA Defenders

We acknowledge that there is extensive literature [15–18] on manipulating the frequency spectrum of AEs, some of which draw inspiration from JPEG compression. However, little attention has been given to the dynamic changes in the AE spectrum and the differences across color channels. Our proposed perspective highlights three key distinctions from existing work. First, previous studies are primarily conducted in the RGB color space, treating all three color channels as a unified entity, which

overlooks the psychovisual characteristic of human vision—specifically, the varying sensitivities to brightness and color. In contrast, our approach accounts for these characteristics by performing channel-wise calculations in the YCbCr color space. Second, while existing research typically examines the spectral characteristics of AEs under a single perturbation, our study assesses the spectrum differences across varying perturbation amplitudes. Third, in addition to analyzing the frequency spectrum of AEs, we also compare it with the frequency spectrum of images processed by SOTA defenders, offering a comprehensive understanding of both the attackers and defenders.

Specifically, we conduct experiments with the AllConv [19] model pre-trained on the Cifar-10 dataset [20], and the calculation can be expressed as:

$$D_z = \mathbb{E}(\mathcal{F}(I[i]_z) - \mathcal{F}(I'[i]_z)), i \in [1, 1000], z \in [1, c] \tag{1}$$

where D_z is the amplitude of spectrum change of the z channel. $I[i] \in \mathbb{R}^{h \times w \times c}$ is a clean image from 1000 randomly selected images and $I'[i]$ is a corresponding AE by an L_2 version of FGSM attacker with three perturbation amplitudes (indicated by parameter $\varepsilon = 0.1, 0.5, 1.0$). $\mathcal{F}$ indicates the FDCT (Forward Discrete Cosine Transform). $\mathbb{E}$ indicates the average over all images.

Figure 2 illustrates the amplitudes of spectrum differences, revealing that the spectral changes of adversarial examples (AEs) predominantly occur in the high-frequency region across the color channels, while the low-frequency components in the upper left corner remain largely unchanged. Notably, as the perturbation amplitude increases (indicated by larger values), the high-frequency variations become more pronounced, indicating a correlation between perturbation amplitude and the extent of spectral change.

Despite the rising perturbation amplitude, the spectrum differences of AEs processed by SOTA defenders remain constant. This observation highlights a key limitation of SOTA defenders: they can only address perturbations within fixed frequency regions. Consequently, this inability to adapt to dynamic perturbations may explain their vulnerability in real-world scenarios.

In summary, this section highlights that as the perturbation amplitude increases, the high-frequency variations in the image also significantly increase. However, existing defense methods are limited to producing fixed spectral changes and thus struggle to effectively counteract these dynamic spectral variations. This limitation renders current defense mechanisms notably vulnerable when faced with continuously changing adversarial perturbations. Consequently, the development of adaptive defense strategies that can respond to dynamic spectral changes becomes crucial for enhancing defense effectiveness under various attack conditions.

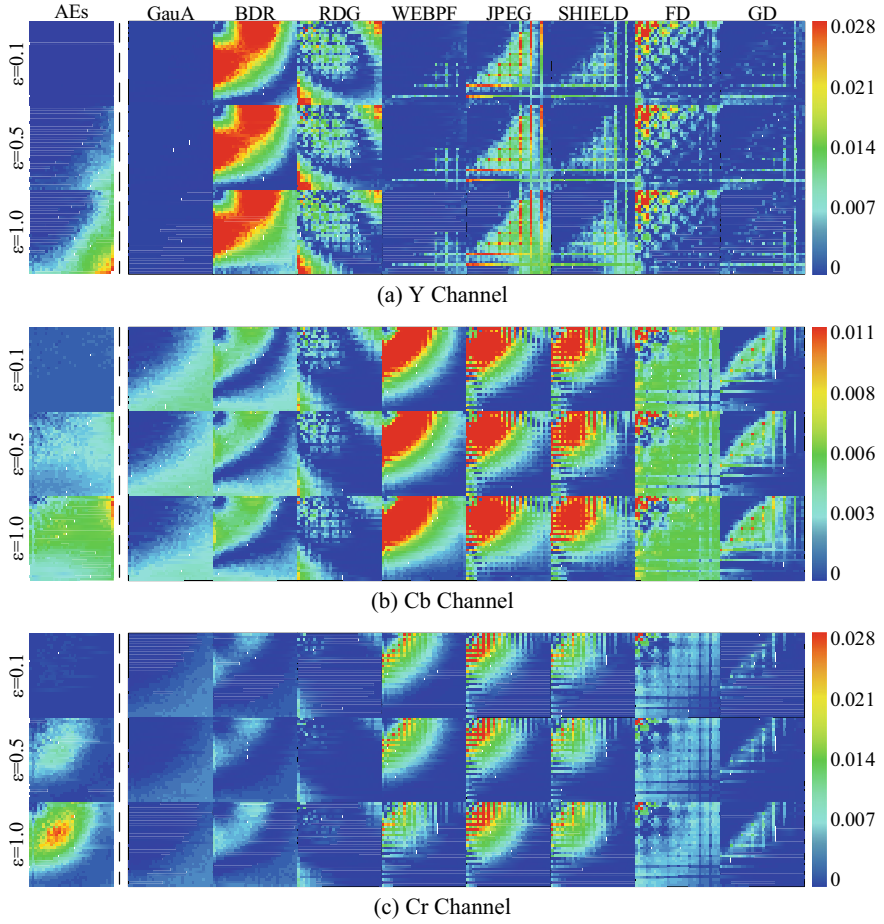


Fig. 2 The first column shows the spectrum difference between clean images and AEs. Other columns show the spectrum difference between AEs before and after processing by each defender. Zero frequency is located in the upper left corner of each subfigure

3 Proposed Adaptive Quantization Table-Guided Image Compression

To address the dynamic spectrum of AEs, we propose an adaptive defender by applying JPEG quantization adapted to the spectral characteristics of the input image. The overall framework consists of following two key steps:

Step 1: Spectrum-based quantization. As shown in algorithm 1, clean images and AEs are converted from the RGB color space to the YCbCr color space (Line 4 in algorithm 1). Each channel of images is cut into grids of a size of 8 * 8 pixels (Line 5–6). Each grid is then converted into the frequency space by the FDCT (Line 7–11).

The absolute difference between the spectrum of AEs and that of clean images is calculated and then averaged image-wisely (Line 12). Here the mean spectrum difference indicates the absolute spectrum variation of AEs. Since the spectrum amplitude at low frequencies is higher than that of high frequencies, for a fair evaluation, we divide the mean spectrum difference by the mean spectrum of clean images (Line 15).

A high relative spectrum difference indicates a large perturbation introduced in the frequency band, which is the potential harmful component and should be suppressed. Referring to the statistical processing in [18], the frequency components with a relative difference less than the threshold are set to 1, while the remaining frequency components are set to 100. The binarized frequency components are then averaged to calculate the quantization table for each channel (Line 16–17).

Algorithm 1: Spectrum-based Quantization	
Input:	Clean images $I^n \in R^{n\times h\times w\times c}$ with image number n, image size $h \times w$ and color channel number c, threshhold $(T_1, ..., T_c) \in R$, key factors $(\varepsilon_1, ..., \varepsilon_n)$
Output:	Quantization tables Q_s
1:	**for** $\varepsilon \in (\varepsilon_1, ..., \varepsilon_n)$ **do**
2:	$Q = zeros(8, 8, c)$
3:	$I_a^n = FGSM_e(I^n)$
4:	$I_a^n = RGB2YCbCr(I_a^n), I^n = RGB2YCbCr(I^n)$
5:	$n_h = h/8, n_w = w/8, n_c = c$;
6:	$\mathcal{G}_{I^n}, \mathcal{G}_{I_a^n} = I^n(p, i, j, k), I_a^n(p, i, j, k)\|(p, i, j, k)$ $\in \{(1, ..., n) \times (1, ..., n_w) \times (1, ..., n_h) \times (1, ..., n_c)\}$ /*Get amplitude difference in frequency domain*/
7:	$S_c, S_a = zeros(n, n_w, n_h, n_c, 8, 8)$
8:	**for** $I^n(p, i, j, k), I_a^n(p, i, j, k)$ in $\mathcal{G}_{I^n}, \mathcal{G}_{I_a^n}$ **do**
9:	$S_c[p, i, j, k, ...] = \log(1 + abs(\mathcal{F}(I^n(p, i, j, k)))$
10:	$S_a[p, i, j, k, ...] = \log(1 + abs(\mathcal{F}(I_a^n(p, i, j, k)))$
11:	**end for**
12:	$S_{ad} = mean(abs(S_a - S_c), axis = 1, 2)$/*Binarization to get quantization table*/
13:	$S_{cm} = mean(S_c, axis = 0, 1, 2)$
14:	**for** k in c **do**
15:	$S_{rd} = S_{ad}[:, k, ...]/S_{cm}[k]$
16:	$S_{rd}[S_{rd} >= T_k] = 100, S_{rd}[S_{rd} < T_k] = 1$
17:	$Q[..., k] = mean(S_{rd}, axis = 0)$
18:	**end for**
19:	$Q_s[\varepsilon] = round(Q)$
20:	**end for**
21:	**return** Q_s

Step 2: Adaptive Adversarial Defense by applying JPEG compression with adaptive quantization tables.

Step 2.1: Adaptive quantization table generation. A regression model is used to estimate the perturbation amplitude based on the spectrum of AEs. Here, our regression model consists of a backbone of the ResNet50 and a full connection layer. The quantization table Q is obtained by looking-up and linear interpolation of tables obtained in Step 1. In addition, we argue that AEs with small perturbations are hard to distinguish from each other in the frequency domain. This may result in poor regression accuracy and thus poor defensive performance for them. Therefore, inspired by the focal loss [21], we design the SEfocal loss to pay more attention to attackers with a small perturbation amplitude:

$$Loss_{SE} = (1 + |\log_{10} \frac{y}{10}|^2) \cdot (a - y)^2 \tag{2}$$

where y is the ground-truth label and a is the model output.

Step 2.2: Adaptive JPEG compression. Similar to step 2.1, we adopt the same color-space conversion, gridding, and FDCT for the input image. Components in each grid are assigned with the corresponding values from the adaptive quantization table. Each grid is then transformed back to the image domain by the IDCT (inverse DCT). Finally, the image is converted back to the RGB color space. In addition to JPEG compression, we empirically add a horizontal flip operation to disturb the carefully crafted perturbation without significantly decreasing the accuracy of clean images.

4 Experiments

4.1 Experimental Setup

Attackers: Standard attackers (FGSM [5], PGD [22], DeepFool [23], and CW [24]), AutoAttack [25], and BPDA [26, 27] are evaluated. All attackers except BPDA are implemented in [10]. Evaluation against FGSM and PGD are based on three perturbations ($\varepsilon = 0.1, 0.5, 1.0$ on Cifar-10 and $\varepsilon = 1, 5, 10$ on ImageNet). The PGD is iterated 10 times. For DeepFool and CW, the default settings are used. For AutoAttack, the maximum iteration is 100 and the same perturbations are used as FGSM. BPDA is implemented following the settings in [18], in which we approximate the obfuscated gradients and perform attacks with PGD iterated 50 times.

Dataset: Both Cifar-10 [20] and ImageNet [28] subset are used. For standard attackers, the full Cifar-10 and 10% images from 100 classes of ImageNet are used for quantization and evaluation. For the auto- and adaptive attacker, following the settings in [11, 12, 17, 18], we randomly select 100 images from Cifar-10 and ImageNet for evaluation.

Implementation details: The quantization tables are generated based on FGSM attackers and tested against all attackers. (1) Experiments on Cifar-10: In step 1, we set $\varepsilon \in \{0.01, 0.1, 0.5, 1.0\}$. In Step 2.1, the training data are generated by FGSM with $\varepsilon \sim N(0, 1/3)$. The Adam optimizer is used with a learning rate of 1e-3. The model is trained with a batch size of 256 and a maximum epoch number of 100. (2) Experiments on ImageNet: In step 1, we set $\varepsilon \in \{0.1, 1, 5, 10\}$. In step 2.1, the training data are generated by FGSM with $\varepsilon \sim N(0, 10/3)$. The learning rate is 1e-7. The batch size is 16 and the gradient is accumulated to obtain an equivalent batch size of 256. Other learning parameters are the same as in previous experiments.

Evaluation metric: Similar to settings in [16–18], the classification accuracy under defending is calculated. The only difference is that the accuracy on clean images and AEs generated by the above attackers are averaged as a metric for the whole defense performance.

4.2 Defending Against Standard Attacks

Figure 3 shows the generated quantization tables for AllConv and the spectrum difference of AEs before and after our defending. As shown in Fig. 3a, the quantization tables manifest a dynamic change across various channels and perturbation amplitudes. Comparing Fig. 3b with Fig. 2, it is obvious that our spectrum difference is consistent with that of AEs in terms of both value and frequency range.

As the defense performance shown in Table 1, on the Cifar-10 benchmark, compared to 'NoDefense', our defender improves the classification accuracy by 40.27% on AllConv, 40.90% on ResNet50 and 40.64% on VGG16. Compared to JPEG compression and its extensions (JPEG, SHIELD, FD, and EffiPre), our

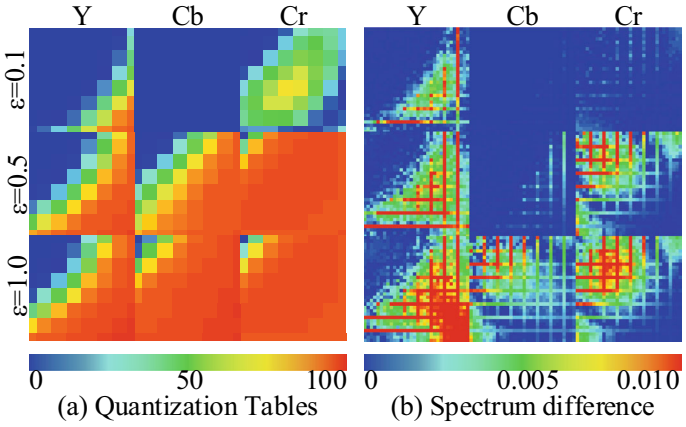


Fig. 3 Quantization tables and the spectrum difference of AEs before and after our defending

Table 1 Defense performance against standard attacks

Method	Cifar-10			ImageNet	
	AllComv	ResNet50	VGGI6	ResNet50	VGG16
NoDefense	38.07	42.53	42.29	24.73	22.31
GauA [10]	49.84	59.25	49.91	33.56	29.51
BDR [11]	49.89	66.57	61.70	41.09	28.42
RDG [12]	50.70	59.03	64.38	50.11	32.40
MFIR [13]	66.31	60.61	59.69	60.58	39.33
JPEG [14, 15]	67.86	78.46	78.05	40.49	35.11
SHIELD [16]	64.12	74.25	76.45	52.87	42.09
FD [17]	55.73	74.52	69.36	42.51	36.60
EffiPre [18]	53.48	67.40	68.72	57.51	38.71
Ours	78.34	83.43	82.93	61.82	50.13

defender improves the defense performance by at least 10.48% on AllConv, 4.97% on ResNet50, and 4.88% on VGG16. Compared to other defenders (GauA, BDR, RDG, and MFIR), our defender achieves a higher defense performance by at least 12.03% on AllConv, 16.86% on ResNet50, and 18.55% on VGG16. In the ImageNet dataset, the mean accuracy gain of our defender over the JPEG and its descendants is at least 4.31% on ResNet50 and 8.04% on VGG16, respectively. Compared to other methods, our defender achieves a higher accuracy by at least 1.24% on ResNet50 and 10.8% on VGG16, confirming the generality of our defender.

4.3 Defending Against Auto- and Adaptive Attacks

Evaluation on auto-attacker: The defended classification results of Allconv on Cifar-10 and those of VGG16 on ImageNet are shown in Table 2. The accuracy of a defenseless AllConv is only 44.75%, which shows the powerful attacking performance of AutoAttack. Defenders like GauA, BDR, RDG, and MFIR only show limited defense performance with a slight accuracy improvement of less than 5% compared to the case of 'NoDefense'. Compression-based defenders like JPEG, SHIELD, FD, and EffiPre can improve the mean accuracy to 69.50%, which is still insufficient in defending the AutoAttack. Our method achieves the highest mean accuracy of 81.75% on AllConv and 48.00% on VGG16, which is at least 12.25 and 10.75% higher than SOTA defenders.

Evaluation on adaptive attacker: After 50 rounds, most defenders show weak defensive performance (Fig. 4), among which the attack success rate (ASR) of GauA, BDR, JPEG, and FD is close to 100% and ASR of SHIELD and MFIR is approximately 55%.

Table 2 Defense performance against AutoAttack

Method	AllCoav (Cifar10)	VGGI6 (lmageNet)	Method	AllConv (Cifar10)	VGGI6 (ImageNet)
CauA [10]	45.50	20.50	JPEG [14, 15]	66.75	29_50
BDR [11]	46.00	20.25	SHIHLLD [16]	69.50	37.25
RDG [12]	49.25	21.50	FD [17]	67.75	30.00
MFIR [13]	49.50	34.00	EffiPre [18]	56.50	31.25
Nodefense	44.75	20.50	Ours	81.75	48.00

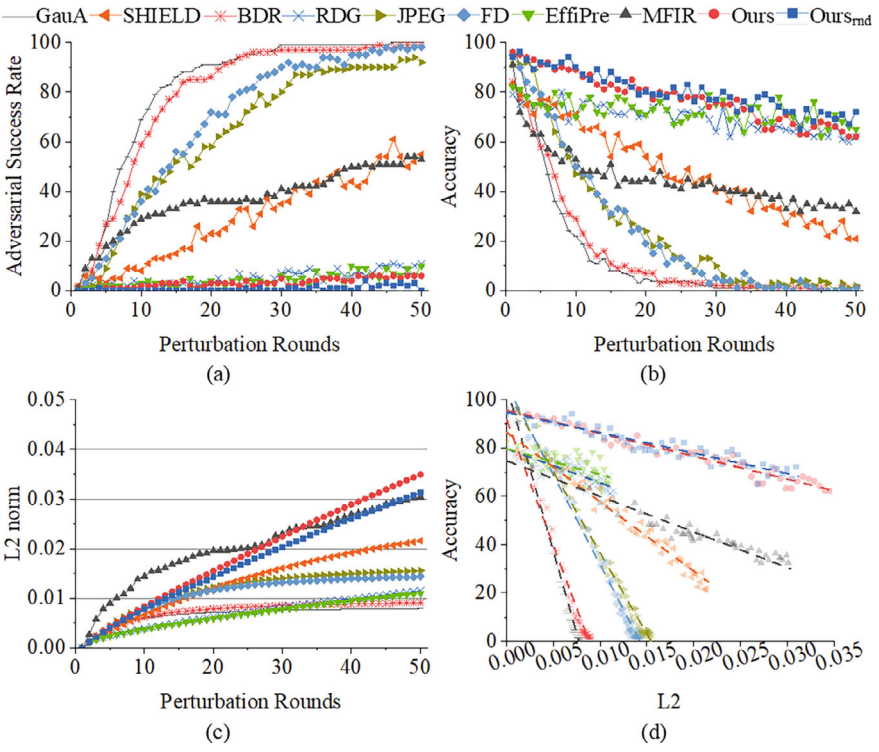


Fig. 4 Defense performance against BPDA. **a** Adversarial success rate. **b** Model accuracy. **c** L_2 norm of adversarial perturbations. **d** L_2 norm of adversarial perturbations versus model accuracy

The RDG and EffiPre show an improvement with an ASR of approximately 10%. Ours achieves the strongest defense with an ASR of 6%. In addition, we extend our method by randomly scaling the estimated ε to increase the difficulty of gradient estimation in BPDA. The new version indicated by Ours$_{rnd}$ achieves the lowest ASR of 0% and the highest accuracy of 72%, which is at least 10% and 7% better than SOTA defenders, respectively.

Besides ASR, the amplitude of perturbation is also an important indicator, with larger values representing an increased difficulty for the BPDA to bypass the defender. As shown in Fig. 4c, after 50 round attacks, the L_2-norm on GauA, BDR, EffiPre, and RDG is about 0.01. The L_2-norm on SHIELD, FD, JPEG, and MFIR is about 0.015–0.030. Our method achieves the highest L_2-norm with 0.031 in the standard version and 0.035 in the random version, which means that our defender is more difficult to bypass by BPDA.

For a comprehensive evaluation of attacking difficulty and accuracy, we scatter both the L_2-norm of perturbation and the classification accuracy in Fig. 4d and draw the trend lines. The trend close to the upper right corner of the figure means a higher accuracy and a more robust defender that is more difficult to bypass. Compared to SOTA defenders, our defender pushes the trend line closer to the upper right corner, especially in the region with a large L_2-norm, which indicates a higher robustness for perturbations with larger amplitudes.

4.4 *Influence of Hyper-Parameters: Thresholds*

Similar to the previous section, the experiments are conducted with the AllConv model on the Cifar-10 dataset. For each attacker of $FGSM_\varepsilon$, $\varepsilon \in \{0.1, 0.5, 1.0\}$, the defense performance is recorded and shown in Fig. 5. What is more, to avoid the impact by estimation errors, the ground truth of ε rather than the estimated ε by the regression model is utilized.

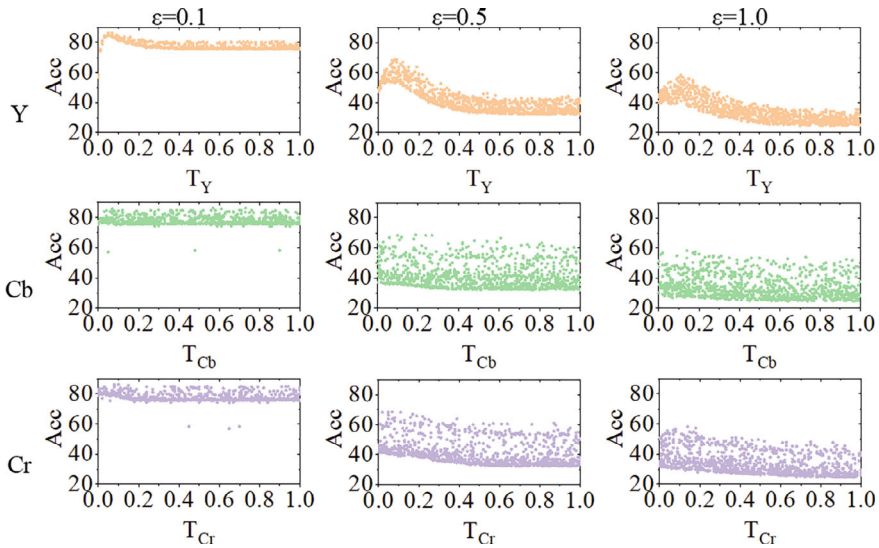


Fig. 5 Defense performance by different thresholds used in quantization table calibration

For each attack of FGSM, the popular hyperopt toolkit is used to search the corresponding optimal thresholds. Specifically, the iteration number is set to 1000. The random search algorithm is utilized. The optimization objective is to maximize the defended classification accuracy of the proposed method within the parameter space of $T_Y \in [0, 1]$, $T_{cb} \in [0, 1]$, $T_{cr} \in [0, 1][0, 1]$.

As shown in Fig. 5, the classification accuracy of different AEs varies greatly. In the Y channel, the accuracy tends to increase at a very small threshold T_Y and decreases with bigger thresholds. The optimal threshold is about 0.1, indicating that when the amplitude change in the Y channel spectrum exceeds 10%, the corresponding component of the spectrum should be suppressed. For both Cb and Cr channels, the accuracy changes slightly and small thresholds achieve slightly higher accuracy. This suggests that the effects of Cb and Cr channels are secondary to the Y channel.

4.5 Transferability of Quantization Tables

To explore the transferability of quantization models, we conduct two experiments on Cifar-10. First, we compare the quantization tables of AllConv, ResNet50, and VGG16. As shown in Fig. 6, the quantization tables of the Y-channel are consistent among all models. For the Cb and Cr channels, they differ slightly. This can be related to the thresholds discussed in Fig. 5. In Fig. 5, the accuracy change in terms of the threshold on the Cb and Cr channels is relatively mild, making it difficult to find an optimal threshold. Different sub-optimal thresholds may lead to different quantization tables. Therefore, a natural question is: can these different quantization tables achieve a similar defense performance? To answer this question, we conduct the second experiment.

In the second experiment, we apply the quantization table of AllConv directly to both VGG16 and ResNet50. Their defense performance is shown in Table 3, where the subscript t indicates using transferred quantization tables, and the bold indicates the best performance. Compared to the case without defenders, transferred tables improve the mean accuracy by 41.15% on the ResNet50 and 41.50% on the VGG16,

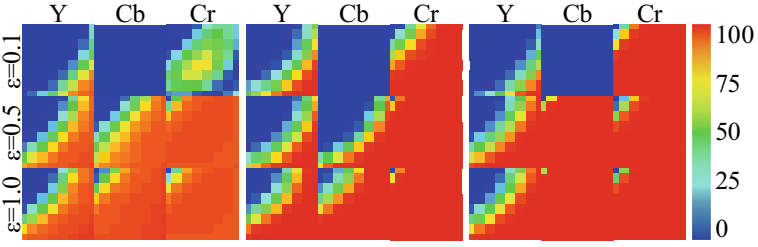


Fig. 6 Quantization tables of AllConv, ResNet50, and VGG16

respectively. Compared with settings with the corresponding quantization tables, transferred tables increase the mean accuracy by 0.24% on the ResNet50 and 0.88% on the VGG16, respectively.

Thus, the results show a similar or even slightly higher defense performance of transferred quantization tables, proving their transferability. This provides a potential way to efficiently obtain the quantization tables, that is, when defending a black-box or complex model, a lightweight alternative model can be used in generating quantization tables.

5 Conclusion

In this paper, we have addressed the critical challenge posed by AEs with varying perturbation amplitudes, which significantly affect the performance of existing defense mechanisms for intelligent vehicles. By investigating the relationship between perturbation amplitude variations and changes in the image frequency spectrum, we have revealed that increasing perturbation amplitudes lead to heightened high-frequency variations in the spectrum. Based on this insight, we have proposed an adaptive quantization table-guided image compression, in contrast to existing methods that rely on fixed quantization tables. Our method learns the image spectrum using a deep learning module, allowing for the adaptive adjustment of compression quantization tables according to spectral changes. This approach significantly enhances defensive effectiveness against various adversarial attacks compared to existing methods that rely on fixed quantization tables. Despite these contributions, our work has limitations, particularly in its lack of integration with actual intelligent vehicles. Future research should focus on conducting real vehicle or hardware-in-the-loop experiments to validate the effectiveness of our proposed method in practical scenarios. This will provide deeper insights into its applicability and robustness in real-world driving environments, ultimately contributing to safer autonomous driving systems.

Table 3 Defense performance of ResNet50 and VGG16 using transferred quantization tables

Model	Method	Mean Accuracy	Clean Image	FGSM-0.1	FGSM-05	FGSM-1.0	PGD-0.1	PGD-0.5	PGD-1.0	Deepfool	CW
ResNet50	NoDefense	42.52	**95.32**	82.00	55.33	49.11	69.32	3.58	3.29	21.03	3.65
	Ours	83.43	91.50	90.07	**78.49**	**70.00**	89.93	**78.92**	**68.88**	89.90	91.11
	$Ours_t$	**83.67**	94.44	**90.87**	78.31	69.80	**90.12**	78.41	67.76	**90.91**	**92.45**
VGGI6	NoDefense	42.31	**93.58**	81.51	55.16	49.36	75.39	10.05	4.71	6.03	5.02
	Ours	82.93	91.46	89.71	79.48	**71.93**	89.59	80.45	72.68	81.96	89.07
	$Ours_t$	**83.81**	92.73	**90.33**	**79.90**	71.57	**90.17**	**82.13**	**73.93**	**82.91**	**90.52**

Acknowledgements Project is supported by the Natural Science Foundation of Chongqing [CSTB2023NSCQ-MSX0063]

References

1. Uikey R, Lone HR, Agarwal A (2024) Indian traffic sign detection and classification through a unified framework. IEEE Trans Intell Transp Syst 25(10):14866–14875
2. Turay T, Vladimirova T (2022) Toward performing image classification and object detection with convolutional neural networks in autonomous driving systems: a survey. IEEE Access 10:14076–14119
3. Im Choi J, Tian Q (2022) Adversarial attack and defense of YOLO detectors in autonomous driving scenarios, pp 1011–1017
4. Qin C et al (2021) Efficient non-targeted attack for deep hashing based image retrieval. IEEE Signal Process Lett 28:1893–1897
5. Goodfellow IJ, Shlens J, Szegedy C (2014) Explaining and harnessing adversarial examples. arXiv preprint arXiv:1412.6572
6. Liang H et al (2022) Adversarial attack and defense: a survey. Electronics 11(8):1283
7. Li Y et al (2022) A review of adversarial attack and defense for classification methods. Am Stat 76(4):329–345
8. Wang N et al (2023) Does physical adversarial example really matter to autonomous driving? Towards system-level effect of adversarial object evasion attack, pp 4412–4423
9. Badjie B, Cec JE, Lio I, Casimiro A (2024) Adversarial attacks and countermeasures on image classification-based deep learning models in autonomous driving systems: a systematic review. ACM Comput Surv
10. Nicolae M et al (2018) Adversarial robustness toolbox v1.2.0. CoRR. 1807.01069
11. Xu W, Evans D, Qi Y (2017) Feature squeezing: detecting adversarial examples in deep neural networks. arXiv preprint arXiv:1704.01155
12. Qiu H et al (2020) Mitigating advanced adversarial attacks with more advanced gradient obfuscation techniques. arXiv preprint arXiv:2005.13712
13. Li F, Du X, Zhang L (2022) Adversarial attacks defense method based on multiple filtering and image rotation. Discr Dyn Nat Soc 2022(1)
14. Dziugaite GK, Ghahramani Z, Roy DM (2016) A study of the effect of JPG compression on adversarial images. arXiv preprint arXiv:1608.00853
15. Aydemir AE, Temizel A, Temizel TT (2018) The effects of JPEG and JPEG2000 compression on attacks using adversarial examples. arXiv preprint arXiv:1803.10418
16. Das N et al (2018) SHIELD: fast, practical defense and vaccination for deep learning using JPEG compression, pp 196–204
17. Liu Z et al (2019) Feature distillation: Dnn-oriented jpeg compression against adversarial examples, pp 860–868
18. Qiu H et al (2021) An efficient preprocessing-based approach to mitigate advanced adversarial attacks. IEEE Trans Comput
19. Springenberg JT et al (2014) Striving for simplicity: the all convolutional net. arXiv preprint arXiv:1412.6806
20. Krizhevsky A (2009) Learning multiple layers of features from tiny images. University of Toronto
21. Lin T et al (2017) Focal loss for dense object detection, pp 2980–2988
22. Madry A et al (2018) Towards deep learning models resistant to adversarial attacks, p 9
23. Dezfooli SMM, Fawzi A, Frossard P (2016) DeepFool: a simple and accurate method to fool deep neural networks, pp 2574–2582
24. Carlini N, Wagner D (2017) Towards evaluating the robustness of neural networks, pp 39–57

25. Croce F, Hein M (2020) Reliable evaluation of adversarial robustness with an ensemble of diverse parameter-free attacks, pp 2206–2216
26. Athalye A, Carlini N, Wagner D (2018) Obfuscated gradients give a false sense of security: circumventing defenses to adversarial examples, pp 274–283
27. Tramer F et al (2020) On adaptive attacks to adversarial example defenses. Adv Neural Inf Process Syst 33:1633–1645
28. Russakovsky O et al (2015) Imagenet large scale visual recognition challenge. Int J Comput Vis 115(3):211–252

Open Access This chapter is licensed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License (http://creativecommons.org/licenses/by-nc-nd/4.0/), which permits any noncommercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if you modified the licensed material. You do not have permission under this license to share adapted material derived from this chapter or parts of it.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.



Hydraulic Characteristics Analysis of Working Device for Typical Working Conditions of Mini Excavators with Electric Drive Excavation



Yifan Zhang, Li Chen, Zhou You, and Yongming Bian

Abstract By adopting high-capacity batteries and high-power electric motors, electric-driven excavators have replaced traditional internal combustion engines and become an important solution to environmental challenges. With the increasing demand for environmental protection, electric-driven mini excavators have received widespread attention for their low-noise and zero-emission characteristics. In this paper, a systematic study is conducted on the hydraulic characteristics of the electric-driven mini excavator under three typical working conditions. By arranging sensors in the key parts of the excavator, and carrying out three typical working conditions of the real machine operation, real-time collection of the data of the hydraulic system of the working device, this paper analyzes the oil pressure changes of the cylinder under each working condition. The research results reveal the characteristics of the hydraulic system under different working conditions, and provide a theoretical basis for the promotion and optimization of the electric drive mini excavator in practical applications.

Keywords Electric mini excavator · Work unit · Hydraulic system

Y. Zhang · Z. You · Y. Bian (✉)
School of Mechanical Engineering, Tongji University, Shanghai, China
e-mail: ymbianmail@163.com

L. Chen
School of Automotive Studies, Tongji University, Shanghai, China

Y. Bian
Shanghai Engineering Research Center for Safety Intelligent Control of Building Machinery, Shanghai, China

© The Author(s) 2026

H. Saman et al. (eds.), *The 9th International Conference on Advances in Construction Machinery and Vehicle Engineering*, Lecture Notes in Mechanical Engineering,
https://doi.org/10.1007/978-981-96-5527-4_116

1 Introduction

Excavators are one of the most commonly used construction machines in engineering and construction, and are mainly used to use buckets to excavate materials and transfer them to designated locations. At present, excavators usually use diesel engines as the power source. Diesel engines have the advantages of high torque, good economic performance and high efficiency, which can provide sufficient power for the operation of excavators. However, with the rapid development of global industry, human beings are facing serious energy shortage and environmental pollution problems. Diesel engines have high fuel consumption, are noisy and emit polluting gases and solid particles, which is not in line with the requirements of today's society for environmental protection [1–3].

In order to respond to the national concept of accelerating the establishment of green, low-carbon and recycling development, electric-driven excavators are gradually entering the public's field of vision. Electric-driven excavators use high-capacity batteries and high-power motors to replace the fuel tank and engine driven by internal combustion engines. Compared with internal combustion engine-driven excavators, electric-driven excavators perform better in terms of efficiency, noise and pollution [4–7].

Both electric and combustion engine-driven excavators rely on hydraulics to drive their work units. Therefore, to determine whether the electric drive excavator can replace the internal combustion engine-driven excavator, it is necessary to analyze its hydraulic system [8–10]. In this paper, we will use the oil pressure data of each part of the electric-driven mini excavator SY19E under typical working conditions to analyze the load capacity of the electric-driven mini excavator and its corresponding performance, in order to reflect the overall performance level of the electric-driven mini excavator.

2 Experimental Method

2.1 Overall Experimental Program

To analyze the hydraulic system of the excavator, six pressure sensors were installed on the hydraulic cylinders of the boom, arm, and bucket. The data acquisition scheme for the pressure sensors is illustrated in Fig. 1.

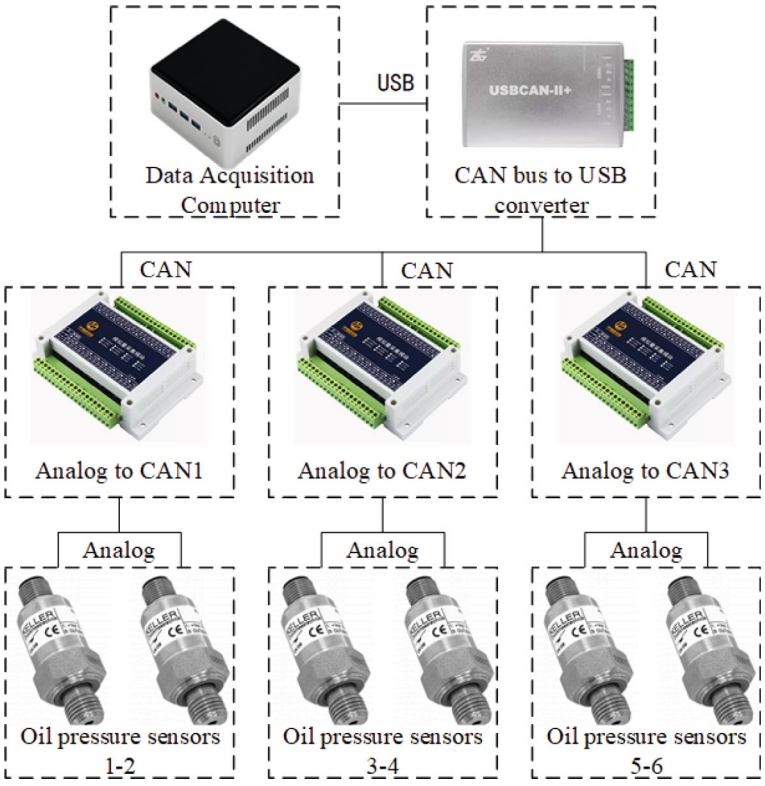


Fig. 1 Electric drive mini excavator test system overall program

2.2 *Construction of Commissioning Box and Sensor Calibration*

Before the formal installation and commissioning of electric drive mini excavation, it is necessary to build a test box for commissioning to determine the feasibility of the scheme, test the use and setting of each module, and calibrate the oil pressure sensor, so as to obtain more accurate data in the experiment. The commissioning box is shown in Fig. 2.

Oil pressure sensor detects 0–40 MPa oil pressure and outputs the value with 4–20 mA current, oil pressure and current are linear correspondence, but the correspondence may be shifted in actual use, so it is necessary to calibrate each oil pressure sensor individually. The calibration process was as follows: the sensor was mounted on the calibration equipment and pressurized to 0, 10, 20 and 30 MPa respectively. Readings were taken and summarized as shown in Table 1.

According to the excavator manual, it is known that the oil pressure of the hydraulic system of the work unit is in the range of 0–25 MPa, so it is reasonable to take the



Fig. 2 The commissioning box

Table 1 Oil pressure sensors calibration table

Sensor number	0 MPa	10 MPa	20 MPa	30 Mpa
1	03D8 (984)	0808 (2056)	0BC3 (3011)	0F1B (3963)
2	0447 (1095)	0807 (2055)	0BC4 (3012)	0F95 (3989)
3	0401 (1025)	081F (2079)	0BDC (3036)	0FA7 (4007)
4	03D8 (984)	0804 (2052)	0BC4 (3012)	0F99 (3993)
5	03DA (986)	080A (2058)	0BC0 (3008)	0F90 (3984)
6	03D8 (984)	0822 (2082)	0BCD (3021)	0F85 (3973)

two points of 10 and 20 MPa to determine the formula for converting the oil pressure readings, and the specific formula is calculated as follows:

$$Y_1 = 0.01X_1 - 10.11 \tag{1}$$

$$Y_2 = 0.01X_2 - 10.12 \tag{2}$$

$$Y_3 = 0.01X_3 - 10.36 \tag{3}$$

$$Y_4 = 0.01X_4 - 10.12 \tag{4}$$

$$Y_5 = 0.01X_5 - 10.08 \tag{5}$$

$$Y_6 = 0.01X_6 - 10.21 \tag{6}$$

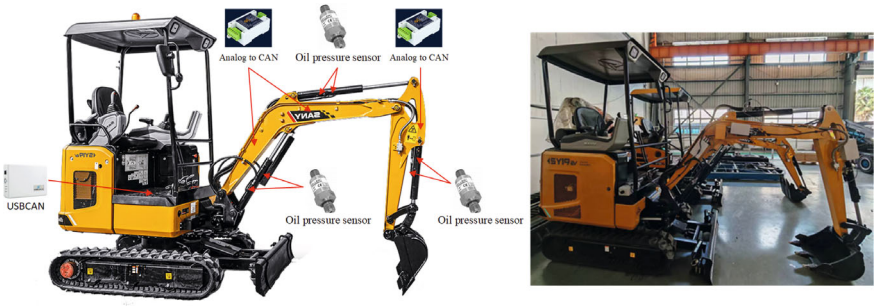


Fig. 3 SY19E electric drive excavator hardware installation diagram

Among them: Y_1–Y_6 are the pressure values of 6 oil pressure sensors, MPa; X_1–X_6 are the upper computer reading decimal values.

2.3 Component Installation

Install all the equipment to the excavator, the installation scheme and installation results are shown in Fig. 3.

After the construction of the hardware to carry out the power test, after the test is correct to carry out the bucket single-action, moving arm single-action, bucket rod single-action three typical conditions of excavation of gravel test, the experimental process is shown in Fig. 4.

2.4 Test Data

After the above experiments, the operating pressures of the cylinders were obtained when the electric-driven mini excavator was working under three typical working conditions, namely, single-acting boom, single-acting dipper arm, and single-acting bucket. The results are shown in Figs. 5, 6 and 7.

2.5 Methods of Analysis

The cylinder size of the excavator has been measured as $D = 55$ mm for the outer diameter and $d = 35$ mm for the rod diameter, from which the following data can be calculated:



Fig. 4 SY19E test process pictures

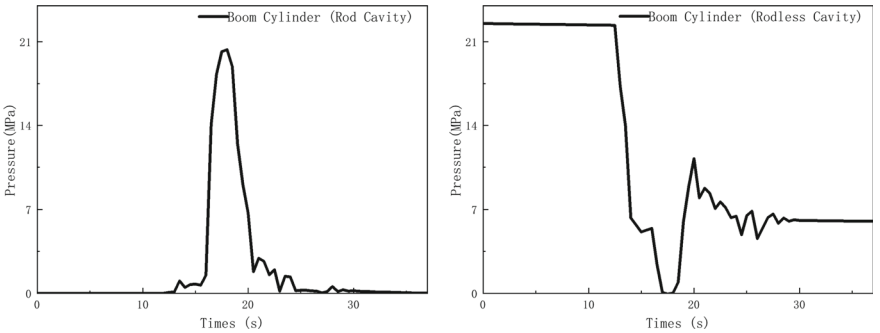


Fig. 5 Test results of the single-acting boom of the SY19E

$$A = \frac{\pi D^2}{4} \tag{7}$$

For each hydraulic cylinder, the force provided by the cylinder can be calculated by the following formula:

$$F = P_1 A_1 - P_2 A_2 \tag{8}$$

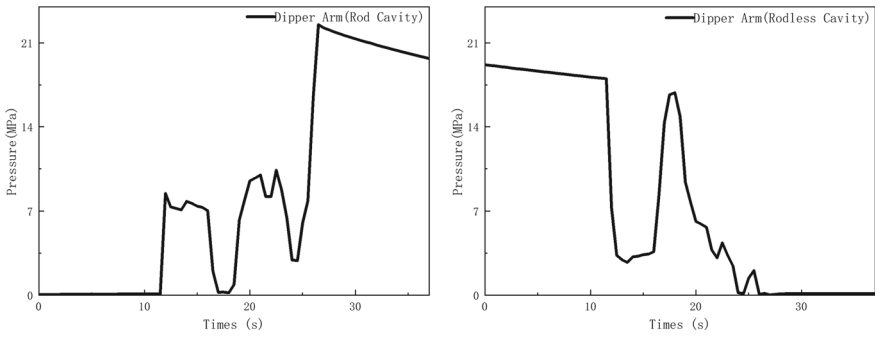


Fig. 6 Test results of the single-acting dipper arm of the SY19E

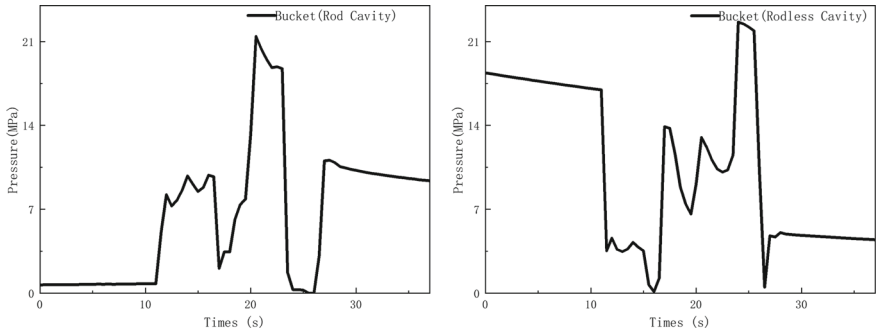


Fig. 7 Test results of the single-acting bucket of the SY19E

In the above equation, P_1, P_2: the pressure of the two chambers of the hydraulic cylinder; A_1, A_2: the area of the two chambers of the hydraulic cylinder; F: hydraulic cylinder load force.

At the same time, we can calculate the output power of the cylinder of the electric drive mini excavator. The first thing needed is the flow rate when the cylinder is working. The cylinder extends and retracts at a steady rate during the test, with a speed v = 0.1 m/s. The flow rate is calculated as follows:

$$q = Av \tag{9}$$

In the above equation, q represents the flow rate in the cylinder; v represents the cylinder speed; and A represents the cylinder area at the corresponding working condition.

For the typical working conditions of an electric drive mini excavator, the input power of the cylinder is determined by the pressure and flow rate. The formula is as follows:

$$P = |P_1 - P_2| \times q \tag{10}$$

Table 2 Cylinder load and input power under three typical working conditions

Working condition	Single-acting boom	Single-acting dipper arm	Single-acting bucket
Maximum load (kN)	53.5	45.5	53.4
Maximum power (W)	5337	4534	5297

In the above equation, P represents the input power of the cylinder, P1 and P2 represent the pressure of the two chambers of the cylinder, and q represents the flow rate in the cylinder.

Based on the above equations, the cylinder load and input power of the electric-driven excavator under three typical working conditions are calculated as shown in Table 2.

3 Conclusion

This experiment involved installing oil pressure sensors on an electric-driven excavator to record hydraulic pressure data of the hydraulic cylinders under three typical operating conditions. Based on the pressure data, the load capacity of the hydraulic cylinders and the input power were calculated for each typical condition. These findings provide valuable support for the study of operational endurance and motor load in electric-driven mini excavators. However, this experiment has certain limitations, such as a small sample size and the restricted applicability of the research to specific construction scenarios. These aspects warrant further investigation in future studies.

Acknowledgements This work supported by the National Key Research and Development Program of China (Grant No. 2023YFF0613200).

References

1. Zhang S, Deng B, Song C (2008) Technology and tendency of wheel excavator. World Sci-Tech R and D 30:336–339
2. Khoa NX, Lim O (2022) A review of the external and internal residual exhaust gas in the internal combustion engine. Energies 15:1208
3. Shen Y, Li X, Wang J, Liu C, Kong W (2023) An extrapolation framework for torque spectrum of excavator internal combustion engine via bivariate diffusion-based kernel density estimation. Proc Inst Mech Eng C J Mech Eng Sci 237:133–147
4. Shen W, Jiang J, Su X, Karimi HR (2015) Control strategy analysis of the hydraulic hybrid excavator. J Franklin Inst 352:541–561
5. Liu B, Yan Z, Ge L, Quan L (2019) Electric drive hydraulic power source characteristics of small hydraulic excavator. Trans Chin Soc Agric Machin 50:387–393
6. Yu Y, Do TC, Park Y, Ahn KK (2021) Energy saving of hybrid hydraulic excavator with innovative powertrain. Energy Convers Manage 244:114447

7. Yuan Z, Ma W, Yin C, Ma S, Cao D (2024) Research on control and energy consumption characteristics of distributed independent electrical-hydraulic actuator control system used on excavators. IEEE Sens J 24(21):34070–34082
8. You Z, Bian Y, Zhang Y, Chen L (2024) Model predictive control of permanent magnet synchronous motor based on optimal switching frequency. In: 2024 9th international conference on control and robotics engineering (ICCRE), pp 235–240
9. Wang PY, Wang QN, Zhang QM, Yang SQ (2010) The analysis of energy-saving mechanism of hybrid excavator. In: Tan HH (ed) Applied mechanics and mechanical engineering, PTS 1–3. vol 29–32. International conference on applied mechanics and mechanical engineering, p 1657
10. Wang F, Wu J, Xu B, Sun Z (2023) A novel hydraulic load-sensing control solution to working functions of mobile machine. Proc Inst Mechan Eng Part I: J Syst Control Eng 237:1147–1160

Open Access This chapter is licensed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License (http://creativecommons.org/licenses/by-nc-nd/4.0/), which permits any noncommercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if you modified the licensed material. You do not have permission under this license to share adapted material derived from this chapter or parts of it.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.



Integral Backstepping Displacement Tracking Control for a Hydraulic Active Suspension Actuator



Lizhe Wu, Dingxuan Zhao, Tao Ni, Tuo Jia, and Qinghe Guo

Abstract The control effectiveness of electro-hydraulic actuators significantly determines the overall performance of hydraulic active suspension systems. However, the electro-hydraulic actuator systems exhibit notable nonlinearity and various uncertainties, posing challenges for achieving high-precision displacement tracking control. This paper first establishes the servo valve flow equation, active suspension actuator flow equation, and the system's dynamic equations based on the electro-hydraulic servo model of the hydraulic active suspension system, providing theoretical support for the control strategy design. Subsequently, an integral backstepping control method is introduced, which constructs a Lyapunov function to ensure the system's asymptotic stability and achieve precise displacement tracking control of the hydraulic active suspension actuator. This method incorporates the integral term into the backstepping control design, enhancing the system's ability to compensate for model uncertainties and external disturbances. Through displacement tracking and error analysis for three different desired trajectories, the results demonstrate that the proposed control strategy exhibits excellent performance in terms of displacement tracking accuracy, disturbance rejection, and system stability. This study provides an effective control scheme to enhance the dynamic response and steady-state accuracy of hydraulic active suspension systems.

Keywords Electro-hydraulic actuator · Hydraulic active suspension · Integral backstepping control · Displacement tracking

L. Wu · D. Zhao (✉) · T. Ni · T. Jia
Yanshan University, Qinhuangdao, China
e-mail: zdx-yw@ysu.edu.cn

Q. Guo
Hubei University of Automotive Technology, Shiyan, China

© The Author(s) 2026

H. Saman et al. (eds.), *The 9th International Conference on Advances in Construction Machinery and Vehicle Engineering*, Lecture Notes in Mechanical Engineering,
https://doi.org/10.1007/978-981-96-5527-4_117

1 Introduction

The vehicle suspension [1] plays a pivotal role in maintaining stability and enhancing off-road performance. Its main functions include effectively reducing shocks caused by road surface irregularities, suppressing vibrations induced by dynamic tire loads, and enhancing the adhesion between the tires and the road. The design of the suspension system structure and its vibration-damping performance largely determine the smoothness and handling stability of the vehicle during operation. Therefore, designing a suspension system with excellent vibration-damping and dynamic response performance is essential for ensuring driving safety and passenger comfort.

From the perspective of vibration reduction, passive suspension systems demonstrate good damping effects only under specific driving conditions, but they often struggle to adapt in complex and variable operating environments, potentially affecting the vehicle's handling performance. Active suspension systems monitor the vehicle's state and road information in real time through sensors, using controllers to adjust the output energy of actuators to counteract shocks caused by uneven road surfaces. This significantly enhances the vehicle's dynamic response and comfort across various operating conditions, effectively compensating for the shortcomings of passive suspensions.

In active suspension systems, hydraulic active suspensions often employ asymmetric hydraulic cylinder structures. Compared to active air suspensions and electromagnetic suspensions, hydraulic systems offer advantages such as compact size, high output force, rapid response, and a high power-to-weight ratio, making them particularly suitable for space-constrained suspension systems. They have gradually been applied in vehicle suspension systems [2]. However, the electro-hydraulic actuator system faces significant nonlinearity and multiple uncertainties, such as pressure nonlinearity in the actuator chamber, flow nonlinearity of the servo valve, and uncertainty in system parameters, posing severe challenges for high-precision control of the actuator. Therefore, the precision of displacement control in hydraulic actuators is crucial for achieving accurate control of active suspension systems and has a decisive impact on the vibration-damping and dynamic response performance of the suspension system.

Integral backstepping control, as a nonlinear control method that combines integral control and backstepping control concepts, exhibits strong robustness against uncertain disturbances and system nonlinearities, making it potentially valuable for displacement tracking control in electro-hydraulic actuators. For instance, Liu et al. [3] proposed an adaptive fuzzy integral backstepping control strategy with command filters to address the trajectory tracking problem of quadrotor UAVs. Haruna et al. [4] investigated the energy efficiency of the widespread application of backstepping control in a class of nonlinear dynamic systems, proposing a switched-step integral backstepping control (SSIBC) scheme to enhance robustness against measurement noise. Chetouani et al. [5] developed a nonlinear integral backstepping control approach for grid-connected variable frequency systems during voltage drops, utilizing particle swarm optimization. Feng et al. [6] introduced an adaptive

integral backstepping control strategy for the precision motion of ultrasonic motors, based on a piezoelectric integral terminal sliding mode technique. Yue et al. [7] proposed a robust adaptive integral backstepping control method for a photoelectric tracking system, employing a modified LuGre friction model. Lastly, Zhao et al. [8] presented an integral backstepping control method to regulate sinusoidal ripple current, with modeling and analysis conducted on a full-bridge DC/DC converter.

Therefore, to further improve the position tracking regulation of the hydraulic active suspension systems, the integral backstepping control method is introduced.

2 Nonlinear Dynamic Model of Hydraulic Active Suspension Actuator

2.1 Schematic Diagram of Hydraulic Active Suspension

The working principle of the system is depicted in Fig. 1. In the system, a displacement sensor is integrated within the actuator to continuously monitor the output displacement of the hydraulic cylinder and provide feedback. The controller employs integral backstepping closed-loop control for the electro-hydraulic servo valve, guided by the magnitude of the positional deviation. Additionally, each chamber of the hydraulic suspension actuator system is equipped with a pressure sensor to monitor the hydraulic fluid pressure in real-time.

Building upon the diagram of the hydraulic suspension control system presented in Fig. 1, this study develops a nonlinear representation of the electro-hydraulic servo mechanism under reasonable assumptions aligned with practical conditions, aiming to deliver a more precise depiction of the system's dynamic characteristics.

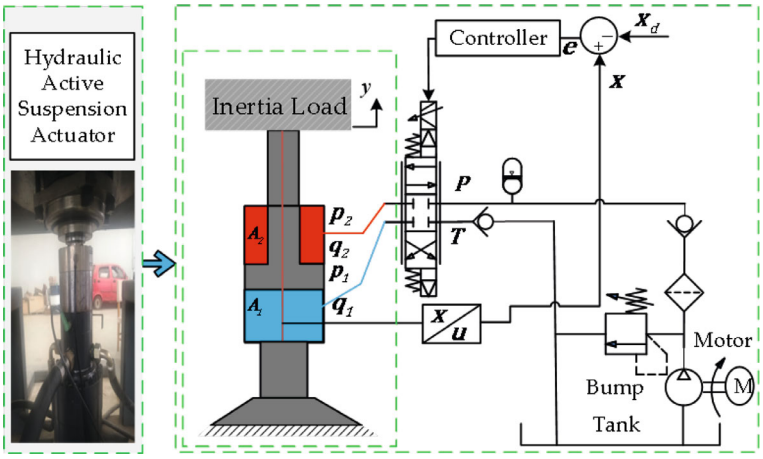


Fig. 1 Schematic of the hydraulic active suspension servo control system

2.2 Dynamic Model

The servo valve's flow equation is formulated as follows:

$$Q_1 = \sqrt{2}k_{q1}x_v\left[s(x_v)\sqrt{P_s - P_1} + s(-x_v)\sqrt{P_1 - P_r}\right] \tag{1}$$

$$Q_2 = \sqrt{2}k_{q2}x_v\left[s(x_v)\sqrt{P_2 - P_r} + s(-x_v)\sqrt{P_s - P_2}\right] \tag{2}$$

where $k_{q1} = C_d W_1\sqrt{\frac{1}{\rho}}$, $k_{q2} = C_d W_2\sqrt{\frac{1}{\rho}}$.

$s(x_v)$ is defined as

$$s(x_v) = \begin{cases} 1, & x_v \geq 0 \\ 0, & x_v < 0 \end{cases} \tag{3}$$

Here, k_{q1} and k_{q2} denote the flow rate coefficients associated with the left and right ends of the servo valve spool displacement. C_d signifies the discharge coefficient. W_1 and W_2 describe the rates of area change at the left and right openings of the throttle orifice within the servo valve assembly. ρ represents the density of the hydraulic oil. P_r refers to the system's return pressure, while P_s indicates the supplied pressure.

In practical control applications, because the frequency of the servo valve is much higher than the system bandwidth, the spool displacement can be approximated as a linear proportional relationship with the control output, represented as follows:

$$x_v = Bu, \quad s(x_v) = s(u) \tag{4}$$

where u is the system's control output; B is the servo valve gain.

Disregarding any leakage from the hydraulic cylinder, the pressure fluctuations within its two chambers can be expressed as follows:

$$\dot{P}_1 = \frac{\beta_{e1}}{V_{01} + A_1 y}(-A_1\dot{y} - C_t P_L + Q_1) \tag{5}$$

$$\dot{P}_2 = \frac{\beta_{e2}}{V_{02} - A_2 y}(A_2\dot{y} + C_t P_L - Q_2) \tag{6}$$

where $V_1 = V_{01} + A_2 y$, $V_2 = V_{02} - A_2 y$, $P_L = P_1 - P_2$. V_{01} represents the initial capacity of the system's left-hand chamber, while V_{02} denotes the initial capacity of its right-hand chamber. β_{e1} and β_{e2} signify the effective compressibility of the hydraulic fluid V_1 and V_2 indicate the control volumes of the left and right cavities, respectively. P_L denotes the pressure differential between the two sides of the hydraulic cylinder.

Additionally, C_t characterizes the internal leakage coefficient of the actuator caused by hydraulic force. Furthermore, Q_1 and Q_2 are the flow equations for the servo valve derived from Eqs. (1) and (2).

Based on Newton's second law, the dynamic equations of the system can be expressed as follows:

$$m\ddot{y} = P_1A_1 - P_2A_2 - mg - B_p\dot{y} \tag{7}$$

where A_1 and A_2 correspond to the effective areas of the piston without the rod and the piston with the rod, respectively; m denotes the mass of the inertial load; B_p signifies gravitational acceleration; g represents the hydraulic fluid's resistance to flow; and y refers to the displacement output of the actuator.

Assumption 1: The servo valve is matched and symmetric, such that $k_{q1} = k_{q2} = k_q$.

Assumption 2: The total compressible volume of the system remains constant during operation, such that $\beta_{e1} = \beta_{e2} = \beta_e$.

Assumption 3: External leakage from the servo valve and actuator is negligible.

Assumption 4: The supply oil pressure P_s is a constant value.

Assumption 5: Electro-hydraulic servo actuators function under standard conditions, necessitating that the pressures within their two chambers adhere to the parameters $0 < P_r < P_1 < P_s$ and $0 < P_r < P_2 < P_s$.

From Eqs. (1) to (4), Q_1, and Q_2 are subsequently represented as follows:

$$Q_1 = NM_1u \tag{8}$$

$$Q_2 = NM_2u \tag{9}$$

where $N = \sqrt{2}k_qB$, $M_1 = \left[s(u)\sqrt{P_s - P_1} + s(-u)\sqrt{P_1 - P_r}\right]$, $M_2 = \left[s(u)\sqrt{P_2 - P_r} + s(-u)\sqrt{P_s - P_2}\right]$.

By defining the state parameter $x = \left[y, \dot{y}, P_1A_1 - P_2A_2\right]$, the derived dynamic Eqs. (1) to (9) are reformulated into the following state-space equations:

$$\begin{cases} \dot{x}_1 = x_2 \\ \dot{x}_2 = \dfrac{1}{m}\left(x_3 - mg - B_px_2\right) \\ \dot{x}_3 = B_3u - G_c - C_tH_u \\ y = x_1 \end{cases} \tag{10}$$

where $B_3 = \left(\frac{A_1E_1}{V_1} + \frac{A_2E_2}{V_2}\right)N\beta_e$, $G_c = \left(\frac{A_1^2}{V_1} + \frac{A_2^2}{V_2}\right)\beta_ex_2$, $H_u = \left(\frac{A_1}{V_1} + \frac{A_2}{V_2}\right)(P_1 - P_2)\beta_e$.

3 Design of Integral Backstepping Displacement Tracking Controller for Hydraulic Active Suspension Actuator and Stability Proof

3.1 Controller Design

Assume that the target regulation for state variables is x_{id} $(i = 1, 2, 3)$, the error is $z_i = x_{\mathrm{id}} - x_i$ $(i = 1, 2, 3)$, and the controller parameters are k_i $(i = 1, 2, 3)$ and k_{pi} $(i = 1, 2, 3)$, meeting the conditions $k_i > 0$ and $k_{pi} > 0$.

Assumption 6 The system reference command signal x_{1d} is a continuous third-order signal, and the desired position, velocity, acceleration, and jerk commands of the system are all bounded.

Step 1: Based on the state Eq. (10) and the state tracking error $z_1 = x_{1d} - x_1$, the dynamic equation is given by

$$\dot{z}_1 = \dot{x}_{1d} - \dot{x}_1 = \dot{x}_{1d} - x_2 \tag{11}$$

Design the virtual controller $x_{2\mathrm{d}}$ as

$$x_{2\mathrm{d}} = \dot{x}_{1\mathrm{d}} + k_1 z_1 + k_{p1}\varphi_1 \tag{12}$$

where $\varphi_1 = \int_0^t z_1(\tau)d\tau$.

Step 2: Based on the state tracking error $z_2 = x_{2d} - x_2$, the dynamic equation of z_2 is given by

$$\dot{z}_2 = \dot{x}_{2d} - \dot{x}_2 = \dot{x}_{2d} - \frac{1}{m}\left(x_3 - mg - B_p x_2\right) \tag{13}$$

Design the virtual controller $x_{3\mathrm{d}}$ as

$$x_{3\mathrm{d}} = mz_1 + m\dot{x}_{2\mathrm{d}} + B_p x_2 + mk_{p2}\varphi_2 + mg + mk_2 z_2 \tag{14}$$

where $\varphi_2 = \int_0^t z_2(\tau)d\tau$.

Step 3: Based on the state tracking error $z_3 = x_{3d} - x_3$, the dynamic equation for z_3 is given as follows:

$$\dot{z}_3 = \dot{x}_{3d} - \dot{x}_3 = \dot{x}_{3d} - (B_3 u - G_c - C_t H_u) \tag{15}$$

The final control law u is given as follows:

$$u = \frac{1}{B_3}\left[k_3 z_3 + G_c + C_t H_u + \dot{x}_{3d} + \frac{e_2}{m} + k_{p3}\varphi_3\right] \tag{16}$$

where $\varphi_3 = \int_0^t z_3(\tau)d\tau$.

3.2 Proof of Stability

Step 1: Define the Lyapunov function

$$V_1(z_1, \varphi_1) = \frac{1}{2}z_1^2 + \frac{1}{2}k_{p1}\varphi_1^2 \tag{17}$$

Taking the derivative of $V_1(z_1, \varphi_1)$ yields:

$$\dot{V}_1(z_1, \varphi_1) = z_1\dot{z}_1 + k_{p1}\varphi_1\dot{\varphi}_1 = z_1(\dot{x}_{1d} - x_2) + k_{p1}\varphi_1 z_1 = -k_1 z_1^2 + z_1 z_2 \tag{18}$$

Step 2: Define the Lyapunov function

$$V_2(z_1, z_2, \varphi_1, \varphi_2) = \frac{1}{2}z_1^2 + \frac{1}{2}k_{p1}\varphi_1^2 + \frac{1}{2}z_2^2 + \frac{1}{2}k_{p2}\varphi_2^2 \tag{19}$$

Taking the derivative of $V_2(z_1, z_2, \varphi_1, \varphi_2)$ yields:

$$\begin{aligned}\dot{V}_2(z_1,\ z_2, \varphi_1, \varphi_2) &= z_1\dot{z}_1 + k_{p1}\varphi_1\dot{\varphi}_1 + z_2\dot{z}_2 + k_{p2}\varphi_2\dot{\varphi}_2 \\ &= -k_1 z_1^2 + z_1 z_2 + z_2\dot{z}_2 + k_{p2}\varphi_2 z_2 \\ &= -k_1 z_1^2 + z_2\left(z_1 + \dot{x}_{2d} - \frac{1}{m}\left(x_3 - mg - B_p x_2\right) + k_{p2}\varphi_2\right) \\ &= -k_1 z_1^2 - k_2 z_2^2 + \frac{z_2 z_3}{m}\end{aligned} \tag{20}$$

Step 3: Define the Lyapunov function

$$V_3(z_1, z_2, z_3,\ \varphi_1,\ \varphi_2,\ \varphi_3) = \frac{1}{2}z_1^2 + \frac{1}{2}z_2^2 + \frac{1}{2}z_3^2 + \frac{1}{2}k_{p1}\varphi_1^2 + \frac{1}{2}k_{p2}\varphi_2^2 + \frac{1}{2}k_{p3}\varphi_3^2 \tag{21}$$

Taking the derivative of $V_3(z_1, z_2, z_3,\ \varphi_1,\ \varphi_2,\ \varphi_3)$ yields:

$$\begin{aligned}\dot{V}_3(z_1, z_2, z_3,\ \varphi_1,\ \varphi_2,\ \varphi_3) &= -k_1 z_1^2 - k_2 z_2^2 + \frac{z_2 z_3}{m} + z_3\dot{z}_3 + k_{p3}\varphi_3\dot{\varphi}_3 \\ &= -k_1 z_1^2 - k_2 z_2^2 + \frac{z_2 z_3}{m} + z_3\dot{z}_3 + k_{p3}\varphi_3 z_3\end{aligned}$$

Table 1 Dynamic parameters

Parameters	Values	Parameters	Values
m	1500 kg	C_d	0.7
A_1	5.67×10^{-3} m^2	$W_1 = W_2$	1.12×10^{-2} m
A_2	1.256×10^{-3} m^2	g	9.8 m/ s^2
β_e	7.0×10^8 pa	ρ	860 kg$\big/$ m^3
B_p	2600 N s/ m	P_s	1.78×10^7 Pa
C_t	3.0×10^{-13} m^3/s$\big/$ pa	P_r	0 Pa
$V_{01} = V_{02}$	1×10^{-3} m^3	N	3.78×10^{-4} m$^4\Big/\left(s \cdot V \cdot \sqrt{N}\right)$

$$\begin{aligned} &= -k_1 z_1^2 - k_2 z_2^2 + z_3\left(\frac{z_2}{m} + \dot{x}_{3d} - B_3 u + G_c + C_t H_u + k_{\mathrm{p}3}\varphi_3\right) \\ &= -k_1 z_1^2 - k_2 z_2^2 - k_3 z_3^2 \end{aligned} \tag{22}$$

According to Lyapunov stability theory, it can be concluded that the proposed control law guarantees the system's asymptotic convergence, ensuring that the entire system tracks the intended path.

4 Simulation Test Verification

4.1 Parameter Settings

To verify the control effectiveness of the integral backstepping method, the setup was implemented in Matlab/Simulink with a sampling time of 1 ms and a simulation duration of 20 s. The desired displacement tracking signals were set to $x_{1d} = 0.1\sin(1.884t)$,$0.1\sin(3.14t) \times \left[1 - \exp(-0.01t^3)\right]$ and $0.1\sin(6.28t)$. Table 1 provides the dynamic parameters of the hydraulic active suspension actuator.

4.2 Simulation Validation and Result Analysis

The parameters for the integral backstepping controller are set as $k_1 = 200, k_2 = 200$, $k_3 = 220, k_{p1} = 1, k_{p2} = 5$, and $k_{p3} = 6$. The displacement tracking results are shown in Figs. 2, 3, and 4, respectively. Figure 5 presents a comparison of the displacement tracking errors at three different frequencies.

Based on the results shown in Figs. 2, 3, and 4, it is evident that the control system achieves effective tracking of the displacement signal at all three tested frequencies. However, as the tracking frequency increases, control performance declines slightly, as indicated by a gradual increase in tracking error. Figure 5 illustrates that, across

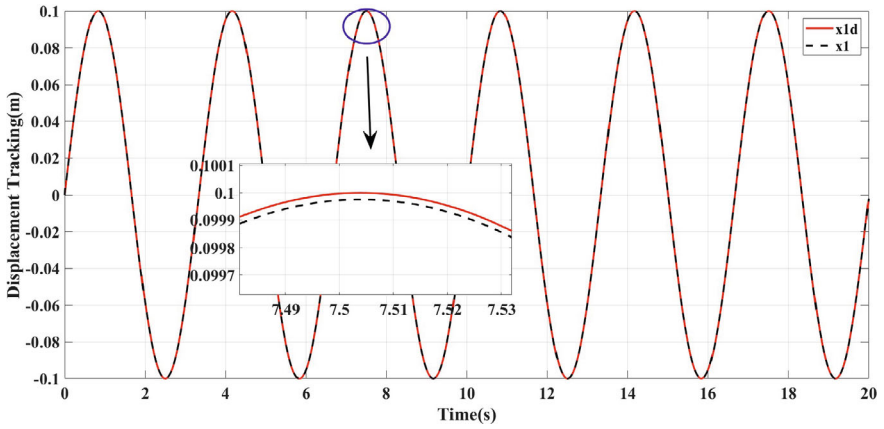


Fig. 2 Desired displacement tracking for integral backstepping control as $0.1\sin(1.884t)$

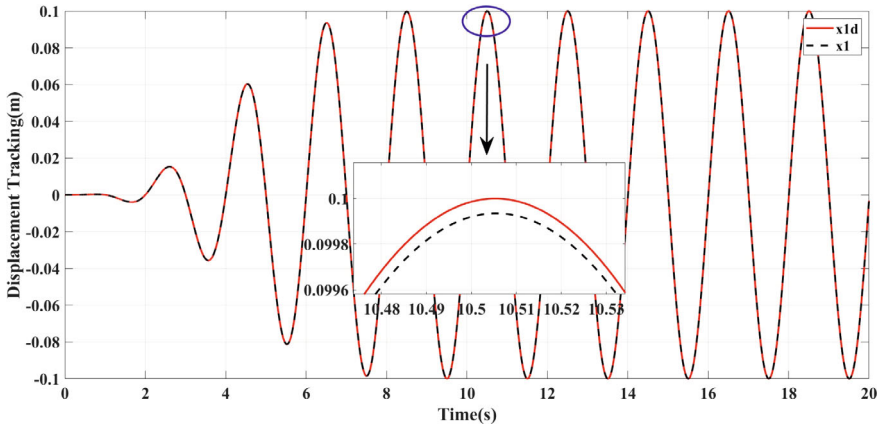


Fig. 3 Desired displacement tracking for integral backstepping control as $0.1\sin(3.14t) \times \left[1 - \exp(-0.01t^3)\right]$

all tested frequencies, the maximum error remains within the range of 10^{-4}(m) once the system reaches steady-state, demonstrating that the integral backstepping control method meets the requirements for tracking accuracy and stability within this frequency range. Therefore, this control method effectively fulfills the displacement tracking needs of the hydraulic active suspension system, verifying its applicability in practical scenarios.

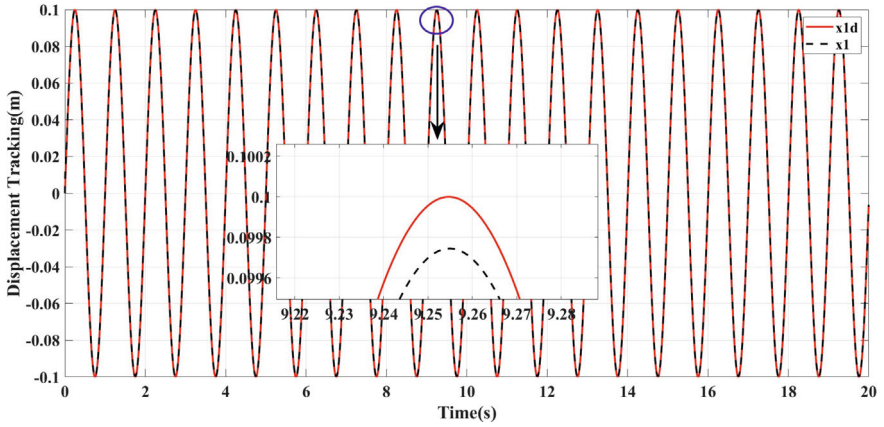


Fig. 4 Desired displacement tracking for integral backstepping control as $0.1\sin(6.28t)$

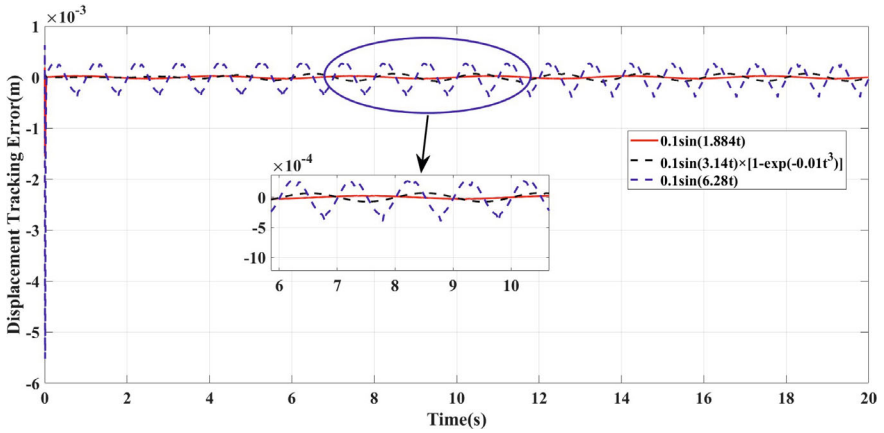


Fig. 5 Displacement tracking error for integral backstepping control

5 Conclusion

Simulation results indicate that the integral backstepping control strategy for the hydraulic active suspension actuator demonstrates excellent tracking accuracy and system stability. This conclusion, derived from data analysis of the simulation results, provides valuable insights for the design and optimization of hydraulic active suspension systems. Future research could explore more advanced control algorithms to enhance system control performance and stability at higher tracking frequencies, thereby further optimizing the overall performance of electro-hydraulic servo active suspension systems under various operating conditions.

Acknowledgements This study was supported by Grant No. U20A20332.

References

1. Hrovat D (1997) Survey of advanced suspension developments and related optimal control applications. Automatica 33(10):1781–1817
2. Cytrynski S, Neerpasch U, Bellmann R et al (2018) The active suspension of the new Mercedes-Benz GLE. ATZ Worldwide 120(2):42–45
3. Liu W et al (2022) Command filter-based adaptive fuzzy integral backstepping control for quadrotor UAV with input saturation. J Frankl Inst 360:484–507
4. Haruna A, Mohamed Z, Efe MÖ, Abdullahi AM (2023) Switched step integral backstepping control for nonlinear motion systems with application to a laboratory helicopter. ISA Trans 141:470–481
5. Chetouani E et al (2024) Nonlinear integral backstepping control based on particle swarm optimization for a grid-connected variable wind energy conversion system during voltage dips. Comput Electr Eng 120:109790
6. Feng Z et al (2020) Integral terminal sliding-mode-based adaptive integral backstepping control for precision motion of a piezoelectric ultrasonic motor. Mechan Syst Signal Process 144:106856
7. Yue F et al (2018) Robust adaptive integral backstepping control for opto-electronic tracking system based on modified LuGre friction model. ISA Trans 80:312–321
8. Zhao Y et al (2023) Integral backstepping control of sinusoidal ripple current charge for battery chargers. In: 2023 7th CAA international conference on vehicular control and intelligence (CVCI), pp 1–6

Open Access This chapter is licensed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License (http://creativecommons.org/licenses/by-nc-nd/4.0/), which permits any noncommercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if you modified the licensed material. You do not have permission under this license to share adapted material derived from this chapter or parts of it.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.



Impact of Friction Block Shape and Arrangement on High-Speed EMU Brake Noise Studied Using Complex Modal Analysis

Bo Yu, Xingchao Li, and Han Zhou

Abstract This paper utilizes the ANSYS Workbench platform to establish a shaft disc brake finite element model, aiming to analyze braking noise. By thoroughly examining the impact of the shape and arrangement of brake friction blocks, extensive simulations reveal significant variations in noise levels among different configurations. These findings offer valuable insights for the design of CR450 EMUs brake pads.

Keywords Complex mode method · Braking noise · Finite element simulation · Mode coupling · Shaft disc brake

1 Introduction

High-speed rail facilitates people's travel but also introduces certain nuisances, with one notable issue being noise pollution resulting from braking. Braking noise is typically generated by the friction between the brake disc and brake pad, producing one or more resonance frequencies during the braking process [1]. During braking, the brake disc functions as a speaker. The frequency of braking noise typically ranges from a few tens of Hz to 20,000 Hz, often categorized as low-frequency and high-frequency noise based on different frequency ranges [2]. This type of braking noise, caused by self-excited vibrations, can lead to fatigue in the braking mechanism, posing potential safety risks for the train and its passengers.

The explanation of the braking noise mechanism can be broadly categorized into two main theories: self-excited vibration and tribology [3]. Researchers typically approach this issue from the perspective of self-excited vibration. In the early study

B. Yu · X. Li
Locomotive & Car Research Institute, China Academy of Railway Sciences Corporation Limited, Beijing, China

H. Zhou (✉)
College of Transportation Engineering, Tongji University, Shanghai, China
e-mail: 2433670@tongji.edu.cn

© The Author(s) 2026

H. Saman et al. (eds.), *The 9th International Conference on Advances in Construction Machinery and Vehicle Engineering*, Lecture Notes in Mechanical Engineering,
https://doi.org/10.1007/978-981-96-5527-4_118

of braking noise (prior to the 1970s), the focus was initially on the friction surface itself, leading to the development of two simpler models: Stick–Slip motion [4] and the negative slope model of friction versus relative sliding speed [5]. Post-1980s, scholars began to consider that structural issues might be the root cause of the noise, thus proposing the Sprag-Slip theory [6]. This theory posits that vibration and friction noise stem from the structural instability of the entire friction system due to self-locking between friction surfaces [7]. Currently, it is widely accepted that the mode coupling mechanism, developed from the two-mode separation theory, is the primary cause of friction noise [8]. In essence, friction noise is a self-excited vibration phenomenon driven by mode coupling induced by friction.

This study utilizes the ANSYS Workbench platform to construct a finite element model of the shaft disc brake. Employing the complex modal analysis method for simulation, the brake noise is then analyzed through the control variable method. The investigation delves into various factors including the shape and arrangement of the brake friction blocks, ultimately leading to the formulation of an optimization strategy.

2 Methods

2.1 Derivation of Complex Modal Analysis Method

To perform a stability analysis of braking noise, the model must first be discretized. This involves defining the contact relationships among the model components and formulating the corresponding motion equations [9]. The dynamics of the braking system can be expressed as:

$$[M]\{\ddot{x}\} + [C]\{\dot{x}\} + [K]\{x\} = \{0\} \tag{1}$$

Here, [M] represents the mass matrix, [C] denotes the damping matrix, [K] indicates the stiffness matrix, and {x} is the displacement vector. The dimensions of each matrix here will change accordingly as the degrees of freedom of the brake disc structure vary.

Due to the friction present in the braking system, the stiffness matrix [K] can be expressed in the following manner:

$$[K] = [K_s] + \mu[K_f] \tag{2}$$

where $[K_s]$ is the structural stiffness matrix, $[K_f]$ is the asymmetric stiffness matrix caused by the friction force, and μ is the friction coefficient. It is important to note that the stiffness matrix [K] in this context is asymmetric. This asymmetric nature of the stiffness matrix directly hampers the use of conventional methods for solving the

equation, and it also results in eigenvalues that are complex numbers, as discussed further below.

Let the solution of the equation be:

$$\{x\} = \{\varphi\}e^{\lambda t} \tag{3}$$

The equation obtained by reverting formula (3) to formula (1) is as follows:

$$(\lambda^2[M] + \lambda[C] + [K])\{\varphi\} = [D(\lambda)]\{\varphi\} = \{0\} \tag{4}$$

where [D(λ)] represents the eigenmatrix of the vibration system, indicating that the equation involves the "quadratic eigenvalues" problem. This equation is a 2n-th order algebraic equation in λ with 2n eigenvalues, and their eigenvectors are:

$$\lambda_i = \alpha_i \pm i\omega_i, (i = 1,2,3,\ldots,2n) \tag{5}$$

All of these eigevectors are complex numbers, and their imaginary part ω_i represents the natural frequency of the system, while their real part α_i represents the damping coefficient. α_i occurs in positive and negative conjugate pairs, that is, at the same natural frequency, two real parts will occur and the absolute values of the two real parts are equal. The vector solution of the displacement of the system can be expressed as follows:

$$x = Ae^{\lambda_i t} = e^{\alpha_i t}(A_1 \cos\omega_i t + A_2 \sin\omega_i t) \tag{6}$$

As evident from the formula, when the real part value is positive, the system becomes unstable as the amplitude index increases over time, leading to potential noise issues. Larger values of the real part correspond to a higher growth rate of amplitude. However, a larger absolute value of the real part does not necessarily equate to greater noise, as noise generated by the same amplitude at different vibrational frequencies is not directly comparable. Nonetheless, numerous experiments indicate that modes with large positive real values often produce braking noise, and the frequency most prone to noise generally aligns with the mode having the maximum real value. Thus, the likelihood of noise occurrence can be assessed by the real value of the mode, where the absolute value of the real part at a given natural frequency correlates positively with the probability of noise generation.

Moreover, focusing solely on material friction without considering material damping may result in an overestimation of unstable modes, which are often not associated with actual braking noise. According to reference [10], excessively conservative predictions of unstable modes can be filtered based on the mode damping ratio (the real part of the complex eigenvalue relative to the vibration frequency). Here, unstable modes with a mode damping ratio less than − 0.001 are extracted.

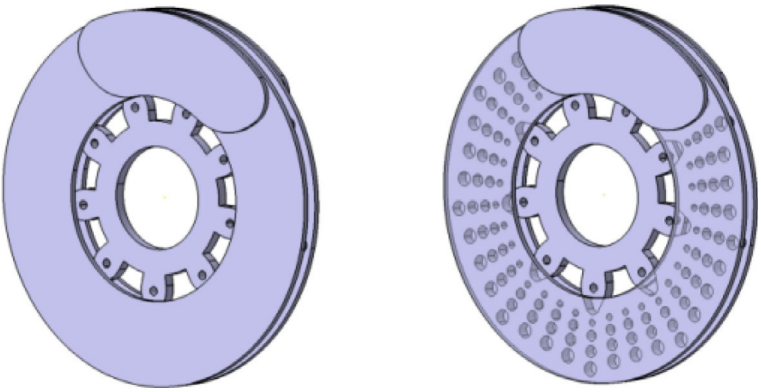


Fig. 1 Geometric model of the brake disc and the brake plate

2.2 Modeling

In this study, the static analysis and mode analysis modules of the ANSYS Workbench platform are employed to investigate brake disc noise, with the CRH series EMU brake disc serving as the focal research object. To streamline calculations and enhance computational speed, the simulation simplifies the setup by excluding the brake clamp and other related mechanisms, focusing solely on constructing the geometric models of the brake disc and the brake plate, as illustrated in Fig. 1.

As depicted in Fig. 2, a total of five distinct shapes of friction blocks are conceived: rotundity, regular hexagon, sector, triangle (chamfered), and ellipse. On each brake plate, 18 friction blocks are symmetrically arrayed as shown in Fig. 3. The finite element simulation model is shown in Fig. 4.

In this study, the 25Cr2MoV alloy is selected for the brake disc, while metallurgical material is chosen for the brake pads. The specific material properties for both components are detailed in Table 1.

After extracting all the obtained unstable modes, the positive real part of the complex eigenvalue is taken as the basis for the possibility of braking noise [11].

3 Results and Discussion

3.1 Different Friction Block Shapes

The brake pieces of five different shapes shown in Figs. 2 and 3 are selected, and the unstable modes of brake noise extracted under the same working condition (friction coefficient $\mu = 0.25$, brake disc speed is 6 rad/s, brake pressure is 1.5 MPa) are drawn in the same picture for comparison, and the following results are obtained.

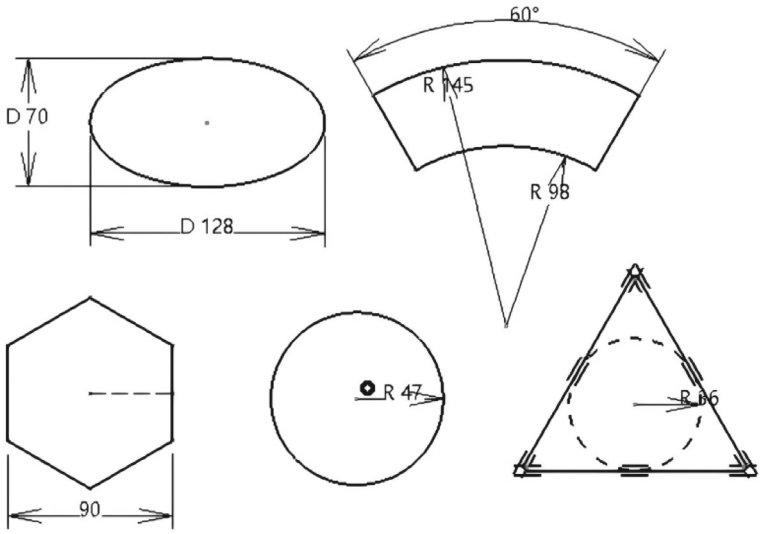


Fig. 2 Geometric dimensions of friction blocks on brake pads

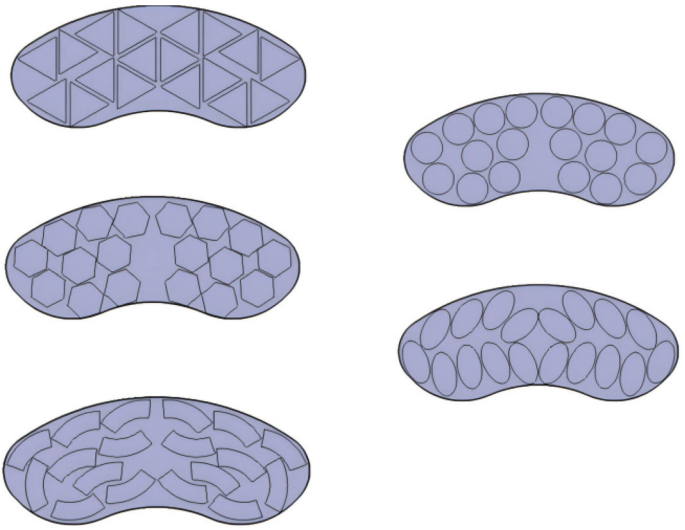


Fig. 3 Arrangement of friction blocks on brake pads

As illustrated in Fig. 5, there is a discernible relationship between brake noise and the shape of the friction blocks on the brake pads. The probability of noise in different brake pads can be inferred by the area enclosed by the points and the coordinate axes in the figure. For instance, the likelihood of brake noise in brake pads featuring hexagonal friction blocks is notably lower compared to those with

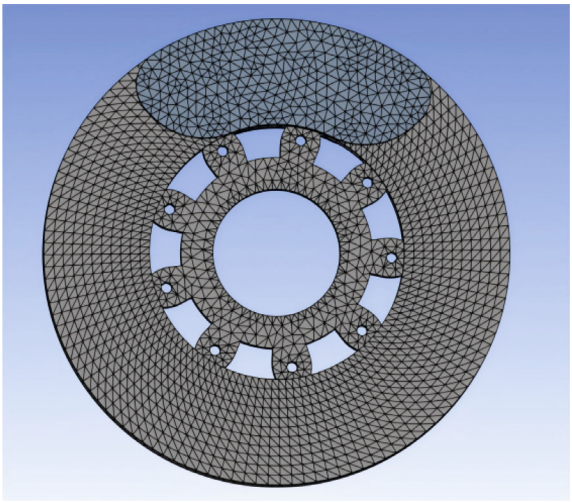


Fig. 4 Finite element model of axle disc braking device

Table 1 Material properties of brake discs and pads

Parameter (unit)	Brake discs (25Cr2MoV)	Pads (metallurgical materials)
Density (kg/m^3)	7620	5250
Young's modulus (GPa)	211	200
Poisson's ratio	0.25	0.28

other shapes. Conversely, brake pads with sector-shaped friction blocks exhibit a significantly higher probability of noise occurrence when compared to pads with other shapes.

3.2 Different Friction Block Arrangement

To elucidate how the arrangement of the friction blocks of the brake pads influences brake noise, the following conditions are described: a friction coefficient of $\mu = 0.25$, hexagonal friction block shape, brake disc speed of 6 rad/s, and brake pressure of 1.5 MPa.

By selecting the unstable modes corresponding to the three complex eigenvalues with the largest real parts and examining their vibration cloud diagrams in Fig. 6, we observe that the 374th mode exhibits characteristics akin to circumferential modes. Meanwhile, the 378th and 566th modes display traits associated with distinct internal and external circles, respectively, and emerge as the most prevalent vibration modes in this simulation. A notable similarity among these three unstable modes is the presence of maximum amplitude at the interface between the brake pads and the brake disc, leading to compromised contact during vibration. This phenomenon constitutes a

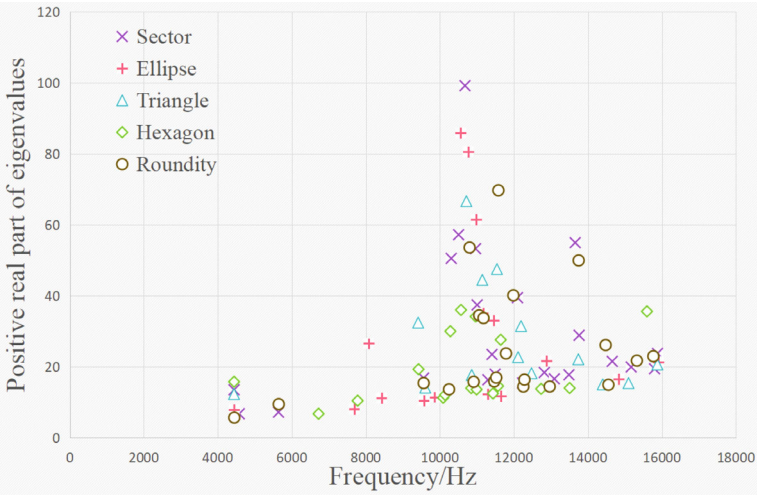


Fig. 5 Distribution of positive real parts of eigenvalues under different friction block shapes

significant contributor to brake screaming. Based on these findings, it can be inferred that strategically arranging the friction blocks at the edges can effectively mitigate the amplitude of the brake pads in those regions to a certain degree.

Consequently, for a brake disc incorporating elliptical friction blocks, the arrangement of these friction blocks will be optimized based on the principle of positioning them towards the periphery of the disc to mitigate edge vibrations. Following the adjustment in Fig. 7, a simulation will be conducted anew to assess the efficacy of this action in potentially reducing the incidence of braking noise.

The repositioned brake pads will be subjected to simulation analysis, during which the unstable mode of brake noise will be extracted under identical operational conditions to those employed in the preceding experiment. The outcomes of this analysis are depicted in the accompanying Fig. 8.

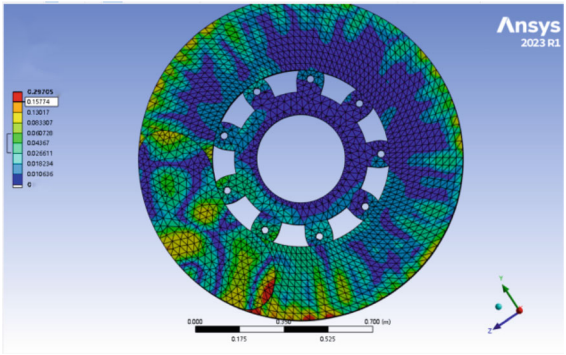
Evidently, the changed elliptical friction block significantly decreases the probability of brake noise, validating the correctness of the inferred friction block layout change. This confirms that altering both the shape and arrangement of friction blocks can effectively reduce the likelihood of friction noise.

The selection of the elliptical friction pad for the experiment is due to its most notable experimental results. In fact, we have also conducted simulation studies on the arrangement of other friction pads, obtaining similar results.

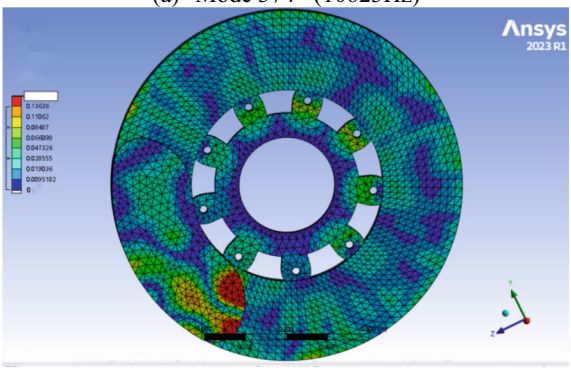
4 Conclusions and Contributions

(1) Braking noise is influenced by friction block shape; hexagonal friction blocks produce the least noise among five shapes.

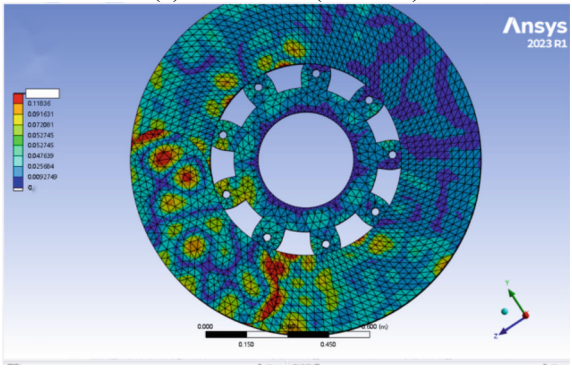
Fig. 6 Three unstable modal vibration cloud maps



(a) Mode 374th (10623Hz)



(b) Mode 378th (10709Hz)



(c) Mode 566th (13472Hz)



Fig. 7 Comparison of the arrangement of friction blocks on brake pads before and after changes

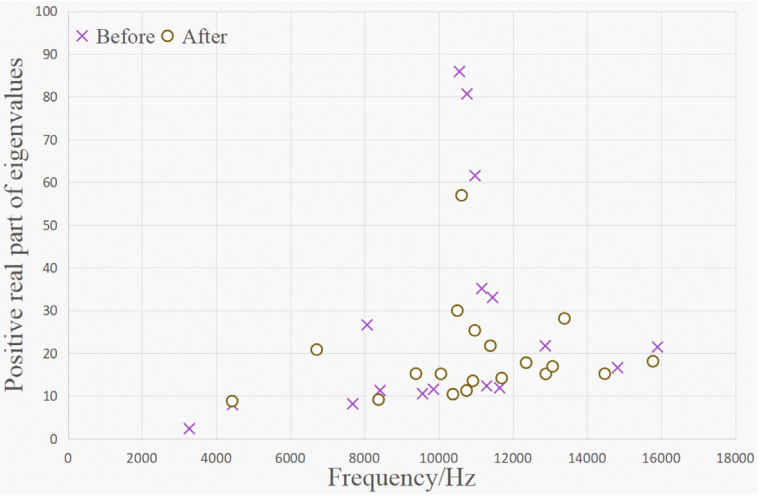


Fig. 8 The distribution of positive real parts of eigenvalues under different layout patterns

(2) Noise is also affected by friction block arrangement. Rearranging elliptical friction blocks to the periphery significantly reduces noise, demonstrating that altering both shape and arrangement can decrease friction noise.

Acknowledgements This work was supported by the Key Projects of China Academy of Railway Sciences Corporation Limited. (Grant No. 2022YJ340), National Key R&D Program of China (Grant No. 2023YFB4301605).

References

1. Yu C (2023) Landscape garden. Brake noise and its handling methods. Time Motors 11:104–106 (in Chinese)
2. Felske A, Hcppe G, Matthai H (1980) A study on drum brake noise by holographic vibration analysis. SAE800221
3. Guan D, Su X (2004) Review, development and review of the brake vibration noise study. Eng Mechan 4:150–155 (in Chinese)
4. Meehan PA, Leslie AC (2021) On the mechanisms, growth, amplitude and mitigation of brake squeal noise. Mech Syst Signal Process 152:107469
5. Chen G, Shi X (2002) Experimental study of the frictional force-relative sliding velocity relation. Lubr Seal 3:44–45 + 48 (in Chinese)
6. Zhang L, Liu L, Meng D (2016) Analysis of automotive brake frictional vibration Sprag-slip theory. Chinese society of automotive engineering. Proceedings of the 2016 China Society of Automotive Engineering Annual Conference. School of Automotive Engineering, Tongji University; Collaborative Innovation Center for Intelligent New Energy Vehicles, Tongji University, vol 7 (in Chinese)
7. Kang J, Nam J (2020) Minimal model for Sprag-slip oscillation as catastrophe-type behavior. Appl Sci 10(8):2748

8. Zhang Q, Yu Z, Liu H, Mo J, Xiang Z, Zhu S, Jin W (2024) Analysis of friction-induced vibration and wear characteristics during high-speed train friction braking process. Tribol Int 196:109701
9. Bouazizi M, Soula M, Lazghab T (2024) Numerical analysis of brake squeal noise—application to the disk brake systems. Noise Vibr Worldwide 55(3):117–122
10. Yan H, Panpan H (2015) Study on wheel brake scream noise based on complex eigenvalue method. Railway Locomot Veh 35(01):54–56 (in Chinese)
11. Schmid D, Gräbner N, von Wagner U (2022) Friction-induced noise in drum brakes: finite-element modeling and experiments with special focus on damping. Arch Appl Mech 92(8):2467–2481

Open Access This chapter is licensed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License (http://creativecommons.org/licenses/by-nc-nd/4.0/), which permits any noncommercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if you modified the licensed material. You do not have permission under this license to share adapted material derived from this chapter or parts of it.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.

