

Bin Xu  
Jianlong Qiu (Eds.)

Communications in Computer and Information Science

2516

# Cognitive Computation and Systems

Third International Conference, ICCCS 2024  
Linyi, China, December 20–22, 2024  
Revised Selected Papers, Part II

Part 2



Springer

# **Communications in Computer and Information Science**

**2516**

## **Series Editors**

Gang Li , *School of Information Technology, Deakin University, Burwood, VIC, Australia*

Joaquim Filipe , *Polytechnic Institute of Setúbal, Setúbal, Portugal*

Zhiwei Xu, *Chinese Academy of Sciences, Beijing, China*



## **Rationale**

The CCIS series is devoted to the publication of proceedings of computer science conferences. Its aim is to efficiently disseminate original research results in informatics in printed and electronic form. While the focus is on publication of peer-reviewed full papers presenting mature work, inclusion of reviewed short papers reporting on work in progress is welcome, too. Besides globally relevant meetings with internationally representative program committees guaranteeing a strict peer-reviewing and paper selection process, conferences run by societies or of high regional or national relevance are also considered for publication.

## **Topics**

The topical scope of CCIS spans the entire spectrum of informatics ranging from foundational topics in the theory of computing to information and communications science and technology and a broad variety of interdisciplinary application fields.

## **Information for Volume Editors and Authors**

Publication in CCIS is free of charge. No royalties are paid, however, we offer registered conference participants temporary free access to the online version of the conference proceedings on SpringerLink (<http://link.springer.com>) by means of an http referrer from the conference website and/or a number of complimentary printed copies, as specified in the official acceptance email of the event.

CCIS proceedings can be published in time for distribution at conferences or as post-proceedings, and delivered in the form of printed books and/or electronically as USBs and/or e-content licenses for accessing proceedings at SpringerLink. Furthermore, CCIS proceedings are included in the CCIS electronic book series hosted in the SpringerLink digital library at <http://link.springer.com/bookseries/7899>. Conferences publishing in CCIS are allowed to use Online Conference Service (OCS) for managing the whole proceedings lifecycle (from submission and reviewing to preparing for publication) free of charge.

## **Publication process**

The language of publication is exclusively English. Authors publishing in CCIS have to sign the Springer CCIS copyright transfer form, however, they are free to use their material published in CCIS for substantially changed, more elaborate subsequent publications elsewhere. For the preparation of the camera-ready papers/files, authors have to strictly adhere to the Springer CCIS Authors' Instructions and are strongly encouraged to use the CCIS LaTeX style files or templates.

## **Abstracting/Indexing**

CCIS is abstracted/indexed in DBLP, Google Scholar, EI-Compendex, Mathematical Reviews, SCImago, Scopus. CCIS volumes are also submitted for the inclusion in ISI Proceedings.

## **How to start**

To start the evaluation of your proposal for inclusion in the CCIS series, please send an e-mail to [ccis@springer.com](mailto:ccis@springer.com)

Bin Xu · Jianlong Qiu  
Editors

# Cognitive Computation and Systems

Third International Conference, ICCCS 2024  
Linyi, China, December 20–22, 2024  
Revised Selected Papers, Part II

*Editors*

Bin Xu  
Northwestern Polytechnical University  
Xi'an, China

Jianlong Qiu  
Linyi University  
Linyi, China

ISSN 1865-0929

ISSN 1865-0937 (electronic)

Communications in Computer and Information Science

ISBN 978-981-96-7437-4

ISBN 978-981-96-7438-1 (eBook)

<https://doi.org/10.1007/978-981-96-7438-1>

© The Editor(s) (if applicable) and The Author(s), under exclusive license  
to Springer Nature Singapore Pte Ltd. 2025

This work is subject to copyright. All rights are solely and exclusively licensed by the Publisher, 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.

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.

# Preface

This volume contains the papers from the Third International Conference on Cognitive Computation and Systems (ICCCS 2024), which was held in Linyi, Shandong, China, on December 20–22, 2024. ICCCS is an international conference on the related fields of cognitive computing and systems which was initiated by the Technical Committee on Cognitive Computing and Systems of the Chinese Association of Automation in 2022. ICCCS 2024 was hosted by the Technical Committee on Cognitive Computing and Systems of the Chinese Association of Automation, and co-organized by Linyi University, Henan Polytechnic University, the Intelligent Robotics Research Center of the Artificial Intelligence Institute of Tsinghua University, and the School of Automation of Northwestern Polytechnical University.

Cognitive computing is an interdisciplinary field of research that focuses on developing intelligent systems that can process and analyze large amounts of complex data in a way that mimics human cognition. It is based on the idea that intelligent systems can be designed to learn, reason, and interact with humans in a more natural way, and that these abilities can be used to solve complex problems that would otherwise be difficult or impossible for humans to solve alone. By enabling agents to understand and process data in a more human-like way, cognitive computing can help organizations make better decisions, improve customer service, develop more effective treatments for diseases, and even predict and prevent natural disasters.

ICCCS aims to bring together experts from different areas of expertise to discuss the state of the art in cognitive computing and intelligent systems, and to present new research results and perspectives on future development. It is an opportunity to promote the research and development of cognitive computing and systems, and provides a face-to-face communication platform for scholars, engineers, teachers, and students engaged in cognitive computing and systems or related fields to promote the application of artificial intelligence technology in industry.

ICCCS 2024 received 155 submissions, all of which were written in English. After a thorough single-blind peer review process, in which each paper received at least 4 reviews, 54 papers were selected for presentation as full papers, resulting in an approximate acceptance rate of 34%. The accepted papers addressed challenging issues in various aspects of cognitive computation and systems, including not only theories and algorithms in cognition, machine learning, computer vision, decision making, etc., but also systems and applications in autonomous vehicles, computer games, intelligent robots, etc.

We would like to thank the members of the Advisory Committee for their guidance, and the members of the Program Committee and additional reviewers for reviewing the papers. We would also like to thank all the speakers and authors, as well as the

participants for their great contributions that made ICCCS 2024 successful. Finally, we thank Springer for their trust and for publishing the proceedings of ICCCS 2024.

December 2024

Bin Xu  
Jianlong Qiu

# Organization

## Conference Committee

### Honorary Chairs

Bo Zhang  
Nanning Zheng  
Deyi Li

Tsinghua University, China  
Xi'an Jiaotong University, China  
CAAI, China

### Advisory Committee Chairs

Qionghai Dai  
Lin Chen  
Fuchun Sun  
Zengguang Hou  
Shusheng Zhang

Tsinghua University, China  
Institute of Biophysics, CAS, China  
Tsinghua University, China  
Institute of Automation, CAS, China  
Linyi University, China

### General Chairs

Bin Xu  
Jianlong Qiu

Northwestern Polytechnical University, China  
Linyi University, China

### Organizing Committee Chairs

Ancai Zhang  
Jiuru Wang  
Huaping Liu  
Zhongyi Chu  
Quanyong Fan

Linyi University, China  
Linyi University, China  
Tsinghua University, China  
Beihang University, China  
Northwestern Polytechnical University, China

**Program Committee Chairs**

|                |                                                     |
|----------------|-----------------------------------------------------|
| Dewen Hu       | National University of Defense Technology,<br>China |
| Cong Wang      | Shandong University, China                          |
| Deqian Fu      | Linyi University, China                             |
| Xiangyong Chen | Linyi University, China                             |
| Bo Zhao        | Beijing Normal University, China                    |

**Program Committee**

|               |                                              |
|---------------|----------------------------------------------|
| Janmin Li     | Tsinghua University, China                   |
| Xiaolin Hu    | Tsinghua University, China                   |
| Wei Qian      | Henan Polytechnic University, China          |
| Xiaohui Liang | Northwestern Polytechnical University, China |
| Weixin Han    | Northwestern Polytechnical University, China |

**Publications Chairs**

|           |                                                                  |
|-----------|------------------------------------------------------------------|
| Quanbo Ge | Nanjing University of Information Science &<br>Technology, China |
| Janmin Li | Tsinghua University, China                                       |
| Rui Zhang | Northwestern Polytechnical University, China                     |

**Finance Chairs**

|           |                         |
|-----------|-------------------------|
| Feng Zhao | Linyi University, China |
| Ming Guo  | Linyi University, China |

## **Registration Chairs**

Xing Wang

Linyi University, China

Kun Zhou

Linyi University, China

## **Local Arrangements Chair**

Guochen Pang

Linyi University, China



## Contents – Part II

### Cognitive Computing And Information Processing

|                                                                                                                                        |    |
|----------------------------------------------------------------------------------------------------------------------------------------|----|
| A Colour Secret Image Sharing Scheme Based on the Theory of Prime Number Distribution .....                                            | 3  |
| <i>Jinqiu Xue and Wenyin Zhang</i>                                                                                                     |    |
| A Novel Design of Radial Basis Function Neural Network Integrating Markov Chain Monte Carlo Clustering Algorithm .....                 | 20 |
| <i>Yunlong Zhu, Zunwei Fu, and Zheng Wang</i>                                                                                          |    |
| A Synthesis of Techniques for Feature Downgrading Processing in IoT Security .....                                                     | 27 |
| <i>Yifang Wang</i>                                                                                                                     |    |
| Anomaly Data Identification and Reconstruction in Wind Power Forecasting Based on MLOF-iForest and RF .....                            | 37 |
| <i>Pengcheng Du and Yu Du</i>                                                                                                          |    |
| Broad Learning System-Enhanced Fuzzy Inference Framework for High-Efficiency Classification .....                                      | 47 |
| <i>Yizhen Wang, Zunwei Fu, and Congcong Zhang</i>                                                                                      |    |
| Design of Radial Basis Function Neural Network with Neural Gas Network Clustering for Classification of Logistics Black Plastics ..... | 55 |
| <i>Hongliang Yu, Kun Zhou, Ming Guo, Xiangyong Chen, and Jianlong Qiu</i>                                                              |    |
| Enhanced Group Method of Data Handling with Roulette Neuron Selection for Time Series Forecasting .....                                | 66 |
| <i>Xixin Wang, Shaoguang Shi, and Congcong Zhang</i>                                                                                   |    |
| FE-ResNet50: Frequency Enhanced Attention Network for sEMG Gesture Recognition .....                                                   | 74 |
| <i>Haozhu Wang, Du Jiang, Juntong Yun, Ying Liu, and Boao Li</i>                                                                       |    |
| Huber Correntropy Kalman Filter .....                                                                                                  | 89 |
| <i>Shuo Wang and Xiaoliang Feng</i>                                                                                                    |    |

**Intelligent Cooperative Control**

A Framework for Visual Target Navigation for Quadcopter Based on Large Language Models in Unknown Environment ..... 99  
*Yunzhuo Liu, Zhaowei Ma, Yuqi Yang, Mengyun Wang, and Yifeng Niu*

Adaptive Output Feedback Control for Nonlinear Time-Delay Systems with Neural Network and Static Gains ..... 111  
*Wenjie Li and Zhengqiang Zhang*

Automatic Ground Collision Avoidance for Aircraft in Uncertain Environments ..... 120  
*Rui Li and Rui Zhang*

Investigations into the Existing State of Distribution Network Security Using Dispersed Generators ..... 134  
 *Jiage Zhang, Chen Lei, and Liping He*

Parkinglot Obstacle Detection System Using Infrastructure-Based Cameras .... 149  
*Yuesheng He, Fei Wang, Zihan Zong, and Ming Yang*

Predefined-Time Consensus for First-Order Heterogeneous Nonlinear Multi-agent Systems ..... 161  
*Shufen Liu, Jianhua Liu, Feng Zhao, and Xiangyong Chen*

Research on Automatic Throttle Speed Stability System Based on Magic Carpet Landing Technology ..... 170  
*Letian Zhao, Hang Chen, Yangyang Zhou, Xiaolei Ma, and Yidi Lei*

The Decentralized Trusted Identity Management in the Collaboration of Supply Chain and Logistics Operations ..... 182  
*Zhangliang Li, Zhengyuan Yue, Zhongli Qiao, Keqing Wang, Ziqi Liu, Jianlong Qiu, and Deqian Fu*

Unify the Research and Design of Power Quality Regulators ..... 198  
*Yaoyao Qin, Chen Lei, Liping He, and Longjing Liu*

**Learning and Systems**

A Zero Trust Continuous Authentication Scheme Based on Keystroke Dynamics ..... 215  
*Ping Gong*

Cognitive-Based Autonomous Orbit of Non-Cooperative Targets ..... 222  
*Jianliang Ma, Lele Zhang, Jian Yang, and Fang Deng*

|                                                                                                                                   |            |
|-----------------------------------------------------------------------------------------------------------------------------------|------------|
| Finite-Time Stabilizing Control of Furuta Pendulum System Based<br>on Disturbance Observer and Integral Sliding Mode Method ..... | 235        |
| <i>Zhujun Wang, Shuli Gong, Ancai Zhang, and Junyao Yu</i>                                                                        |            |
| Method of Heading Estimation Based on Underwater Acoustic Sensors .....                                                           | 246        |
| <i>Qinglin Wen, Chun Jia, and Jiachang Zhang</i>                                                                                  |            |
| Optimal Cooperative Control for Constrained Multi-agent Systems Using<br>Adaptive Dynamic Programming Method .....                | 255        |
| <i>Zijie Guo, Wenshuai Lin, Xiaohong Zheng, and Zilong Zhang</i>                                                                  |            |
| Pareto-Wise Ranking Generator for Multi-objective Coevolutionary<br>Generative Adversarial Networks .....                         | 268        |
| <i>Weifeng Guo, Yijun Fu, Ying Bi, and Jing Liang</i>                                                                             |            |
| Research Progress on Control Technology of Tiltrotor Aircraft .....                                                               | 280        |
| <i>Zhang Junhong, Li Hao, Zhao Letian, Liu Long, and Zhao Hai</i>                                                                 |            |
| Robust and Non-fragile Control for Non-uniform Sampled-Data Systems<br>Under Try-Once-Discard Protocol .....                      | 292        |
| <i>Jiyuan Zhang, Shiwei Cai, and Jinling Liang</i>                                                                                |            |
| Tracking Control of 2-DOF Continuum Manipulator Using Fuzzy PID<br>Method .....                                                   | 304        |
| <i>Youchuan Wang, Junfeng Sun, Ancai Zhang, Wenbo Zheng,<br/>and Haichuan Yang</i>                                                |            |
| <b>Author Index .....</b>                                                                                                         | <b>313</b> |

# Contents – Part I

## Cognitive Computing And Information Processing

|                                                                                                                                              |    |
|----------------------------------------------------------------------------------------------------------------------------------------------|----|
| Iris Recognition Based on Asynchronous Genetic Particle Swarm Optimization .....                                                             | 3  |
| <i>Du Yining and Li Aiju</i>                                                                                                                 |    |
| Joint Torque Sensor-Based Distributed Coordinated Control of Dual-Arm Reconfigurable Manipulators .....                                      | 15 |
| <i>Bing Ma, Hangwei Zhang, Ximing Yao, Yuan Yao, Tianjiao An, and Qiang Pan</i>                                                              |    |
| MAMF-ResNet: Multi-Attention Mechanisms Fusion Network for Brain Tumor Classification .....                                                  | 27 |
| <i>Qingyun Huo, Meng Yang, Yuxuan Li, Yunyu Wang, Xin Zhang, and Mingtao Liu</i>                                                             |    |
| Noise Sensitive Relation Aware Cross-Lingual Entity Alignment .....                                                                          | 39 |
| <i>Yunlong Tian, Shuanglong Yao, Liqiu Shan, and Xing Wang</i>                                                                               |    |
| Origin Classification of Rice Wine Based on an Electronic Nose and Convolutional Neural Network .....                                        | 54 |
| <i>Wenqi Sun, Ancai Zhang, Guangyuan Pan, and Wenbo Zheng</i>                                                                                |    |
| Parkinglot Obstacle Detection System Using Infrastrue-Based Cameras .....                                                                    | 63 |
| <i>Yuesheng He, Fei Wang, Zihan Zong, and Ming Yang</i>                                                                                      |    |
| Self-tuning Control of Manipulator Based on Fuzzy Linear Active Disturbance Rejection Control and Particle Swarm Optimization Algorithm .... | 75 |
| <i>Bo Tao, Zhuxiang Chen, Du Jiang, and Juntong Yun</i>                                                                                      |    |
| The Neural Network of College Students Research on the Intelligent Scoring System .....                                                      | 88 |
| <i>Yang Ji and Huang Shen</i>                                                                                                                |    |
| Uniform Light Field Control System Based on IoT Technology - Solving Eye Discomfort Caused by Changes in Light Intensity .....               | 96 |
| <i>Haobin Yang, Haiying Zhang, and Xunyu Qiao</i>                                                                                            |    |

**Intelligent Cooperative Control**

Attention-Fused Vision Mamba Model for Remote Sensing Image  
Detection and Segmentation ..... 117  
*Zhao Yusen, Wu Shan, and Tian Liang*

Control Allocation Based Fault-Tolerant Attitude Control of Hypersonic  
Reentry Vehicle ..... 131  
*Yuan Zhang and Weixin Han*

Cooperative Multi-target Enclosing Control of Multi-robots Based  
on Virtual Target Points ..... 144  
*Fang Fang Zhang, Zhen Yu Zhu, and Ji Xian Gao*

Event-Triggered Secure Consensus Control of Nonlinear Multiagent  
Systems with hybrid DoS attacks ..... 152  
*Yingxin Zhang, Huixia Cui, and Senping Jia*

FastLSLO: An Efficient LiDAR Odometry Based on Improved Lie Group  
B-Splines ..... 165  
*Xinyang Tang, Wei Yuan, Chenxi Yang, Chunxiang Wang, Bing Wang,  
and Ming Yang*

Improved BP Neural Network Based Deck Motion Prediction and Landing  
Control for Carrier-Based Aircrafts ..... 178  
*Zhaoxing Li and Xingzhao Zhang*

Joint Sequencing and Merging Optimization for Airplanes and Helicopters .... 186  
*Dawei Wang, Yi Lyu, Ken Chen, Yiman Zhang, and Chengcheng Wu*

Research on UAV SINS/GNSS Integrated Navigation Error Model  
for Transpolar Flight ..... 198  
*Guoqiang Zhang, Qi Zhou, and Jinjiang Wang*

Technical Research on Helicopter Blind Landing Under Degraded Visual  
Environments (DVE) ..... 214  
*Yanwei Du, Mingdong Qi, and Heng Zhang*

**Learning and Systems**

A Contactless Demonstration Learning Method for Robotic Systems  
Based on Dynamic Parametric Regression Modeling ..... 231  
*Meng Li, Jinzhu Peng, Jixian Gao, Nan Zhao, Yaonan Wang,  
and Mingkuo Wu*

|                                                                                                                                              |            |
|----------------------------------------------------------------------------------------------------------------------------------------------|------------|
| <b>A Multi-scale Fusion and Dynamic Upsampling Model for Road Surface<br/>Snow Detection</b> .....                                           | <b>243</b> |
| <i>Lipeng Du, Xinhao Zhou, Qili Chen, Tong Wang, Lin Zhao,<br/>and Guangyuan Pan</i>                                                         |            |
| <b>A Scoring System for Single and Parallel Bars Actions Based<br/>on Multi-view 3D Pose Estimation</b> .....                                | <b>254</b> |
| <i>Yuntong Kang, Mingwei Cao, Haoran Yao, Sen Qiu, and Zhelong Wang</i>                                                                      |            |
| <b>Adaptive PID Controller for Industrial Process Based on Reinforcement<br/>Learning</b> .....                                              | <b>265</b> |
| <i>Yicong Yang and Qili Chen</i>                                                                                                             |            |
| <b>Computational Analysis of Synaptic Plasticity in Echo State Network</b> .....                                                             | <b>272</b> |
| <i>Xinyu Shen, Shaoqi Cheng, Fanjun Li, and Jiayue Feng</i>                                                                                  |            |
| <b>LEHR: LLM-Driven Evolutionary Hybrid Rewards for Multi-agent<br/>Reinforcement Learning</b> .....                                         | <b>283</b> |
| <i>Yuan Wei, Xiaohan Shan, and Jianmin Li</i>                                                                                                |            |
| <b>Magnetic Core Loss Prediction: A Data-Driven NGO-GRU Model</b> .....                                                                      | <b>296</b> |
| <i>Ningning Hu, Yongqiang Mao, Lanmei Cong, Zhaohui Zhang,<br/>and Ziyue Han</i>                                                             |            |
| <b>Research on Method of Control Surface Jamming Fault Injection in Fly<br/>Test of Fly-by-Wire Aircraft Based on Multiple Control</b> ..... | <b>308</b> |
| <i>Lei Ming, Xie Qingping, and Li Yajing</i>                                                                                                 |            |
| <b>Tactile-Based Manipulation for Wire Following</b> .....                                                                                   | <b>319</b> |
| <i>Jiazhen Cai, Jing Cui, and Zhongyi Chu</i>                                                                                                |            |
| <b>Author Index</b> .....                                                                                                                    | <b>333</b> |

# **Cognitive Computing And Information Processing**



# A Colour Secret Image Sharing Scheme Based on the Theory of Prime Number Distribution

Jinqiu Xue and Wenyin Zhang<sup>(✉)</sup>

School of Information Science and Engineering, Linyi University, Shuangling Street,  
Linyi 276000, Shandong, China  
zhangwenyin@lyu.edu.cn

**Abstract.** Secret image sharing is a technique for protecting secret information that is fault-tolerant by splitting a secret image into multiple shadow images. In this paper, we address the lack of research on current colour secret image sharing schemes and propose a new scheme in combination with the theory of prime number distribution. The scheme deals directly with the three components of each pixel value of the secret image. The three components of the pixel value are first mapped to the sequence number of the approximate pixel set in which they are located, then a specific function is selected to perform the mapping operation on the RGB intensity values using mathematical transformations, and finally a  $k - 1$  polynomial is used to generate the shared shares. When reconstructing, a value from the approximate pixel set where the pixel value is located will be selected randomly as the restored value. Unlike most image visual share schemes, this algorithm does not require the use of random permutations to convert a secret image into a random image, thus making the scheme simpler and more reliable. In addition, the shared images generated by this scheme are smaller in size than the secret images. Finally, the secret image can be recovered through the set of eligible participants.

**Keywords:** secret image sharing · theory of prime number distribution · color image · smaller shared images

## 1 Introduction

With the development of computer Internet technology, a large amount of information is stored and disseminated in different forms in the network. As a commonly used information carrier, images have been widely used in fields such as communication, medicine, military and commerce. The issue of security and integrity of images in the process of transmission, storage and restoration has received higher and higher attention. In modern cryptography, secret image sharing technology is a hotspot that has received much attention and is an extension of secret sharing in image processing and application, through which the integrity of the image as well as the security of the image can be maintained when sharing secret images.

The concept of secret sharing was first proposed independently by Blakley [1] and Shamir [2] respectively, the main idea of the scheme is to split the secret into  $n$  parts and



distribute them to different participants for management, only  $k$  or more participants cooperating to satisfy the threshold can reconstruct the secret, while less than  $k$  participants cooperating cannot derive any useful information about the secret. There are various ways to implement the construction of  $(k, n)$ -threshold secret sharing schemes [3, 4], which have been extensively studied in recent years as an important way to maintain key security.

Currently, most of the researches mainly focus on grayscale images, while most of the images used in daily life are in color. In order to realize the practical value of secret image sharing, we need to effectively apply the secret sharing scheme to color images. The traditional color image secret sharing scheme is to share and recover the R, G and B color planes separately [5]. Wu et al. [6] proposed two Color-black-and-white visual cryptography schemes (CBW-VCS). CBWVCS is determined by adding one or two column vectors to the base matrix of traditional VCS and constructing a generalized  $(k, n)$  threshold. This scheme solves the pixel expansion problem, but the recovered image is a superimposed image of binary secret image and color shadow image, which is not universal for various application scenarios and requires multiple reconstructions, and it cannot achieve direct recovery of lossless secret image. Wang et al. [7] constructed an extended Visual Cryptography system which can embed two secret images on two color cover images. To embed two secret images in a halftoning process, constrained error diffusion was designed. By flipping the overlapping layers, the two secret images are recovered and the two secret images do not interfere with each other. However, the recovered secret image is still the superimposed image of color secret image and color shadow image, and the PSNR value of the reconstructed secret image is low. Shankar et al. [8] developed a Secure  $(k, k)$  Multiple Secret Sharing (SKMSS) scheme using hybrid optimal SIMON cipher. The proposed SKMSS method constructs a set of securely generated noisy components based on performing hashing and grouping ciphers on the secret image itself. After encrypting the share by a hybrid optimal SIMON cipher based on the noisy image, the share is created and sent securely. Pan et al. [9] proposed two visual cipher schemes capable of fully recovering secret color images. The two schemes belong to the  $(n, n)$  threshold scheme. The first scheme generates meaningless shares. The second scheme generates  $n - 1$  meaningful shares and one meaningless share. These  $n-1$  meaningful shares are color QR codes that can be decoded with a general decoder instead of a standard decoder. Both schemes are pixel unscalable. Although the shares in the second scheme are color QR codes, they are not of high quality. The size of the color QR code affects the decoding speed. Li et al. [10] proposed a reversible and lossless secret image sharing scheme with authentication features specifically designed for color images. Noise-like shadows are generated by combining the original secret image and the authentication image, which are then embedded into the color image using enhanced steganography. During the recovery process, participants can non-destructively extract shares and complete the authentication phase without the need for a trusted third party. The authentication results can be easily verified by visual inspection. Ultimately, the qualified shading is used to reconstruct the secret image, ensuring the secure transmission of confidential image data. Xiong et al. [11] proposed a new reversible secret image sharing and authentication scheme for embedding palette images into true color images. In this scheme, the pixels of the palette secret image and the bits of the cover pixels are

used as coefficients of a polynomial. The sharing is embedded into the true color cover image to generate an intermediate steganographic image. Authentication information is then derived from the intermediate steganographic image and hidden in the cover image. A final steganographic image similar to the cover image is obtained and sent to the authorized participants. At the receiver side, once the  $k$  steganographic images are authenticated, the secret image and cover image can be recovered losslessly for the  $(k, n)$  threshold scheme.

In this paper, we propose a color secret image sharing scheme based on the theory of prime number distribution. The scheme directly processes the three RGB components of each pixel in a color image by mapping the RGB component value of each pixel to the sequence number of the set of approximate pixels in which it resides. These sequence numbers are then processed using mathematical transformations that jointly map the RGB components of each secret pixel to a new value. Finally, the shared shares are generated using a specific  $k-1$  polynomial, thus ensuring that only participants who fulfill the eligibility conditions can recover the original image. It is worth noting that the method operates directly on each pixel component of the color image, avoiding the loss of information caused by first converting the color image to grayscale in traditional methods. From the experimental results, it can be seen that the algorithm exhibits a high degree of randomness in generating shared images, which significantly improves the security of image sharing. Meanwhile, the shared image generated by this algorithm is smaller than the original secret image and takes less time to generate the shared image, which makes storage and transmission more efficient.

The rest of the paper is organized as follows: in Sect. 2, the prime number distribution theorem is introduced. Section 3, 4 describes the background of the proposed scheme in this paper as well as the specific algorithmic implementation. The security analysis of the scheme is presented in Sect. 5. Experimental results and analysis are given in Sect. 6. Finally, Sect. 7 summarizes the paper.

## 2 Prime Number Distribution Theorem

A prime number is a natural number that is divisible only by 1 and itself. The distribution of prime numbers is regular in the aggregate, but locally, the occurrence of prime numbers appears to be random. The prime number theorem is the central theorem of the theory of prime number distribution, a proposition about the number of prime numbers, reflecting the asymptotic distribution of prime numbers in the positive integers is centered on Riemann's formula, with Gauss's formula as the upper limit of the normal distribution, which can also be called the theorem of the normal distribution of prime numbers.

In the prime number theorem, for a positive real number  $x$ , define  $\pi(x)$  to represent the prime counting function, the value of  $\pi(x)$  is the number of primes that is not greater than  $x$ . As  $x \rightarrow \infty$ , the asymptotic expression holds:

$$\pi(x) \simeq \frac{x}{\ln x} \simeq \int_2^x \frac{dt}{\ln t} \quad (1)$$

where  $\pi(x)$  is also an increasing step function of  $x$ , reflecting the infinity of the number of prime numbers. This also leads to an equivalent formulation of Euclid's theorem:

$$\lim_{x \rightarrow \infty} \pi(x) = \infty \quad (2)$$

In this paper, the variation of image pixel values is studied mainly based on the theory of prime number distribution. The number of prime numbers contained in each natural number from 0 to 255 is divided into 54 pixel sets according to an increasing law, ensuring that there is only one prime number in each pixel set. In particular, 0,1 and 2 are combined into one pixel set as shown in Table 1.

**Table 1.** Approximate pixel set

| pixel values    | serial numbers | pixel values       | serial numbers |
|-----------------|----------------|--------------------|----------------|
| 0, 1, 2         | 0              | 107, 108           | 27             |
| 3, 4            | 1              | 109, 110, 111, 112 | 28             |
| 5, 6            | 2              | 113–126            | 29             |
| 7, 8, 9, 10     | 3              | 127, 128, 129, 130 | 30             |
| 11, 12          | 4              | 131–136            | 31             |
| 13, 14, 15, 16  | 5              | 137, 138           | 32             |
| 17, 18          | 6              | 139–148            | 33             |
| 19, 20, 21, 22  | 7              | 149, 150           | 34             |
| 23–28           | 8              | 151–156            | 35             |
| 29, 30          | 9              | 157–162            | 36             |
| 31–36           | 10             | 163, 164, 165, 166 | 37             |
| 37, 38, 39, 40  | 11             | 167–172            | 38             |
| 41, 42          | 12             | 173–178            | 39             |
| 43, 44, 45, 46  | 13             | 179, 180           | 40             |
| 47–52           | 14             | 181–190            | 41             |
| 53–58           | 15             | 191, 192           | 42             |
| 59, 60          | 16             | 193, 194, 195, 196 | 43             |
| 61–66           | 17             | 197, 198           | 44             |
| 67, 68, 69, 70  | 18             | 199–210            | 45             |
| 71, 72          | 19             | 211–222            | 46             |
| 73–78           | 20             | 223, 224, 225, 226 | 47             |
| 79, 80, 81, 82  | 21             | 227, 228           | 48             |
| 83–88           | 22             | 229, 230,231, 232  | 49             |
| 89–96           | 23             | 233–238            | 50             |
| 97, 98, 99, 100 | 24             | 239, 240           | 51             |
| 101, 102        | 25             | 241–250            | 52             |
| 103–106         | 26             | 251–255            | 53             |

### 3 Background

The polynomial interpolation based threshold secret sharing scheme was first proposed by Shamir, the basic idea of the scheme is based on the fact that for  $k$  points in the plane, there exists a unique  $k - 1$  polynomial. Using this fact, Shamir divides the secret  $S$  into  $n$  shares  $S_1, S_2, \dots, S_n$  and distributes them to  $n$  participants  $P_1, P_2, \dots, P_n$  such that it takes any  $k$  ( $1 < k \leq n$ ) or more participants to reconstruct the original secret data  $S$ . If the number of collected shares is less than  $k$ , no obtain any additional information about the secret data  $S$ . Shamir defines a  $k - 1$  polynomial as follows:

$$h(x) = c_0 + c_1x + c_2x^2 + \dots + c_{k-1}x^{k-1} \pmod{p} \quad (3)$$

where the coefficient  $c_0$  is secret ( $c_0 = S$ ),  $p$  is a prime number greater than  $n$ , and the other coefficients  $c_1, c_2, \dots, c_{k-1}$  are randomly chosen elements from the number field  $F_p$ . The value of the function  $h(x)$  is modulo  $p$  computed. Finally, for each participant  $P_i$ , the points  $(x_i, h(x_i))$  on the polynomial are given as shares, where  $x_i \in F_p$  and for all  $i \neq j$ , there is  $x_i \neq x_j$ . During the reconstruction phase, when any  $k$  or more participants are gathered together, they will be able to reconstruct the above polynomials using Lagrange interpolation formulas, as follows:

$$\phi(x) = \sum_{i=1}^k h(x_i) \prod_{\substack{j=1 \\ j \neq i}}^k \frac{x - x_j}{x_i - x_j} \pmod{p} \quad (4)$$

Then, the value of  $\phi(x)$  at  $x = 0$  is computed to obtain the secret  $c_0 = S$  as follows:

$$S = c_0 = \phi(0) = \sum_{i=1}^k h(x_i) \prod_{\substack{j=1 \\ j \neq i}}^k \frac{-x_j}{x_i - x_j} \pmod{p} \quad (5)$$

Thien and Lin [12] proposed a new idea based on the secret sharing scheme. In Thien and Lin's scheme, the secret image is first processed by substitution and then a secret sharing stage is performed to obtain the shadow image. Shamir in his scheme uses the first coefficient of a  $k - 1$  degree polynomial as the secret and the remaining  $k - 1$  coefficients are chosen randomly from the interval  $[0, p - 1]$ . Thien and Lin do not use random elements as the remaining of the polynomial's  $k - 1$  coefficients, but instead use the next  $k - 1$  neighbouring pixel values as coefficients. In this way, the shadow size is reduced to  $1/k$  times the size of the secret image.

In this section, based on the above scheme, a new visual split storage scheme for colour images is proposed, which has the following advantages:

- (i) It can satisfy the case of a large number of participants, which solves the situation that the number of participants is larger than 256 or the selected prime number in most schemes and cannot be shared.
- (ii) It does not require preprocessing such as chaos mapping or random permutation of the secret image, which saves the computational overhead of the scheme.
- (iii) The ability to generate shadow images with smaller size saves space for storage and transmission of shadow images.

## 4 The Proposed Scheme

### 4.1 Double Mapping Encryption Algorithm

The most commonly used 24-bit RGB colour images have three 8-bit colour channels, red (r), green (g) and blue (b). The pixel intensity of each colour component varies from 0 to 255. Since an RGB colour image is represented as an RGB triad (r, g, b), the combination of these three colour components has a total of  $256 \times 256 \times 256$ , 16777216 ( $= 2^{24}$ ) different colour values.

Firstly, according to the approximate pixel set shown in Table 1 in Sect. 2, the three components of each pixel value of the colour secret image are mapped to the sequence number of the pixel value where they are located, e.g. mapping (26, 62, 199) results in (8, 17, 45). So the mapped colour image has a total of 157464 ( $= 54 \times 54 \times 54$ ) different colour combinations, which saves overhead for subsequent calculations.

Then, define a new mapping  $f$ ,  $f(r, g, b) = 54^2 * r + 54 * g + b$ . Clearly  $f$  is single map. Moreover for any  $x$  ( $0 \leq x \leq 157464$ ), there exists a unique combination (r, g, b) such that  $f(r, g, b) = x$ .

Based on the above double mapping algorithm, combined with the polynomial difference based sharing scheme of Thien and Lin, the following  $k - 1$  polynomials are defined:

$$h(x) = c_0 + c_1x + c_2x^2 + \dots + c_{k-2}x^{k-2} + c_{k-1}x^{k-1} \pmod{157477} \quad (6)$$

where  $k - 1$  pixels of the secret image are used to assign  $k - 1$  coefficients of the polynomial, the coefficients  $c_1, c_2, \dots, c_{k-1}$  are the values of the colour image pixel values after bimapping, and the coefficients  $c_0$  ( $0 \leq c_0 < 157477$ ) are chosen randomly to introduce randomness in the shaded image. The prime number  $p$  is chosen to be the smallest prime number larger than 157464 (Table 2).

In Algorithm 1, it is important to note that:

1.  $k - 1$  pixel values are selected as one input at a time in step 1, corresponding to one pixel value for each shadow image, so the shadow image is  $1/(k - 1)$  of the secret image. The size of the shadow image is  $M \times N/(k - 1)$  since for the shadow images are arranged starting from columns.
2. To enhance security, the first coefficient  $c_0$  is randomly assigned in step 6.
3. steps 9–10 ensure that all shared values are not greater than 157464, thus avoiding the problem of overflow of pixel value components in steps 15–17.
4. since each component of the pixel value of the generated shaded image is less than 54 and the image is darker, step 14 is added for randomly generating a parameter to add a value to each component of the pixel value of the coloured image to make the generated share visible.

### 4.2 Reconstruction Process

In order to successfully reconstruct the secret image, at least  $k$  shadow images are required. Therefore, without loss of generality, assume  $k$  shadow images  $SC_1, SC_2, \dots, SC_k$ . From these shared images, a series of polynomials are reconstructed using Lagrange interpolation, and the original secret image is obtained by recovering the polynomial coefficients in turn. The steps used in the reconstruction stage are as follows.

**Table 2.** Colour image sharing algorithm

---

Algorithm 1: Colour image sharing scheme based on prime distribution theory

---

Input: a colour image  $S$  of size  $M \times N$ , parameters  $k$  and  $n$ , and prime number  $P = 157477$ .

Output:  $n$  shaded images  $SC_1, SC_2, \dots, SC_n$ .

1. Select  $k-1$  pixel values  $v_1, v_2, \dots, v_{k-1}$  sequentially from the secret image.
  2. Each component of these  $k-1$  pixel values is mapped to the sequence number of the approximate pixel set in which it is located, and the updated image values as  $v_1', v_2', \dots, v_{k-1}'$ .
  3. **for** ( $i = 1, i < k, i++$ ) **do**
  4.      $c_i = v_1'(r) * 54^2 + v_1'(g) * 54 + v_1'(b)$
  5. **end**
  6. Randomly generate a number as a parameter  $c_0$  ( $0 \leq c_0 < 157477$ )
  7. **for** ( $t=1, t < n+1, t++$ ) **do**
  8.      $f(t) = c_0 + c_1x + c_2x^2 + \dots + c_{k-2}x^{k-2} + c_{k-1}x^{k-1} \pmod{157477}$
  9.     **if** ( $f(t) > 157464$ ) **then**
  10.         **goto** step 6
  11.     **end**
  12. **end**
  13. **for** ( $t=1, t < n+1, t++$ ) **do**
  14.     Randomly generated parameters  $o$  ( $0 \leq o \leq 202$ )
  15.      $SC_t(i,j)_R = f(t) // (54 * 54) + o$
  16.      $SC_t(i,j)_G = f(t) // 54 \% 54 + o$
  17.      $SC_t(i,j)_B = f(t) \% 54 + o$
  18. **end**
- 

1. a pixel value is selected from the  $k$  shaded images in turn, the parameter  $o$  is subtracted for each component of the pixel value, and then  $k$  values are obtained according to the function  $f(r, g, b) = 54^2 * r + 54 * g + b$  to construct a  $k - 1$  st order polynomial, and the polynomial coefficients  $c_1, c_2, \dots, c_{k-1}$  are computed.
2. according to the inverse process of steps 3–4 in Algorithm 1, the sequence number of the approximate pixel set where each component of the  $k - 1$  pixel values of the secret image is located can be obtained.
3. according to the sequence number of each component, any one of the values in the set of approximate pixels in which it is located is selected as the recovered pixel value.
4. after traversing all the shadow images, the original secret image can be recovered. From the recovery algorithm, it can be concluded that the algorithm is symmetrical between the process of dividing the memory and the process of reconstruction, and the algorithm is lossy due to the fact that any one of the values in the set of approximate pixels is selected as the recovered pixel value, but the quality of the recovered image is within the acceptable range.

## 5 Security Analysis

In order to verify that no information about the secret is revealed when there are less than  $k$  participants, this section demonstrates what happens when  $k - 1$  participants will participate in the reconstruction. Suppose that the first  $k - 1$  participants have the shared shares  $SC_1 = [1, SC_1(i, j)]$ ,  $SC_2 = [2, SC_2(i, j)]$ , ...,  $SC_{k-1} = [1, SC_{k-1}(i, j)]$ .

For  $k - 1$  participants, the secret polynomial is unknown. However, the number of times the polynomial is known is  $k - 1$ . Therefore, for each fragment, the unknown polynomial is assumed to be of the form  $d_0 + d_1x + d_2x^2 + \dots + d_{k-1}x^{k-1} \pmod{p}$ . The goal of the participants was to find the unique  $d_0, d_1, d_2, \dots, d_{k-2}$  to reconstruct the secret pixel of the fragment.

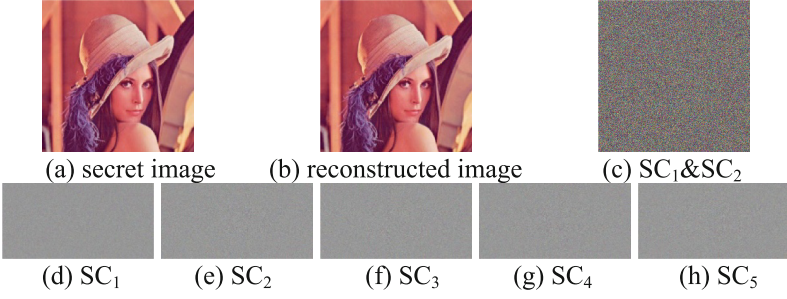
Participants can then construct a  $k - 1$  equation consisting of  $k$  unknowns  $d_i, i = 0, 1, \dots, k - 1$  according to the following formula:

$$\begin{cases} d_0 + d_1 \times 1 + d_2 \times 1^2 + \dots + d_{k-1} \times 1^{k-1} \equiv_p SC_1(i, j) \\ d_0 + d_1 \times 2 + d_2 \times 2^2 + \dots + d_{k-1} \times 2^{k-1} \equiv_p SC_2(i, j) \\ \dots \\ d_0 + d_1 \times (k - 1) + d_2 \times (k - 1)^2 + \dots + d_{k-1} \times (k - 1)^{k-1} \equiv_p SC_{k-1}(i, j) \end{cases} \quad (7)$$

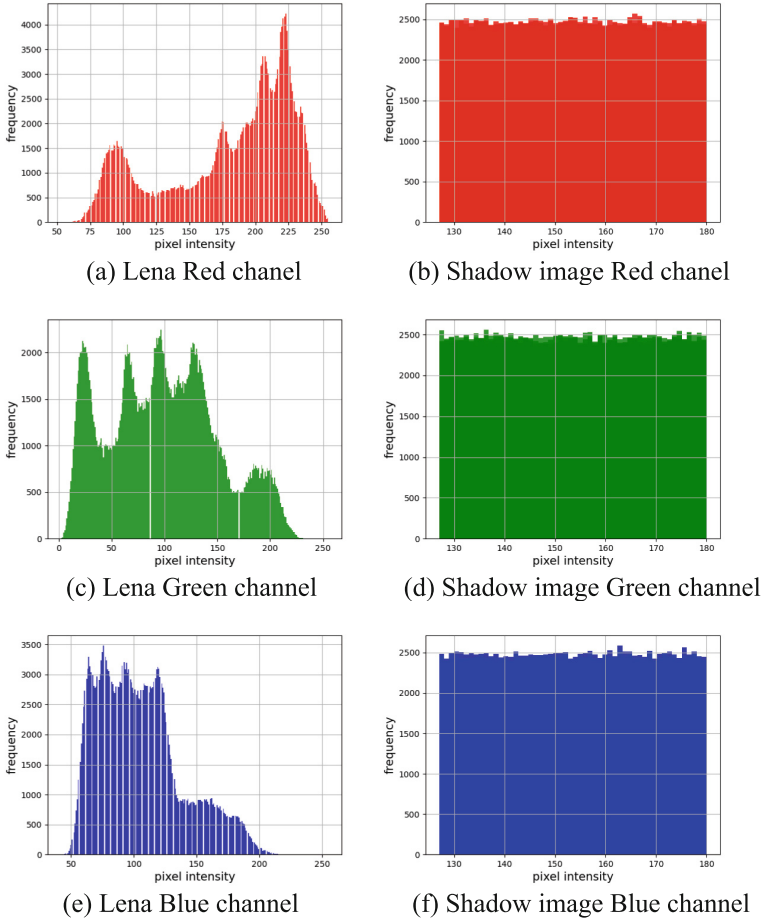
where there are  $k$  unknowns and the number of equations is  $k - 1$ . So it is necessary to guess the share of the  $k$ th participant correctly to obtain a unique solution to the system of equations. Since  $0 \leq h(x) \leq 157464$ , the probability of getting the polynomial result correctly is  $1/157464$ . Hence the probability of correctly guessing the entire secret image of size  $M \times N$  is  $\left(\frac{1}{157464}\right)^{\frac{M \times N}{k-1}}$ . So if any  $k - 1$  shared shares are obtained, it is almost impossible to uniquely determine the correct secret message. On the other hand, any  $k$  or more participants can uniquely recover the secret data. Therefore, it can be concluded that the secret image sharing scheme proposed in this chapter is computationally secure.

## 6 Experimental Results and Analysis

In this section, a colour image Lena of size  $512 \times 512$  is used as the secret image as shown in Fig. 1(a). The threshold parameter  $k = 3$ ,  $n = 5$  is chosen and the value of random parameter  $o$  is 127. Thus, the size of each shaded image is  $1/2$  the size of the secret image as shown in Fig. 1(d)-(h). By collecting any three of the five shared shares, the original image can be recovered, as shown in Fig. 1(b), although the scheme is lossy, it cannot be discerned through the human eye, and the peak signal-to-noise ratio of the recovered image to the original secret image reaches 38.16 dB, with a structural similarity of 0.96. Meanwhile, Fig. 1(c) is the result of the combination of the shaded image 1 ( $SC_1$ ) and the shaded Image 2 ( $SC_2$ ) after reconstruction, demonstrating that a number of shared shares smaller than the threshold will not recover the secret image. The simulation results in this section demonstrate the effectiveness of the proposed image visual sharing scheme.



**Fig. 1.** Experimental results of the (3, 5) threshold image sharing scheme



**Fig. 2.** Histogram analysis results



## 6.1 Histogram Analysis

The histogram of a colour image shows the distribution of its pixel intensities, and the difference in bar heights shows the variation in the frequency of each pixel in the image. For a truly random image, these values are expected to be uniformly distributed in the interval  $[0, 255]$ . In order to ensure the security of the proposed scheme, we expect the shaded image to have similar characteristics as the random image. Figure 2(a), (c), (e) shows the histograms of the red channel, green channel and blue channel of the secret image Lena. Figure 2(b), (d), (f) on the other hand, combines each component of the five shaded image pixels into one histogram, and since the parameter  $\alpha$  is 127, the pixel intensity interval changes from the original  $[0, 53]$  to  $[127, 180]$ . As can be seen in Fig. 2, the histogram corresponding to the shaded image is almost flat, and the height difference between the highest and lowest bars is small. The heights of these bars are almost the same, reflecting the uniform distribution of pixel values. Based on these results, it can be concluded that the shaded image shows no meaningful information that could reveal the secret image.

## 6.2 Pixel Correlation Analysis

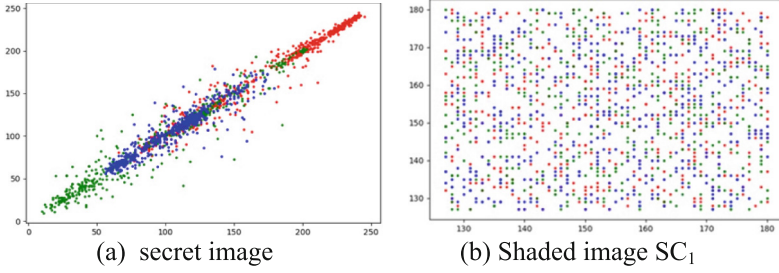
In addition to histogram analysis, correlation analysis is another metric used to test the randomness of an image. It reflects the degree of correlation between two variables. Highly correlated pixels have an optimal value of 1 for the correlation coefficient, on the other hand, uncorrelated pixels have a correlation coefficient very close to zero.

In order to calculate the correlation coefficient of a colour image, firstly, two sequences  $\{x_j\}_{j=1}^N$  and  $\{y_j\}_{j=1}^N$  of  $N$  pairs of neighbouring pixels are randomly selected, and then, the correlation coefficient of the two sequences is derived by using the following equation:

$$\rho_{x,y} = \frac{\frac{1}{N} \sum_{i=1}^N (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\left(\frac{1}{N} \sum_{i=1}^N (x_i - \bar{x})^2\right) \left(\frac{1}{N} \sum_{i=1}^N (y_i - \bar{y})^2\right)}} \quad (8)$$

where  $\bar{x}$  and  $\bar{y}$  denote the mean values of  $x$  and  $y$ , respectively. In meaningful raw secret images, the correlation between neighbouring pixels is relatively high. In order to prevent being statistically analysed, the correlation between adjacent pixels must be reduced. For this reason, there is a need to propose a robust encryption algorithm to generate a random image in which the correlation between neighbouring pixels should be small or no correlation for improving security.

In the experiment, a row of pixel values is randomly selected from the original secret image in Fig. 1(a) and the shadow image in Fig. 1(d), and the correlation coefficients in the three directions of Horizontal, Vertical and Diagonal are observed, as shown in Table 3. From the data in the table, it can be seen that the secret image is characterised by strong correlation between neighbouring pixels, while the correlation coefficient of neighbouring pixels in the shadow image is closer to zero, which indicates that the proposed algorithm perfectly eliminates the correlation between neighbouring pixels in the shadow image and satisfies the prerequisites of the secret sharing scheme.



**Fig. 3.** Scatterplots of secret image and shadow image

Figure 3 shows a scatter plot of pixel pairs for the secret image and a shaded image. x-axis values represent the values of the pixels in a particular row, and the y-axis represents the values of the neighbouring pixels to the right of the pixel corresponding to the x-axis. The scatterplot visualises the relationship between the values of neighbouring pixels in an image. If two neighbouring pixel values are close, the point will be close to the diagonal and if the difference is large, the point will be far from the diagonal. It can be observed that in the meaningful secret image, almost all pixels accumulate along the line  $y = x$ , which indicates a high correlation between the values of  $x$  and  $y$ . The scatterplot of the shadow image shows that the pixels are close to the diagonal. However, the points on the scatter plot of the shaded image are randomly distributed with little or no correlation between the values of neighbouring pixels. This indicates that the image does not provide any meaningful information and looks like a random image.

### 6.3 Entropy Analysis

In order to measure the unpredictability of the information in an image, it can be evaluated using Shannon information entropy. It is used to calculate the amount of data contained in a digital image. A high value of entropy for a completely random image indicates that less information is revealed. For a given source  $m$ , the entropy  $E(m)$  is given by:

$$E(m) = - \sum_{i=0}^{n-1} P(m_i) \times \log_2 P(m_i) \quad (9)$$

where the  $P(m_i)$  value is the probability of occurrence of a given symbol  $m_i$  in a given source  $m$ . Applied to the image domain, the symbol  $m_i$  is the pixel value. For grayscale or colour images, the number of grey levels is  $n = 256$ . The ideal entropy value of a randomly generated image,  $E(m)$  is 8. From the point of view of security, the entropy value of the shadow image obtained by the share generation algorithm should preferably be close to the ideal value of 8. The entropy results for the colour components, R, G, and B, of both the original secret image and the shadow image are shown in Table 4. It is clear from the tabular data that the entropy values of the R, G, and B components of the shadow image are very close to the value 5.754. Since the algorithm generates only 54 different values for each component of the pixel values of the shadow image, it can be assumed that 5.754 ( $2^{5.754} \approx 54$ ) is the ideal value of the entropy of the shadow image. Thus, our proposed scheme generates statistically random shares, illustrating the unpredictability of these shadow images.

**Table 3.** Correlation coefficient of pixel values of secret image and shadow image

| Image           | Component | Horizontal    | Vertical      | Diagonal      |
|-----------------|-----------|---------------|---------------|---------------|
| Secret image    | R         | 0.9843202446  | 0.9922728755  | 0.9749304159  |
|                 | G         | 0.9767414573  | 0.9876747850  | 0.9636965953  |
|                 | B         | 0.9600837800  | 0.9757288531  | 0.9423053376  |
| SC <sub>1</sub> | R         | −0.0015753976 | 0.0014529780  | −0.0007672264 |
|                 | G         | −0.0014562086 | −0.0025840600 | 0.0022319272  |
|                 | B         | −0.0050985438 | −0.0060153474 | −0.0011226487 |
| SC <sub>2</sub> | R         | 0.0011808172  | 0.0011158157  | 0.0016152525  |
|                 | G         | −0.0008833490 | 0.0001167807  | 0.0017657444  |
|                 | B         | 0.0020872883  | 0.0061039862  | −0.0026227192 |
| SC <sub>3</sub> | R         | 0.0031013061  | 0.0022077958  | 0.0036787304  |
|                 | G         | −0.0006259095 | 0.0015829644  | −0.0042509128 |
|                 | B         | −0.0002434784 | 0.0011754490  | −0.0051891249 |
| SC <sub>4</sub> | R         | 0.0019265673  | 0.0049449935  | 0.0024980655  |
|                 | G         | 0.0040619898  | 0.0009630449  | −0.0042391799 |
|                 | B         | 0.0009245454  | −0.0026276089 | −0.0027307871 |
| SC <sub>5</sub> | R         | 0.0004803473  | 0.0035977616  | 0.0003895992  |
|                 | G         | −0.0058247297 | −0.0017924039 | 0.0013101470  |
|                 | B         | −0.0037391158 | 0.0007805913  | 0.0011641809  |

**Table 4.** Entropy values of each component of secret image and shadow images

| Image           | Component    |              |              |
|-----------------|--------------|--------------|--------------|
|                 | R            | G            | B            |
| Secret image    | 7.2381961853 | 7.5860556603 | 6.9440202728 |
| SC <sub>1</sub> | 5.7544773236 | 5.7545622630 | 5.7546067756 |
| SC <sub>2</sub> | 5.7545931252 | 5.7546720888 | 5.7544979204 |
| SC <sub>3</sub> | 5.7547036853 | 5.7546235836 | 5.7546743805 |
| SC <sub>4</sub> | 5.7545790551 | 5.7546160259 | 5.7546273126 |
| SC <sub>5</sub> | 5.7545713310 | 5.7545760794 | 5.7545803016 |

## 6.4 Similarity Analysis

In this subsection, the similarity between the shadow images generated by the proposed secret image sharing scheme will be analysed. The Number of Pixels Change Rate (NPCR) and the Uniform Average Change Intensity (UACI) [13] are the two most

commonly used similarity metrics. They aim to measure the randomness between two random images. Suppose  $P_1$  and  $P_2$  are two images of size  $M \times N$ . Then, the NPCR and UACI between  $P_1$  and  $P_2$  are as follows:

$$\text{NPCR} = \frac{\sum_{i,j} D(i,j)}{M \times N}, D(i,j) = \begin{cases} 1, & P_1(i,j) \neq P_2(i,j) \\ 0, & \text{Others} \end{cases} \quad (10)$$

$$\text{UACI} = \frac{1}{M \times N} \sum_{i,j} \frac{|P_1(i,j) - P_2(i,j)|}{255} \quad (11)$$

Due to the special nature of the generated shadow images, it is necessary to replace 255 with 53 in the denominator of the formula to obtain the UACI metric applicable to this algorithm. The five experimentally generated shadow images (Fig. 1(d)–(h)) were compared with each other, and Table 5 was obtained by averaging the 10 sets of calculated data.

**Table 5.** NPCR and UACI values

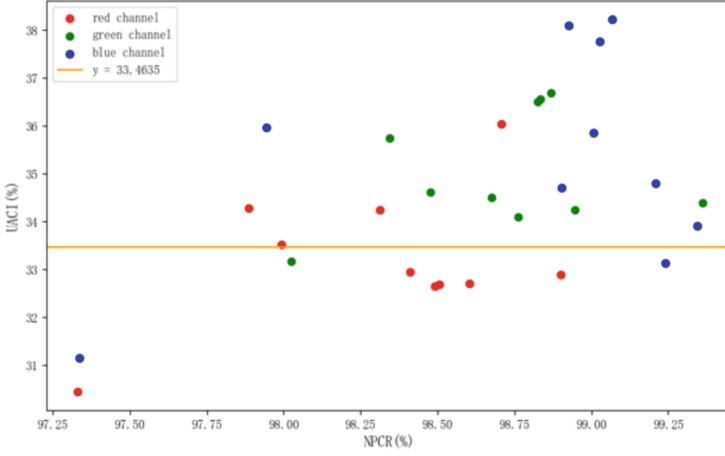
| Channel | Index            |                  |
|---------|------------------|------------------|
|         | Average NPCR (%) | Average UACI (%) |
| R       | 98.3141          | 33.2466          |
| G       | 98.7123          | 35.0567          |
| B       | 98.8001          | 35.3659          |

The experimental data show that almost 98.3% of the pixels are different between the shaded images, and the UACI values of the R, G, and B components in the shaded images are all greater than 33.2%. The theoretical ideal values of NPCR and UACI are 99.6094% and 33.4635%, respectively. The compression of the shadow image pixels results in the NPCR failing to reach the ideal value, but the UACI basically reaches the ideal state.

Therefore, the shadow images generated by the proposed scheme are more random, acquiring one shadow image does not reveal any information about other shadow images, and the proposed scheme has relatively high encryption effect.

To further verify the robustness of the shadow images generated by the scheme, the NPCR values and UACI values of each shadow image are compared with those of other shadow images, and these comparisons are analysed as data points. Figure 4 illustrates the distribution of a total of 30 data points, providing a visual basis for assessing the randomness and security of the shadow images.

In the figure, the orange horizontal line represents the pass line of UACI, which is used to assess whether the intensity of pixel change meets the expected standard. By analysing the distribution of data points in the graph, it can be found that there are differences in the performance of different colour channels. The red channel has a relatively low UACI value and fails to achieve the desired pixel change intensity, while the green and blue channels have significantly higher pixel change intensities, indicating



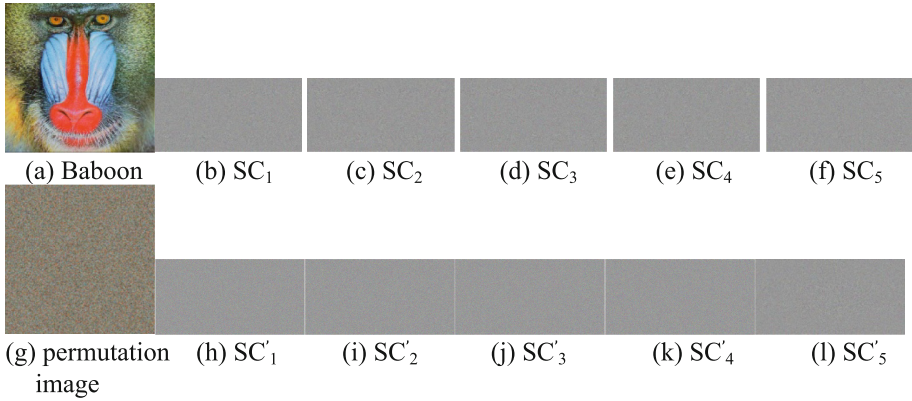
**Fig. 4.** NPCR values and UACI values of shadow images

that the shadow images generated in these channels have better randomness. About two-thirds of the data points in the figure lie above the UACI pass line, suggesting that most of the shading images have variation intensities within acceptable limits and meet the security requirements for image sharing. Although some points fail to reach the pass line, especially concentrated in the red channel, the overall results still show that the scheme has better performance in generating random shadow images.

## 6.5 Pre-processing Analysis

In traditional image visual memory splitting algorithms, the original secret image is usually not directly usable for generating a shared share, but needs to undergo an additional preprocessing step to convert it into a replacement image. This preprocessing process increases the computational overhead of the algorithm and reduces the overall efficiency.

Figure 5 shows the shared images generated from the secret image and the permutation image, respectively, where (b), (c), (d), (e), (f) are the five shared images generated from the original secret image, and (h), (i), (j), (k), (l) are the shared images generated from the permutation image for the secret image. From the data in Table 6, it can be seen that the information entropy and correlation coefficient of the shadow images generated by this algorithm before and after image substitution are almost unchanged. This indicates that the scheme proposed in this paper does not need to convert the secret image into a random image and can use the original secret image directly for share generation. This not only dramatically simplifies the process of the algorithm, but also reduces the computational complexity, making the scheme more efficient in practical applications. By eliminating the image replacement operation, this scheme achieves the optimisation of the share image generation process, while maintaining high information security and randomness of the share image.



**Fig. 5.** Shares from the secret image and the permuted image

**Table 6.** Entropy and Correlation Coefficient of shadow images generated before and after permuting the secret image

|     |              | Entropy-R | Entropy-G | Entropy-B | Correlation |
|-----|--------------|-----------|-----------|-----------|-------------|
| SC1 | Non-Permuted | 5.754585  | 5.754567  | 5.754632  | −0.000574   |
|     | Permuted     | 5.754534  | 5.754542  | 5.754636  | −0.001390   |
| SC2 | Non-Permuted | 5.754584  | 5.754628  | 5.754663  | −0.000313   |
|     | Permuted     | 5.754665  | 5.754628  | 5.754554  | 0.002736    |
| SC3 | Non-Permuted | 5.754675  | 5.754558  | 5.754609  | −0.001050   |
|     | Permuted     | 5.754500  | 5.754650  | 5.754518  | −0.000703   |
| SC4 | Non-Permuted | 5.754580  | 5.754611  | 5.754574  | 0.001062    |
|     | Permuted     | 5.754650  | 5.754570  | 5.754607  | −0.001428   |
| SC5 | Non-Permuted | 5.754604  | 5.754653  | 5.754611  | −0.000417   |
|     | Permuted     | 5.754615  | 5.754660  | 5.754646  | −0.002667   |

## 6.6 Comparison of Other Schemes

In this section, the proposed scheme is compared with the classical colour image visual partitioning scheme and the results are shown in Table 7. Through the comparison, it can be concluded that the image visual split storage algorithm based on the prime distribution theory has advantages in several aspects. Firstly, the algorithm is able to generate smaller shadow image shares, which reduces the cost of storage and transmission. In addition, the scheme does not require complex preprocessing operations such as substitution, which simplifies the algorithm process and improves the operability and ease of use of the algorithm. More critically, the algorithm maintains a high level of effectiveness with a majority of participants sharing, and the reconstruction quality of the images remains within acceptable limits despite lossy recovery.

**Table 7.** Comparison of different schemes

| scheme          | Lossless recovery | Share size      | preprocessing | Works if no. of participants > 157464 |
|-----------------|-------------------|-----------------|---------------|---------------------------------------|
| Xiong [14]      | No                | $\frac{1}{k}$   | Yes           | No                                    |
| Prasetyo [15]   | No                | 1               | No            | Yes                                   |
| Wang [16]       | Yes               | $\frac{1}{k-1}$ | Yes           | No                                    |
| Proposed Scheme | No                | $\frac{1}{k-1}$ | No            | Yes                                   |

Meanwhile, the algorithm proposed in this chapter also has obvious advantages in terms of computational efficiency. In the literature [17], the large prime number 16777259 is used to perform modulo operation on the pixel values of the mapped secret image, while the scheme in this chapter only needs to perform modulo operation on 157477, which significantly reduces the computational complexity and saves the time overhead. This improvement not only makes the algorithm more efficient, but also reduces the demand for computational resources, making it more suitable for various practical application environments.

In addition, in order to avoid the problem of pixel value overflow, this scheme will check the result after the polynomial modelling operation to remove the invalid shared values in the range of [157465, 157476]. The actual available shared value range is [0, 157476] and the probability of generating an invalid shared value within this range can be calculated by the following equation:

$$P_{(\text{Invalid shared value})} = \frac{157477 - 157465}{157477} \times 100\% \approx 0.000762\%$$

It can be seen that the probability of generating an invalid shared value during the sharing operation is extremely low, about 0.00762%. This result indicates that the number of pixels that need to be regenerated with shared values during the whole sharing process is extremely limited, and thus hardly affects the sharing efficiency significantly.

This low-probability invalid value generation characteristic ensures the stability and efficiency of the scheme when processing colour images. Compared with other algorithms that may require complex preprocessing or multiple iterations, this scheme improves the sharing efficiency while still maintaining effective protection against outliers, thus enhancing the robustness and reliability of the image visual memory splitting scheme in this chapter.

## 7 Conclusion

In this chapter, a new visual memory splitting scheme for color images is proposed, which utilizes an approximate pixel set and a set function  $f$  for the bi-mapping operation, allowing the parameters of the modulo operation to be drastically reduced. Through this optimization, the algorithm achieves a significant improvement in time efficiency,

reduces the consumption of computational resources, and improves its feasibility in practical applications. Although the algorithm has a certain degree of loss on the reconstructed secret image, its advantage is that the size of the generated shadow image is significantly smaller than that of the original secret image, a feature that not only reduces the storage and transmission costs, but also increases the flexibility of the scheme in practical applications. The experimental results further validate the performance advantages of the split-storage algorithm, showing that it outperforms the traditional scheme in handling the task of color image split-storage.

The scheme shows obvious advantages in terms of time efficiency, the size of the generated shadow image and the complexity of the algorithmic process, providing an efficient and practical solution for color secret image sharing.

## References

1. Shamir, A.: How to share a secret. *Commun. ACM* **22**(11), 612–613 (1979)
2. Blakley, G.R.: Safeguarding cryptographic keys. In: *International Workshop on Managing Requirements Knowledge*, p. 313. IEEE Computer Society (1979)
3. Hu, F., Li, W., Yu, N.:  $(k, n)$  threshold secret image sharing scheme based on Chinese remainder theorem with authenticability. *Multimed. Tools Appl.* **83**(14), 40713–40732 (2024)
4. Li, F., Hu, H., Zhu, S., et al.: A verifiable  $(k, n)$ -threshold dynamic quantum secret sharing scheme. *Quantum Inf. Process.* **21**(7), 259 (2022)
5. Yan, X., Liu, X., Yang, C.N.: An enhanced threshold visual secret sharing based on random grids. *J. Real-Time Image Process.* **14**, 61–73 (2018)
6. Wu, X., Yang, C.N.: Probabilistic color visual cryptography schemes for black and white secret images. *J. Vis. Commun. Image Represent.* **70**, 102793 (2020)
7. Wang, L., Yan, B., Yang, H.M., et al.: Flip extended visual cryptography for gray-scale and color cover images. *Symmetry* **13**(1), 65 (2020)
8. Shankar, K., Taniar, D., Yang, E., et al.: Secure and optimal secret sharing scheme for color images. *Mathematics* **9**(19), 2360 (2021)
9. Pan, J.S., Liu, T., Yang, H.M., et al.: Visual cryptography scheme for secret color images with color QR codes. *J. Vis. Commun. Image Represent.* **82**, 103405 (2022)
10. Li, W., Peng, C., Tan, W., et al.: A reversible and lossless secret image sharing scheme with authentication for color images. *J. King Saud Univ.-Comput. Inf. Sci.* **35**(10), 101854 (2023)
11. Xiong, L., Ding, R., Yang, C.N., et al.: Invertible secret image sharing with authentication for embedding color palette image into true color image. *IEEE Trans. Circuits Syst. Video Technol.* (2024)
12. Thien, C., Lin, J.: Secret image sharing. *Comput. Graphics* **26**(5), 765–770 (2002)
13. Wu, Y., Noonan, J.P., Agaian, S.: NPCR and UACI randomness tests for image encryption. *Cyber J.: Multidisc. J. Sci. Technol. J. Sel. Areas Telecommun. (JSAT)* **1**(2), 31–38 (2011)
14. Xiong, L., Zhong, X., Yang, C.N.: DWT-SISA: a secure and effective discrete wavelet transform-based secret image sharing with authentication. *Signal Process.* **173**, 107571 (2020)
15. Prasetyo, H., Hsia, C.: Lossless progressive secret sharing for grayscale and color images. *Multimed. Tools Appl.* **78**, 24837–24862 (2019)
16. Wang, R., Yang, G., Yan, X., et al.: Secret image sharing in the encrypted domain. *J. Vis. Commun. Image Represent.* **98**, 104013 (2024)
17. Sardar, K.M., Adhikari, A.: A new lossless secret color image sharing scheme with small shadow size. *J. Vis. Commun. Image Represent.* **68**, 102768 (2020)





# A Novel Design of Radial Basis Function Neural Network Integrating Markov Chain Monte Carlo Clustering Algorithm

Yunlong Zhu<sup>1,2</sup>, Zunwei Fu<sup>1,2,3</sup>(✉), and Zheng Wang<sup>1,2,3</sup>

<sup>1</sup> Linyi University, Linyi, China

zhuyunlong@lyu.edu.cn, {zwfu, wangzheng}@suwon.ac.kr

<sup>2</sup> Research Center for Big Data and Artificial Intelligence, Linyi University, Linyi 276005, China

<sup>3</sup> College of Information Communication Technology, University of Suwon, Hwaseong-si 16499, South Korea

**Abstract.** This paper presents a novel approach to enhance the performance of Radial Basis Function Neural Networks (RBFNNs) for regression tasks through an innovative integration of Markov Chain Monte Carlo (MCMC) sampling with Fuzzy C-Means clustering. Traditional RBFNNs, while effective for various pattern recognition tasks, often face challenges in optimal placement of their radial basis functions. We address this limitation by introducing a hybrid methodology that combines the robust exploration capabilities of MCMC with the flexibility of fuzzy clustering. Our approach dynamically adjusts RBF centers based on both data distribution and regression performance metrics, utilizing cross-entropy loss as a key indicator for optimization. The proposed framework incorporates a comprehensive learning strategy that includes strategic initialization of RBF centers, adaptive parameter adjustment, and weighted least square error estimation for coefficient optimization. Experimental results across multiple benchmark datasets demonstrate significant improvements in regression tasks compared to conventional RBFNNs implementations, while maintaining computational efficiency.

**Keywords:** Radial basis function networks · Second Keyword · Third Keyword

## 1 Introduction

Artificial Neural Networks (ANNs) have emerged as powerful tools in pattern recognition and machine learning, with Radial Basis Function Neural Networks (RBFNNs) representing a distinctive class of neural networks [1]. Unlike traditional multilayer perceptron's, RBFNNs feature a unique three-layer architecture: an input layer that receives data, a hidden layer with radial basis function neurons, and an output layer that generates the network's response. The distinguishing characteristic of RBFNNs lies in their hidden layer, where each neuron implements a radial basis function, typically a Gaussian function, that responds to a local region of the input space.

The fundamental principle of RBFNNs involves transforming input data through radial basis functions, which act as local receptive fields in the input space. Each RBF neuron produces its highest output when the input is near its center, with the response decreasing as the input moves away from the center [2, 3]. This localized response property makes RBFNNs particularly effective at function approximation.

A fundamental challenge in RBFNNs design lies in the optimal positioning of Radial Basis Functions (RBFs) across the input domain. Traditional approaches typically employ unsupervised clustering techniques to position these functions by examining data distribution patterns [5]. However, this conventional methodology presents a limitation: while the network's errors heavily depend on RBF positioning, the standard distribution-based placement doesn't directly address regression performance. Our research introduces an innovative refinement approach that considers both data distribution patterns and regression error information to optimize RBF placement.

Conventional clustering methods often pursue objectives that don't align perfectly with the specific needs of RBFNN-based regression systems. Our approach begins with standard clustering for initial RBF placement but distinguishes itself by incorporating an iterative refinement process that leverages additional performance metrics to enhance regression accuracy.

The cornerstone of our methodology is the utilization of cross-entropy loss to derive crucial performance indicators for each data instance. These indicators provide valuable insights into the network's regression effectiveness, where lower cross-entropy values indicate superior regression outcomes compared to instances with higher values. We particularly focus on addressing data points with elevated cross-entropy values during RBF positioning. Our novel approach implements a Markov Chain Monte Carlo (MCMC) based Fuzzy C-Means clustering algorithm, which integrates sophisticated MCMC sampling techniques with fuzzy clustering principles. This hybrid method enables comprehensive exploration of potential solutions while maintaining the adaptability of fuzzy logic, ultimately leading to more precise RBF placement that accounts for individual data point performance.

The optimization of polynomial coefficient vectors represents another critical aspect of RBFNNs development. We address this challenge through the application of Least Square Error (LSE) estimation, where the weighting factors are derived from cross-entropy loss calculations. The effectiveness of our proposed MFCM-RBFNNs is thoroughly evaluated using diverse machine learning datasets, demonstrating its practical applicability and performance advantages.

The main contributions of this paper can be summarized as follows:

- We propose an innovative clustering method that combines Markov Chain Monte Carlo sampling with FCM clustering for RBF center placement. This hybrid approach offers more robust and accurate center positioning compared to traditional clustering methods, as it can better explore the solution space while avoiding local optima
- We develop a comprehensive framework that dynamically adjusts RBF parameters based on regression performance. This framework incorporates cross-entropy loss information to guide the optimization process, ensuring that the network architecture adapts to both data distribution and regression requirements.

## 2 Radial Basis Function Optimized by Fuzzy C-Means Clustering and MCMC-Fuzzy C-Means Clustering

### 2.1 Fuzzy C-Means Clustering Based Radial Basis Function Neural Network

Let's consider a dataset  $X = \{x_1, x_2, \dots, x_N\}$ , where  $x_i \in \mathbb{R}^p$  is the input vector and  $y_i \in \mathbb{R}$  is the corresponding output for the  $i_{th}$  sample, where  $p$  stands for the dimensionality of each input pattern and  $N$  is the number of data. The goal is to approximate the function  $f(X)$  such that  $\hat{y}(X) \approx y$  for all samples. Apply FCM clustering on the training dataset to identify the centers of the radial basis functions. Unlike k-means, FCM assigns membership values  $u_{ij}$  that indicate the degree to which data point  $x$  belongs to cluster  $j$ , Minimize the following objective function to determine the cluster centers  $c_j$ :

$$J = \sum_{i=1}^n \sum_{j=1}^K u_{ij}^m \|x_i - c_j\|^2 \quad (1)$$

where  $m$  is the fuzzification parameter,  $K$  is the number of clusters, and  $u_{ij}$  are the membership values,  $\|\cdot\|$  represents the Euclidean distance,  $m$  is the fuzzification parameter.

The optimization of the objective function, denoted as (1), is executed through an iterative process. This process comprises two primary steps: the determination of membership degrees  $u_{ij}$  for each data point and the computation of cluster prototypes. These steps are performed alternately and recursively until convergence criteria are met or a predetermined number of iterations is reached. The membership degrees quantify the extent to which each data point belongs to each cluster, while the cluster prototypes represent the central tendencies or representative points of the clusters. This iterative approach allows for the progressive refinement of both the cluster assignments and the cluster characteristics, ultimately leading to an optimal or near-optimal solution that minimizes the objective function.

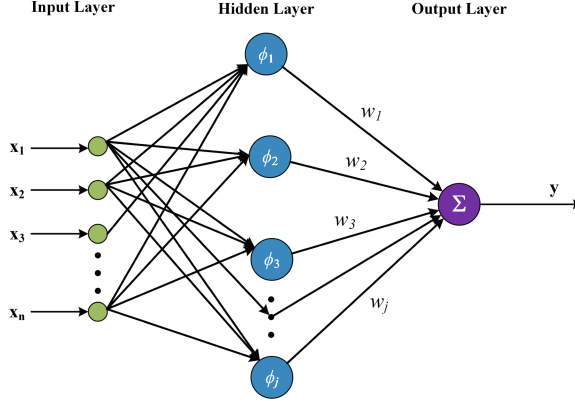
The membership degree in a  $j$ -th cluster is calculated as

$$u_{ij} = \frac{1}{\sum_{k=1}^K \left( \frac{\|x_i - c_j\|}{\|x_i - c_k\|} \right)^{\frac{2}{m-1}}} \quad (2)$$

The cluster centers  $c_j$  are updated based on the current membership values using the formula:

$$c_j = \frac{\sum_{i=1}^n u_{ij}^m x_i}{\sum_{i=1}^n u_{ij}^m} \quad (3)$$

The membership values  $u_{ij}$  and cluster centers  $c_j$  are iteratively updated until convergence, typically when the change in cluster centers between iterations is smaller than a predefined threshold (Fig. 1).



**Fig. 1.** Radial Basis Function Neural Networks (RBFNNs).

In a regression task, the goal of RBFNNs is to approximate an unknown function  $f: \mathbb{R}^p \rightarrow \mathbb{R}$  that maps an input vector  $X \in \mathbb{R}^p$  to a scalar output  $y \in \mathbb{R}$ , such that  $f(X) + \epsilon$ , where  $\epsilon$  is a noise term, RBFNNs accomplishes this by learning the parameters of a model consisting of weighted sums of radial basis functions. For each cluster center  $\mathbf{c}_j$ , calculate the radial basis function value  $\phi_j(\mathbf{x}_i)$  for a given input  $\mathbf{x}_i$  using a Gaussian function:

$$\phi_j(\mathbf{x}_i) = \exp\left(-\frac{\|\mathbf{x}_i - \mathbf{c}_j\|^2}{2\sigma_j^2}\right) \quad (4)$$

where  $\sigma_j$  is the width of the basis function for cluster  $j$ , often calculated based on the spread of points within that cluster.

Construct the hidden-to-output layer connection weights using a linear model. The output  $y$  is given as a weighted sum of the radial basis function activations:

$$y(\mathbf{x}) = \sum_{j=1}^K w_j \phi_j(\mathbf{x}) \quad (5)$$

where  $w_j$  are the weights associated with each radial basis function.

When applying the Least Squares Estimation (LSE) technique, this method minimizes the error between the predicted output  $y(\mathbf{x})$  and the actual target values  $y_{true}$ . The error  $E$  is expressed as:

$$E = \sum_{i=1}^n \left( y_{true} - \sum_{j=1}^K w_j \phi_j(x_i) \right)^2 \quad (6)$$

## 2.2 Markov Chain Monte Carlo Method Based-Fuzzy C-Means Clustering

In the following sections, we develop the MFCM model and introduce a Markov-Chain Monte-Carlo (MCMC) method for parameter inference. MCMC methods work by constructing a Markov chain that has the desired probability distribution as its equilibrium distribution. In our context, we employ MCMC to estimate the parameters of the MFCM model, allowing us to handle complex posterior distributions and incorporate prior knowledge into our inference process [6].

The MFCM model is predicated on a joint likelihood function, with its objective function formulated as the negative logarithm of this joint likelihood. This formulation adheres to the principle of maximum likelihood estimation, where minimizing the negative log-likelihood is equivalent to maximizing the likelihood itself [7]

$$\begin{aligned}
 Q(\mathbf{X}, \mathbf{U}, \mathbf{K}, \mathbf{C}) = & \frac{1}{K} \sum_{i=1}^N \sum_{j=1}^K u_{ij}^m \| \mathbf{x}_n - \mathbf{c}_j \|^2 \\
 & - \frac{2}{K} \sum_{i=1}^N \sum_{j=1}^K (\alpha_j - 1) \log(u_{ij}) \\
 & + \frac{1}{K} \sum_{j=1}^K (\mathbf{c}_j - \boldsymbol{\mu}_c)^T \boldsymbol{\Sigma}_c^{-1} (\mathbf{c}_j - \boldsymbol{\mu}_c) + \text{Penalty}(K). \quad (7)
 \end{aligned}$$

where the  $\mathbf{U}$  is the membership degree matrix,  $K$  is the number of clusters,  $\mathbf{C}$  is the cluster,  $\alpha$  is the Dirichlet parameter,  $\boldsymbol{\mu}$  is the mean of the input data and  $\boldsymbol{\Sigma}_c$  is the covariance matrix.

## 3 Experimental Results

### 3.1 Experimental Settings

The operational parameters of the proposed MFCM-RBFNNs architecture are established with specific values outlined in Table 1. These configurations were determined through extensive experimentation, aiming to explore the model's capabilities across diverse regression tasks while maintaining practical computational requirements. The selected numerical parameters emerged from methodical experimentation, during which learning dynamics and convergence characteristics were carefully analyzed.

To assess the performance of the algorithms, we record both the Root Mean Squared Error (RMSE) across five runs. The RMSE is calculated using the formula  $RMSE = \sqrt{1/n_{test} \sum_{i=1}^{n_{test}} (y_{true} - \hat{y}(x_i))^2}$ , where  $n_{test}$  is the number of test samples,  $y_{true}$  is the true output for sample  $i$ , and  $\hat{y}(x_i)$  is the predicted output.

### 3.2 Performance Demonstration

In this subsection, we will discuss the outcomes of numerical experiments conducted using the proposed MFCM-RBFNN model, as well as several baseline models. These

**Table 1.** Parameter settings for proposed MFCM-RBFNNs

| Parameter                     | Value            |
|-------------------------------|------------------|
| Fuzzification Coefficient     | 1.1, 2.0, 3.0    |
| Number of Clusters            | 2, 3, 4, 5, 6, 7 |
| Dirichlet parameter           | 0.95, 2.0, 3.0   |
| Number of sampling iterations | 50, 100, 150     |

**Table 2.** Comparison results of conventional RBFNNs and proposed MFCM-RBFNNs

| Datasets  | RBFNNs              |                     | MFCM-RBFNNs (Proposed) |                                       |
|-----------|---------------------|---------------------|------------------------|---------------------------------------|
|           | Training            | Testing             | Training               | Testing                               |
| AutoMPG6  | $0.0785 \pm 0.0024$ | $0.0812 \pm 0.0023$ | $0.0658 \pm 0.0041$    | <b><math>0.0785 \pm 0.0027</math></b> |
| Compactiv | $0.0903 \pm 0.0015$ | $0.0921 \pm 0.0047$ | $0.0782 \pm 0.0019$    | <b><math>0.0791 \pm 0.0035</math></b> |
| Ele-2     | $0.0168 \pm 0.0002$ | $0.0171 \pm 0.0002$ | $0.0145 \pm 0.0004$    | <b><math>0.0140 \pm 0.0003</math></b> |
| Laser     | $0.0679 \pm 0.0012$ | $0.0708 \pm 0.0021$ | $0.0568 \pm 0.0032$    | <b><math>0.0585 \pm 0.0035</math></b> |
| Wankara   | $0.0218 \pm 0.0024$ | $0.0241 \pm 0.0001$ | $0.0185 \pm 0.0003$    | <b><math>0.0209 \pm 0.0002</math></b> |

experiments were carried out on five real-world datasets sourced from UCI [8] and KEEL [9].

Table 2. Shows the results of our numerical experiments demonstrate that the MFCM-RBFNN model significantly outperforms the baseline models across all five datasets. This enhanced performance can be attributed to the model's ability to effectively capture complex patterns in the data through its integration of fuzzy clustering and radial basis functions. By leveraging the strengths of MFCM, the MFCM-RBFNN model exhibits improved robustness and adaptability, leading to more accurate predictions and a reduction in error rates. These findings highlight the model's potential for applications in various real-world scenarios, underscoring its effectiveness in handling diverse datasets.

## 4 Conclusions

In conclusion, the proposed RBFNN model enhanced with Markov Chain Monte Carlo-based Fuzzy C-Means (MFCM-RBFNN) demonstrates significant improvements in performance across various datasets. By integrating the strengths of radial basis functions and MCMC with fuzzy clustering, MFCM provides a more nuanced understanding of the data distribution, allowing for the identification of optimal RBF locations that better represent the underlying patterns in the dataset. This strategic initial positioning enhances the overall performance of the model, ensuring that the RBFs are well-placed to capture the complexities of the data accurately. This research paves the way for further exploration and refinement of hybrid models that leverage the benefits of both fuzzy clustering and advanced sampling techniques.

## References

1. Li, W., Hori, Y.: An algorithm for extracting fuzzy rules based on RBF neural network. *IEEE Trans. Industr. Electron.* **53**(4), 1269–1276 (2006)
2. Pedrycz, W.: Conditional fuzzy clustering in the design of radial basis function neural networks. *IEEE Trans. Neural Netw.* **9**(4), 601–612 (1998)
3. Park, B.J., Pedrycz, W., Oh, S.K.: Polynomial-based radial basis function neural networks (P-RBFNNs) and their application to pattern classification. *Appl. Intell.* **32**(1), 27–46 (2008)
4. Billings, S.A., Zheng, G.L.: Radial basis function network configuration using genetic algorithms. *Neural Netw.* **8**(6), 877–890 (1995)
5. Liu, Y., Zhang, Y., Yu, Z., Zeng, M.: Incremental supervised locally linear embedding for machinery fault diagnosis. *Eng. Appl. Artif. Intell.* **50**, 60–70 (2016)
6. Gu, X., Wang, S.: Bayesian Takagi–Sugeno–Kang fuzzy model and its joint learning of structure identification and parameter estimation. *IEEE Trans. Industr. Inf.* **14**(12), 5327–5337 (2018)
7. Glenn, T.C., Zare, A., Gader, P.D.: Bayesian fuzzy clustering. *IEEE Trans. Fuzzy Syst.* **23**(5), 1545–1561 (2014)
8. The UCI Machine Learning Repository. <https://archive.ics.uci.edu>
9. KEEL Software and KEEL Datasets (2017). <http://sci2s.ugr.es/keel>



# A Synthesis of Techniques for Feature Downgrading Processing in IoT Security

Yifang Wang<sup>(✉)</sup>

Qingdao Huanghai University, Qingdao 266555, China  
Wangyf11@qdhhc.edu.cn

**Abstract.** With the wide application of IoT devices, the proliferation of vulnerability information poses a serious challenge to cyber security. Vulnerability classification, as an important application of natural language processing technology in cybersecurity, aims to extract key features from vulnerability descriptions and classify them to help security managers quickly identify and fix vulnerabilities. Traditional dimensionality reduction methods such as TF-IDF, despite reducing algorithmic complexity, have limited understanding of semantics and context and are prone to ignore some important information. In recent years, statistical methods such as information gain and chi-square test have been introduced to improve feature extraction, but there are still problems such as high-dimensional sparse data dependency. The rise of deep learning techniques, especially the application of word embedding and pre-trained language models, has dramatically improved the accuracy and robustness of vulnerability classification and alleviated the limitations of traditional methods. Combined with artificial intelligence techniques, vulnerability classification is developing in a more efficient and intelligent direction, providing strong support for IoT cybersecurity.

**Keywords:** Dimensionality Reduction · NLP · Vulnerability classification

## 1 Introduction

With the massive number of devices accessing the Internet of Things (IoT), all kinds of vulnerability information continue to emerge, bringing increasingly severe challenges to network security. These vulnerabilities may become the entrance for hacker attacks, thus threatening personal privacy, data security and even the stability of the whole IoT system [1]. Therefore, effective management and application of vulnerability information and timely mastery of the latest vulnerability situation are crucial to guarantee network security. Vulnerability classification [2] can discover vulnerability information in a timely manner and help network security managers to quickly identify vulnerabilities so that timely measures can be taken to prevent and repair vulnerabilities.

Vulnerability classification is actually the application of natural language processing techniques to the field of cyber security [3] to help vulnerability analysis. Vulnerability databases usually store vulnerability versions, numbers, vulnerability descriptions and other information, in which the vulnerability information described in natural language



contains key features to distinguish the vulnerability categories. Therefore, the essence of vulnerability classification is to extract key features from vulnerability descriptions and categorize this information into predefined categories. Dimensionality reduction processing is a key step in vulnerability classification, which aims to convert natural language text into structured features that can be processed by the model. However, traditional dimensionality reduction processing methods, such as the TF-IDF algorithm [4], reduce the complexity of the algorithm by filtering keywords that have a low frequency in the corpus and a high contribution to the content of the document, but may ignore words that are more common in the full text but have a high impact on a particular paragraph.

Other statistical methods, although capable of filtering the contribution of classification results by measuring features, are still limited by the inherent problems of statistical methods. With the rise of deep learning technologies, natural language processing [5] capabilities have been significantly improved, providing new solution ideas for vulnerability classification tasks. Deep learning-based feature extraction methods, such as word embedding techniques and pre-trained language models, are able to capture richer contextual semantics and inter-word relationships, thus significantly improving the accuracy and robustness of classification. The introduction of these methods not only effectively alleviates the limitations of traditional methods, but also promotes the development of vulnerability classification in a more efficient and intelligent direction.

## 2 Definition and Related Techniques of Dimensionality Reduction Processing

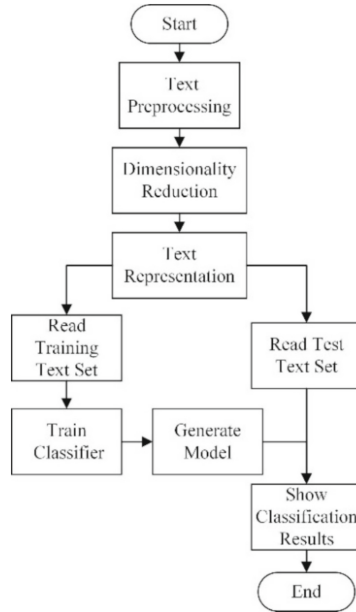
### 2.1 Basic Definition

Dimensionality reduction processing is a core step in data processing [6] and pattern recognition, where the goal is to extract meaningful features from raw data and transform the data into a form that can be understood and processed by machine learning models. These features can be regarded as the core patterns or attributes of the data, reflecting important information about the data, and are a key factor in improving the performance of algorithms.

Dimensionality reduction technology is widely used in image processing, speech processing and text processing, etc. Through dimensionality reduction processing, data dimensionality can be reduced, algorithm operation efficiency can be improved, important information can be retained, helping the model to more accurately identify data patterns, and improving classification, clustering and prediction capabilities.

Vulnerability classification is the process of categorizing and tagging some known vulnerability information according to some formulated system or criteria, and then categorizing the vulnerability information that needs to be classified into one or more categories. In this process, most of the work is automated through advanced technologies such as machine learning and natural language processing. The specific vulnerability classification process is shown in Fig. 1, where the vulnerability text is first preprocessed, and then data dimensionality reduction is performed to obtain low-dimensional data and vector representation. Next, the vector-represented dataset is divided into a training

set and a test set, and the training set is input into the classifier for training, and the appropriate classifier parameters are obtained through iterative adjustment, and then the test set is input into the classifier to obtain the classification results.



**Fig. 1.** Vulnerability classification process.

Vulnerability Information usually exists in an unstructured form, and analyzing it directly using these data may face computational difficulties. Therefore, the core task of feature extraction lies in converting unstructured text into a numerical form that can be handled by models, such as vectors or matrices. This process is not only the basis of text analysis, but also an important means to improve model performance.

The first task of text feature extraction is to transform the text content into a numerical representation that can be understood by the model. For example, Bag of Words (BoW) is one of the most common representations, which generates a high-dimensional sparse vector for reflecting the overall content of the text by counting the frequency of occurrence of words in the document. In addition, TF-IDF (Word Frequency-Inverse Document Frequency) further improves the BoW method by reducing the weight of high-frequency common words and highlighting rare but important keywords.

In addition to representing text content, extracting key patterns in text is also an important goal of feature extraction. These patterns may include words, phrases, syntactic structures, or even higher-order semantic relationships. In sentiment analysis tasks, extracting sentiment vocabulary and intonation patterns, on the other hand, can better determine the sentiment tendency of a text. The introduction of deep learning techniques further enhances the pattern capturing capability. For example, Convolutional Neural Networks (CNN) are able to extract local patterns in text, such as key phrases,

while Recurrent Neural Networks (RNN) and their variants (e.g., LSTM, GRU) are good at capturing sequential relationships in text.

Text data often contains noisy information that is irrelevant to the analysis, such as stop words (e.g., “the”, “and”, etc.), punctuation marks, or spelling errors. If these data are used directly for modeling, it may increase the computational burden or even interfere with the model’s judgment. The feature extraction process can significantly reduce the noise interference by deactivating word filtering, stemming (e.g., Porter or Lancaster algorithms), and spelling correction. In addition, feature selection techniques (e.g., methods based on information gain or mutual information) can filter out the most valuable features for the target task, thus further improving the efficiency and performance of the model.

2.2 Related Technologies

**Feature Extraction.** The process of feature extraction is a crucial task that runs through various stages of text processing, aiming to transform raw text data into numerical feature representations suitable for processing by machine learning or deep learning models. This process usually needs to be refined at multiple levels, including text preprocessing, feature generation, feature selection, and context-sensitive feature extraction.

In the text preprocessing [7] stage, the data is cleaned and normalized to ensure a good foundation for subsequent feature extraction. This segment includes operations such as word separation, deactivation removal, and word shape normalization, which aims to transform complex text into controllable basic units, making subsequent processing more efficient.

Segmentation is the process of breaking down text into its smallest basic units (such as words or phrases) so that they become the basis for subsequent processing. In English, word segmentation usually relies on spaces or punctuation marks to divide words; while in Chinese, due to the lack of clear word boundaries, it usually needs to be accomplished with the help of a segmentation tool (e.g., Jieba). The accuracy of segmentation is crucial to the effectiveness of feature extraction. For example, In Table 1, the results of the segmentation of a particular piece of vulnerability information are presented.

Table 1. Comparison of results before and after word segmentation

| Before                                         | After                                                    |
|------------------------------------------------|----------------------------------------------------------|
| Text = “F5 BIG-IP是一款集成了网络流量管理、负载均衡等功能的应用交付平台。” | F5/ /BIG-/IP/是/一款/集成/了/网络流量/管理/、/负载/均衡/等/功能/的/应用/交付/平台/。 |

Discontinued words are those words that occur frequently but contribute less to the semantics of the text (e.g., “the”, “is” in English, “的” in Chinese, “了” in English, ‘的’ and ‘了’ in Chinese). By removing these words, the data noise can be reduced and the efficiency and quality of feature extraction can be improved. After filtering the

**Table 2.** Comparison of results before and after Comparison before and after filtering of deactivated words

| Before                                                            | After                                      |
|-------------------------------------------------------------------|--------------------------------------------|
| F5/ /BIG/-/IP/是/一款/集成/了/网络流量/管<br>理/、/负载/均衡/等/功能/的/应用/交付/平台/<br>。 | F5/BIG/IP/集成/网络流量/管理/负载/均衡/<br>功能/应用/交付/平台 |

deactivated words for the partitioning results obtained in Table 1, the results shown in Table 2 can be obtained.

In order to solve the problem of feature dispersion of words due to word form changes, different forms of words can be unified into a root or basic form by stem extraction or word form reduction. For example, the words “studying” and “studies” are united as “study” through morphological normalization, thus reducing redundant features.

**Feature Generation.** The next session of feature generation [8] represents the pre-processed text data into a numerical form. The patterns in the language are abstracted into vectorized representations by employing methods such as Bag of Words (BoW) model, TF-IDF or Topic Modeling, which facilitates the model to understand the patterns, semantics and statistical properties in the text.

BoW [9] modeling generates a fixed-length vector by mapping words in a text to a fixed vocabulary (dictionary) and counting the frequency of each word in the text. The method ignores word order and syntax and focuses only on whether words occur and how many times they occur.

Term frequency-inverse document frequency (TF-IDF)) is a weighting technique widely used in the field of text mining [10], which is used to quantify the importance of a word for a particular document in a document collection or corpus. In text categorization tasks, TF-IDF is used as a feature vector of a document to help the model identify the keywords in the document and aid in document categorization.

When processing text data, TF-IDF algorithm has the advantages of simplicity, effectiveness, and wide applicability, so it is widely used. However, TF-IDF mainly focuses on the importance of words in the whole document, ignoring the order and contextual relationship between words. Therefore, TF-IDF usually needs to be paired with other methods to compensate for its limitations and improve the effectiveness and accuracy of text processing.

**Feature Selection.** Feature selection is a key step in filtering the generated high-dimensional features, aiming at removing redundant information and retaining only the most important features for the target task. This both improves model efficiency and helps reduce computational complexity. Commonly used methods include information gain, chi-square test.

Information Gain (IG) is an important concept in machine learning and information theory [11] and is used to measure how informative a feature is for a given task. It works by selecting the most informative feature by measuring the feature’s ability to discriminate between classification tasks.

The IG is calculated based on entropy, which is regarded as a measure of the degree of chaos or uncertainty in a system [12]. A system is usually said to have high entropy if it has many possible states and each state has the same probability of occurring. The greater the information gain, the more significant the feature is in distinguishing the classification, and therefore it is preferentially retained. For example, in text categorization, the word “virus” has a higher information gain for texts on the topic of computer security, while “file” may contribute less.

The chi-square test is a statistical method for measuring the correlation between features and categories and assessing their significance by calculating the difference between the actual distribution of features and the expected distribution. The stronger the correlation between features and categories, the larger the chi-square value, and the more significant the contribution of features to the classification task. For example, in spam categorization, the term “winning the lottery” has a high chi-square value with the spam category, which indicates that this term is important for distinguishing spam.

Finally, entering the context-sensitive feature extraction stage, modern natural language processing techniques such as word embedding (Word2Vec) or pre-trained language models based on deep learning (BERT, GPT) are able to capture richer contextual semantics, dynamically generate feature representations, and overcome the limitations of traditional methods that cannot handle contextual information.

Word embedding is to map words into a low-dimensional dense vector space to capture semantic similarities and contextual relationships. For example, there is a semantic translation relationship between the vectors of “king” and “queen”.

Pre-trained language models are deep learning-based pre-trained models capable of generating context-sensitive feature representations that not only consider the semantics of the words themselves, but also dynamically adjust the representations with the context. For example, in the sentence “The bank has a beautiful view of the river”, BERT can accurately distinguish whether “bank” is a financial institution or a geographical location.

Through these multi-level processing steps, feature extraction can transform unstructured text into structured and information-rich numerical representations, providing a solid foundation for a variety of text analysis tasks, while achieving significant improvements in efficiency and accuracy.

### 3 Research on Key Technologies

With the wide application of high-dimensional vulnerability datasets, the problem of processing high-dimensional data has gradually emerged, and the increase of computational complexity and storage space poses a serious challenge to the efficiency of classification models. Therefore, how to process datasets for dimensionality reduction has become the focus of research. Feature selection and feature extraction are the two main methods for dimensionality reduction, and researchers have carried out extensive exploration and innovation in these two aspects.

**Feature Selection Related Research.** Research in feature selection has focused on filtering the most relevant features from raw data so as to improve the classification ability and computational efficiency of the model while reducing the dimensionality of

the data. Researchers have made significant progress by improving the feature screening mechanism and increasing the consideration of data characteristics.

Yang et al. [13] proposed a feature selection algorithm for document frequency in 1997, which realizes the purpose of reducing the dimension of the feature set by removing the features that are below a certain threshold, which is simple and easy to operate, but is not effective when dealing with complex nonlinear data. Nevertheless, this algorithm lays the foundation for feature selection and provides useful inspiration for subsequent research.

In order to address the shortcomings of the document frequency-based algorithm in nonlinear data processing, subsequent researchers have attempted to incorporate more data correlation analysis methods in the feature selection process to further explore the complex relationship between features and target labels. Zhenhai Zhang et al. [14] proposed an information entropy-based multi-label feature selection algorithm in 2013, which quantifies the contribution of features to target classification by calculating the information gain between features and label sets.

The algorithm performs well in dealing with direct associations between features and labels, but has limitations in dealing with interconnections between features. When there are highly correlated features in the feature set, the algorithm may not be able to accurately identify and fully utilize the relationships between these features, thus limiting its effectiveness in complex data scenarios.

Zhou et al. [15] proposed a feature selection method based on mutual information of correlation coefficients in 2022, which successfully filtered out features with higher information content by calculating the absolute value of correlation coefficients between features. By calculating the correlation metrics between features, such as mutual information and correlation coefficient, this method can effectively identify redundant features and retain those features that have an independent and important role in the classification task. This mechanism of comprehensively considering the relationship between features can further optimize the quality of the feature set while ensuring the validity of the features, and significantly improve the performance and generalization ability of the model, thus showing greater advantages in the processing of high-dimensional datasets.

Alsahaf et al. [16] in 2022 proposed an innovative feature selection framework in conjunction with the tree augmentation model, focusing on addressing the impact of highly redundant feature sets on classification performance. The method achieves more accurate feature selection by re-weighting the samples to enhance the model's focus on key features. This innovative design combining sample weighting and tree enhancement model not only provides a new idea for the optimization of feature selection methods, but also injects new vitality into the field of high-dimensional data processing and further promotes its wide application in text classification, image recognition, and other complex tasks.

**Feature Extraction Related Research.** Compared to feature selection, feature extraction captures richer information in the data by generating new features. These methods seek to retain key patterns in the data while reducing its dimensionality to improve the performance of classification and clustering tasks. Researchers have continued to improve the performance of feature extraction algorithms by introducing techniques such as clustering, statistical analysis and deep learning.

Lewis et al. first applied the mutual nearest neighbor clustering method to text feature extraction in 1992, where feature words were grouped and clustered by calculating the similarity between words. This method provides important theoretical and practical insights for text feature extraction, and its main advantage lies in its ability to reduce redundant features through similarity analysis, and to a certain extent improves the model's understanding of the thematic structure of the text. However, with the increase of data dimensions and the widespread existence of nonlinear features, the applicability of the method in high-dimensional complex data scenarios appears insufficient [17]. To cope with this problem, researchers have tried to combine statistical analysis methods on this basis to further improve the generality and performance of the algorithm.

Rahman et al. [18] proposed a feature extraction method combining principal component analysis (PCA) and t-statistic in 2020, focusing on the analysis and dimensionality reduction of electroencephalographic emotion signals (EEG). The method reduces the dimensionality of signal features through PCA, which reduces redundant data, while the t-statistic is utilized to filter out features with significant differentiation. In a series of experiments, the method demonstrated high accuracy and stability, effectively improving the performance of the classification model. However, the study mainly focuses on commonly used signal features and fails to fully explore the potential complex information in the data, which limits the performance of the algorithm in more complex application scenarios.

To further optimize the feature extraction performance, researchers have begun to explore the fusion strategy of multiple algorithms. Li et al. [19] in 2022 proposed an innovative approach combining Weighted Least Squares (WLS), Parametric Time-Frequency Analysis (PTFA), and Multi-Synchronous Compression Transform (IMSST). By integrating the advantages of these three techniques, the method demonstrates excellent performance in feature extraction, especially when dealing with non-smooth signals. Compared with the traditional methods, the algorithm is able to gather signal energy more efficiently, while improving feature discrimination and model robustness.

Ahmed et al. [20] in 2022 proposed a hybrid feature extraction method that combines differential entropy features and convolutional neural network (CNN) models to capture complex nonlinear features. This method utilizes deep learning to extract text global information while capturing local features through differential entropy, significantly improving the classification performance of emotion EEG datasets.

## 4 Discussion

Feature extraction and feature selection, as key aspects of data processing and machine learning, although have made significant progress in theoretical research and application practice, still face many challenges and urgent problems when dealing with complex real-world scenarios. These problems not only limit the application scope of existing methods, but also expose the inadequacy of current research in terms of theoretical depth and practical applicability.

First of all, insufficient modeling of correlation between features is a common problem. Most of the existing methods focus only on the importance assessment of a single feature and ignore the potential interactions and redundancy that may exist between

features. When dealing with data with highly redundant or strongly correlated features, this neglect may lead to the loss of key features or incomplete information extraction, which in turn affects the performance of the model.

In addition, the challenge of collaborative modeling of multimodal data is becoming more and more prominent. In modern applications, data sources are often diverse (e.g., a combination of text, image, and speech), and how to effectively fuse the complementary information between these modalities in feature extraction has become an important issue.

Finally, insufficient integration of domain knowledge and insufficient generalizability and robustness are also of concern. In specific application scenarios, domain knowledge can provide effective guidance for feature selection and extraction, but this has been neglected in many studies. Meanwhile, many algorithms have unstable performance on cross-domain data, and are even more prone to fail in the face of noisy or anomalous data.

To address these challenges, future research needs to pay more attention to modeling interactions between features in algorithm design, improve the handling of nonlinear data, develop feature fusion methods for multimodal data, and optimize the computational efficiency and robustness of algorithms. Through these efforts, feature extraction and selection techniques will show greater potential for application in more complex and diverse scenarios.

## 5 Conclusion

Vulnerability classification, as an intersection of natural language processing and cyber security, has made significant progress in recent years. However, traditional dimensionality reduction methods still have problems such as insufficient semantic understanding and limited ability to process high-dimensional sparse data when facing complex and diverse vulnerability information. The introduction of deep learning techniques has injected new vitality into this field, which can significantly enhance the ability to capture semantic and contextual information in vulnerability descriptions and improve the accuracy and efficiency of classification through methods such as word embedding and pre-trained language models. The development of these techniques not only makes up for the limitations of traditional methods, but also provides a more intelligent solution for security in the IoT environment. In the future, comprehensive research combining deep learning and security management system will further promote the application of vulnerability classification technology and provide comprehensive support for network security protection.

## References

1. Bansal, B., Jenipher, V.N., Jain, R., et al.: Big data architecture for network security. *Cyber Secur. Netw. Secur.* 233–267 (2022)
2. Chakraborty, S., Krishna, R., Ding, Y., et al.: Deep learning based vulnerability detection: are we there yet. *IEEE Trans. Softw. Eng.* (2021)



3. Kumar, R., Goyal, R.: On cloud security requirements, threats, vulnerabilities and counter-measures: a survey. *Comput. Sci. Rev.* **33**, 1–48 (2019)
4. Grootendorst, M.: BERTopic: neural topic modeling with a class-based TF-IDF procedure. arXiv preprint [arXiv:2203.05794](https://arxiv.org/abs/2203.05794) (2022)
5. Lyu, Q., Apidianaki, M., Callison-Burch, C.: Towards faithful model explanation in NLP: a survey. *Comput. Linguist.* 1–70 (2024)
6. Aminizadeh, S., Heidari, A., Toumaj, S., et al.: The applications of machine learning techniques in medical data processing based on distributed computing and the internet of things. *Comput. Methods Program. Biomed.* 107745 (2023)
7. Lubis, A.R., Safitri, H.R., Lubis, M., et al.: Implementation of preprocessing in text summarization techniques for Indonesian language documents using the flax T5 approach. In: 2023 11th International Conference on Cyber and IT Service Management (CITSM), pp. 1–6. IEEE (2023)
8. Petit, G., Popescu, A., Schindler, H., et al.: FeTrIL: feature translation for exemplar-free class-incremental learning. In: *Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision*, pp. 3911–3920 (2023)
9. Dai, S., Li, K., Luo, Z., et al.: AI-based NLP section discusses the application and effect of bag-of-words models and TF-IDF in NLP tasks. *J. Artif. Intell. Gener. Sci. (JAIGS)* **5**(1), 13–21 (2024). ISSN 3006-4023
10. Zhou, H.: Research of text classification based on TF-IDF and CNN-LSTM. *J. Phys. Conf. Ser.* **2171**(1), 012021 (2022)
11. Hou, Y., Zhang, J., Cheng, J., et al.: Measuring and improving the use of graph information in graph neural networks. arXiv preprint [arXiv:2206.13170](https://arxiv.org/abs/2206.13170) (2022)
12. Liu, X., Zhang, J., Pei, Z.: Machine learning for high-entropy alloys: progress, challenges and opportunities. *Prog. Mater. Sci.* **131**, 101018 (2023)
13. Yang, Y., Pedersen, J.O.: A comparative study on feature selection in text categorization (1997)
14. Zhang, Z., Li, S., Li, Z., et al.: A class of multi-label feature selection algorithms based on information entropy. *Comput. Res. Dev.* **50**(6), 1177–1184 (2013)
15. Zhou, H., Wang, X., Zhu, R.: Feature selection based on mutual information with correlation coefficient. *Appl. Intell.* 1–18 (2022)
16. Alsahaf, A., Petkov, N., Shenoy, V., et al.: A framework for feature selection through boosting. *Expert Syst. Appl.* **187**, 115895 (2022)
17. Lewis, D.: Feature selection and feature extraction for text categorization. In: *Human Language Technology. Association for Computational Linguistics* (1992)
18. Rahman, M.A., Hossain, M.F., Hossain, M., et al.: Employing PCA and t-statistical approach for feature extraction and classification of emotion from multichannel EEG signal. *Egypt. Inform. J.* **21**(1), 23–35 (2020)
19. Li, X., Zhao, H., Yu, L., et al.: Feature extraction using parameterized multisynchrosqueezing transform. *IEEE Sens. J.* **22**(14), 14263–14272 (2022)
20. Ahmed, M.Z.I., Sinha, N., Phadikar, S., et al.: Automated feature extraction on AsMap for emotion classification using EEG. *Sensors* **22**(6), 2346 (2022)



# Anomaly Data Identification and Reconstruction in Wind Power Forecasting Based on MLOF-iForest and RF

Pengcheng Du<sup>1</sup> and Yu Du<sup>2</sup>(✉)

<sup>1</sup> School of Electrical Engineering, Guangxi University, Nanning 530004, China  
<sup>2</sup> School of Automation and Electrical Engineering, Linyi University, Linyi 276000, China  
pcdu2023@yeah.net

**Abstract.** This study proposes an anomaly data cleaning method for wind power generation data, addressing issues caused by equipment failures, unstable weather conditions, and data collection module errors. To overcome the limitations of the traditional LOF algorithm—which primarily detects locally sparse outliers but struggles to distinguish anomalies in global datasets where there is no clear distinction between normal and anomalous data—an iForest-LOF weighted fusion algorithm is introduced. This method enhances the identification of outliers in the entire dataset. Moreover, Mahalanobis distance is employed to optimize the LOF algorithm, with the covariance matrix improving the consideration of feature correlations, thereby increasing the accuracy of anomaly detection. After anomaly detection with MLOF-iForest, the identified anomalous data is treated as a missing value imputation problem. The RF algorithm is then applied to explore the implicit relationships between weather features and wind power generation, thus addressing the limitations of traditional KNN algorithms, which struggle to reconstruct large continuous segments of anomalous data by only considering neighboring outliers. Case studies demonstrate that the proposed wind power data recognition and reconstruction method effectively cleans the data, providing a high-quality dataset for subsequent wind power forecasting.

**Keywords:** Wind power prediction · Data Fusion · Machine Learning · Random Forest · Outlier Detection

## 1 Introduction

With the rapid expansion of the wind power market and the growing maturity of big data analytics, the value of wind power operational data has gained significant attention. Research on wind power data has become a key area of interest [1]. By analyzing the measured wind speed and power data from a specific turbine, the turbine's wind speed-power curve can be derived. This curve is essential for evaluating turbine design, site selection, power generation capacity, and estimating energy loss [2]. However, these measured curves often exhibit chaotic distributions and contain numerous outliers [3], primarily due to two factors: 1) faults or interference during data collection, which leads

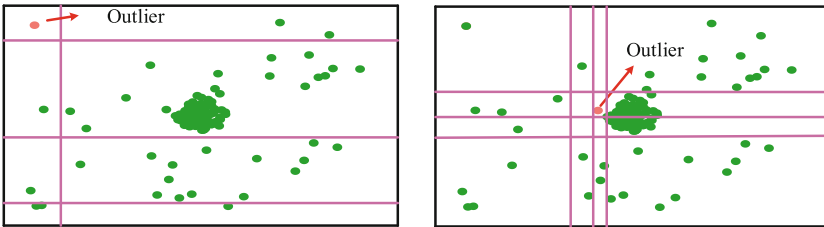
to abnormal or missing data [4]; and 2) data anomalies arising from wind curtailment due to limited wind power acceptance capacity [5]. As a result, the measured wind speed-power curves may not accurately reflect the turbine's true operating conditions [6]. Identifying and addressing anomalies in raw wind speed and power data collected by SCADA systems is therefore crucial.

This paper proposes an enhanced LOF-iForest algorithm for anomaly detection in wind power data, addressing the limitations of the traditional LOF algorithm, which often focuses too much on local anomalies while neglecting global data patterns. By integrating the iForest algorithm, the method better captures deeper structures within the data's global distribution. Given the high volatility of wind power sequences, the method combines the strengths of MLOF and iForest in a weighted ensemble approach, reducing the impact of fluctuations on anomaly detection and improving efficiency for volatile datasets.

Additionally, recognizing the strong correlation between wind power and wind speed, this paper applies the Mahalanobis distance to optimize distance calculations between data points, enhancing feature correlation measurement. This improves the model's ability to detect relationships between features, leading to more accurate distance estimations and better generalization. Finally, anomalies identified by the MLOF-iForest algorithm are treated as missing values and reconstructed using the Random Forest (RF) algorithm, overcoming the limitations of traditional KNN algorithms, which struggle with large, contiguous regions of anomalous data. Case studies demonstrate that the proposed method effectively cleanses wind power data, yielding a high-quality dataset that enhances prediction accuracy.

## 2 Wind Power Anomaly Detection in Wind Farms Based on MLOF-iForest

Compared to algorithms like LOF, which detect wind power anomalies based on density characteristics, the Isolation Forest (iForest) algorithm combines multiple base learners to mitigate the errors of individual learners by compensating for their respective shortcomings, thus enhancing the overall performance of the anomaly detection model. Given a wind power dataset  $X$ , the anomaly detection process using iForest is divided into two main steps:



**Fig. 1.** Normal and outlier data points in 2D space.

**(1) Construct an isolated forest  $F$ ,  $s$  is the number of internal binary trees**

$$F = \{f_1, f_2, \dots, f_i\} \quad (1)$$

where  $f_i$  is a binary tree,  $i = 1, 2, \dots, s$ .

Step 1:  $m$  data points are randomly selected from the dataset  $X = \{x_1, x_2, \dots, x_n\}$  to form a subspace.

$D = \{x_1, x_2, \dots, x_m\}$ , which is then used as the training data for the root node of a binary tree.

Step 2: A data feature  $\alpha$  is randomly selected from subspace  $D$ , and a splitting point  $\tau$  is randomly chosen within the sample range of that feature, where  $\tau$  lies within the interval  $[\alpha_{min}, \alpha_{max}]$ . In the case of a two-dimensional space, the process by which the Isolation Forest algorithm separates normal points from outliers is illustrated in Fig. 1.

Steps 1 and 2 are recursively applied to the child nodes of both branches until the data can no longer be further split or the tree depth reaches the predefined limit, resulting in the construction of a complete Isolation Forest  $F$ .

**(2) The anomaly degree  $\delta$  of the data point  $x$  in dataset  $X$  is calculated and assessed.**

$$\delta(x, m) = 2^{-E[L(x)]/c(m)} \quad (2)$$

where,  $L(x)$  denotes the node depth of the data point  $x$  in the binary tree,  $E[L(x)]$  represents the average path length of  $x$  across  $s$  trees, and  $c(m)$  refers to the average path length of a tree constructed from  $m$  points, as expressed in Eq. (3).

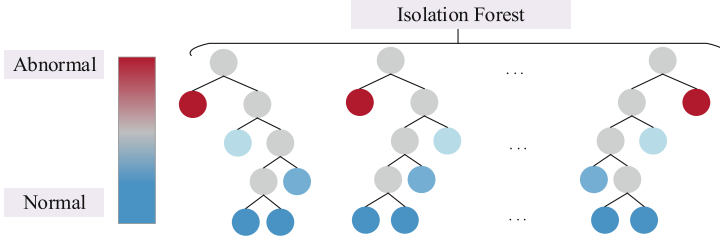
$$c(m) = 2H(m-1) - [2(m-1)/m] \quad (3)$$

$$H(\varphi) = \ln \varphi + A, A \text{ is Euler constant} \quad (4)$$

As illustrated in Fig. 2, during the construction of iForest, most outliers are identified in the early stages of hyperplane partitioning, offering high computational efficiency for high-dimensional datasets. However, iForest constructs isolation trees by randomly selecting features and data points, focusing on the degree of separation in the overall data distribution, which results in poor performance when detecting local anomalies. Additionally, the performance and stability of the iForest algorithm are strongly influenced by the number of isolation trees. With a small number of trees, the overall distribution of the data is inadequately represented, leading to a decrease in anomaly detection accuracy.

Since LOF primarily focuses on identifying anomaly clusters with lower density in local regions of the dataset, it often overlooks the global data distribution characteristics. This study utilizes Mahalanobis distance to optimize the calculation of distances between data points. Given the strong volatility of wind power sequences, and by combining the anomaly scores from both the LOF and iForest algorithms, an improved LOF-iForest-based method for identifying wind power anomalies is proposed. The main steps are as follows:

- 1) Data points that are clearly anomalous are initially filtered using physical rule-based methods. The anomaly score  $S_{iForest}$  is then calculated based on the global distribution of the wind power data using the iForest algorithm.



**Fig. 2.** Diagram of iForest algorithm

The isolation forest model is constructed with 100 binary trees. For each data point in the dataset, the path length within each tree (i.e., the depth from the root node to the data point) is calculated by traversing all trees. The average path length across all trees is then computed. Anomalous data points are identified by comparing their anomaly score to a predefined threshold, with data exceeding the threshold marked as outliers.

- 2) The LOF value  $S_{MLOF}$ , considering Mahalanobis distance, is calculated based on the neighborhood information of wind power data points.

In this study, Mahalanobis distance is used to quantify the relationships between the attributes of the data via covariance distance. Let the Euclidean distance between data points  $x_i$  and  $x_j$  be denoted as  $dist_o$ . The sample distance based on Mahalanobis distance in the LOF algorithm is expressed in Eq. (5).

$$dist(x_i, x_j) = dist_o \times \sqrt{(X - Y)^T S^{-1} (X - Y)} \quad (5)$$

where,  $T$  represents transpose,  $S$  is the covariance matrix,  $X$  and  $Y$  are the feature objects belonging to  $x_j$  and  $x_i$  in the sample respectively.

- 3) The  $S_{iForest}$  values obtained in step 2 are normalized. This normalization ensures that, when fused with the anomaly scores calculated by the iForest algorithm, the LOF values do not dominate the fusion process.

$$S_{MLOF\_nor} = \frac{S_{MLOF}(x_i) - \min(S_{MLOF}(x_i))}{\max(S_{MLOF}(x_i)) - \min(S_{MLOF}(x_i))} \quad (6)$$

- 4) The weighted fusion anomaly score  $S$  is calculated by combining the anomaly scores from iForest with the LOF values, as shown in Eq. (7). The anomaly scores of all data points are then output and sorted in descending order. The proportion of anomalous data points within the entire dataset is set to 0.15, and the final set of anomalies is obtained.

$$S = (1 - \omega) * S_{iForest} + \omega * S_{MLOF} \quad (7)$$

where,  $\omega$  represents the fusion weight,  $\omega$  is set to 0.6. Given the higher randomness of iForest, a greater weight is assigned to the LOF values to more effectively capture the multiple local anomaly clusters in the wind power and wind speed scatter plots.

### 3 Reconstruction of Anomalous Power Data in Wind Farms Based on RF

After identifying anomalous data, the most convenient and efficient approach is typically to remove it. However, excessive removal of anomaly data can lead to the loss of valuable information, thereby diminishing the quality of the dataset. This results in the model's inability to learn important features effectively, reducing its performance during both training and prediction, ultimately affecting prediction accuracy. Therefore, after anomaly detection, missing value reconstruction must be performed to preserve the integrity of the dataset and provide a foundation for subsequent research.

For conventional electrical data missing value reconstruction, statistical methods have typically provided satisfactory results. However, for wind power data, which exhibits strong volatility and nonlinear characteristics, these methods perform poorly. Traditional statistical approaches fail to capture the correlations between wind power and meteorological data, thus missing the opportunity to uncover deep, hidden relationships for accurate imputation. This study, therefore, employs Random Forest (RF), which is capable of handling nonlinear distribution features and considering the interrelationships among different features, to reconstruct anomalous wind power data. Let the wind power dataset be denoted as  $D$ , which includes feature variables such as wind speed, wind direction, temperature, humidity, and pressure, along with the target variable, power. The specific steps for anomaly reconstruction are outlined as follows:

The original dataset is partitioned into two subsets: one containing samples with power values and the other containing samples with missing values. The anomaly dataset is designated as the test set  $D_{\text{test}}$  for the Random Forest model, while the remaining normal data is used as the training set  $D_{\text{train}}$ . The two datasets are represented by Eqs. (8) and (9).

$$D_{\text{train}} = (X_1, Y), (X_2, Y), (X_3, Y), (X_4, Y), (X_5, Y) \quad (8)$$

$$D_{\text{test}} = (X_{\text{miss}_1}, Y), (X_{\text{miss}_2}, Y), (X_{\text{miss}_3}, Y), (X_{\text{miss}_4}, Y), (X_{\text{miss}_5}, Y) \quad (9)$$

where,  $X_1 \sim X_5$  represent the five feature variables: wind speed, wind direction, temperature, humidity, and air pressure, respectively.  $Y$  denotes the set of normal power data samples, and  $X_{\text{miss}_i}$  refers to the  $i$ -th feature of the  $j$ -th sample, where the power value is missing.

2) Use the training set  $D_{\text{train}}$  to build a random forest model  $RF_{\text{model}}$ , the process of which is shown in Fig. 3.

1.  $K$  sample subsets are randomly drawn from the training set with replacement.

$$D_{\text{train}_1}, \dots, D_{\text{train}_k} = \text{BootstrapSampling}(D_{\text{train}}, k) \quad (10)$$

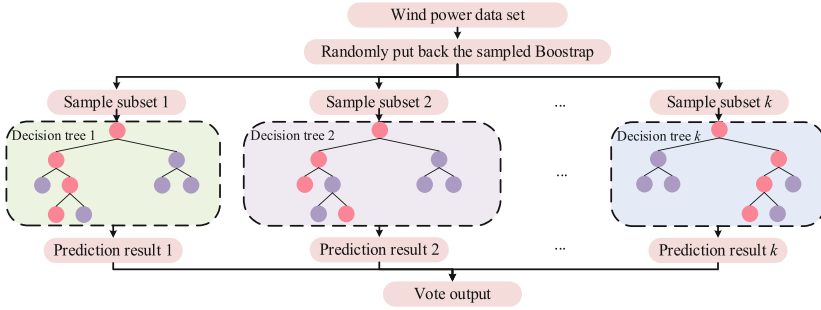
2. Random feature selection:  $\alpha$  features are randomly selected from each training sample set to form the feature set, where  $\alpha$  is smaller than the total number of features,  $A$ , in the original data.

3. Generate multiple decision trees: Repeat the process of generating  $g$  decision trees based on the  $\alpha$  selected features. Due to the random selection of features, each decision tree introduces variability and increases the diversity of the model. Therefore, pruning is not performed on the generated trees.
4. Sample prediction: The generated decision trees are applied to make predictions on the test samples, and all results are recorded.
5. Voting decision: The voting results from all decision trees are counted, and the model with the highest number of votes is selected as the final decision.

3) The trained random forest model is used to predict the power values of the samples in the test set, yielding the fitting results of the random forest and subsequently reconstructing the wind power anomaly data one by one.

$$\hat{Y}_{miss} = RF_{model}(X_{miss}) \quad (11)$$

$$D_{test\_filled} = \{(X_{miss\_1}, \hat{Y}_{miss}), (X_{miss\_2}, \hat{Y}_{miss}) \cdots, (X_{miss\_5}, \hat{Y}_{miss})\} \quad (12)$$

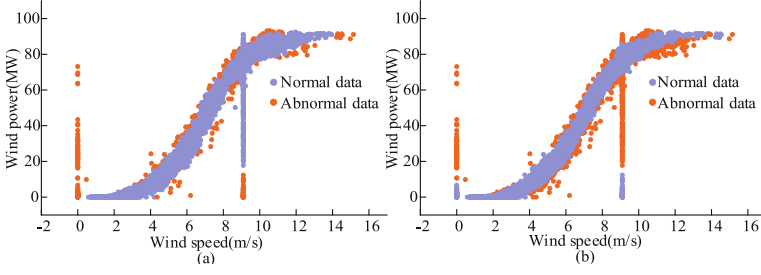


**Fig. 3.** Diagram of random forest

## 4 Case Study

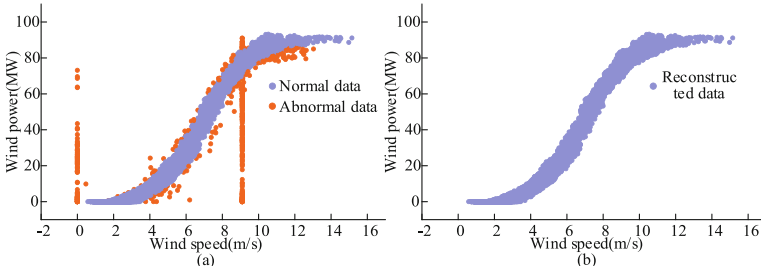
To ensure the accuracy and completeness of the data input into the prediction model, the original dataset undergoes data cleaning. First, the method of combining the Mahalanobis distance-optimized LOF algorithm and the iForest algorithm is used to identify abnormal wind power data. Second, an anomaly dataset is constructed based on the identification results, and the reconstruction of abnormal data is treated as a missing value imputation problem, utilizing Random Forest (RF) to reconstruct the data within a reasonable range. To clearly demonstrate the effectiveness of the data cleaning process, the results of the data preprocessing are comprehensively evaluated from three perspectives: the wind power-wind speed scatter plot, the wind power output line chart, and the wind power prediction error calculation.

Figure 4(a) and 4(b) present the abnormal wind power data recognition results from the RANSAC algorithm and the single LOF algorithm. From the figures, it is evident that



**Fig. 4.** NSAC algorithm/LOF algorithm anomalous data identification results

RANSAC successfully identifies a larger number of discrete data points and wind turbine downtime data points. However, it also mistakenly identifies some wind speed sensor data points as anomalies. This occurs because abnormal wind power data points are often clustered or sparsely distributed locally, and RANSAC is not effective at identifying anomalies that are closely attached to normal data points. The algorithm is primarily designed for identifying linear sequence anomalies and tends to fall into local optima when dealing with nonlinear sequences. In addition, the single LOF algorithm places too much emphasis on local characteristics, making it difficult to identify vertical anomaly clusters, and it mistakenly classifies many normal data points as anomalies.



**Fig. 5.** MLOF-iForest algorithm anomalous data identification results and RF anomaly data reconstruction results

Figure 5(a) shows the results of abnormal wind power data identification using the combined MLOF and iForest algorithms. It can be observed that the proposed method demonstrates high efficiency in identifying various types of abnormal wind power data, including wind speed sensor anomalies, turbine downtime data, and other complex discrete data. Compared to the RANSAC and single LOF algorithms, the proposed algorithm more effectively identifies wind speed sensor anomalies and exhibits better handling of discrete data.

Figure 5(b) illustrates the results after reconstructing the abnormal data identified by the MLOF-iForest algorithm using Random Forest. As shown in Fig. 5(b), the reconstructed data closely approximates the ideal wind power curve. Table 1 shows that the MLOF-iForest algorithm has a lower removal rate than the LOF algorithm, causing less disruption to the original data while achieving better identification performance.

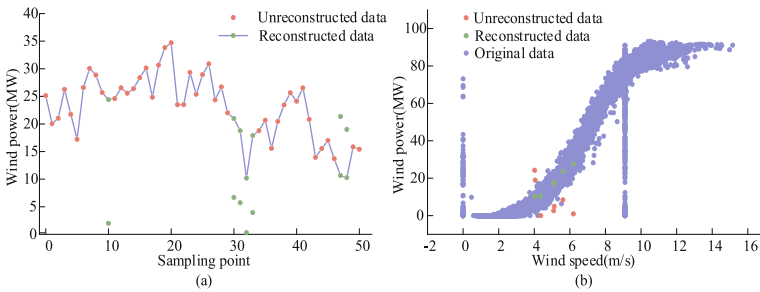


Additionally, the computation time is significantly reduced compared to the RANSAC algorithm, which requires fitting polynomial coefficients.

**Table 1.** The errors of three methods of wind power anomaly data identification

| Abnormal data identification method | Rejection rate (%) | Calculation time (s) |
|-------------------------------------|--------------------|----------------------|
| RANSAC                              | 8.09               | 29.26                |
| LOF                                 | 12.87              | 18.44                |
| MLOF-iForest                        | 10.3               | 21.57                |

Figure 6(a) shows a segment of wind power data before and after anomaly reconstruction. It can be observed that, aside from the wind farm shutdown anomalies (where power is zero), which are reconstructed within a reasonable range, the discrete anomalies caused by issues such as data collection system failures or extreme weather events show reduced fluctuations after reconstruction via the Random Forest (RF) method. The reconstructed curve appears smoother and more stable. Furthermore, Fig. 6(b) presents the reconstructed wind power vs. wind speed 2D scatter plot. As seen, the anomalies are reconstructed within the reasonable fluctuation range of the wind power curve. These results indicate that the combined approach, which integrates anomaly detection based on the MLOF-iForest algorithm and anomaly reconstruction via the RF algorithm, effectively cleanses the wind power data.



**Fig. 6.** Wind power curve after reconstruction of anomalous data and Scatter distribution before and after anomaly data reconstruction.

Finally, to investigate the impact of the proposed anomaly data detection and reconstruction method on wind power forecasting accuracy, training is performed on both the preprocessed and raw wind power datasets, followed by prediction on the original dataset. The prediction errors between the two datasets are then compared. The forecasting models used are commonly employed neural networks for time series data: CNN, LSTM, BiLSTM, and TCN. Table 2 presents the error evaluation metrics for the four forecasting models before and after data preprocessing. The results show a slight improvement in the accuracy of all wind power forecasting models. Specifically, the MAE decreases by up

to 6.34%, the RMSE decreases by up to 6.76%, and the  $R^2$  increases by 0.52%. Based on the analysis of these performance evaluation metrics, the proposed MLOF-iForest-based anomaly data detection method, along with the RF-based anomaly data reconstruction method, effectively improves the quality of the wind power dataset, thereby enhancing the accuracy of wind power forecasting.

**Table 2.** Prediction errors of four models before/after identification and reconstruction of outlier data

|                      | Prediction model | MAE(MW) | RMSE(MW) | $R^2$  |
|----------------------|------------------|---------|----------|--------|
| Pre-data processing  | CNN              | 1.6785  | 2.6483   | 0.9416 |
|                      | LSTM             | 1.5098  | 2.4089   | 0.9528 |
|                      | TCN              | 1.5754  | 2.3927   | 0.9515 |
|                      | BiLSTM           | 1.3796  | 2.1854   | 0.9598 |
| Post-data processing | CNN              | 1.5948  | 2.5899   | 0.9441 |
|                      | LSTM             | 1.4232  | 2.2959   | 0.9578 |
|                      | TCN              | 1.5173  | 2.3572   | 0.9531 |
|                      | BiLSTM           | 1.2921  | 2.0376   | 0.9634 |

## 5 Summary and Prospect

This paper addresses the issue of increased volatility in wind power time series due to outlier data accumulated in wind farm SCADA systems, proposing an anomaly detection method based on MLOF-iForest and an anomaly data reconstruction method using RF. First, the proposed method efficiently handles dense outlier data clusters by leveraging the local density characteristics of LOF, and enhances the calculation of inter-point distances using Mahalanobis distance, which considers the correlation between wind power and wind speed features, thus improving detection accuracy. Next, iForest is applied to identify and filter sparse outliers, creating an anomaly dataset. Finally, given the nonlinear distribution characteristics of wind power data, RF is employed to effectively mine the implicit relationships between power and weather variables in the normal data, completing the reconstruction of the outlier data. Experimental results show that the proposed MLOF-iForest and RF-based anomaly data identification and reconstruction method can effectively reduce the average absolute error of prediction models by approximately 4.99%.

**Acknowledgment.** This work was supported by the Innovation Project of Guangxi Graduate Education (YCBZ2024033).

## References

1. Zhu, A., Zhao, Q., Shi, Z., Yang, T., Zhou, L., Zeng, B.: A novel combined model based on advanced optimization algorithm, and deep learning model for abnormal wind speed identification and reconstruction. *Energy* **312**, 133510 (2024). ISSN 0360-5442
2. Tian, Z., Wang, J.: Variable frequency wind speed trend prediction system based on combined neural network and improved multi-objective optimization algorithm. *Energy* **254**(Part A), 124249 (2022). ISSN 0360-5442
3. Zhao, J., Guo, Y., Xiao, X., Wang, J., Chi, D., Guo, Z.: Multi-step wind speed and power forecasts based on a WRF simulation and an optimized association method. *Appl. Energy* **197**, 183–202 (2017). ISSN 0306-2619
4. Lucheroni, C., Boland, J., Ragno, C.: Scenario generation and probabilistic forecasting analysis of spatio-temporal wind speed series with multivariate autoregressive volatility models. *Appl. Energy* **239**, 1226–1241 (2019). ISSN 0306-2619
5. Li, J., Song, Z., Wang, X., Wang, Y., Jia, Y.: A novel offshore wind farm typhoon wind speed prediction model based on PSO–Bi-LSTM improved by VMD. *Energy* **251**, 123848 (2022). ISSN 0360-5442
6. Pang, J., Li, C.: Context-aware feature reconstruction for class-incremental anomaly detection and localization. *Neural Netw.* **181**, 106788 (2025). ISSN 0893-6080



# Broad Learning System-Enhanced Fuzzy Inference Framework for High-Efficiency Classification

Yizhen Wang<sup>1,2</sup>, Zunwei Fu<sup>3</sup>, and Congcong Zhang<sup>1,2</sup>(✉)

<sup>1</sup> School of Mathematics and Statistics, Linyi University, Linyi 276005, China  
zhangcc@lyu.edu.cn

<sup>2</sup> Shandong Data Open Innovative Application Laboratory, Linyi 276005, China

<sup>3</sup> Department of Computer Science, The University of Suwon, Hwaseong-si,  
Gyeonggi-do 18323, South Korea

**Abstract.** In this paper, a novel Takagi-Sugeno-Kang fuzzy system is implemented through the introduction of a Broad Learning System. This system not only enhances the accuracy and interpretability of neuro-fuzzy models but also retains the fast computation characteristics of the Broad Learning System. Initially, the traditional TSK fuzzy system is used to partition the input evenly, resulting in fuzzy sets with appropriate labels. In the consequent part of the fuzzy system's If-Then rules, the Broad Learning System is used in place of the traditional first-order polynomial. The pseudoinverse is employed to solve the optimal weights for the output layer of the BLS, ensuring rapid and efficient computation without the need for iterative optimization. The proposed BLS-based fuzzy inference system was tested on standard benchmark datasets and its performance was compared against various advanced non-fuzzy and neuro-fuzzy approaches. The results demonstrate that the proposed model outperforms other related models, effectively handling high-dimensional data while maintaining the interpretability of fuzzy rules, thus avoiding the computational complexity associated with rule explosion, and showing promising potential for real-world applications.

**Keywords:** Takagi-Sugeno-Kang Fuzzy System · Broad Learning System · Feature Layer Integration · Retrograde Perspective

## 1 Introduction

Fuzzy inference systems have achieved significant advancements in rule extraction tasks and are widely used to address reasoning problems involving ambiguous and uncertain information. These systems are extensively applied in prediction, control, decision-making, and diagnosis [1–3], establishing themselves as a crucial research area in AI-driven prediction. Compared to traditional exact logic inference systems, fuzzy inference systems are better suited for tackling complex real-world problems, enabling the extraction of more information from data.

Among various fuzzy inference methods, the TSK fuzzy system stands out due to its strong interpretability and low computational complexity [4]. It has found widespread applications in fields like speech processing, wireless system control, time series prediction, regression, and classification. However, when handling multiple inputs, TSK fuzzy systems often face the challenge of a large number of rules, leading to the well-known ‘rule explosion phenomenon’.

In contrast, the Broad Learning System proposed by Chen and Liu [5] demonstrates robust learning capabilities and efficient training by expanding feature nodes and augmentation nodes extensively. Broad Learning System employs pseudo-inverse operations for rapid model training and relies on a single-layer planar network, thus avoiding the complexity of deep network cascading structures. However, due to its large number of nodes and weights, BLS can be difficult to interpret, making it challenging to understand the specific role of each node and its contribution to the final output.

To combine the interpretability of TSK fuzzy systems with the computational efficiency and high-dimensional data processing capability of the Broad Learning System, we propose a novel approach named the Broad Learning System-Enhanced Fuzzy Inference Learning System. This model combines the strong learning and fast computation of BLS with the interpretability of fuzzy rules, creating a streamlined network structure that enhances predictive performance. The proposed approach not only effectively manages high-dimensional data but also maintains the clarity of fuzzy rules, thereby mitigating the computational complexity associated with rule explosion.

## 2 Architecture of the Broad Learning System-Enhanced Fuzzy Inference Framework

The system design depicted in Fig. 1 follows the fundamental architecture of fuzzy systems. In the proposed model, input partitioning is conducted using the TSK fuzzy system, where the inputs are initially transformed into fuzzy membership values. These fuzzy memberships form the input matrix of the model, representing the degree of belongingness of each input to the defined fuzzy sets. The fuzzy rule base in the model retains the classic IF-THEN structure, with fuzzy implications in the antecedent part.

### 2.1 Fuzzy Membership Mapping

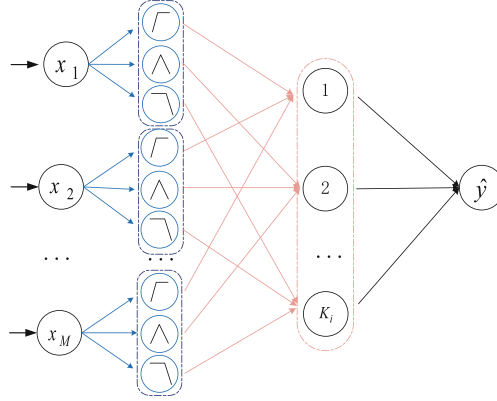
Suppose the system input vector is  $x = (x_1, x_2, \dots, x_M)$ . The  $k$ -th fuzzy rule in the TSK fuzzy system can be expressed as:

$$R_k : \text{If } x_1 \text{ is } A_1^k \text{ and } x_2 \text{ is } A_2^k \text{ and } \dots \text{ and } x_M \text{ is } A_M^k, \text{ Then } y = f_k(x) \quad (1)$$

where  $A_i^k$  represents the fuzzy set corresponding to input variable  $x_i$  in the  $k$ -th rule,  $f_k(x)$  is the polynomial function defining the output for the  $k$ -th rule.

The calculation formula of fuzzy membership is:

$$\mu_{A_i^k}(x_i) = \begin{cases} 0, & \text{if } x_i \leq a_i^k \text{ or } x_i \geq c_i^k \\ \frac{x_i - a_i^k}{b_i^k - a_i^k}, & \text{if } a_i^k < x_i \leq b_i^k \\ \frac{c_i^k - x_i}{c_i^k - b_i^k}, & \text{if } b_i^k < x_i < c_i^k \end{cases} \quad (2)$$



**Fig. 1.** Architecture of proposed model

where  $a_i^k, b_i^k, c_i^k$  are the left endpoint, middle point and right endpoint of the triangular membership function in the  $k$ -th rule respectively.  $\mu_{A_i^k}(x_i)$  is the membership degree of the input  $x_i$  to the corresponding fuzzy set  $A_i^k$ .

By organizing the membership degrees of all input variables into a matrix, we obtain the membership matrix  $\mu$ , which serves as the input to the feature layer of the BLS:

$$\mu = [\mu_{A_1}(x_1) \ \mu_{A_2}(x_2) \ \dots \ \mu_{A_M}(x_M)]. \quad (3)$$

## 2.2 Feature Layer Integration

In the feature layer of the width learning system, the input matrix is the membership matrix  $\mu$ . The output of the feature layer is calculated by performing a linear transformation on the membership matrix:

$$Z = \mu \cdot W_f \quad (4)$$

where  $W_f = \text{rand}(d, n_f) * 2 - 1$  is the weight matrix of the feature layer, Obtained by random initialization.

## 2.3 Pseudo-inverse Weight Optimization

The output of the feature layer is used to compute the final weight matrix using the pseudo-inverse method, which allows for efficient training and optimal fitting of the target outputs:

$$\beta = (\Theta^T \cdot \Theta + I \cdot C)^{-1} \cdot \Theta^T \cdot Y \quad (5)$$

where  $C$  is the regularization parameter used to prevent overfitting.

The intermediate output is obtained by linear combination of feature layer output and pseudo-inverse weights:

$$\hat{y}_{BLS} = \Theta \cdot \beta. \quad (6)$$

The output of each fuzzy rule is obtained by multiplying the intermediate output and the rule activation:

$$f_k(x) = \hat{y}_{\text{BLS}} \cdot \mu_k(x). \quad (7)$$

The final output is the weighted average of all fuzzy rule outputs, i.e. the defuzzification result:

$$\hat{y} = \sum_{k=1}^K \tilde{\mu}_k(x) f_k(x), \quad (8)$$

where  $\tilde{\mu}_k(x)$  denotes the firing strength of the  $k$ th fuzzy rule with respect to input  $x$  by

$$\begin{aligned} \tilde{\mu}_k(x) &= \frac{\mu_k(x)}{\sum_{k=1}^K \mu_k(x)}, \\ \mu_k(x) &= \prod_{i=1}^M \mu_{A_i^k}(x_i). \end{aligned} \quad (9)$$

### 3 Experimental Studies

This section presents a comprehensive comparison of the proposed model against a range of representative non-fuzzy and fuzzy rule-based models, based on experiments conducted on benchmark datasets for classification tasks. Specifically, we selected the following methods for comparison:

- 1) Support Vector Machine (SVM), utilizing a Gaussian kernel;
- 2) Naïve Bayes classifier;
- 3) C4.5 decision tree;
- 4) Random Forest;
- 5) TSK Fuzzy System (FS-FCSVM).

The experimental results for these models are directly referenced from the studies in [6] and [7] to ensure a fair comparison.

#### 3.1 Experimental Setting

To thoroughly assess the performance of the proposed network, 13 benchmark datasets from machine learning were selected, obtained from the KEEL and UCI Machine Learning Repositories. These datasets include diverse variable types such as integers, real numbers, and categorical variables, as shown in Table 1. During data preprocessing, each feature was standardized using the mean and standard deviation from the training set to enhance model training effectiveness and stability.

Table 2 summarizes the characteristics of these algorithms, including model structure, activation functions, and learning rate strategies. We configured representative

**Table 1.** Properties of classification machine learning datasets.

| Datasets | No. of features | No. of samples | No. of classes |
|----------|-----------------|----------------|----------------|
| Abalone  | 8               | 4174           | 3              |
| Balance  | 4               | 625            | 3              |
| Bupa     | 6               | 345            | 2              |
| German   | 24              | 1000           | 2              |
| Haberman | 5               | 132            | 3              |
| Heart    | 13              | 270            | 2              |
| Iris     | 4               | 150            | 3              |
| Pima     | 8               | 768            | 2              |
| Saheart  | 10              | 462            | 2              |
| Seeds    | 7               | 210            | 3              |
| Segement | 18              | 2310           | 7              |
| Vehicle  | 18              | 846            | 4              |
| Zoo      | 16              | 101            | 7              |

**Table 2.** Classifier parameter settings for performance comparison.

| Algorithm     | Parameter settings                                                                                                                                                                                                                                                                                                          |
|---------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Naïve Bayes   | Default hyperparameter                                                                                                                                                                                                                                                                                                      |
| SVM           | Regularization parameter $C \in \{10^{-3}, \dots, 10^0, \dots, 10^3\}$ ,<br>Gaussian kernel width parameter $\sigma \in \{10^{-3}, \dots, 10^0, \dots, 10^3\}$                                                                                                                                                              |
| C4.5          | Confidence = 0:20, Sample size = 2, Threshold = 10                                                                                                                                                                                                                                                                          |
| Random Forest | Batchsize = 100,<br>MaxDepth = 5–10,<br>Minimum sample leaf = 10–20,<br>Minimum sample split = 5–10,<br>Bootstrap = True,<br>obb_score = True                                                                                                                                                                               |
| FS-FCSVM      | Number of Clusters $c = 2-10$<br>Fuzzification Parameter $m = 2$<br>Maximum Number of Iterations = 100<br>Convergence Threshold = $1 \times 10^{-5}$<br>Regularization Parameter $\lambda \in \{10^{-1}, 10^{-2}, 10^{-3}, 10^{-4}\}$<br>Gaussian kernel width parameter $\sigma \in \{10^{-3}, \dots, 10^0, \dots, 10^3\}$ |



model parameters through a rigorous experimental iteration process, involving precise learning rate adjustments to ensure convergence stability and efficiency, aiming to optimize performance across various environments.

To ensure a thorough evaluation, we applied five-fold cross-validation by randomly dividing each dataset into five equal parts, with each subset representing 20% of the data. Four subsets were used for training, and the remaining one was reserved for testing. This approach effectively assesses the model's generalization ability and reduces bias introduced by data partitioning.

For performance evaluation, we report the classification accuracy (CA) in terms of its mean and standard deviation, along with the average number of fuzzy rules, to ensure result reliability and consistency. The formula for CA is as follows:

$$CA(\%) = \frac{1}{N} \left[ \sum_{k=1}^N \Gamma(y_k, \hat{y}_i) \right] \times 100, \quad \Gamma = \begin{cases} 1, & \text{if } y_k = \hat{y}_i \\ 0, & \text{otherwise} \end{cases} \quad (10)$$

where  $y_k$  denotes the actual class label, and  $\hat{y}_i$  represents the class label predicted by the proposed classifier.

### 3.2 Machine Learning Datasets

To conduct a comprehensive comparative analysis on benchmark datasets, we compared the proposed model with the classical TSK model and the FS-FCSVM model, which is based on fuzzy support vector machines. Table 3 presents the average classification rates of TSK, FS-FCSVM, and FC-BLS across 5 datasets. The experimental results show that FC-BLS achieves a higher average classification rate compared to both TSK and FS-FCSVM on all datasets.

**Table 3.** Results comparison of fuzzy rule-based models.

| Datasets | TSK   |      | FS-FCSVM [7] |      | FC-BLS       |      |
|----------|-------|------|--------------|------|--------------|------|
|          | Mean  | STD  | Mean         | STD  | Mean         | STD  |
| Balance  | 78.09 | 1.21 | 90.56        | 0.52 | <b>91.20</b> | 0.03 |
| Haberman | 52.20 | 0.76 | 73.77        | 0.01 | <b>74.56</b> | 0.56 |
| Iris     | 84.46 | 0.47 | 95.22        | 0.13 | <b>96.67</b> | 0.08 |
| Segment  | 59.32 | 1.69 | 66.59        | 0.31 | <b>94.20</b> | 1.54 |
| Vehicle  | 52.48 | 0.54 | 58.92        | 0.08 | <b>79.56</b> | 1.22 |

Table 4 provides a comparison between the proposed method and traditional classifiers, presenting results as mean classification rates with standard deviations in parentheses. FC-BLS outperforms the other four classifiers on most datasets, exhibiting higher accuracy. Specifically, it achieves a 10–20% accuracy improvement over Naïve Bayes on datasets like Abalone, Bupa, and Heart. The lower standard deviation of FC-BLS underscores its stability and consistent performance, confirming both higher accuracy and greater reliability compared to traditional models.

**Table 4.** Results comparison of non-fuzzy rule-based models.

| Datasets | SVM [8]         | Naïve Bayes [8] | C4.5 [8]        | Random Forest [8]      | FC-BLS                 |
|----------|-----------------|-----------------|-----------------|------------------------|------------------------|
| Abalone  | 63.82<br>(0.42) | 57.72<br>(0.47) | 60.33<br>(0.45) | 64.06<br>(0.39)        | <b>66.92</b><br>(0.32) |
| Balance  | 79.38<br>(0.18) | 78.56<br>(0.21) | 81.03<br>(0.02) | 81.63<br>(0.03)        | <b>91.20</b><br>(0.03) |
| Bupa     | 68.03<br>(0.03) | 63.99<br>(0.03) | 62.05<br>(0.05) | 63.55<br>(0.27)        | <b>73.91</b><br>(0.14) |
| German   | 71.54<br>(0.51) | 66.85<br>(0.68) | 67.70<br>(0.14) | 76.72<br>(0.16)        | <b>76.90</b><br>(0.27) |
| Haberman | 73.58<br>(0.63) | 64.82<br>(0.02) | 72.72<br>(0.11) | 74.26<br>(0.39)        | <b>74.56</b><br>(0.56) |
| Heart    | 53.32<br>(0.38) | 61.96<br>(2.15) | 55.26<br>(0.03) | 63.25<br>(0.12)        | <b>84.44</b><br>(0.04) |
| Iris     | 96.70<br>(0.03) | 95.17<br>(0.02) | 96.70<br>(0.02) | <b>98.00</b><br>(0.02) | 96.67<br>(0.36)        |
| Pima     | 74.62<br>(0.28) | 76.18<br>(0.96) | 71.23<br>(0.31) | 66.41<br>(1.81)        | <b>77.61</b><br>(0.65) |
| Saheart  | 70.08<br>(0.45) | 68.92<br>(0.03) | 72.65<br>(0.02) | 66.18<br>(0.61)        | <b>74.00</b><br>(0.19) |
| Seeds    | 90.82<br>(0.04) | 86.51<br>(0.13) | 93.37<br>(0.02) | 92.46<br>(0.98)        | <b>95.71</b><br>(0.33) |
| Segement | 92.42<br>(0.04) | 93.83<br>(1.02) | 92.30<br>(0.12) | 92.38<br>(0.83)        | <b>94.20</b><br>(1.92) |
| Vehicle  | 81.20<br>(0.01) | 79.05<br>(0.24) | 70.83<br>(0.72) | <b>83.96</b><br>(0.20) | 79.56<br>(1.22)        |
| Zoo      | 91.53<br>(0.04) | 94.14<br>(1.02) | 93.45<br>(0.12) | 72.38<br>(0.54)        | <b>97.05</b><br>(0.30) |

## 4 Conclusion

This paper introduces the Broad Learning System-Enhanced Fuzzy Inference Framework, a novel approach designed to address classification challenges in TSK fuzzy systems. By integrating the Broad Learning System in place of the traditional first-order polynomial and utilizing the pseudoinverse for optimal coefficient computation, FC-BLS achieves both computational efficiency and reduced structural complexity. Experimental results demonstrate that FC-BLS consistently outperforms existing fuzzy and neuro-fuzzy models, particularly in managing high-dimensional data while preserving interpretability and avoiding rule explosion. These advantages make FC-BLS a compelling choice for real-world applications requiring a balance of accuracy and simplicity. Future work may explore enhancing model efficiency, scalability, and extending its application to more complex and diverse datasets.

## References

1. Zhao, J., Lin, C.M.: Wavelet-TSK-type fuzzy cerebellar model neural network for uncertain nonlinear systems. *IEEE Trans. Fuzzy Syst.* **27**(3), 549–558 (2018)
2. Huang, H., Yang, C., Chen, C.L.P.: Optimal robot-environment interaction under broad fuzzy neural adaptive control. *IEEE Trans. Cybern.* **51**(7), 3824–3835 (2021)
3. Lian, Z., Shi, P., Lim, C.-C., Yuan, X.: Fuzzy-model-based lateral control for networked autonomous vehicle systems under hybrid cyber-attacks. *IEEE Trans. Cybern.* **53**(4), 2600–2609 (2023)
4. Chang, P.-C., Liu, C.-H.: A TSK type fuzzy rule-based system for stock price prediction. *Expert Syst. Appl.* **34**(1), 135–144 (2008)
5. Chen, C.L.P., Liu, Z.: Broad learning system: an effective and efficient incremental learning system without the need for deep architecture. *IEEE trans. Neural Netw. Learn. Syst.* **29**(1), 10–24 (2017)
6. Fazzolari, M., Alcal, R., Herrera, F.: A multi-objective evolutionary method for learning granularities based on fuzzy discretization to improve the accuracy-complexity trade-off of fuzzy rule-based classification systems: D-MOFARC algorithm. *Appl. Soft Comput.* **24**, 470–481 (2014)
7. Zhang, Y., Ishibuchi, H., Wang, S.: Deep Takagi-Sugeno-Kang fuzzy classifier with shared linguistic fuzzy rules. *IEEE Trans. Fuzzy Syst.* **26**(3), 1535–1549 (2017)
8. Frank, E., Hall, M.A., Witten, I.H.: *Data Mining: Practical Machine Learning Tools and Techniques*, 4th edn. Morgan Kaufmann, San Mateo (2016)



# Design of Radial Basis Function Neural Network with Neural Gas Network Clustering for Classification of Logistics Black Plastics

Hongliang Yu, Kun Zhou<sup>(✉)</sup>, Ming Guo, Xiangyong Chen, and Jianlong Qiu

School of Automation and Electrical Engineering, Linyi University,  
Linyi 27600, China  
zhoukun@lyu.edu.cn

**Abstract.** With the increasing dependence of the logistics industry on plastics, the amount of plastic waste has gradually risen, leading to environmental problems and resource wastage. Due to the high absorption of visible light and near-infrared wavelengths by black plastics, most common optical technologies struggle to effectively sort black plastics, and traditional manual classification methods are inefficient and inaccurate. To address this issue, we propose an efficient algorithm for the identification of black plastics. First, Raman spectroscopy is used to analyze black plastic and obtain relevant data. Then, principal component analysis (PCA) is applied to reduce the dimensionality of the spectral data and extract effective features. Next, a classification algorithm combining Radial Basis Function Neural Network (RBFNN) and Neural Gas Network (NGN) clustering is designed. This algorithm integrates the adaptive learning ability and pattern recognition capability of neural networks, enabling efficient classification of black plastics. Experiments were conducted using a Raman dataset, and the proposed algorithm was compared with three methods: RBFNN based on MinMax, Fuzzy C-Means (FCM), and Hard C-Means Clustering (HCM). Additionally, experiments were also performed on the publicly available datasets of Glass, Abalone, MyHog, and Page. The experimental results show that the proposed algorithm demonstrates superior performance across all datasets.

**Keywords:** PCA · Radial Basis Function Neural Network · Raman Spectroscopy · Logistics Black Plastic · Neural Gas Network Clustering

## 1 Introduction

The widespread use of plastic materials in various sectors of modern society has led to a continuous increase in demand, while also causing the accumulation of plastic waste, which in turn results in environmental pollution. Due to its excellent applicability, plastic is widely used in industries such as packaging, construction, automotive, electronics, and healthcare, driving the rise in plastic

demand [1]. In the process of plastic waste recycling, classification based on resin type is a critical step [2]. However, most plastic waste classification worldwide still relies on manual labor, which faces issues such as slow processing speeds, high labor intensity, and low efficiency [3]. To address these problems, various new technologies have been developed and introduced.

In 1995, Scott proposed a recognition method based on Near-Infrared Radiation (NIR) technology to distinguish between Polyethylene Terephthalate (PET) and Polyvinyl Chloride (PVC) resins [4]. In 2001, Edward and Sommer introduced a classification method that utilized NIR transmission spectra to categorize plastic resins into PET and PVC [5]. In recent years, several new methods have been developed for the classification of post-consumer black plastic waste. Rozenstein et al. (2017) demonstrated that Mid-Infrared (MIR) technology can be used to identify a variety of black plastics, such as PET, PP, PS, PLA, and PE [6]. Compared to the Near-Infrared (NIR) region, the MIR region is less affected by the surface morphology and color (including black) of plastics, making it more reliable for plastic classification [7–10]. However, the effectiveness of MIR technology still requires further research [8]. Shameem et al. (2017) showcased a machine learning approach that combines Laser-Induced Breakdown Spectroscopy (LIBS) and Raman Spectroscopy to identify different plastics (such as PET, PE, PP, and PS) in black complex mixtures [9]. Küter et al. (2018) introduced a method for black plastic identification based on Terahertz technology and machine learning, which is capable of detecting subtle differences in black plastics [10]. Although these technologies show potential, they are still in the developmental stage and have relatively high costs [11]. In comparison, methods combining Attenuated Total Reflectance Fourier Transform Infrared Spectroscopy (ATR-FTIR) and deep learning for plastic identification have achieved a high recognition rate [12]. Kun Zhou et al. (2022) developed a black plastic classification method based on laser spectroscopy and deep learning, employing 1D and 2D Convolutional Neural Network (CNN) classifiers. They further enhanced classification accuracy through peak selection and data reconstruction techniques [13]. Additionally, Kun Zhou et al. (2022) introduced a novel two-stream fuzzy neural network (TSFNN) framework, which integrates fuzzy logic with transfer learning. By leveraging the strengths of fuzzy rule-based radial basis function neural networks (FRBFNNs) and CNNs, this approach achieved notable improvements in classification performance [14]. In a related study, Attilio Sbrana et al. (2023) tackled the challenge of classifying low-reflectivity black plastics using Near-Infrared Hyperspectral Imaging combined with deep learning techniques [15]. G. Bonifazi et al. (2024) proposed and validated an innovative method that merges Laser-Induced Fluorescence (LIF) with machine learning to identify black plastics in packaging waste [16].

This paper classifies black plastic based on spectral data obtained using a Raman spectrometer. To improve classification accuracy, a method called NGN-RBFNN that combining Radial Basis Function Neural Network (RBFNN) and Neural Gas Network (NGN) clustering is proposed. The main contributions are as follows:

- Raman data contains a large amount of redundant information. Therefore, Principal Component Analysis (PCA) is applied for dimensionality reduction to extract representative and effective feature.
- RBFNN is used for the identification of black plastics (such as PET, PP, PS, etc.). RBFNN demonstrates excellent performance in plastic identification due to its powerful nonlinear mapping ability and local response characteristics. With a simple structure and minimal parameter adjustment, RBFNN offers good adaptability and can effectively handle the recognition tasks of various types of plastics.
- The combination of NGN clustering and RBFNN enhances the generalization ability and robustness of the network. The self-organizing learning of the NGN provides a good initialization for the RBFNN, helping it better capture the intrinsic structure of the data and improving classification performance.

## 2 Preprocessing Algorithm and Classification Algorithm Structure

### 2.1 PCA

PCA is a commonly used dimensionality reduction technique aimed at reducing the dimensionality of large datasets. In this study, since the spectral data collected by the Raman spectrometer has high dimensionality and contains redundant information, the PCA method is applied for dimensionality reduction. This allows for the extraction of the main feature components of the data while preserving as much variance as possible from the original data. The following is a brief overview of the application of PCA [17, 18]:

1. First, compute the mean of the dataset. Let the training set be  $X_1, X_2, \dots, X_m$ .

$$\Psi = \frac{1}{M} \sum_{n=1}^M X_n \quad (1)$$

2. Next, subtract the computed mean from the data points to center the data.

$$\Phi_i = X_i - \Psi \quad (2)$$

3. Then, construct the covariance matrix of the data, which helps in understanding the correlations between different variables.

$$C = \frac{1}{M} \sum_{n=1}^M \Phi_n \Phi_n^T = A \cdot A^T \quad (3)$$

4. Compute the eigenvectors and eigenvalues of the covariance matrix, which are key components of PCA.
5. Select the most important principal components based on the eigenvalues and form the eigenvector matrix. These principal components best represent the original data.

$$W = [u_1, u_2, \dots, u_m] \quad (4)$$

6. Finally, use these eigenvectors to construct a new dataset that retains most of the information while effectively reducing the dimensionality.

$$Y = W^T(X_{new} - \Psi) \quad (5)$$

## 2.2 Neural Gas Network

The approach introduced by Martinetz et al. [19], known as the Neural Gas Algorithm, draws an analogy between the motion of cluster centers in the data space and the Brownian motion of gas molecules within a confined space.

The Neural Gas Algorithm can be outlined in the following steps:

- (1) Network Initialization: Begin by selecting  $N$ , the number of clusters, given an input dataset  $X$  with dimensions  $a \times b$ , where  $a$  represents the total number of data points and  $b$  denotes the number of features. Randomly initialize the weight vectors  $\omega = \{\omega_1, \omega_2, \omega_N\}$ , each with a dimension of  $1 \times n$  in the input space. Define the network parameters:  $t_{\max}$  (the maximum number of iterations),  $\eta_{initial}$  and  $\eta_{final}$  (the starting and ending rates of learning, typically set to 0.3 and 0.003, respectively), and  $\lambda_{initial}$  and  $\lambda_{final}$  (the initial and final decay constants, such as 5 and 0.05).
- (2) Input Vector Selection: A randomly chosen input vector  $x_j$  (of size  $1 \times n$ ) is introduced to the network.
- (3) Distance Calculation: The Euclidean distance between the input vector  $x_j$  and the weight vectors of each neuron is computed.
- (4) Neighborhood Ranking: The weight vectors  $w_i$  are sorted based on their distance to the input vector  $x_j$ . The rank  $k_i$  of each weight  $w_i$  is determined according to its position in the sorted list, with  $k_i$  taking values from 1 to  $N$ .
- (5) Weight Update: the following rule is applied to update the neuronal weights:

$$\omega_i(t+1) = \omega_i(t) + \eta(t)h_\lambda(k_i)(x_j - \omega_i) \quad (6)$$

where  $t$  is the training time;  $\eta(t) = \eta_{initial}(\eta_{final}/\eta_{initial})^{t/t_{\max}}$ ,  $\eta \in [0, 1]$  is the learning rate;  $h_\lambda(k_i) = e^{-(k_i/\lambda)}$  is the neighborhood function; and  $\lambda(t) = \lambda_{initial}(\lambda_{final}/\lambda_{initial})_{\max}$ ,  $\lambda \in [0, \infty]$  is the decay constant.

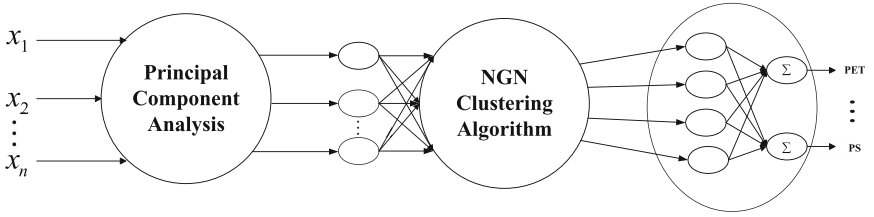
- (6) Iteration: Steps 2 through 5 should be repeated until the maximum number of iterations ( $t_{\max}$ ) is reached.
- (7) Assignment of Objects: At the end of each iteration, assign each data point to the neuron with the weight vector closest to it.

For effective self-organization, the training should commence with a high initial learning rate, which is progressively reduced over the course of each adaptation. The corresponding weight vectors are adjusted to move closer to the new input vector as it is introduced to the network (Step 5). The adjustments are made based on the ranking of the weight vectors or nodes, determined by their distances from the input vector. Initially, the node positions may be disordered,

but after multiple iterations, the weights will gradually distribute evenly across the input space. The Neural Gas method is based on a “competitive learning” mechanism, employing a “winner-takes-most” strategy. After several iterations, the node positions will no longer depend on the initial setup, resembling the diffusion process of gas molecules. As the training progresses, the nodes gradually stabilize, ultimately forming the desired topological structure.

### 2.3 Radial Basis Function Neural Network Integrated with Neural Gas Clustering

The structure of the proposed NGN-RBFNN is shown in Fig. 1. NGN-RBFNN is used as a classifier, where the Neural Gas Network clustering is employed to analyze the patterns of the input data and determine the center positions of the radial basis functions in the RBFNN.



**Fig. 1.** Structure of NGN-RBFNN

Typically, Gaussian-type RBFs are used as receptive fields.

$$\varphi_i(x) = \exp\left(-\frac{\|x - v_i\|^2}{2\sigma_i^2}\right) \quad (7)$$

The topology of the RBFNN is shown in Fig. 2.

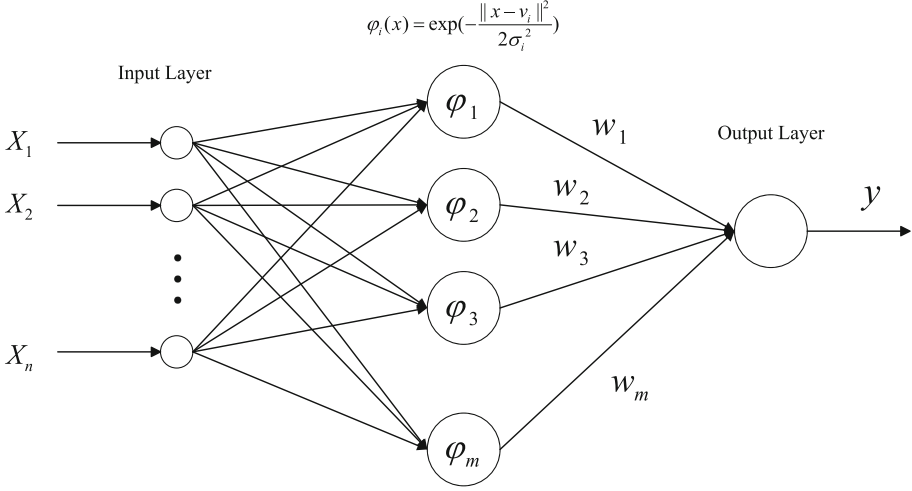
In this case, the activation level of the  $i$ th node in the hidden layer is represented by  $\varphi_i(x)$ . Similarly,  $\sigma_i$  denotes the expansion of the  $i$ th receptive field, and  $v_i$  represents the center of the  $i$ th basis function (receptive field). The proposed NGN-RBFNN differs from the traditional RBFNN in two significant ways. First, Neural Gas Network (NGN) clustering is used to determine the receptive field prototypes in the NGN-RBFNN. Neural gas networks can effectively reveal the structure within data and perform unsupervised clustering, which in turn enhances the performance of RBFNN in classification tasks. Second, in the NGN-RBFNN network structure, the type of connection (i.e., weight) between the hidden layer and the output layer is chosen to be a linear function or a quadratic polynomial, rather than being restricted to certain specific numerical ranges. This design makes the network more flexible and adaptable.

## 3 Experimental Study

### 3.1 Datasets

In this section, the datasets used in the experiments are introduced as follow:



**Fig. 2.** Structure of RBFNN

**Raman:** Raman spectroscopy obtains molecular vibration information by measuring the frequency shift of scattered light, which is used to reveal the structure and composition of substances. The Raman dataset used in this study was collected by a laboratory Raman spectrometer and contains 300 samples, with 3001 spectral features per sample. The dataset is divided into three categories: PET, PP and PS.

**Glass:** The Glass dataset is used to determine the type of glass based on its chemical composition, which helps in identifying its potential applications, such as in construction or vehicle windows. The dataset contains 214 records, with 10 key features per record, involving six categories.

**Abalone:** The Abalone Age dataset is used to predict the lifespan of abalones, with a focus on analyzing their annular markings. The dataset contains 8 features, such as sex, diameter, and weight. Data preprocessing includes encoding gender and applying standardization or normalization techniques.

**MyHog:** The MyHog dataset is based on the Yale Face Database A and is designed for face recognition. Features are extracted through gradient enhancement and refined edge texture extraction. The dataset consists of 1500 records, each with 257 features, forming a three-class classification problem.

**Page:** The Page Blocks Classification dataset, containing 5,473 instances, is used for page layout analysis and document classification. Each instance is defined by 10 numerical features, such as area, aspect ratio, and black pixel percentage, enabling classification into five categories: text, image, table, title, and footer. It serves as a benchmark for evaluating machine learning algorithms in document layout analysis.

3.2 Experimental Results

To validate the effectiveness of the algorithm, we selects the RBFNN based on MinMax (MM-RBFNN), RBFNN based on HCM (HRBFNN), and RBFNN based on FCM (FRBFNN) algorithms for comparative experiments, and uses a five-fold cross-validation strategy. In the experiment, accuracy and standard deviation are used as evaluation metrics. The experimental results on the Raman dataset are shown in Table 1.

Table 1. Performance of Different Algorithms on the Raman Dataset

| Method             | TR            | STD   | TE            | STD   |
|--------------------|---------------|-------|---------------|-------|
| MM-RBFNN           | 96.666        | 0.883 | 93.666        | 2.173 |
| HRBFNN             | 96.500        | 1.200 | 94.666        | 1.394 |
| FRBFNN             | 96.000        | 0.912 | 94.666        | 0.745 |
| Proposed NGN-RBFNN | <b>96.666</b> | 1.250 | <b>95.000</b> | 1.178 |

The comparison results with MM-RBFNN, HRBFNN, and FRBFNN methods indicate that the proposed NGN-RBFNN algorithm performs better on the Raman dataset. As shown in Table 1, the average recognition rate (TR) of the NGN-RBFNN algorithm on the training dataset is 96.67%, demonstrating strong recognition capability. Additionally, its standard deviation (1.25) is relatively small. For the test dataset, the NGN-RBFNN algorithm achieves an average recognition rate (TE) of 95%, which shows that the algorithm not only performs excellently on the training dataset but also maintains good generalization ability on the testing dataset.

To verify the performance of the algorithm on other datasets, experiments were also conducted on the Glass, Abalone, and MyHog datasets, with the experimental results shown in Table 2, Table 3, and Table 4, respectively.

Table 2. Performance of Different Algorithms on the Glass Dataset

| Method             | TR     | STD   | TE            | STD          |
|--------------------|--------|-------|---------------|--------------|
| MM-RBFNN           | 83.179 | 2.368 | 65.415        | 4.510        |
| HRBFNN             | 84.698 | 2.138 | 60.708        | 5.457        |
| FRBFNN             | 85.983 | 1.404 | 65.880        | 4.286        |
| Proposed NGN-RBFNN | 85.278 | 1.359 | <b>67.740</b> | <b>3.328</b> |

Based on the results in Tables 2, 3, and 4, it can be seen that the NGN-RBFNN algorithm outperforms other comparison methods across different datasets, particularly in terms of classification accuracy and Standard Deviation.

**Table 3.** Performance of Different Algorithms on the Abalone Dataset

| Method             | TR     | STD   | TE     | STD   |
|--------------------|--------|-------|--------|-------|
| MM-RBFNN           | 72.925 | 1.437 | 69.400 | 6.561 |
| HRBFNN             | 72.625 | 1.231 | 70.000 | 6.680 |
| FRBFNN             | 71.250 | 1.416 | 68.100 | 7.074 |
| Proposed NGN-RBFNN | 72.750 | 1.034 | 69.800 | 6.750 |

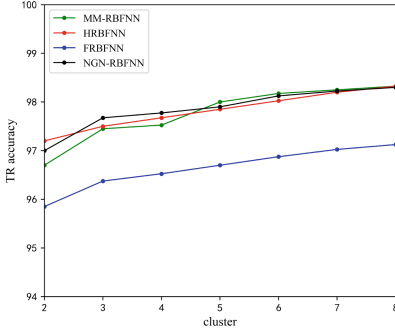
**Table 4.** Performance of Different Algorithms on the MyHog Dataset

| Method             | TR     | STD   | TE     | STD   |
|--------------------|--------|-------|--------|-------|
| MM-RBFNN           | 99.133 | 0.573 | 88.800 | 4.482 |
| HRBFNN             | 99.466 | 0.182 | 89.733 | 2.326 |
| FRBFNN             | 99.050 | 0.191 | 91.466 | 2.022 |
| Proposed NGN-RBFNN | 99.450 | 0.172 | 90.800 | 3.819 |

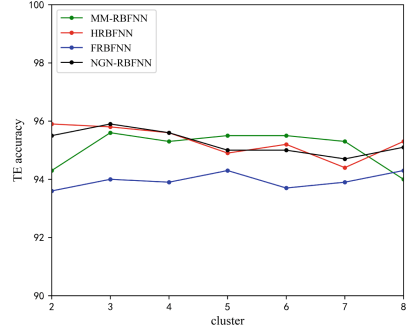
Specifically, on the Glass dataset, NGN-RBFNN achieved a testing dataset accuracy of 67.74%, which is significantly higher than other methods (MM-RBFNN at 65.42%, HRBFNN at 60.71%, FRBFNN at 65.88%). Moreover, its standard deviation (3.24) is relatively small. On the Abalone dataset, NGN-RBFNN attained a testing dataset accuracy of 69.8%, again outperforming FRBFNN (68.1%) and MM-RBFNN (69.4%), although it is slightly lower than HRBFNN (70.0%). Finally, on the Myhog dataset, NGN-RBFNN achieved a testing dataset accuracy of 90.8%, which is slightly higher than HRBFNN’s 89.73%. In summary, the NGN-RBFNN algorithm consistently demonstrates high accuracy and low error across all test datasets, indicating its robust generalization ability and effectiveness in handling complex datasets.

To compare the performance of different clustering algorithms under various cluster numbers, a series of experiments were designed and tested on the widely used Page dataset. The number of clusters was set to range from 2 to 8, covering a variety of scenarios from small to large-scale cases. The experimental results for the training dataset are shown in Fig. 3, which displays the recognition accuracy of each algorithm at different cluster numbers. Figure 4 presents the results for the testing dataset, further validating the generalization ability of each algorithm in practical applications.

From Figs. 3 and 4, it can be observed that, under different clustering configurations, the NGN-RBFNN algorithm consistently maintains a high accuracy on the training dataset, with its performance typically being no lower than or better than the other comparison algorithms. This result highlights the effectiveness of the NGN-RBFNN algorithm in extracting features from the training data and performing precise clustering. Furthermore, under most clustering configurations, the NGN-RBFNN algorithm also demonstrates excellent accuracy on the testing dataset, showcasing its outstanding generalization ability and its capac-



**Fig. 3.** Comparison of Accuracy of Different Algorithms with Varying Number of Clusters on the Training data



**Fig. 4.** Comparison of Accuracy of Different Algorithms with Varying Number of Clusters on the Testing data

ity to effectively transfer the knowledge learned during training to the testing dataset.

In summary, the NGN-RBFNN algorithm outperforms other comparison algorithms (MM-RBFNN, HRBFNN, and FRBFNN) across multiple datasets (Raman dataset, Glass dataset, Abalone dataset, Myhog dataset, and Page dataset). The algorithm achieves high recognition rates on the training dataset and maintains strong recognition performance on the testing dataset as well. At the same time, the standard deviation of the NGN-RBFNN algorithm is relatively low. Considering the training recognition rate, testing recognition rate, the NGN-RBFNN algorithm shows excellent performance across different datasets, giving it a significant advantage in practical applications.

## 4 Conclusion

This paper proposes a black plastic classification algorithm based on PCA, NGN clustering, and RBFNN. First, PCA is used for extracting the main feature components of the data while preserving as much of the original data variance as possible. Subsequently, NGN clusters the data by simulating the movement of gas molecules, providing initial cluster centers for the RBFNN. Finally, the RBFNN performs classification based on the cluster information provided by NGN. Experimental results show that the NGN-RBFNN system achieves a training recognition rate of 96.67% and a testing recognition rate of 95% on the Raman dataset, outperforming other comparison algorithms. Moreover, the model demonstrates high accuracy and low standard deviation on both the training and testing datasets, indicating strong generalization ability.

The NGN-RBFNN system provides an effective solution for black plastic classification, contributing to the development of a circular economy, reducing environmental pollution, and promoting sustainable development. Future research

could further explore more advanced clustering algorithms, optimize the network structure, expand the dataset, and validate the system's industrial application performance in real recycling plants, offering more efficient technological support for the classification and recycling of black plastics.

**Acknowledgments.** This research was supported in part by the National Natural Science Foundation of China under Grants 62173175, 61877033, 62406132, 61903170 and 62173175, in part by the Natural Science Foundation of Shandong Province under Grants ZR2024MF032 and ZR2024QF039, in part by the Young Elite Scientists Sponsorship Program of Shandong Province under Grants SDAST2024QTB023.

**Disclosure of Interests.** The authors declare that they have no conflicts of interest regarding the publication of this paper.

## References

1. Brunner, S., Fomin, P., Kargel, Ch.: Automated sorting of polymer flakes: fluorescence labeling and development of a measurement system prototype. *Waste Manag.* **38**(1), 49–60 (2015)
2. Gundupalli, S.P., Hait, S., Thakur, A.: A review on automated sorting of source-separated municipal solid waste for recycling. *Waste Manag.* **60**(1), 56–74 (2017)
3. Niaounakis, M.: *Biopolymers: Reuse, Recycling, and Disposal*, 2nd edn. William Andrew, New York (2013)
4. Scott, D.M.: A two-color near-infrared sensor for sorting recycled plastic waste. *Meas. Sci. Technol.* **6**(1), 156–159 (1995)
5. Edward, J., Sommer, J.R.: Method and apparatus for near-infrared sorting of recycled plastic waste. US Patent 2001/0045518, A1 (2001)
6. Turner, A.: Black plastics: linear and circular economies, hazardous additives and marine pollution. *Environ. Int.* **117**, 308–318 (2018)
7. Vázquez-Guardado, A., Money, M., McKinney, N., Chanda, D.: Multi-spectral infrared spectroscopy for robust plastic identification. *Appl. Opt.* **54**(24), 7396–7405 (2015)
8. Signoret, C., Caro-Bretelle, A.S., Lopez-Cuesta, J.M., Ienny, P., Perrin, D.: Alterations of plastics spectra in MIR and the potential impacts on identification towards recycling. *Resour. Conserv. Recycl.* **161**, 104980 (2020)
9. Neo, E., Yeo, Z., Low, J., Goodship, V., Debattista, K.: A review on chemometric techniques with infrared, Raman and laser-induced breakdown spectroscopy for sorting plastic waste in the recycling industry. *Resour. Conserv. Recycl.* **180**, 106217 (2022)
10. Becker, W., Sachsenheimer, K., Klemenz, M.: Detection of black plastics in the middle infrared spectrum (MIR) using photon up-conversion technique for polymer recycling purposes. *Polymers* **9**, 435 (2017)
11. Küter, A., Reible, S., Geibig, T., Nüßler, D., Pohl, N.: THz imaging for recycling of black plastics. *TM Tech. Mess.* **85**(3), 191–201 (2018)
12. Shameem, K.M., et al.: A hybrid LIBS-Raman system combined with chemometrics: an efficient tool for plastic identification and sorting. *Anal. Bioanal. Chem.* **409**(13), 3299–3308 (2017)

13. Zhou, K., Oh, S.K., Pedrycz, W., et al.: Design of data feature-driven 1D/2D convolutional neural networks classifier for recycling black plastic wastes through laser spectroscopy. *Adv. Eng. Inform.* **53**, 101695 (2022)
14. Zhou, K., Oh, S.K., Qiu, J., et al.: Reinforced two-stream fuzzy neural networks architecture realized with the aid of one-dimensional/two-dimensional data features. *IEEE Trans. Fuzzy Syst.* **31**(3), 707–721 (2022)
15. Sbrana, A., de Almeida, A.G., de Oliveira, A.M., et al.: Plastic classification with NIR hyperspectral images and deep learning. *IEEE Sens. Lett.* **7**(1), 1–4 (2022)
16. Bonifazi, G., Capobianco, G., Cucuzza, P., et al.: Black plastic waste classification by laser-induced fluorescence technique combined with machine learning approaches. *Waste Biomass Valoriz.* **15**(3), 1641–1652 (2024)
17. Smith, L.: A tutorial on principal components analysis. Cornell University, USA, 65 (2002)
18. Yoo, S.-H., Oh, S.-K., Pedrycz, W.: Design of face recognition algorithm using PCA-LDA combined for hybrid data preprocessing and polynomial-based RBF neural networks: design and its application. *Expert Syst. Appl.* **40**(5), 1451–1466 (2013)
19. Martinetz, M., Berkovich, S., Schulten, K.: “Neural-gas” network for vector quantization and its application to time series prediction. *IEEE Trans. Neural Netw.* **4**, 558–569 (1993)



# Enhanced Group Method of Data Handling with Roulette Neuron Selection for Time Series Forecasting

Xixin Wang<sup>1</sup> , Shaoguang Shi<sup>1,2</sup> , and Congcong Zhang<sup>1</sup>

<sup>1</sup> School of Mathematics and Statistics, Linyi University, Linyi 276005, China  
shishaoguang@lyu.edu.cn

<sup>2</sup> School of Mathematics and Statistics, Shandong Normal University, Jinan 250014,  
People's Republic of China

**Abstract.** Time series forecasting is a pivotal tool across various fields, including finance, energy, and healthcare, where accurate predictions can significantly influence decision-making processes. In this study, we introduce an Enhanced Group Method of Data Handling (GMDH) model, which is integrated with a Roulette Neuron Selection Algorithm to bolster forecasting accuracy and efficiency. The enhancements to the GMDH framework are designed to adeptly address the non-linear relationships commonly found in time series data. This is achieved through the introduction of a roulette selection mechanism, which optimizes the neuron selection process during the network training phase. The adaptive nature of the roulette selection strategy is key to this model, as it enhances the diversity of neuron selection and mitigates the risk of premature convergence. Such improvements lead to the development of more robust forecasting models that are capable of delivering superior performance. Through extensive experimental evaluations on benchmark time series datasets, our proposed approach demonstrates clear superiority over traditional GMDH models as well as other contemporary state-of-the-art methods. This superiority is evident in both predictive accuracy and computational efficiency. The findings from our research suggest that incorporating a roulette-based selection mechanism into the GMDH framework provides a promising avenue for tackling complex time series forecasting challenges. By effectively managing the intricacies of non-linear data patterns and enhancing model adaptability, this approach sets a new standard for precision and resource efficiency in the field of time series forecasting.

**Keywords:** Time Series Analysis · Neural Networks · Roulette Neuron Selection · Group Method of Data Handling (GMDH)

## 1 Introduction

A significant volume of data is generated from diverse sources, such as sensors, machinery. Time series data represent one of the most prevalent types, found

commonly in everyday situations. The task of forecasting time series, which entails predicting future values over a specified duration, is crucial for real-world applications including economic trend analysis, climate forecasting, machinery maintenance, air quality tracking, and power consumption prediction [1, 2].

Traditional statistical inference techniques typically utilize mathematical equations to analyze and describe trends based on historical data for the purpose of forecasting future demand. These statistic models include Autoregressive (AR), Moving Average (MA), and their combination, ARIMA. However, since time series often exhibit nonlinear and long-term dynamic characteristics in reality, this undoubtedly presents significant challenges for time series analysis.

GMDH (Group Method of Data Handling) is a self-organizing data modeling method introduced by Soviet scientist Alexey G. Ivakhnenko in the 1960s. It generates optimal models using polynomial functions and a stepwise selection process, effectively managing complex nonlinear systems. GMDH has advantages such as self-organization, the ability to handle nonlinear time series, and strong robustness. Despite its advantages, GMDH often risks overfitting, which can be critical in certain situations.

Therefore, this paper proposes a novel method to mitigate the overfitting risk in GMDH models by using a roulette wheel optimization algorithm to replace the traditional selection of neuron count and iteration layers. The effectiveness of this approach was validated using several public datasets. Experimental results clearly demonstrate that the modified GMDH model shows significant improvement over traditional statistical models and the conventional GMDH [3–5].

The originality of this study lies in the innovative selection of a method from the field of optimization algorithms to modify the approach used in GMDH. We propose a GMDH network structure optimized by a roulette wheel selection (RWS) algorithm. This proposed method effectively reduces the risk of overfitting in GMDH, improves model accuracy, and ensures high precision in time series prediction.

## 2 The Group Method of Data Handling

In this section, we begin by introducing the background of GMDH, and then we delve into the algorithmic process and its topological structure.

### 2.1 Introduction to GMDH

In GMDH models, Kolmogorov-Gabor polynomials are typically used to link inputs and outputs, expressed as [5–8]:

$$y = a_0 + \sum_{i=1}^m a_i x_i + \sum_{i=1}^m \sum_{j=1}^m a_{ij} x_i x_j + \sum_{i=1}^m \sum_{j=1}^m \sum_{k=1}^m a_{ijk} x_i x_j x_k + \cdots \quad (1)$$



where  $x$  represents input vectors  $x = \{x_1, x_2, \dots, x_m\}$  and  $A = (a_0, a_1, \dots, a_m)$  are weight variables. In the context of random process theory, Kolmogorov-Gabor polynomials have the capability to approximate any observations of a stationary random sequence. These polynomials can be computed using either adaptive methods or Gaussian normal equations. However, two main issues arise: 1. Input sets are often incomplete in most cases, making vector matrices non-invertible. 2. Increasing data volume escalates computation time and space requirements.

Thus, reducing feature selection is necessary. Ivakhnenko addressed this by refining Kolmogorov-Gabor polynomials and developing the GMDH algorithm, using lower-order polynomials to approximate Kolmogorov-Gabor polynomials for handling large datasets. He demonstrated that a second-order polynomial can reconstruct Kolmogorov-Gabor polynomials under certain conditions:

$$y = a_0 + a_1x_i + a_2x_j + a_3x_ix_k + a_4x_i^2 + a_5x_j^2 \quad (2)$$

This reduces computational load while improving precision for perceptron types by classifying information into “useful” and “useless”. GMDH is categorized as a self-organizing method, with Ivakhnenko emphasizing the necessity and importance of the self-organization process for classifying all information.

## 2.2 The Iterative Multilayered GMDH Algorithm

In this multilayered algorithm, the iteration rule remains consistent across all series. For example, a specific description is used in the first series, the form

$$y = a_0 + a_1x_i + a_2x_j + a_3x_ix_j + a_4x_i^2 + a_5x_j^2 \quad (3)$$

is used in the first series, a particular description

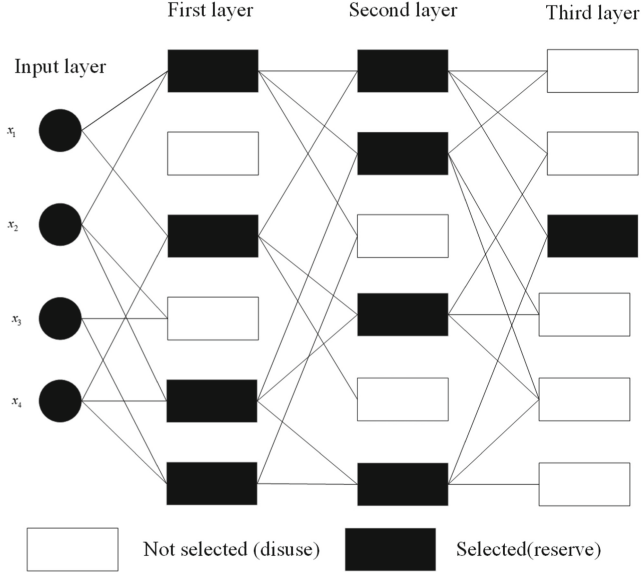
$$z = b_0 + b_1y_i + b_2y_j + b_3y_iy_j + b_4y_i^2 + b_5y_j^2 \quad (4)$$

in the second series, a particular description

$$w = c_0 + c_1z_i + c_2z_j + c_3z_iz_j + c_4z_i^2 + c_5z_j^2 \quad (5)$$

In this iterative approach, the output values from one series are utilized as input for the subsequent series, continuing in this manner. However, this method might miss certain models, potentially resulting in a multilayer error. To address this, it's crucial to assess the convergence of the model's self-organization process. This ensures that the results align with those obtained from regression analysis using identical sorting criteria.

### 2.3 Topology Diagram of GMDH



**Fig. 1.** The Architecture of the GMDH Neural Network Method

According to Eq. (2), for a bivariate polynomial,  $\frac{m \times (m-1)}{2}$  candidate neurons will be constructed, where  $m$  is the number of neurons in the previous layer. If inputs are allowed from the previous layer and the input layer,  $m$  will be the sum of neurons from both layers [9] (Fig. 1).

GMDH neural network methodology is hierarchical, consisting of layers of neurons. The number of neurons in each layer depends on the number of inputs. Assuming the number of inputs entering a layer is  $p$ , the number of neurons in that layer becomes:

$$h = \frac{p \times (p-1)}{2} \quad (6)$$

This Eq. (6) takes into account all pairwise combinations of inputs; however, not all layers are required to contain  $h$  neurons.

The GMDH neural network method is designed to self-organize by selecting input variables that significantly impact the classification process. During training, the network dynamically chooses neuron connections to optimize its performance. This adaptability allows the GMDH method to automatically determine the optimal number of layers, aiming to achieve maximum accuracy while minimizing the risk of overfitting. By doing so, the network can efficiently model complex relationships within the data, maintaining a benefit between model complexity and generalization.

### 3 Enhanced Group Method of Data Handling

Due to the traditional GMDH’s use of minimum variance selection for neuron selection, the results obtained are generally superior on the training set. However, this method of neuron selection can easily lead to overfitting in small time series, resulting in suboptimal predictive performance. To mitigate overfitting, we have employed an alternative neuron selection approach based on segmenting and using optimization algorithms to select neurons from different segments. This modification allows for improved prediction performance even in the case of large datasets with poor data quality. In this paper, we employ roulette wheel selection to determine probabilities based on RMSE values from the test set, and similarly select the number of layers. This approach benefits by allowing the number of neurons and layers to be determined probabilistically rather than pre-specified, helping to reduce overfitting and enhance model accuracy.

#### 3.1 Roulette Wheel Selection

Roulette Wheel Selection is a widely used technique in genetic algorithms to choose individuals for the subsequent generation. When the wheel is “spun”, individuals with greater fitness have a higher probability of being selected, ensuring that the best-performing individuals are more likely to contribute their genes to the next generation. This method helps maintain diversity while promoting the evolution of optimal solutions over successive generations [10].

Mathematically, if the fitness of individual  $i$  is denoted as  $f_i$ , and there are  $N$  individuals in the population, the probability  $P_i$  of selecting individual  $i$  is given by:

$$P_i = \frac{f_i}{\sum_{j=1}^N f_j} \quad (7)$$

## 4 Experiment

In this section, we assess and compare the effectiveness of the proposed forecasting model within the framework of large-scale time-series analysis. We utilize five distinct datasets obtained from the following <https://github.com/q1hai/Time-series-datasets.git> are considered. Due to conference requirements, the model was fitted to several datasets including MDT, GE, JD, OC, and TI (Table 1).

#### 4.1 Calculate the Forecasting Indices

To test the Modify Group Method of Data (MGMDH) model, the Root Mean Square Error (RMSE) and Mean Square Error (MSE) were calculated as Eq. (8) [11].

$$TFI/EFI = \begin{cases} \sqrt{\frac{1}{N} \sum_{k=1}^N (y_k - \hat{y}_k)^2} & \text{(RMSE)} \\ \frac{1}{N} \sum_{k=1}^N (y_k - \hat{y}_k)^2 & \text{(MSE)} \end{cases} \quad (8)$$

**Table 1.** Datasets

| Dataset type                     | Description                                                                                                                                                                                          |
|----------------------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Minimum Daily Temperatures (MDT) | The dataset consists of 3,650 daily minimum temperature records collected over a ten-year period (1981–1990) in Melbourne. This dataset exhibits clear seasonal effects and significant fluctuations |
| Gree Electric Stock (GE)         | The dataset contains 4,000 records of Gree Electric stock data, covering the period from 1999 to 2015                                                                                                |
| JD Stock (JD)                    | The dataset consists of 1,349 records of JD stock data, covering the period from May 22, 2014, to December 3, 2019                                                                                   |
| Occupancy (OC)                   | The dataset contains 17,600 records of check-in data, covering the period from October 4 to December 14 in 2016                                                                                      |
| Temperatures in Italian (TI)     | The dataset consists of 8,991 hourly temperature records collected in a city in Italy, covering the period from March 10, 2004, to April 4, 2005                                                     |

$TFI$  represents the forecasting index for the training set, while  $EFI$  represents the forecasting index for the testing set.  $y_k$  denotes the actual output, and  $\hat{y}_k$  represents the output predicted by the model.

#### 4.2 Experiment: Temperatures in Italian ( $TI$ )

To illustrate the experimental setup, the Temperatures in Italian ( $TI$ ) dataset is used as an example. The methodology of the model is introduced, with the time interval ( $n$ ) set to 50 in this experiment.

*Datatransformation:* The  $X_{TI} = \{x_1, x_2, \dots, x_{8991}\}$  will be transformed into  $D_{TI} = \{d_1, d_2, \dots, d_{8991}\} = \{(x_1, x_2, \dots, x_{51}), (x_2, x_3, \dots, x_{52}), \dots, (x_{8941}, x_{8942}, \dots, x_{8991})\}$ . The first 50 data are used as input for an example, and the last data is used as output [15, 16].

**ARMA**<sup>1</sup>, **SVR**<sup>2</sup>, **LR**<sup>3</sup>, conventional **GMDH**<sup>4</sup> and **MGMDH**<sup>5</sup> are introduced in Table 2. The index (RMSE) is emphasized in bold.

<sup>1</sup> ARMA: Auto-Regressive Moving Average.

<sup>2</sup> SVR: Support Vector Regression.

<sup>3</sup> LR: Linear Regression.

<sup>4</sup> GMDH: Group method of data handling.

<sup>5</sup> Modified GMDH.

**Table 2.** Experimental results of MGMDH compared to other prediction models

| Dataset | ARMA [12]        | SVR [14]         | LR [14]                            | GMDH              | MGMDH                               |
|---------|------------------|------------------|------------------------------------|-------------------|-------------------------------------|
| MDT     | $2.747 \pm 0.07$ | $2.421 \pm 0.03$ | $2.417 \pm 0.03$                   | $2.49 \pm 0.01$   | <b><math>2.413 \pm 0.02</math></b>  |
| GE      | $0.587 \pm 0.04$ | $0.484 \pm 0.02$ | $0.485 \pm 0.02$                   | $0.57 \pm 0.02$   | <b><math>0.472 \pm 0.01</math></b>  |
| JD      | $0.845 \pm 0.03$ | $0.781 \pm 0.04$ | $0.779 \pm 0.04$                   | $0.80 \pm 0.01$   | <b><math>0.7024 \pm 0.02</math></b> |
| OC      | $78.95 \pm 1.38$ | $42.82 \pm 5.15$ | <b><math>42.06 \pm 3.49</math></b> | $70.19 \pm 12.25$ | $68.083 \pm 5.02$                   |
| TI      | $1.592 \pm 0.02$ | $1.171 \pm 0.04$ | $1.068 \pm 0.04$                   | $1.32 \pm 0.02$   | <b><math>1.027 \pm 0.02</math></b>  |

### 4.3 Conclusion Summary

From Table 2: (1) The proposed modified model outperforms other classical time series prediction models (4/5). (2) Compared to the traditional GMDH model, the MGMDH model shows significant improvements in prediction accuracy and overfitting resistance.

Specifically, when applied to large-scale time series predictions, the stochastic nature of the roulette wheel optimization algorithm helps counteract overfitting, enhancing the model’s predictive power. This optimization represents a significant improvement over traditional selection methods, enabling the model to be applied to a wider range of datasets with varying quality [17, 18].

## References

1. De Gooijer, J.G., Hyndman, R.J.: 25 years of time series forecasting. *Int. J. Forecast.* **22**(3), 443–473 (2006)
2. Box, G., Pierce, D.A.: Distribution of residual autocorrelations in autoregressive-integrated moving average time series models. *J. Am. Stat. Assoc.* **65**(332), 1509–1526 (1970)
3. Anastasakis, L., Mort, N.: The development of self-organization techniques in modelling: a review of the group method of data handling (GMDH). Department of Automatic Control Systems Engineering (2001)
4. Mittal, S.: A survey on modeling and improving reliability of DNN algorithms and accelerators. *J. Syst. Archit.* **104** (2020)
5. Ivakhnenko, A.G.: Polynomial theory of complex systems. *IEEE Trans. Syst. Man Cybern.* **4**, 364–378 (1971)
6. Liu, Z., Wang, Y., Vaidya, S., et al.: KAN: Kolmogorov-Arnold networks. *arXiv preprint arXiv:2404.19756* (2024)
7. Ivakhnenko, A.G., Ivakhnenko, G.A.: The review of problems solvable by algorithms of the group method of data handling (GMDH). *Pattern Recogn. Image Anal. c/c of raspoznavaniye obrazov i analiz izobrazhenii* **5**, 527–535 (1995)
8. Farlow, S.J.: The GMDH Algorithm. *Self-organizing Methods in Modeling*, pp. 1–24. CRC Press (2020)
9. Jinlian, L., Zhang, Y., Lu, J.: Long and medium term power load forecasting based on a combination model of GMDH, PSO and LSSVM. In: 2017 29th Chinese control and decision conference (CCDC). IEEE (2017)

10. Marmolin, H.: Subjective MSE measures. *IEEE Trans. Syst. Man Cybern.* **16**(3), 486–489 (1986)
11. Hochreiter, S.: Long Short-Term Memory. *Neural Computation*. MIT-Press (1997)
12. Tiao, G.C., Grupe, M.R.: Hidden periodic autoregressive-moving average models in time series data. *Biometrika* **67**(2), 365–373 (1980)
13. Krizhevsky, A., Sutskever, I., Hinton, G.E.: ImageNet classification with deep convolutional neural networks. In: *Advances in Neural Information Processing Systems*, vol. 25 (2012)
14. He, K., Zhang, X., Ren, S., et al.: Deep residual learning for image recognition. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 770–778 (2016)
15. Zhou, K., et al.: Design of hierarchical neural networks using deep LSTM and self-organizing dynamical fuzzy-neural network architecture. *IEEE Trans. Fuzzy Syst.* (2024)
16. Zhou, K., et al.: A self-organizing deep network architecture designed based on LSTM network via elitism-driven roulette-wheel selection for time-series forecasting. *Knowl.-Based Syst.* **289**, 111481 (2024)
17. Shritharanyaa, J.P., et al.: Enhancing the performance of extreme learning machine technique using optimization algorithms for embedded workload characterization. *Alex. Eng. J.* **108** (2024)
18. Xie, L., et al.: Detection of setting time during cement hydration using ground penetrating radar. *J. Build. Eng.* **60** (2022)



# FE-ResNet50: Frequency Enhanced Attention Network for sEMG Gesture Recognition

Haozhu Wang<sup>1</sup>(✉), Du Jiang<sup>1,3</sup>, Juntong Yun<sup>1,2</sup>, Ying Liu<sup>4</sup>, and Boao Li<sup>1</sup>

<sup>1</sup> Key Laboratory of Metallurgical Equipment and Control Technology of Ministry of Education, Wuhan University of Science and Technology, Wuhan 430081, China  
2164734468@qq.com

<sup>2</sup> Research Center for Biomimetic Robot and Intelligent Measurement and Control, Wuhan University of Science and Technology, Wuhan 430081, China

<sup>3</sup> Hubei Key Laboratory of Mechanical Transmission and Manufacturing Engineering, Wuhan University of Science and Technology, Wuhan 430081, China

<sup>4</sup> School of Mechanical Engineering, Hubei Engineering University, Xiaogan 432000, China

**Abstract.** In recent years, gesture recognition based on surface electromyography (sEMG) signals has attracted the attention of researchers due to its non-invasive and high efficiency. However, the traditional gesture recognition methods still have limitations in the accuracy of recognizing complex gesture scenarios. In addition, most existing methods only model in the time or spatial dimension, and cannot fully utilize the frequency information of the signal. Although the common attention mechanism (such as SE module) can enhance the importance of channel features, it is limited to global average pooling and easily loses the spatial information within the channel. To address these issues, this paper proposes a hand gesture recognition framework named FE-ResNet50, which combines frequency-enhanced channel attention mechanism (FECAM) with ResNet50 backbone network. The network places the channel attention mechanism before the residual layer of ResNet50 network, while capturing spatial features of hand gesture signals through the deep network structure of ResNet50. By training on a self-built dataset, the network achieves an average classification accuracy of 92.77%. The validation on the NinaProDB1 public dataset shows that the network achieves an accuracy of 88.93%. Experimental results show that the network can effectively improve the performance of sEMG-based gesture recognition, has good robustness and generalization ability, and shows good performance and application potential in actual gesture recognition tasks.

**Keywords:** sEMG · Gesture recognition · FE-ResNet50

## 1 Introduction

The sEMG signal is a physiological signal closely related to human movement. By analyzing the surface electromyography (sEMG) signals generated by the concentrated muscles of the hand and forearm, it is possible to accurately decode the user's movement intention and provide rich gesture characteristics [1–3]. Often, even physically disabled

patients with partial loss of hand function have forearm muscles and nervous systems that can still function normally and generate EMG signals, and these advantages have gradually made the hand an ideal object for sEMG research.

However, although sEMG signals show a wide range of research and application potential in the field of hand gesture recognition, their weak and noise-prone nature makes them challenging to recognize accurately [4]. Currently, sEMG-based gesture recognition methods are mainly divided into traditional machine learning methods and deep learning methods. Traditional methods usually include signal preprocessing (e.g. denoising and filtering), feature extraction and classifier modelling, but the manual feature extraction process is complex and the classification accuracy is limited; deep learning methods use models such as convolutional neural networks (CNNs) to automatically extract features through the training of a large amount of data and gradually optimized the model parameters, which significantly improves the recognition accuracy [5–7]. Nevertheless, most CNN-based gesture recognition methods fail to make full use of the time series information of sEMG signals, and there are still some limitations in the improvement of accuracy [8].

To address the above problems, this paper proposes a novel network, FE-ResNet50, for use in gesture recognition tasks. The method can effectively strengthen the correlation between the channel features and suppress the noise interference by embedding the FECAM attention mechanism in the ResNet50 network and combining it with the Discrete Cosine Transform (DCT) to model the frequency-domain features of the sEMG signals [9]. Meanwhile, the FECAM module can better capture the time-frequency characteristics of the sEMG signal by dynamically adjusting the channel weights. By combining FECAM with ResNet50 residual network, the accuracy and robustness of the network in classifying sparse gesture EMG signals can be further improved. The main contributions of this paper are as follows:

- (1) A dataset containing 12 individuals with 7 different gestures is constructed by collecting surface EMG signals.
- (2) A novel surface EMG signal gesture recognition method, FE-ResNet50, is proposed, which solves the problems and limitations of the traditional SE module, such as the inability to effectively model the frequency domain information, the single allocation of weights between channels, and the easy loss of spatio-temporal feature details.
- (3) The model is validated by the self-collected dataset and the NinaProDB1 public dataset [10], which visually demonstrates the performance of the proposed model and the effectiveness of the improvement.

The main structure of this paper is as follows. In Sect. 2, the acquisition of surface EMG signals, pre-processing, feature extraction as well as the generation of feature maps and the specific process and structure of the FE-ResNet50 network used for classification are described. In Sect. 3, a multi-faceted evaluation experiment of the network performance is carried out with self-acquired datasets and public datasets. Section 4 summarizes the main findings of this paper.



2 Method

2.1 Signal Acquisition

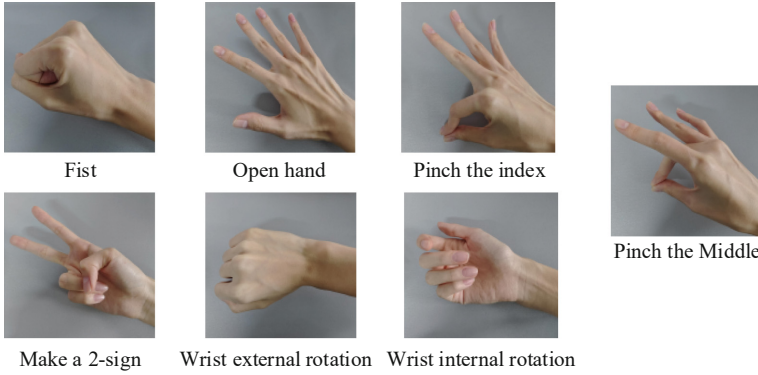
In this study, we recruited 12 volunteers (10 males and 2 females, aged 22–26 years) to participate in an experiment to collect electromyographic signals from the forearm. All subjects have no history of muscle disease and are proficient in operating and applying electromyographic signal acquisition equipment., as shown in Table 1.

Table 1. General Information about Subjects.

| Subject | Gender | Age | Height (cm) | Weight (Kg) | BMI   |
|---------|--------|-----|-------------|-------------|-------|
| h1      | female | 22  | 171         | 50          | 17.09 |
| h2      | male   | 24  | 180         | 78          | 24.07 |
| h3      | male   | 22  | 181         | 70          | 21.37 |
| h4      | male   | 23  | 177         | 60          | 19.15 |
| h5      | male   | 23  | 176         | 66          | 21.31 |
| h6      | female | 23  | 159         | 46          | 18.2  |
| h7      | male   | 26  | 176         | 60          | 19.37 |
| h8      | male   | 22  | 165         | 50          | 18.4  |
| h9      | male   | 22  | 176         | 72          | 23.2  |
| h10     | male   | 22  | 174         | 85          | 28.1  |
| h11     | male   | 22  | 172         | 83          | 28    |
| h12     | male   | 22  | 174         | 70          | 23.1  |

In the experiment, each movement was set to be repeated in 10 groups, each group performed three gestures, each gesture movement lasted for 5 s with a 3-s rest to prevent muscle fatigue, and each subject had a half-minute rest period (during which they did not take off the Myo arm band) before switching to a different movement for acquisition. A total of 7 common gestures in office work were collected, mainly including clenched fist (G1), open hand (G2), pinching index finger (G3), pinching middle finger (G4), make a 2-sign (G5), wrist external rotation (G6) and wrist internal rotation (G7) which are as shown in Fig. 1.

The equipment used for the experiment was an innovative 8-channel MYO (gesture-controlled arm band) arm band from the Canadian start-up company Thalmic Labs, which has a sampling rate of 200 Hz, and which connects wirelessly to other electronic products via a low-power Bluetooth device. Prior to the experiment, subjects were informed of the detailed experimental procedure and potential risks and signed an informed consent form. During the experiment, the wristband should be kept worn on the middle of the forearm and proper skin preparation should be performed to remove excess oil and sweat before wearing. During the acquisition process, the wearer needs to keep the arm stable, keep the amplitude, strength and duration of each gesture movement basically the



**Fig. 1.** Seven Gestures Movements.

same, and make the corresponding gestures according to the prompts. The collected data will be transmitted to the host computer via Bluetooth for the next step of processing. The acquisition device is shown in Fig. 2. The 8-channel raw surface EMG signal map acquired by Myo is shown in Fig. 3.



**Fig. 2.** MYO Acquisition Equipment

## 2.2 sEMG Preprocessing

Since EMG signals are susceptible to interference from ambient noise during acquisition, such as the inherent noise of the acquisition equipment and the industrial frequency noise. Therefore, in order to improve the accuracy of the subsequent feature extraction and classification, it is necessary to carry out preliminary filtering and noise reduction and related pre-processing on the original sEMG signal after acquisition.

**Filtering and Noise Reduction.** The input raw EMG signal is firstly rectified by full-wave rectification and the rectified data is low-pass filtered using a second-order Butterworth filter. The full-wave rectification allows all the energy to be concentrated in the positive semi-axis, which facilitates the subsequent envelope extraction and feature analysis, while the low-pass filtering serves to suppress the high-frequency noise



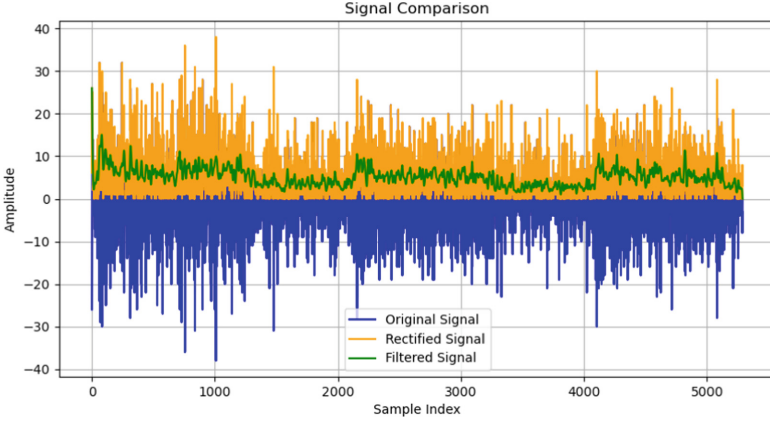
**Fig. 3.** 8-Channel Raw Surface EMG Signals Acquired by Myo.

in the signal while retaining the low-frequency components. For the sEMG signal, the low-frequency part contains important action potential information and signal energy characteristics, and the rectified waveform can be further smoothed by low-pass filtering. In this paper, a second-order Butterworth filter is used, whose frequency response range is set between 5 Hz and 100 Hz to preserve the main frequency band of the EMG signal, while effectively suppressing the low-frequency baseline shift and high-frequency noise. The transfer function of the filter is shown in Eq. (1).

$$H(\omega) = \frac{1}{\sqrt{1 + \left(\frac{\omega}{\omega_c}\right)^{2n}}} \quad (1)$$

Among them,  $\omega_c$  is the passband edge frequency of the filter,  $n$  is the order of the filter. The filtering effect is shown in Fig. 4, It can be seen that the processed signal energy is mainly concentrated in the effective frequency band and some of the additional noise was clearly suppressed.

**Active Segment Detection.** EMG signals acquired during exercise can usually be classified into active and resting segments, of which the active segment signals are the key part used for action classification. In order to extract action features more effectively, the detection process of active segments is carried out in this paper. The core of this process is to identify and effectively differentiate between the active and resting parts of the signal in order to capture the action information more accurately. Specifically, we determine the threshold value by calculating the average of the peak and valley values of the active segment signals in each acquisition channel. This calculation provides us with a quantitative criterion for distinguishing between active and stationary states in the signal. When the signal amplitude is above the threshold, it is considered as an action potential and vice versa as a resting potential. The calculation of the threshold value, the



**Fig. 4.** Comparison of The Effect before and after Filtering.

expression of the action segment recognition function, and the formula for processing the resting signal are shown in Eq. (2)–(4).

$$TH_c = \frac{1}{n} \sum_{i=1}^n (\max(sEMG_c(i)) - \min(sEMG_c(i))) \quad (2)$$

$$y_c(i) = \begin{cases} 1, & \text{if } sEMG_c(i) > TH_c \\ 0, & \text{if } sEMG_c(i) \leq TH_c \end{cases} \quad (3)$$

$$S_c(i) = y_c(i) \cdot sEMG_c(i) \quad (4)$$

where  $TH_c$  is the action threshold of the channel  $c$ ,  $n$  is the number of active segments, and  $sEMG_c(i)$  represents the signal value at sample point  $i$  of channel  $c$ ,  $S_c(i)$  is the action segment signal of channel  $c$ .

### 2.3 Feature Extraction and sEMG Feature Image Construction

Feature extraction of surface electromyographic signals (sEMG) consists of two main types of methods: time domain and frequency domain. Time-domain features are mainly used to describe the trend of the signal over time, and the commonly used methods include integral absolute value (IEMG), absolute average value (MAV), root mean square (RMS), zero crossing rate (ZC) and waveform length (WL). Frequency-domain features, on the other hand, focus on the spectral information of the signal, mainly including mean frequency (MNF), median frequency (MDF) and specific spectral energy (TTP), which are effective in capturing the frequency variations related to the type of exercise or muscle fatigue, and reflecting the energy distribution of the signal and its main frequency components [11].

In this study, a combination of two time-domain features and one frequency-domain feature was chosen to process the sEMG signal and generate the feature map. Among the time-domain features, the integral absolute value (IEMG) reflects the overall energy of muscle activity by accumulating the absolute value of the signal, and the absolute mean (MAV) measures the steady-state intensity of muscle activity by calculating the average of the absolute values of the signal. For the selection of frequency domain features, the mean frequency (MNF) was used in this study. The main frequency components of the signal are derived by calculating the weighted average of the power spectrum, thus reflecting the spectral characteristics of the signal.

The combination of IEMG, MAV and MNF can fully present the characteristics of the sEMG signal. IEMG and MAV capture the amplitude information of the muscle activity in terms of overall intensity and average intensity, respectively, whereas MNF complements the spectral information to capture the frequencies that are not easy to be detected in the time-domain features. These features combine time-domain and frequency-domain approaches to comprehensively characterize the signal from multiple dimensions such as energy, intensity and frequency, which ensures the rationality and effectiveness of the signal processing and thus improves the robustness and accuracy of the gesture classification. The following are the formulas for the four methods in Eq. (5)–(7).

$$IEMG = \sum_{n=1}^N |x_n| \quad (5)$$

$$MAV = \frac{1}{N} \sum_{n=1}^N |x_n| \quad (6)$$

$$MNF = \frac{1}{M} \sum_{j=1}^M P_j \quad (7)$$

where  $x_n$  represents the magnitude of the  $n$ th point in the time series (in mv),  $N$  is the total number of sampling points on the time series,  $P_j$  represents the spectral power of the EMG signal in the  $j$ th frequency band, and  $M$  represents the total length of the frequency band (in HZ).

After completing the selection of features and the generation of feature maps, a time window is further used to segment the whole signal into multiple segments. The segmented time window contains the local features of the signal, which can enhance the model's ability to analyze the action continuity and timing relationship, and also increase the number of data samples to provide more training data for the model, which helps to alleviate the overfitting problem and improve the classification performance. In this paper, the length of the time window is chosen to be 300 ms and the step size is 100 ms.

## 2.4 SEMG Feature Images

In order to further capture the continuity and dynamic change characteristics of the signal, a fixed-length sliding window is usually used to extract features by gradually sliding on

the signal with a certain step size. This method can make the extracted features have time-smoothing and redundant information, and at the same time can ensure that the short-term change characteristics of muscle activity are captured to better reflect the local change rule of the signal. The specific process of generating feature images is as follows, firstly, the three previously extracted features are reorganized and arranged into a feature matrix to prepare for the generation of feature images; secondly, the feature values in the matrix are uniformly normalized to ensure that the different features are consistent with the data range of the acquisition channel; finally, the normalized feature matrix is converted into an image format, and the final feature image is of the size of ( $C \times H \times W$ ), which can be directly input into the model for classification.

## 2.5 FE-ResNet50 Classification Network

In this study, the FE-ResNet50 network was used as the backbone network in the gesture recognition task, and by placing the FECAM attention mechanism before the four residual blocks of the ResNet50 network, it ensured that the features extracted from the original input were enhanced before they were fed into the residual blocks. The modified structure not only maintains the integrity of the ResNet50 architecture, but also enhances the network's ability to learn from subtle details with aggregated contextual information, which is crucial for the gesture recognition task.

**FECAM Module.** The adopted FECAM module converts the input feature maps from time domain (spatial domain) to frequency domain representation by DCT before generating the attentional weights of each channel, through which the global and local frequency features in the EMG signals are fully extracted by this frequency domain transformation. The high frequency captures the spatial variations and the low frequency responds to the overall distribution trend, which makes up for the shortcomings of the SE attention mechanism. Meanwhile, compared with the traditional Fourier method used to transform the signal from the time domain to the frequency domain, FECAM can better solve the problem of high-frequency noise in the signal boundary by the symmetric expansion in the discrete cosine transform. Finally, the extracted frequency-domain features are used to generate attention weights for each channel through global average pooling and convolutional layers, which are weighted to the original input feature maps to dynamically assign weights, suppress interference noise, and highlight discriminative frequency components. The introduction of the FECAM not only enhances the model's ability to model features of complex EMG signals, but also solves the problem that channel dependencies are difficult to be captured. The following is the specific implementation and calculation process of the FECAM attention mechanism.

The last two dimensions (spatial dimensions) of the input feature map  $x$ , i.e., height and width, are firstly obtained, and then converted into frequency domain representation by the two-dimensional discrete cosine transform (DCT-II) as shown in Eq. (8).

$$X_{u,v} = \alpha(u)\alpha(v) \sum_{x=0}^{H-1} \sum_{y=0}^{W-1} X_{x,y} \cdot \left( \frac{\pi(2x+1)u}{2H} \right) \cdot \left( \frac{\pi(2y+1)v}{2W} \right) \quad (8)$$

In Eq.  $X_{x,y}$  is the value of the input feature map at the spatial location  $(x, y)$ ;  $X_{u,v}$  is the frequency component of the feature map in the frequency domain at the location  $(u, v)$ ; and  $\alpha(k)$  is the normalisation factor, which is used to ensure consistent energy. It is taken in the manner shown in Eq. (9).

$$\alpha(k) = \begin{cases} \sqrt{\frac{1}{N}}, & \text{if } k = 0 \\ \sqrt{\frac{2}{N}}, & \text{otherwise} \end{cases} \quad (9)$$

In the code, the 2D discrete cosine transform is implemented by matrix multiplication of Eq. (10).

$$D_N @ x @ D_M^T \quad (10)$$

In Eq.  $D_N$  is the DCT basis matrix in the row direction with size  $(H, H)$ ,  $D_M$  is the DCT basis matrix in the column direction with size  $(W, W)$ , and  $x$  is the two-dimensional spatial data of the input feature map on each channel. Through this calculation, the spatial dimension of the input feature map,  $H \times W$ , is converted to the corresponding frequency representation,  $H \times W$ , and the final frequency-domain representation, which still has the shape  $(64, 64, 16, 16)$ , is obtained.

After obtaining the frequency domain representation of the output feature map  $x_{dct}$  it is globally average pooled, and the spatial dimension  $H \times W$  of each channel is compressed to 1 to obtain the average frequency domain feature of each channel, and the global average pooling is given by Eq. (11).

$$avg\_out[c] = \frac{1}{H \cdot W} \sum_{u=0}^{H-1} \sum_{v=0}^{W-1} x_{dct}[c, u, v] \quad (11)$$

The final shape of  $avg\_out$  is  $(B, C, 1, 1) = (32, 64, 1, 1)$ , and each value represents the global feature of the corresponding channel in the frequency domain.

After pooling the channel attention weights, the two convolutional layers generate weights on  $avg\_out$  according to Eq. (12).

$$fc1_{out}[c'] = \sum_{c=1}^C avg_{out}[c] \cdot W_{fc1}[c, c'] \quad (12)$$

where  $W_{fc1}$  is the convolution kernel parameter with the shape of  $(C/r, C)$ , where  $r = reduction$  is the channel reduction factor, which is used to reduce the amount of computation. The final result shape is  $(B, C/r, 1, 1) = (32, 4, 1, 1)$ .

After convolution by the Relu activation function, a non-linear relationship is introduced, calculated as shown in Eq. (13).

$$relu_{out}[c'] = \max(0, fc1_{out}[c']) \quad (13)$$

Use the  $fc2$  convolutional layer again to recover the original number of channels, as calculated in Eq. (14).

$$fc2_{out}[c] = \sum_{c'=1}^{\frac{C}{r}} relu_{out}[c'] \cdot W_{fc2}[c', c] \quad (14)$$

The shape of the output result is  $(B, C, 1, 1) = (32, 64, 1, 1)$ .

Finally, the *Sigmoid* activation function is used to normalize the attentional weights into the interval  $[0, 1]$  and multiply each channel of the original input feature map  $x$  with its corresponding attentional weight for feature enhancement of the channels. Calculated as in Eq. (15)–(16).

$$\text{attention}[c] = \frac{1}{1 + e^{-fc2_{out}[c]}} \quad (15)$$

$$x'[B, C, H, W] = x[B, C, H, W] \cdot \text{attention}[B, C, 1, 1] \quad (16)$$

The final feature map shape after feature enhancement is  $(B, C, H, W) = (32, 64, 16, 16)$ . The specific FECAM attention mechanism flowchart is shown in Fig. 5.

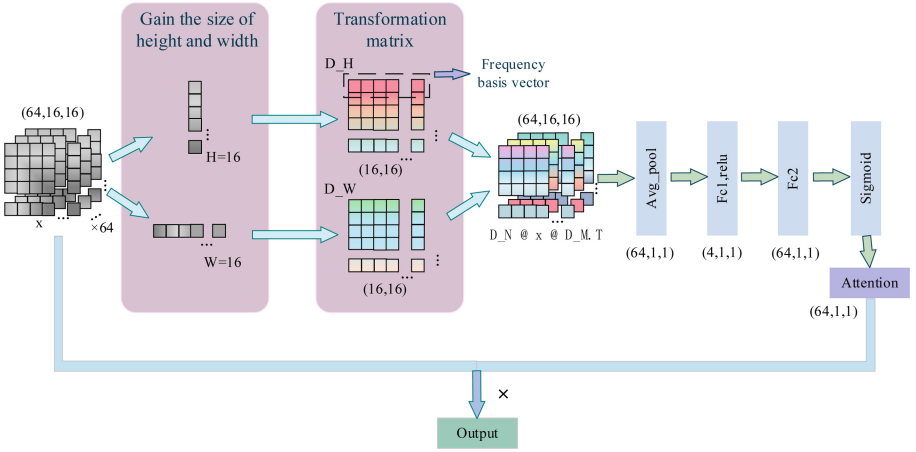
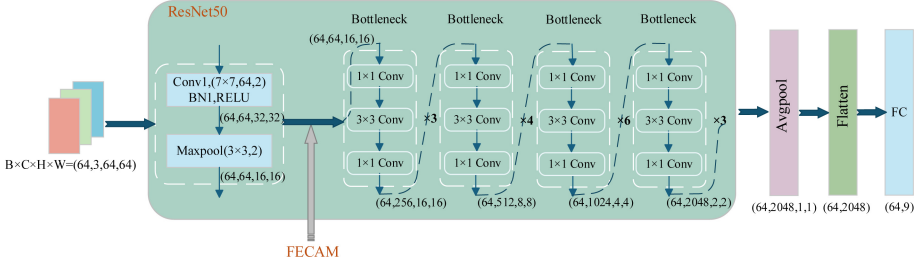


Fig. 5. FECAM Module Flowchart.

**FE-ResNet50 Network.** The process of gesture recognition using the proposed backbone network FE-ResNet50 is mainly divided into, firstly, preliminary convolution, Batch Normalization (BN), Relu activation function and Max Pooling are performed on the input image before the four residual layers, which are aimed at the preliminary pre-processing and feature extraction of the input data and laying the foundation for the subsequent input to the residual layers for deeper feature extraction; secondly, we choose to embed the FECAM module before the four main residual layers of ResNet50 to dynamically adjust the channel weights of the input feature maps. After that, the residual module of the ResNet50 backbone network extracts the deep features of the gesture signals, and at the same time alleviates the problem of gradient vanishing to ensure the efficient training of the deep network. The four residual layers in the residual module are composed of (3, 4, 6, 3) basic Bottleneck residual blocks; finally, after multi-level feature extraction, the features are mapped to the gesture categories for classification by Global Average Pooling and fully connected layers. The network structure is shown in Fig. 6.





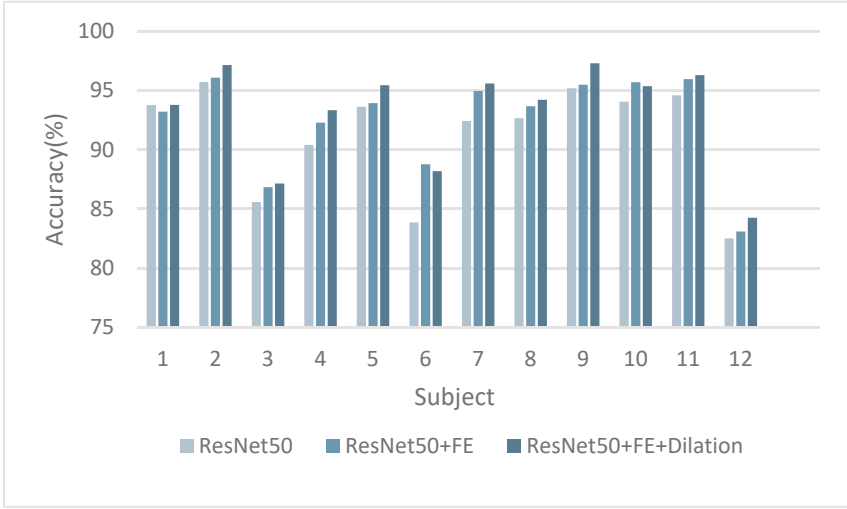
**Fig. 6.** FE-ResNet50 General Framework.

### 3 Experiments and Results

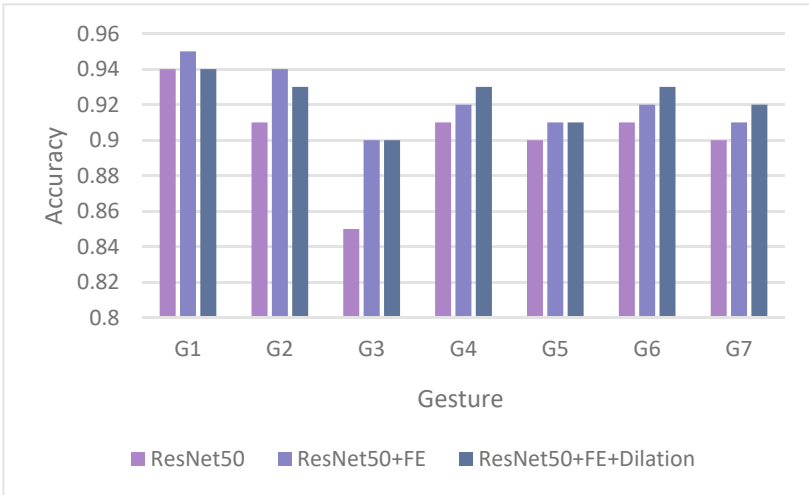
In order to validate the accuracy and effectiveness of the proposed FE-Resnet50 model for gesture recognition, this paper conducts multiple comparative experiments and analyses on the self-collected EMG signal dataset and the public dataset NinaProDB1 respectively. In the self-collected dataset, there were data from a total of 12 subjects, and the seven gesture signals collected by each subject are finally divided into training, validation and test sets according to the ratio of 7:2:1 in terms of the number of repetitions, and then inputted into the network for validation. This division not only ensures that the amount of data in the training phase is sufficient, but also provides reliable validation and testing support for the generalization performance of the model.

Firstly, the classification accuracies of the three models, ResNet50, FE-ResNet50 and FE-ResNet50 with inflated convolution in the residual block, are compared in terms of different collectors and different gestures. The results are shown in Fig. 7 and Fig. 8, and it can be seen that the model with the addition of the FECAM attention mechanism module demonstrates a certain degree of accuracy improvement in both cross-individual and different-gesture experiments. The model incorporating the FECAM module models the frequency domain features through the Discrete Cosine Transform (DCT), which reinforces the global frequency domain information of the signal, while dynamically adjusting the channel weights to highlight key features and suppress irrelevant features. By further expanding the sensory field of the model by combining the expansion convolution in ResNet50, the long-time dependent information in the sEMG signals can be effectively captured, thus improving the classification ability of complex gestures.

Confusion experiments were also conducted on the final improved model, as shown in Fig. 9 below. Through the confusion matrix experiment, we can intuitively observe the strength of the model's prediction ability for each gesture category, and we can see that the model's prediction effect is the worst for the open palm and the pinch index finger (i.e., categories 2 and 3), which is due to the fact that the force generated when doing these two actions is more similar, which is prone to cause misjudge; the gesture model of the make a 2-sign has the best prediction effect, with an accuracy rate of 97.06%. Overall, the model predicts most of the gestures with good accuracy.



**Fig. 7.** Classification Accuracy for 12 Subjects under Different Models.



**Fig. 8.** Classification Accuracy of 7 Gesture Actions under Different Models.

Through the ablation experiments in Table 2, it can be seen that when the model was integrated with the FECAM attention mechanism, the average gesture recognition accuracy of the model increased from 90.12% to 92.40%, a change that directly reflects the effectiveness of FECAM in enhancing the model's ability to capture key features. While the final model combining the inflated convolution with FECAM achieves an accuracy of 92.77%, which is an improvement of 2.63% compared to the original model. The effectiveness of the added module is verified.

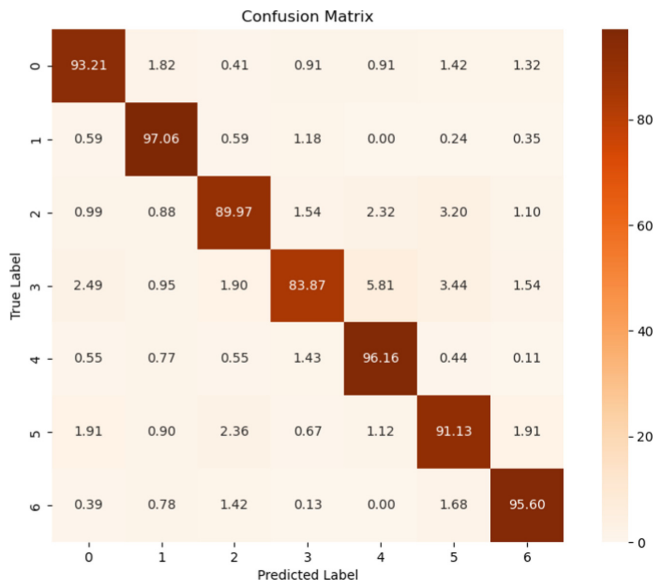


Fig. 9. Confusion Matrix for the FE-ResNet50 Model.

Table 2. Average Gesture Recognition Accuracy Ablation Experiment Results

| Model              | Average Gesture Recognition Accuracy Rate/% |
|--------------------|---------------------------------------------|
| ResNet50           | 90.12                                       |
| + FECAM            | 92.40                                       |
| + Dilation + FECAM | 92.77                                       |

In Table 3, the performance of the improved model on the public dataset NinaProDB1 is evaluated against other existing methods. The experimental results show that the classification accuracy of the FE-ResNet50 model proposed in this paper is 88.93% on this dataset, which is significantly better than traditional machine learning methods (e.g., ELM, 75.11%) and early deep learning models (e.g., Geng\_Net, 77.80%). In addition, the accuracy of FE-ResNet50 still has about 3.93% improvement compared with the latest MSFusionNet. This indicates that the model proposed in this paper can not only show excellent performance on self-constructed datasets, but also has strong generalization ability on public datasets.

**Table 3.** Recognition Accuracy of Different Methods on the NinaProDB1 Dataset

| Algorithms                         | Accuracy (%) |
|------------------------------------|--------------|
| CNN (Du et al.) [12]               | 79.50        |
| ELM (Cene et al.) [13]             | 75.11        |
| Atzori_Net (Atzori et al.) [14]    | 66.67        |
| Geng_Net (Geng et al.) [15]        | 77.80        |
| MSFusionNet (Wei et al., 2019) [6] | 85.00        |
| FE-ResNet50 (Ours)                 | 88.93        |

## 4 Conclusion

In this paper, a gesture recognition method for EMG signals based on FE-ResNet50 network is proposed, which can effectively solve the problems of under-utilization of frequency-domain features, single division of channel weights, and loss of timing information in the existing methods by introducing the FECAM frequency-domain augmented channel attention mechanism. The model combines the ResNet50 deep residual network, which has strong feature extraction capability, to further improve the accuracy of the classification of gesture signals. In the specific implementation, the FECAM module transforms the input signal in the frequency domain via the discrete cosine transform to capture global and local frequency features, in addition to further expanding the receptive field by embedding the dilated convolution in the ResNet50 network to capture the long-term dependence of EMG signals. In order to achieve efficient classification of signals, a combination of IEMG, MAV and MNF time-frequency domain features is also used to generate feature images, which provides rich input information for the model. In the experimental part, FE-ResNet50 shows excellent performance on both self-constructed and public datasets. The experiments on the self-constructed dataset show that the attention model combined with dilated convolution still increases the accuracy by 2.63% on the ResNet50 model, which has already achieved more than 90% classification accuracy, which verifies the effectiveness and adaptability of the added module. Experiments on the public dataset NinaProDB1 show that the classification accuracy of FE-ResNet50 reaches 88.93%, which is an improvement compared with both traditional and current mainstream methods. All these results show that the network proposed in this paper not only performs well in the experimental environment, but also has good robustness and generalization ability.

**Acknowledgments.** This project is supported by the higher education teaching reformation project of Hubei province of China (2022231, 2022216) and department of science and technology of Hubei province of China(2024DJC073). ‘The 14th Five Year Plan’ Hubei Provincial advantaged characteristic disciplines (groups) project of Wuhan University of Science and Technology (2023C0401).

**Disclosure of Interests.** The authors have no competing interests to declare that are relevant to the content of this article.

## References

1. Staudenmann, D., Roeleveld, K.: Methodological aspects of SEMG recordings for force estimation—a tutorial and review. *J. Electromyogr. Kinesiol.* **20**(3), 375–387 (2010). <https://doi.org/10.1016/j.jelekin.2009.08.005>
2. Li, K., Zhang, J., Wang, L.: A review of the key technologies for sEMG-based human-robot interaction systems. *Biomed. Signal Process. Control* **62**, 102074 (2020). <https://doi.org/10.1016/j.bspc.2020.102074>
3. Song, T., Yan, Z., Guo, S.: Review of sEMG for robot control: techniques and applications. *Appl. Sci.* **13**(17), 9546 (2023). <https://doi.org/10.3390/APP13179546>
4. Ding, Z., Yang, C., Tian, Z.: sEMG-based gesture recognition with convolution neural networks. *Sustainability*, **10**(6), 1865 (2018). <https://doi.org/10.3390/su10061865>
5. Tsinganos, P., Cornelis, B., Cornelis, J.: Improved gesture recognition based on sEMG signals and TCN. In: *ICASSP 2019–2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pp. 1169–1173. IEEE (2019). <https://doi.org/10.1109/ICASSP.2019.8683239>
6. Wei, W., Wong, Y., Du, Y.: A multi-stream convolutional neural network for sEMG-based gesture recognition in muscle-computer interface. *Pattern Recogn. Lett.* **119**, 131–138 (2019). <https://doi.org/10.1016/j.patrec.2017.12.005>
7. Kumar, D., Ganesh, A.: A critical review on hand gesture recognition using sEMG: challenges, application, process and techniques. In: *Journal of Physics: Conference Series*, p. 012075. IOP Publishing (2022). <https://doi.org/10.1088/1742-6596/2327/1/012075>
8. Qi, J., Jiang, G., Li, G.: Intelligent human-computer interaction based on surface EMG gesture recognition. *IEEE Access* **7**, 61378–61387 (2019). <https://doi.org/10.1109/ACCESS.2019.2914728>
9. Nguyen, P. T. T., Su, S. F.: A frequency-based attention neural network and subject-adaptive transfer learning for sEMG hand gesture classification. *IEEE Robot. Autom. Lett.* (2024). <https://doi.org/10.1109/LRA.2024.3433748>
10. Chang, J., Phinyomark, A.: Assessment of EMG benchmark data for gesture recognition using the NinaPro database. In: *2020 42nd Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC)*, pp. 3339–3342. IEEE (2020). <https://doi.org/10.1109/EMBC44109.2020.9175260>
11. Sharma, A., Mittal, A., Singh, S.: Hand gesture recognition using image processing and feature extraction techniques. *Procedia Comput. Sci.* **173**, 181–190 (2020). <https://doi.org/10.1016/j.procs.2020.06.022>
12. Du, Y., Jin, W., Wei, W.: Surface EMG-based inter-session gesture recognition enhanced by deep domain adaptation. *Sensors* **17**(3), 458 (2017). <https://doi.org/10.3390/s17030458>
13. Cene, V.H., Tosin, M.: Open database for accurate upper-limb intent detection using electromyography and reliable extreme learning machines. *Sensors* **19**(8), 1864 (2019). <https://doi.org/10.3390/s19081864>
14. Atzori, M., Cognolato, M.: Deep learning with convolutional neural networks applied to electromyography data: a resource for the classification of movements for prosthetic hands. *Front. Neurobot.* **10**, 9 (2016). <https://doi.org/10.3389/fnbot.2016.00009>
15. Geng, W., Du, Y., Jin, W.: Gesture recognition by instantaneous surface EMG images. *Sci. Rep.* **6**(1), 36571 (2016). <https://doi.org/10.1038/srep36571>



# Huber Correntropy Kalman Filter

Shuo Wang and Xiaoliang Feng<sup>(✉)</sup>

School of Electrical Engineering, Shanghai Dianji University, Shanghai 201306, China  
fengxl2002@163.com

**Abstract.** In this paper, based on the information theoretic learning theory, a novel non-Gaussian filter is proposed in the frame of Kalman filter under Huber correntropy loss function, which is short for HCKF (Huber correntropy Kalman filter). Firstly, a novel Huber correntropy loss function is defined. Secondly, HCKF is designed to solve the Huber correntropy loss function by using a fixed-point iterative algorithm and its superior performance is shown by compared Huber correntropy loss function with the L2 loss function, the Huber loss function, and the Gaussian kernel function. The validity of the HCKF is verified by comparing it with three representative filtering methods in the final simulation.

**Keywords:** non-Gaussian noise · linear systems · Huber correntropy loss function · fixed-point iteration · information theoretic learning

## 1 Introduction

According to the MMSE criterion [1], the Kalman filter (KF) is the optimal state estimation method for Gaussian linear systems. However, for the systems disturbed by non-Gaussian noises, the performance of the Kalman filter decreases dramatically.

A number of non-Gaussian filtering methods were studied, including particle filtering (PF), Student's t-distribution filtering, etc. Particle filter achieves approximate state estimation by characterizing the distribution of states in terms of random sampling points through Monte Carlo sampling techniques [2, 3]. Student's t distribution filter achieves higher filtering accuracy by modeling the noise as a Student's t distribution and exploiting the thick-tailed nature of this distribution compared to a Gaussian distribution [4]. Maximum correntropy Kalman filter (MCKF) measures the size of the error through the Gaussian kernel function, and iteratively adjusts the filter gain to improve the filtering accuracy [5–10]. The properties of the Gaussian kernel function cause the weights to decrease rapidly to zero when the error increases, which makes the maximum correntropy Kalman filter more robust to extreme outliers. However, it can lead to the occurrence of singularity phenomena, when the absolute error larger than a certain extent. It deteriorates the estimation accuracy. Improved Kalman filter based on Huber loss function (HKF) improves the filtering accuracy by minimizing robust loss function [11–14]. Coincidentally, maximizing the entropy value can be translated into minimizing the robust loss function [15]. On average, the filtering method with the Huber loss function takes fewer iterations than the maximum correntropy Kalman filter. This is because

the Huber loss function does not adjust the gain when the absolute error is less than a set threshold. But Kalman filter based on Huber loss function shows weak robustness to extreme outliers.

In this paper, Huber correntropy loss function is defined, to balance the filtering accuracy and the real-time performance. To solve the new loss function, a novel non-Gaussian filtering method is designed in the frame of Kalman filter, with the help of fixed-point iteration theory. It is named Huber correntropy Kalman filter and short of HCKF. The final simulation results show the advantages of HCKF.

## 2 Problem Formulation

Consider the following linear state space model

$$\mathbf{x}_k = \mathbf{F}_k \mathbf{x}_{k-1} + \mathbf{w}_{k,k-1} \quad (1)$$

$$\mathbf{y}_k = \mathbf{H}_k \mathbf{x}_k + \mathbf{v}_k \quad (2)$$

in which,  $\mathbf{x}_k \in \mathbb{R}^n$  is system state,  $\mathbf{y}_k \in \mathbb{R}^m$  is measurement,  $\mathbf{F}_k \in \mathbb{R}^{n \times n}$  is the state transfer matrix, and  $\mathbf{H}_k \in \mathbb{R}^{m \times n}$  is measurement matrix.  $\mathbf{w}_{k,k-1} \in \mathbb{R}^n$  is non-Gaussian process noise with covariance  $\mathbf{Q}_{k,k-1}$ ,  $\mathbf{v}_k \in \mathbb{R}^m$  is non-Gaussian measurement noise with covariance  $\mathbf{R}_k$ .

For state estimation of above non-Gaussian linear system, Some effective filtering methods have been proposed. However, in these existing methods, it is difficult to balance filtering accuracy and the real-time performance. In this paper, a novel non-Gaussian filtering method named HCKF with better performance of real-time and filtering accuracy is studied.

## 3 Huber Correntropy Kalman Filter

### 3.1 Huber Correntropy Loss Function

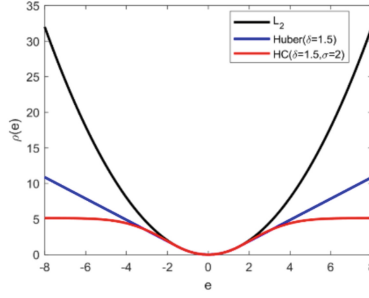
The mathematical expression of Huber correntropy loss function is

$$\rho(e) = \begin{cases} 0.5e^2, & |e| \leq \delta \\ \Phi(e), & |e| > \delta \end{cases} \quad (3)$$

where  $\delta$  is the set threshold,  $\Phi(e) = \frac{2\sigma^2 + \delta^2}{2} \left[ 1 - \frac{2\sigma^2}{2\sigma^2 + \delta^2} \exp(-\frac{e^2 - \delta^2}{2\sigma^2}) \right]$ ,  $\sigma$  is the kernel size of  $\Phi(e)$ .

In Huber correntropy loss function, when the absolute error is less than the set threshold, the loss function behaves as  $L_2$  loss function. And when the absolute error exceeds the set threshold, the  $L_2$  loss function is replaced by the Gaussian kernel function.

Figure 1 shows the trends of the  $L_2$  loss function, the Huber loss function, the Huber correntropy loss function with errors. From the figure, it can be seen that as the error increases,  $L_2$  loss function grows the fastest, Huber loss function is the second fastest, and Huber correntropy loss function is the slowest, which can reflect that Huber correntropy loss function is able to effectively suppress the large absolute error compared with the other two loss function, and has the strongest robustness.



**Fig. 1.** Relationship between different loss functions and errors

### 3.2 Huber Correntropy Kalman Filter

For the system shown in (1) and (2), the state prediction vector  $\hat{\mathbf{x}}_{k|k-1}$  and the prediction error covariance matrix  $\mathbf{P}_{k|k-1}$  at  $k$  are

$$\begin{cases} \hat{\mathbf{x}}_{k|k-1} = \mathbf{F}_k \hat{\mathbf{x}}_{k-1|k-1} \\ \mathbf{P}_{k|k-1} = \mathbf{F}_k \mathbf{P}_{k-1|k-1} \mathbf{F}_k^T + \mathbf{Q}_{k,k-1} \end{cases} \quad (4)$$

The following cost function of Huber correntropy loss function is defined by

$$\mathbf{J}(\mathbf{x}_k) = \frac{1}{n+m} \sum_{i=1}^{n+m} \rho(e_i), \quad (5)$$

where  $e_i = d_i - \mathbf{w}_i \mathbf{x}_k$ ,  $e_i$  is the  $i^{\text{th}}$  element of  $\mathbf{e}_k$ ,  $\mathbf{w}_i$  is the  $i^{\text{th}}$  row of  $\mathbf{W}_k$ ,  $d_i$  is the  $i^{\text{th}}$  element of  $\mathbf{D}_k$ , and

$$\begin{cases} \mathbf{D}_k = (\mathbf{B}_k)^{-1} [\hat{\mathbf{x}}_{k|k-1}, \mathbf{y}_k]^T \\ \mathbf{W}_k = (\mathbf{B}_k)^{-1} [\mathbf{I}, \mathbf{H}_k]^T \\ \mathbf{e}_k = (\mathbf{B}_k)^{-1} \mathbf{\Omega}_k \\ \mathbf{\Omega}_k = [\hat{\mathbf{x}}_{k|k-1} - \mathbf{x}_k, \mathbf{v}_k]^T \\ \mathbf{B}_k = \text{diag}(\text{chol}(\mathbf{P}_{k|k-1}), \text{chol}(\mathbf{R}_k)) \end{cases} \quad (6)$$

$\text{chol}(\ast)$  expresses Cholesky decomposition.

The optimal estimate of  $\mathbf{x}_k$  is obtained by solving the minimum value of (5), i.e.,

$$\hat{\mathbf{x}}_k = \arg \min_{\mathbf{x}_k} \sum_{i=1}^{n+m} \rho(e_i). \quad (7)$$

The solution takes the following form

$$\mathbf{x}_k = \hat{\mathbf{x}}_{k|k-1} + \tilde{\mathbf{K}}_k (\mathbf{y}_k - \mathbf{H}_k \hat{\mathbf{x}}_{k|k-1}), \quad (8)$$



where

$$\begin{cases} \tilde{\mathbf{K}}_k = \tilde{\mathbf{P}}_{k|k-1} \mathbf{H}_k^T (\mathbf{H}_k \tilde{\mathbf{P}}_{k|k-1} \mathbf{H}_k^T + \tilde{\mathbf{R}}_k)^{-1} \\ \tilde{\mathbf{P}}_{k|k-1} = \text{chol}(\mathbf{P}_{k|k-1}) (\mathbf{M}_x)^{-1} (\text{chol}(\mathbf{P}_{k|k-1}))^T \\ \tilde{\mathbf{R}}_k = \text{chol}(\mathbf{R}_k) (\mathbf{M}_y)^{-1} (\text{chol}(\mathbf{R}_k))^T \\ \mathbf{M}_x = \text{diag}(\frac{\partial \rho(e_1)}{\partial \mathbf{x}_k} \frac{1}{e_1}, \dots, \frac{\partial \rho(e_n)}{\partial \mathbf{x}_k} \frac{1}{e_n}) \\ \mathbf{M}_y = \text{diag}(\frac{\partial \rho(e_{n+1})}{\partial \mathbf{x}_k} \frac{1}{e_{n+1}}, \dots, \frac{\partial \rho(e_{n+m})}{\partial \mathbf{x}_k} \frac{1}{e_{n+m}}) \\ e_i = d_i - \mathbf{w}_i \mathbf{x}_k \end{cases} \quad (9)$$

From (9), it is clear that (8) is a fixed-point equation. It can be solved by HCKF, which is summarized as follows.

- 1) Choose appropriate parameters  $\delta$  and  $\sigma$  and a small positive number  $\varepsilon$ . Denote  $\hat{\mathbf{x}}_{k|k}^t$  as the state estimate at the  $t^{\text{th}}$  iteration. Let  $\hat{\mathbf{x}}_{k|k}^0 = \mathbf{x}_{k|k-1}$ .
- 2) Use (4) to obtain  $\hat{\mathbf{x}}_{k|k-1}$  and  $\mathbf{P}_{k|k-1}$ .
- 3) Use (10)–(16) to compute  $\hat{\mathbf{x}}_{k|k}^t$ .

$$\hat{\mathbf{x}}_{k|k}^t = \hat{\mathbf{x}}_{k|k-1} + \bar{\mathbf{K}}_k (\mathbf{y}_k - \mathbf{H}_k \hat{\mathbf{x}}_{k|k-1}), \quad (10)$$

with

$$\bar{\mathbf{K}}_k = \bar{\mathbf{P}}_{k|k-1} \mathbf{H}_k^T (\mathbf{H}_k \bar{\mathbf{P}}_{k|k-1} \mathbf{H}_k^T + \bar{\mathbf{R}}_k)^{-1}, \quad (11)$$

$$\bar{\mathbf{P}}_{k|k-1} = \text{chol}(\mathbf{P}_{k|k-1}) (\mathbf{M}_x)^{-1} (\text{chol}(\mathbf{P}_{k|k-1}))^T, \quad (12)$$

$$\bar{\mathbf{R}}_k = \text{chol}(\mathbf{R}_k) (\mathbf{M}_y)^{-1} (\text{chol}(\mathbf{R}_k))^T, \quad (13)$$

$$\mathbf{M}_x = \text{diag}(\frac{\partial \rho(e_1)}{\partial \mathbf{x}_k} \frac{1}{e_1}, \dots, \frac{\partial \rho(e_n)}{\partial \mathbf{x}_k} \frac{1}{e_n}), \quad (14)$$

$$\mathbf{M}_y = \text{diag}(\frac{\partial \rho(e_{n+1})}{\partial \mathbf{x}_k} \frac{1}{e_{n+1}}, \dots, \frac{\partial \rho(e_{n+m})}{\partial \mathbf{x}_k} \frac{1}{e_{n+m}}), \quad (15)$$

$$e_i = d_i - \mathbf{w}_i \mathbf{x}_{k|k}^{t-1}. \quad (16)$$

- 4) If (17) is valid, set  $\hat{\mathbf{x}}_{k|k} = \hat{\mathbf{x}}_{k|k}^t$  and move on to step 5). Otherwise, let  $t = t + 1$  and return to 3).

$$\frac{\|\hat{\mathbf{x}}_{k|k}^t - \hat{\mathbf{x}}_{k|k}^{t-1}\|}{\|\hat{\mathbf{x}}_{k|k}^{t-1}\|} \leq \varepsilon. \quad (17)$$

- 5) Use (18) update the posterior covariance matrix. Let  $k = k + 1$  and go back to 2).

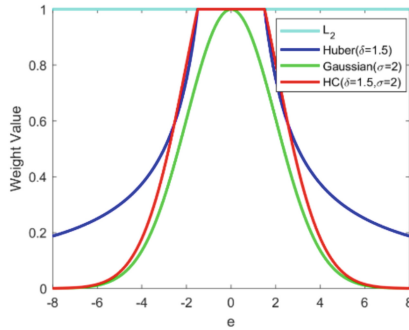
$$\mathbf{P}_{k|k} = (\mathbf{I} - \bar{\mathbf{K}}_k \mathbf{H}_k) \mathbf{P}_{k|k-1} (\mathbf{I} - \bar{\mathbf{K}}_k \mathbf{H}_k)^T + \bar{\mathbf{K}}_k \mathbf{R}_k \bar{\mathbf{K}}_k^T. \quad (18)$$

### 3.3 Performance Metric Analysis

Table 1 shows four kinds of loss functions. By solving these loss functions, four different filtering methods can be obtained: KF, HKF, MCKF and HCKF, respectively. The third column in the Table 1 shows the relevant weighting functions for the prediction error covariance and the measurement noise covariance. These functions are utilized to modify the filtering gains in the four filtering methods mentioned above. Performances of different filtering methods can be analyzed by weighting functions in the next content. Figure 2 shows the variation of different weighting function values with error.

**Table 1.** Four loss functions and their weighting functions

| Name of loss functions | Loss functions $\rho(e)$                                                                         | Weighting functions $\eta(e)$                                                                               |
|------------------------|--------------------------------------------------------------------------------------------------|-------------------------------------------------------------------------------------------------------------|
| $L_2$                  | $0.5e^2$                                                                                         | 1                                                                                                           |
| Huber                  | $\begin{cases} 0.5e^2, &  e  \leq \delta \\ \delta( e  - 0.5\delta), &  e  > \delta \end{cases}$ | $\begin{cases} 1, &  e  \leq \delta \\ \delta/ e , &  e  > \delta \end{cases}$                              |
| Gaussian kernel        | $\exp(-\frac{e^2}{2\sigma^2})$                                                                   | $\exp(-\frac{e^2}{2\sigma^2})$                                                                              |
| Huber correntropy      | $\begin{cases} 0.5e^2, &  e  \leq \delta \\ \Phi(e), &  e  > \delta \end{cases}$                 | $\begin{cases} 1, &  e  \leq \delta \\ \exp(-\frac{e^2 - \delta^2}{2\sigma^2}), &  e  > \delta \end{cases}$ |



**Fig. 2.** Relationship between different weighting functions and errors

Figure 2 shows that the weight function of the  $L_2$  loss function always remains the same value, no matter how the error changes, which leads to filtering accuracy of corresponding Kalman filter performs very bad to outliers. As the error increases, the trend of the weight function of Huber correntropy loss function is between the Gaussian kernel loss function and the Huber loss function. When the error is small, the weight value decreases slightly, and when the error increases, the weight decreases drastically. The weight value of Huber correntropy loss function slows down to zero compared to the Gaussian kernel function, which makes it possible to improve the filtering accuracy

as well as avoid the appearance of singularity phenomena. In addition, when the error decreases to threshold  $\delta$ , Huber correntropy loss function, as the Huber loss function, stops changing the weights. It improves the real-time performance of the algorithm by speeding up the convergence of the fixed-point iterations.

## 4 Simulation

The uniform circular motion model is utilized to verify the effectiveness of HCKF.

$$\begin{cases} \begin{bmatrix} x_k^1 \\ x_k^2 \end{bmatrix} = \begin{bmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{bmatrix} \begin{bmatrix} x_{k-1}^1 \\ x_{k-1}^2 \end{bmatrix} + \mathbf{w}_{k,k-1} \\ \mathbf{y}_k = [0.5 \ 0.5] \begin{bmatrix} x_k^1 \\ x_k^2 \end{bmatrix} + \mathbf{v}_k \end{cases}, \quad (19)$$

where  $\theta = \pi/18$ , and simulation steps  $N$  is taken as 100. Initial state  $\mathbf{x}_0 = [10, 10]^T$ ,  $\mathbf{P}_{0|0} = \text{diag}(0.6, 0.6)$ .

The proposed algorithm's filtering performance under non-Gaussian noise is to be shown. For this, the process noise and the observation noise takes

$$\begin{cases} \mathbf{w}_{k,k-1} \sim \text{Laplace}(0, 0.04) \\ \mathbf{v}_k \sim \text{St}(0, 5) \end{cases} \quad (20)$$

In this simulation, KF, MCKF with different parameters, HKF, and HCKF proposed in this paper are compared. The mean absolute error (MAE), average number of iterations (ANI), and average CPU running time (ART) under 2000 Monte Carlo and  $\varepsilon = 10^{-6}$ . Simulation results are shown in Tables 2 and 3, Figs. 3 and 4.

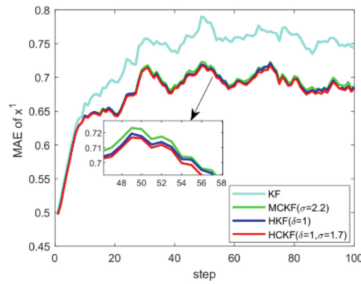
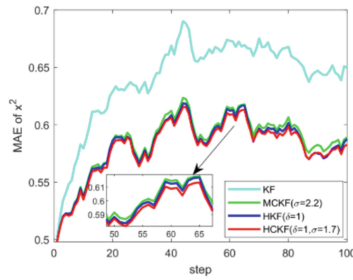
**Table 2.** MAEs of  $x^1$  and  $x^2$

| Filter                             | MAE of $x^1$ | MAE of $x^2$ |
|------------------------------------|--------------|--------------|
| KF                                 | 0.72992      | 0.64074      |
| MCKF( $\sigma = 2$ )               | 0.68045      | 0.58585      |
| MCKF( $\sigma = 2.2$ )             | 0.68002      | 0.58528      |
| MCKF( $\sigma = 2.4$ )             | 0.68011      | 0.58534      |
| HKF( $\delta = 1$ )                | 0.67750      | 0.58238      |
| HCKF( $\delta = 1, \sigma = 1.5$ ) | 0.67622      | 0.58094      |
| HCKF( $\delta = 1, \sigma = 1.7$ ) | 0.67583      | 0.58040      |
| HCKF( $\delta = 1, \sigma = 2$ )   | 0.67625      | 0.58081      |

As in Table 2, and Table 3, although KF has the shortest computation time, its filtering error is the largest. The proposed HCKF ( $\delta = 1, \sigma = 1.7$ ) has the highest

**Table 3.** ANIs and ARTs of different filters

| Filter                             | ANI    | ART( $10^{-5}s$ ) |
|------------------------------------|--------|-------------------|
| KF                                 | 0      | 1.340             |
| MCKF( $\sigma = 2$ )               | 3.2057 | 10.522            |
| MCKF( $\sigma = 2.2$ )             | 3.1504 | 10.918            |
| MCKF( $\sigma = 2.4$ )             | 3.1036 | 10.408            |
| HKF( $\delta = 1$ )                | 2.4981 | 9.539             |
| HCKF( $\delta = 1, \sigma = 1.5$ ) | 2.4814 | 8.928             |
| HCKF( $\delta = 1, \sigma = 1.7$ ) | 2.4571 | 8.829             |
| HCKF( $\delta = 1, \sigma = 2$ )   | 2.4373 | 8.673             |

**Fig. 3.** MAE of  $x^1$  of different methods**Fig. 4.** MAE of  $x^2$  of different methods

filtering accuracy among all methods, the shortest ART and least ANI except KF. The filtering accuracy, ANI and ART of HCKF with other parameters are also better than MCKF and HKF. Therefore, HCKF has better performances both in filtering accuracy and real-time.

## 5 Conclusion

In this paper, a novel loss function called Huber correntropy is proposed and applied in the Kalman filter framework to cope with the non-Gaussian linear filtering problem. A novel Huber correntropy loss function is defined, under which, Huber correntropy Kalman filter is derived, and its performance is analyzed. The proposed method is simulated and verified to perform better than MCKF and HKF. In the future, filtering problems of non-Gaussian nonlinear systems will be studied under Huber correntropy loss function.

## References

1. Kalman, R.E.: A new approach to linear filtering and prediction problems. *J. Basic Eng.* **82**, 35 (1960)
2. Ahwiadi, M., et al.: An adaptive particle filter technique for system state estimation and prognosis. *IEEE Trans. Instrum. Meas.* **69**, 6756–6765 (2020)
3. Huang, P., et al.: Improved particle filter based on kernel density estimation. In: 2021 IEEE 4th International Conference on Electronics Technology (ICET), pp. 1015–1019 (2021)
4. Huang, Y., et al.: A novel robust student's t-based Kalman filter. *IEEE Trans. Aerosp. Electron. Syst.* **53**(3), 1545–1554 (2017)
5. Chen, B., et al.: Maximum correntropy Kalman filter. *Automatica* **76**, 70–77 (2017)
6. Izanloo, R., et al.: Kalman filtering based on the maximum correntropy criterion in the presence of non-Gaussian noise. In: 2016 Annual Conference on Information Science and Systems (CISS) (2016)
7. Chen, B., et al.: Convergence of a fixed-point algorithm under maximum correntropy criterion. *IEEE Signal Process. Lett.* **22**(10), 1723–1727 (2015)
8. Singh, R.K., et al.: Maximum correntropy polynomial chaos Kalman filter for underwater navigation. *Digit. Signal Process.* **155**, 104774 (2024)
9. Liu, S., et al.: Robust maximum correntropy criterion based square-root rotating lattice Kalman filter. *SIViP* **18**(8–9), 6041–6053 (2024)
10. Song, X., et al.: Maximum correntropy Kalman filter for linear discrete-time systems with intermittent observations and non-gaussian noise. *IEEE Trans. Circuits Syst. II Express Briefs* **71**(6), 3246–3250 (2024)
11. Sun, W., et al.: RCKF cooperative navigation algorithm for tightly coupled vehicle ad hoc networks based on Huber M estimation. *IEEE Access* **9**, 139888–139895 (2021)
12. Elsis, M., et al.: Robust Kalman filter for position estimation of automated guided vehicles under cyberattacks. *IEEE Trans. Instrum. Meas.* **72**, 1–12 (2023)
13. Sahl, S., et al.: Robust cubature Kalman filter for moving-target tracking with missing measurements. *Sensors* **24**(2), 392 (2024)
14. Zhao, J., et al.: A study on the model of robust fractional-order extended Kalman filtering with gross error. *GPS Solutions* **28**, 2 (2024)
15. Principe, J. C.: *Information Theoretic Learning: Renyi's Entropy and Kernel Perspectives*. Springer, New York (2010)

# **Intelligent Cooperative Control**



# A Framework for Visual Target Navigation for Quadcopter Based on Large Language Models in Unknown Environment

Yunzhuo Liu, Zhaowei Ma, Yuqi Yang, Mengyun Wang, and Yifeng Niu()

College of Intelligence Science and Technology, National University of Defense  
Technology, Changsha 410073, China

niuyifeng@nudt.edu.cn

**Abstract.** This paper proposes a framework for visual target navigation in unmanned aerial vehicles (UAVs) that leverages LLMs. By utilizing the powerful language understanding capabilities and prior knowledge of LLMs, this framework enhances UAVs' comprehension of natural language instructions and their ability to plan tasks. Initially, the method integrates task instructions, contextual backgrounds, and the UAV's functional library through designed prompt information. Subsequently, LLM serves as a planner, generating long-term plans that align with the task instructions and selecting corresponding functional modules based on the UAV's capabilities to execute control. Leveraging the strong reasoning abilities of LLMs, we derive a sequence for searching rooms. Then, based on visual observations and prior probabilities, we determine the locations of sub-goals. Through a path planning algorithm, the UAV progressively searches forward until the target is discovered. Experimental results demonstrate the effectiveness of the proposed method in addressing visual target navigation problems for UAV.

**Keywords:** LLM · Visual Target Navigation · Task Planning · UAV

## 1 Introduction

Compared to ground robots, quadcopters have advantages such as rapid response, high maneuverability, and more comprehensive vision when performing search tasks in unknown environments. In unknown environment, the quadrotor needs to perceive the environment through sensors such as cameras and perform search planning. If a drone is needed to replace a person to search in an unknown and dangerous environment, the person needs to clearly input specific search tasks including destinations and target objects. However, it is often not easy for robots to locate any specific target based on human needs in an unknown environment [1]. This leads to a classic problem in robotics visual target navigation, which involves autonomous search tasks in unknown environments and

requires robots to possess language understanding and semantic reasoning abilities [2,3]. Traditional approaches typically involve constructing an environmental map based on visual observations and then planning the robot’s navigation path based on the constructed geometric map [4–6]. Due to the lack of prior knowledge of the scene, these methods often require searching large areas. Learning-based methods directly optimize navigation strategies, using deep reinforcement learning to find objects of specified categories in unknown environments based on currently observable images [7–9]. However, these methods focus on how to explore the environment and navigate the robot after obtaining accurate target information, neglecting how to extract key information from user-defined instructions and whether prior information from the physical world can be utilized to enhance the robot’s efficiency.

Humans can easily and quickly determine the approximate location of a target in an unknown environment, thanks to strong prior knowledge and logical reasoning abilities about the physical world. For example, if asked to find a water cup, we usually know it is often located near a water dispenser, meaning we only need to search in the living room rather than the kitchen or bedroom. If robots could possess such ability, it would save a considerable amount of time. At the same time, human’s task instructions are often complex, and even a simple instruction may require robot to perform numerous operations. How robots can quickly and effectively understand human’s intentions and execute them is also an issue that needs to be addressed [10]. This poses high demands on a robot’s information extraction capabilities and level of intelligence.

The emergence of large language models, such as BERT [11], GPT-3 [12], and GPT-4 [13], has made this idea feasible. Their impressive performance upon release amazed people with their language understanding capabilities, enabling them to effectively comprehend human intentions and provide corresponding responses. Additionally, thanks to their vast prior knowledge of the physical world, they can offer suggestions and assistance based on their own knowledge. Consequently, recent researches have begun to integrate LLMs with robots, exploring ways to leverage the powerful abilities of LLMs in the field of robotics to enhance their versatility and intelligence. This includes utilizing LLMs for robots to understand complex human instructions and plan tasks [14,15], converting specified targets in natural language into structured planning language to generate operable plans for agents [16,17], writing Python code to manipulate robotic arms [18], and performing motion planning for autonomous vehicles [19]. Recent studies have also applied LLMs to visual target navigation tasks, demonstrating their effectiveness in this context. Xie et al. [20] improved the path generation of Rapidly-exploring Random Tree (RRT) [25] by introducing LLM for semantic relevance measurement to achieve outdoor navigation for UAVs. Chen et al. [21] utilized the prior knowledge of LLM for semantic inference to obtain scene labels, achieving 3D scene understanding for robots. Yu et al. [22] introduced LLM into the problem of visual target navigation, enabling the use of language to find relevant boundaries from semantic maps as long-term goals and effectively explore the environment. However, these researches have over-



looked the understanding of human task instructions, while the most prominent capability of large language models is natural language processing.

The focus of this study is how to generate long-term plans based on human’s task instructions to achieve UAV control in search tasks, while using the prior knowledge of LLM as the basis for selecting search directions. Specifically, we first generate a long-term plan using LLM, then use visual language model to generate scene descriptions of the surrounding environment, and subsequently generate real-time sub-goal points based on these descriptions to guide the UAV in searching for the target. Additionally, the long-term plans we generate are composed of Python code, enabling effective UAV control.

Our contributions are:

1. We designed a hierarchical prompt structure that integrates scene prior information, task backgrounds, UAV functional library information, and output instance information into the prompt. Experimental validation was conducted under different task forms, exploring and obtaining prompt descriptions with multi-task generalization.
2. Based on traditional task processes, we introduced large models to improve task generalization capabilities. Meanwhile, we designed a function library that aligns with UAV control interfaces for downstream tasks, enabling large models to control UAVs.

This paper proposes a new method for combining large language models with UAV visual target navigation tasks. We will discuss this method in Sect. 2 and conduct experiments to demonstrate its effectiveness in Sect. 3.

## 2 Method

For the problem of UAV visual target navigation based on natural language instructions in unknown environments, a UAV task planning method is proposed. In this paper, we embed LLM into the planning process as a search order planner and a long-term task planner. We also establish a UAV function library based on task requirements, matching the steps of the task plan to generate executable control code.

Figure 1 illustrates the overall framework of our work, we use a scenario where a drone explores and searches for target inside a house as an example. The human provides the drone with task instructions containing the target location. The drone inputs the user instructions, task description, and UAV function library into LLM to generate a task plan that aligns with the user instructions and invokes the matching functions in the library to control the UAV. Meanwhile, this method is extensible and generalizable. Firstly, it can modify the functions of each submodule for different downstream tasks to meet various task requirements. Secondly, the prompt input for large models and the UAV function library can be customized and expanded according to the task context, forming a versatile execution unit.

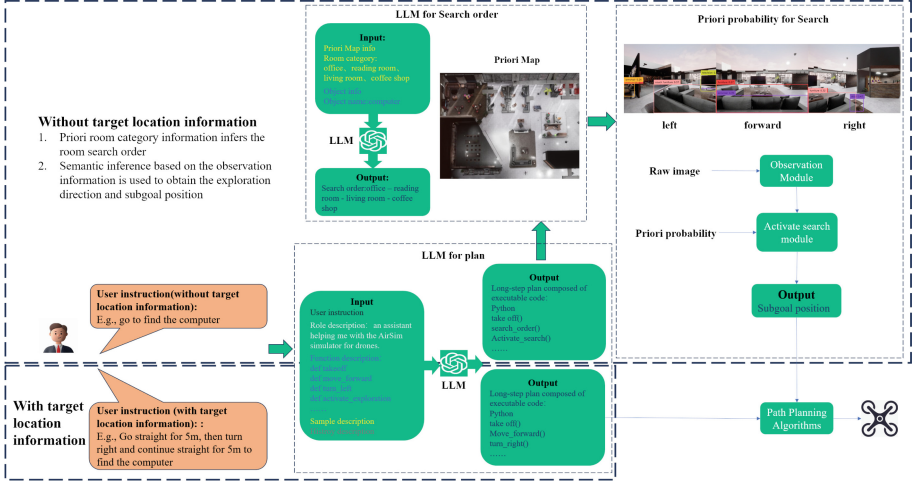


Fig. 1. system overview of our work

## 2.1 LLM for Plan

LLMs are typically utilized as question-answering tools to respond to various user inquiries. However, enabling them to truly understand user needs poses a challenge. A common solution is to standardize user expressions through prompt engineering. Therefore, designing an effective prompt based on task requirements is crucial for leveraging LLMs effectively. We directly utilize powerful LLM as reasoning modules, designing a standardized prompt to infer all information and generate high-level plans. Specifically, by providing LLM with the premises and basis for reasoning, and then based on the currently obtained key information, we formulate a long-term executable plan in conjunction with their own capabilities.

In this paper, GPT-3.5-turbo provided by OpenAI is selected as the task planner, and referring to the work of Vemprala et al. [26], the elements of a good prompt should include:

**Task Command.** The original task instruction provided by the human, which in this paper is in English, such as movement, target detection and search for target.

**Role Description.** The role played by LLM, clarifying the task requirements and context for the drone. Our experiments are conducted based on AirSim, so the role of LLM is an assistant helping me with the AirSim simulator for drones. We hope that LLM provides Python code to fulfill task requirements.

**Function Description.** Since the text output of LLM cannot directly control the drone, we hope that the direct output of the large model are API functions

related to drone control, i.e., letting LLM generate Python functions based on the API function library to control the drone. API functions are mainly divided into High-Level actions and Low-Level actions.

**Low-Level actions:** Low-level actions refer to the most basic behaviors, i.e., simple callable functions whose input parameters are determined by the system. As shown in Fig. 2, they include many fundamental behaviors, such as drone control, vision tools, user interfaces, and LLM queries.

**High-Level actions:** High-level actions typically consist of multiple functions, i.e., nested functions, selectively incorporating many low-level actions. As shown in Fig. 2, they include many high-level behaviors, such as path planning, active exploration, and target recognition.

```

1. High-Level action:
• ob_objects(obj_name_list) - runs the detection model on the obj_name_list and returns three variables - phrases, which is a list of the names of the objects detected in the scene.final_obj_list, which is a list of the specific information of the objects detected in the scene such as obj_name, camera_distance, depth_distance, angel_degree, center_x and center_y.annotated_frame, which is a cv2 image with the bounding box of each detected object.
• RRT_fly(forward_distance,right_distance,obj_name) - flies the drone to the target area to search for the target object, which is represented by the forward distance and right distance of the drone,the forward distance corresponds to the parameter forward_distance and the right distance corresponds to the parameter right_distance.The target object is represented by obj_name.
• active_explore(obj_name) - Active exploration to search for objects when not informed of the approximate location of the target object to be searched for.The target object is represented by obj_name.
.....

2. Low-Level action:
• takeoff() - takes off the drone.
• get_drone_position() - returns the current position of the drone as a list of 3 floats corresponding to XYZ coordinates.
• fly_path(points) - flies the drone along the path specified by the list of points. Each point is again a list of 3 floats corresponding to X, Y, Z coordinates.
• turn_left() - Rotates 10° to the left from the current angle.
• get_image() - renders an image from the drone's front camera
.....

```

**Fig. 2.** API function library example

Theoretically, we can follow the approach outlined in paper [18] to use LLM directly to generate Python code for controlling the drone. However, due to the hallucination phenomenon of LLMs, this may result in the generation of non-executable code, leading to loss of control over the drone, which is highly risky in practical operations. Therefore, we use LLM to combine the rich skill set in the prompts, thereby simplifying the plan and enhancing its quality, while also reducing the service cost of LLM.

**Sample Description.** Examples of “prompt-response” from LLM, clearly demonstrating the expected outcome of the reasoning process in the form of examples, to better enable LLM to understand human’s intent and further standardize their output.

## 2.2 Priori Information for Search

Without a cue indicating the target location, the UAV cannot quickly determine the target’s position and inevitably needs to conduct a purposeless search. Humans can easily and quickly ascertain the approximate location of a target in an unknown environment. This is attributed to humans’ robust prior knowledge

of the physical world and logical reasoning abilities. Inspired by this, we leverage the extensive prior knowledge of LLMs by providing them with all room types and the names of the targets to be searched in the environment. This information is used to generate a search sequence for the rooms, i.e., the drone should prioritize searching rooms where the target is most likely to be present. This paper also selects GPT-3.5-turbo provided by OpenAI as the Q&A tool for the search sequence, with the prompt construction similar to that in Sect. 2.1. Since there is no need for an output function, the prompt does not include Function Description.

During task execution, the UAV obtains scene descriptions through the visual perception module and then matches the visual observations with room categories to determine the search direction, always prioritizing searching the room where the target is most likely to appear.

### 2.2.1 Observation Module

To achieve autonomous navigation in unknown environments, it is crucial to perceive raw observations from the environment and extract vital information for downstream reasoning. The required perceptive abilities consist of two aspects: one is the scene recognition capability, which refers to the ability to recognize any objects within the scene. The second is mapping capability, which refers to the ability to locate objects within the scene and use this as a basis for path planning.

To achieve the aforementioned two capabilities, two locally deployed visual language models are integrated into the framework:

1. Recognize-anything Model [28]: This model is used to construct scene descriptions, specifically the names of objects within the scene. It takes RGB images as input and outputs the names of the objects in the image.
2. GroundingDINO Model [29]: This model is used to obtain the names of objects within the scene and their corresponding bounding boxes. It takes RGB images and a list of object names as input, and outputs the bounding boxes for the respective object names.

The module utilizes the two aforementioned models to obtain the names of objects within the scene and their bounding boxes, and combines information from the depth image. It calculates the positions of the objects through coordinate transformation. The primary function of this module is to use RGB and depth images to construct a semantic map, recording the discovered objects and their corresponding positions.

### 2.2.2 Active Search Strategy

The problem of visual target navigation needs to solve the problem of which place the agent should explore first when the environmental map has not been fully established. Firstly, we utilize LLM to determine a preliminary room exploration sequence based on the search target and prior map information. This approach acts as pruning, avoiding unnecessary exploration processes while serving as a

decision-making basis for the next stage. During the search process, observations are conducted in the front, left, and right directions, and the likelihood rankings of room categories in these three directions are obtained based on prior probabilities to determine the location of the next sub-goal. Finally, path planning is employed to search in the direction of the sub-goal.

To quantify the likelihood of room categories, we measure it by the prior probability of object *a* appearing in room *b*. Specifically, we construct a “room category-object name” prior probability list based on 3,000 images and their recognition results [27], which includes most common room and object names. Prior probabilities, i.e., the prior probability of object *a* in room *b*, are calculated using statistical methods. After obtaining scene descriptions in the front, left, and right directions, entropy is calculated using the following formula to obtain the likelihood rankings  $o_1, o_2, o_3$  of room labels in each direction. To mitigate the influence of uninformative and ubiquitous objects (such as doors and windows), we calculate the entropy for each category, as suggested in [22].

$$H_{O_i} = - \sum_{t_j \in L_T} p(t_j | o_i) \log p(t_j | o_i) \quad (1)$$

where  $o_i \in L_O$  represents the target category and  $t_j \in L_T$  represents the scene category. The probability of each target category appearing in the corresponding scene is normalized to obtain  $p(t_j | o_i)$ . From the formula, it can be seen that objects with greater information have lower  $H_{O_i}$ . Therefore, by sorting  $H_{O_i}$ , we can infer the room labels and rank the possibilities of other labels. Ultimately, we obtain the likelihood rankings of room labels in the three directions and determine which direction to proceed based on the room exploration sequence outlined in Sect. 2.2.

### 3 Experiments

In this section, we design a series of simulation experiments to evaluate the performance of the proposed method, aiming to answer the following question: 1) Can the framework stably and efficiently complete visual target navigation tasks? 2) Can the framework effectively match with multiple algorithms?

#### 3.1 Experimental Setup

We validate our proposed method by setting up six task instructions ranging from simple to complex, as shown in the Table 1. Task 1 is the most basic instruction, used to verify the feasibility of controlling a drone using a large language model. Task 2–4 are for verifying the functionality of each submodule. Task 5–6 are advanced search tasks involving the cooperative work among various submodules, used to verify the feasibility of our proposed method in visual target navigation tasks. For each task, we conducted 25 experiments separately to calculate and analyze task completion.

**Table 1.** six tasks

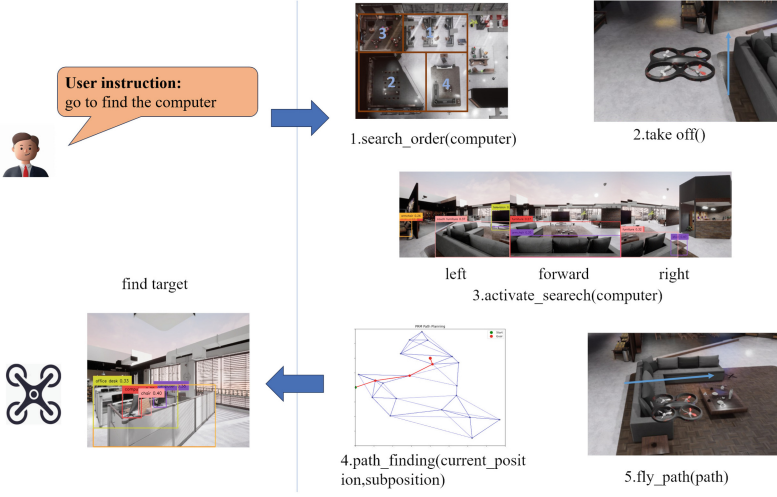
| Task ID | Task Name                                                 | Task Instruction                                                          |
|---------|-----------------------------------------------------------|---------------------------------------------------------------------------|
| 1       | Fly forward                                               | Fly 5 m forward                                                           |
| 2       | Path planning                                             | Fly to (10,5) using path planning algorithms                              |
| 3       | Detect and location                                       | Tell me what you see now and give their locations                         |
| 4       | Search order                                              | Suppose you will search for a computer, give me the room search order     |
| 5       | Search for the target with target location information    | The computer is located within a 10m area in front of you, go and find it |
| 6       | Search for the target without target location information | Go to find the computer                                                   |

Due to the occasional instability in the output of LLM, which can lead to undefined or unexecutable generated code, and an increasing error rate as task complexity increases, we evaluate performance using both code executability and task success rates. Code executability is directly assessed by whether the generated code can run normally; if it cannot, the code is considered unexecutable and the task is deemed a failure. For Tasks 1–2, success rate is evaluated based on whether the target location is reached. For Task 3, success rate is evaluated based on accurate target recognition and localization. For Task 4, success rate is evaluated through the comparison of the room search order with the actual target location. For Tasks 5–6, success is directly evaluated based on whether the target is found.

### 3.2 Simulation and Results

We validated our method in the open-source, cross-platform simulator AirSim, which is based on the Unreal Engine (UE) and supports software-in-the-loop simulation using flight controllers such as PX4 and ArduPilot. It also supports hardware-in-the-loop simulation based on PX4. This simulation platform is suitable for various applications and research areas related to machine learning and artificial intelligence. The Fig. 3 shows an example of searching for a computer without location hint. Based on prior room category information, a pre-determined search order is generated. Then, based on visual observations, the degree of matching with room categories is used to generate sub-goal points. Through path planning algorithms, the UAV gradually approaches and finally find the target.

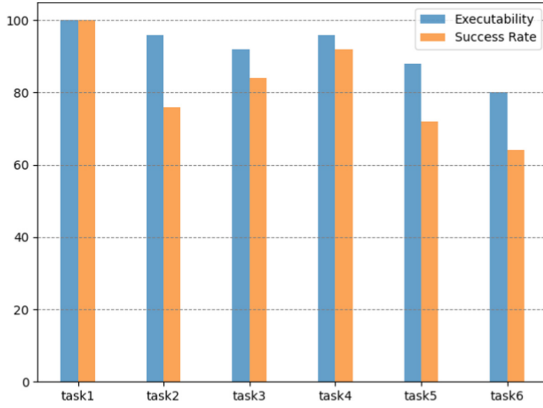
Tasks' success rates are shown in the Fig. 4. The results indicate that our proposed method can effectively execute the aforementioned tasks. As task complexity increases, LLM occasionally makes errors leading to unexecutable generated code. This is due to the increasing number of library functions and other



**Fig. 3.** Example of searching for computer

necessary functions that need to be used, which leads to issues including but not limited to incorrect input parameter formats for functions and mismatches between the output of necessary functions and the format of library functions. The success rates of Tasks 5–6 are influenced by many factors, including code unexecutability due to instability in LLM’s outputs, incomplete API functions leading to collisions with the environment, and misdetections by the perception module causing changes in task progress, such as misdetections due to perspective deviations and collisions due to misjudgments of obstacle size and position. These issues can be observed through the success rates of each submodule, but they can be addressed by replacing corresponding submodules to improve their capabilities and thus increase success rates.

**Ablation Experiments.** To verify the practicality and versatility of our proposed method, we conducted ablation experiments by replacing different path planning algorithms, such as Rapidly-exploring Random Tree (RRT) [25], Probabilistic Roadmap (PRM) [24], and Carrot-Chasing [23]. The results in Table 2 show that our proposed method can effectively apply to various path planning algorithms, and different algorithms have an impact on success rates. The Carrot-chasing algorithm often collides with obstacles due to its lack of obstacle avoidance capabilities, resulting in a low success rate. On the other hand, the RRT algorithm and PRM algorithm possess certain obstacle avoidance abilities, but their success rates are also not high due to limitations in mapping accuracy and algorithm parameter settings.



**Fig. 4.** Executability and Success rate of six tasks

**Table 2.** Executability and Success Rate of different algorithms

| Path Planning Algorithm | Task Name | Executability | Success Rate |
|-------------------------|-----------|---------------|--------------|
| Baseline(RRT)           | Task2     | 96%           | 76%          |
|                         | Task5     | 88%           | 72%          |
|                         | Task6     | 80%           | 64%          |
| PRM                     | Task2     | 100%          | 72%          |
|                         | Task5     | 92%           | 64%          |
|                         | Task6     | 84%           | 60%          |
| Carrot-chasing          | Task2     | 100%          | 48%          |
|                         | Task5     | 88%           | 20%          |
|                         | Task6     | 76%           | 12%          |

## 4 Conclusion

We propose a framework that combines large language models with visual target navigation tasks. As a planner, LLM can receive natural language instructions and directly return a plan based on task requirements. As a pre-planner, LLM can well understand the task context and generate a pre-determined search order to guide the search direction. Experimental results show that this method can effectively be applied to visual target navigation problems, achieving efficient operation from high-level planning to low-level control under any natural language instruction. In the future, our work can achieve a complete closed-loop control process by using reinforcement learning or utilizing LLM itself as a checker for plan generation, thereby reducing the uncertainty in large language model outputs and improving the stability of the framework.



**Acknowledgement.** This work was supported in part by the National Natural Science Foundation of China under Grant 62303481, and in part by the Postgraduate Scientific Research Innovation Project of Hunan Province under Grant CX20240005.

## References

1. Chen, J., Li, G., Kumar, S., et al.: How to not train your dragon: training-free embodied object goal navigation with semantic frontiers. arXiv preprint [arXiv:2305.16925](#) (2023)
2. Chaplot, D., Gandhi, D., Gupta, A., et al.: Object goal navigation using goal-oriented semantic exploration. In: *Advances in Neural Information Processing Systems*, vol. 33, pp. 4247–4258 (2020)
3. Gervet, T., Chintala, S., Batra, D., et al.: Navigating to objects in the real world. *Sci. Robot.* **8**(79), eadf6991 (2023)
4. Yamauchi, B.: A frontier-based approach for autonomous exploration. In: *Proceedings 1997 IEEE International Symposium on Computational Intelligence in Robotics and Automation, CIRA 1997. Towards New Computational Principles for Robotics and Automation*, pp. 146–151. IEEE (1997)
5. Dharmadhikari, M., Dang, T., Solanka, L., et al.: Motion primitives-based path planning for fast and agile exploration using aerial robots. In: *2020 IEEE International Conference on Robotics and Automation (ICRA)*, pp. 179–185 (2020)
6. Wang, C., Zhu, D., Li, T., et al.: Efficient autonomous robotic exploration with semantic road map in indoor environments. *IEEE Robot. Autom. Lett.* **4**(3), 2989–2996 (2019)
7. Yang, W., Wang, X., Farhadi, A., et al.: Visual semantic navigation using scene priors. In: *7th International Conference on Learning Representations (ICLR)*, pp. 1–14 (2019)
8. Druon, R., Yoshiyasu, Y., Kanazaki, A., et al.: Visual object search by learning spatial context. *IEEE Robot. Autom. Lett.* **5**(2), 1279–1286 (2020)
9. Ye, J., Batra, D., Das, A., et al.: Auxiliary tasks and exploration enable object-goal navigation. In: *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pp. 16117–16126 (2021)
10. Stone, A., Xiao, T., Lu, Y., et al.: Open-world object manipulation using pre-trained vision-language models. arXiv preprint [arXiv:2303.00905](#) (2023)
11. Devlin, J., Chang, M.W., Lee, K., et al.: BERT: pre-training of deep bidirectional transformers for language understanding. arXiv preprint [arXiv:1810.04805](#) (2018)
12. Brown, T., Mann, B., Ryder, N., et al.: Language models are few-shot learners. In: *Advances in Neural Information Processing Systems*, vol. 33, pp. 1877–1901 (2020)
13. Achiam, J., Adler, S., Agarwal, S., et al.: GPT-4 technical report. arXiv preprint [arXiv:2303.08774](#) (2023)
14. Liu, B., Jiang, Y., Zhang, X., et al.: LLM+P: empowering large language models with optimal planning proficiency. arXiv preprint [arXiv:2304.11477](#) (2023)
15. Naoki, W., Atsushi, K., Kazuhiro, S., et al.: ChatGPT empowered long-step robot control in various environments: a case application. *IEEE Access* **11**, 95060–95078 (2023)
16. Silver, T., Hariprasad, V., Shuttleworth, R.S., et al.: PDDL planning with pre-trained large language models. In: *NeurIPS 2022 Foundation Models for Decision Making Workshop* (2022)

17. Xie, Y., Yu, C., Zhu, T., et al.: Translating natural language to planning goals with large language models. arXiv preprint [arXiv:2302.05128](#) (2023)
18. Liang, J., Huang, W., Xia, F., et al.: Code as policies: language model programs for embodied control. In: 2023 IEEE International Conference on Robotics and Automation (ICRA), pp. 9493–9500. IEEE (2023)
19. Mao, J., Qian, Y., Zhao, H., Wang, Y., et al.: GPT-driver: learning to drive with GPT. arXiv preprint [arXiv:2310.01415](#) (2023)
20. Xie, Q., Zhang, T., Xu, K., et al.: Reasoning about the unseen for efficient outdoor object navigation. arXiv preprint [arXiv:2309.10103](#) (2023)
21. Chen, W., Hu, S., Talak, R., et al.: Leveraging large (visual) language models for robot 3D scene understanding. arXiv preprint [arXiv:2209.05629](#) (2023)
22. Yu, B., Kasaei, H., Cao, M.: L3MVN: leveraging large language models for visual target navigation. In: 2023 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), pp. 3554–3560. IEEE (2023)
23. Bhadani, R.: Path planning of unmanned system using carrot-chasing algorithm. arXiv preprint [arXiv:2012.13227](#) (2020)
24. LaValle, S.: Rapidly-exploring random trees: a new tool for path planning. Research report 9811 (1998)
25. Kavraki, L.E., Svestka, P., Latombe, J.C., et al.: Probabilistic roadmaps for path planning in high-dimensional configuration spaces. IEEE Trans. Robot. Autom. **12**(4), 566–580 (1996)
26. Vemprala, S.H., Bonatti, R., Bucker, A., et al.: ChatGPT for robotics: design principles and model abilities. IEEE Access (2024)
27. Zhou, B., Zhao, H., Puig, X., et al.: Scene parsing through ADE20K dataset. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 633–641 (2017)
28. Zhang, Y., Huang, X., Ma, J., et al.: Recognize anything: a strong image tagging model. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pp. 1724–1732 (2024)
29. Liu, S., Zeng, Z., Ren, T., et al.: Grounding dino: marrying dino with grounded pre-training for open-set object detection. arXiv preprint [arXiv:2303.05499](#) (2023)



# Adaptive Output Feedback Control for Nonlinear Time-Delay Systems with Neural Network and Static Gains

Wenjie Li<sup>1</sup>  and Zhengqiang Zhang<sup>2</sup>  

<sup>1</sup> Qufu Normal University, Qufu 273165, China

<sup>2</sup> Qufu Normal University, Rizhao 276826, China

qufuzzq@126.com

**Abstract.** This paper solves the adaptive output feedback control problem for a class of nonlinear time-delay system with neural network and static gains. First, a reduced-order state observer is designed through the variable substitution. Then, the unknown continuous functions and system states are suppressed by the neural networks and static gains. To stabilize the system, an output feedback controller is designed by Lyapunov-Krasovskii functional. Finally, the effectiveness of this control scheme is verified by a simulation example.

**Keywords:** Time-delay system · adaptive control · output feedback control · neural network · static gain

## 1 Introduction

Nonlinear characteristics exist in many practical control systems, which makes the design of controllers complicated or even impossible [1]. Since the backstepping technique can eliminate the constraint that system uncertainties need to meet the matching condition, adaptive control can estimate and suppress the influence of system uncertainties while ensuring control performance, these two control tools are widely used in control fields such as mechanical equipment manufacturing and chemical process [2, 3]. However, time delay is not considered in the above work.

Actually, time delay commonly emerges in practical systems, which will cause the control performance to deteriorate, and even the instability of the system [4]. Hence, the control and stabilization problem of nonlinear time-delay system is drawing more and more attention and many accomplishments have been made. The most commonly used control method for uncertain time-delay nonlinear systems is to use a radial basis function neural network (RBF NN) to compensate for the system adaptively [5, 6]. However, the above work are based on the measurable state, which is not easy to meet. Therefore, output feedback control is regarded as a more universal control scheme, and reduced-order observer as a powerful tool for estimating unmeasurable states, such as [7].

Since the introduction of static gain function does not add unnecessary subsystems and effectively offsets the influence of unmeasurable system states, the static gain function method has achieved some results in control field [8]. Subsequently, scholars made a series of improvements to an explicit expression of static gain function [9]. In this paper, we are devoted to solving the problem of adaptive output feedback control based on static gain. The contributions are i) In order to reduce the error caused by the approximation of unknown functions by neural networks, static gains are introduced to handle the unknown system states. ii) New Lyapunov-Krasovskii functional embedded with static gains are designed to compensate for unknown time delays.

## 2 The Problem Statement and Preliminaries

### 2.1 System Statement

In this section, the nonlinear time-delay system is described as

$$\begin{cases} \dot{x}_i = x_{i+1} + \varphi_i(\bar{x}_i) + \psi_i(\bar{x}_i(t - \chi_i(t))), \\ \dot{x}_n = u + \varphi_n(\bar{x}_n) + \psi_n(\bar{x}_n(t - \chi_n(t))), \\ x(t) = \phi(t), \quad -\chi' \leq t \leq 0, \\ y = x_1, \end{cases} \quad (1)$$

where  $x = [x_1, x_2, \dots, x_n]^T \in R^n$  is the state,  $y(t)$  is the output,  $u(t)$  is the controller.  $\varphi_i(\cdot)$ ,  $\psi_i(\cdot)$  are unknown continuous functions, and satisfy  $\varphi_i(0) = \psi_i(0) = 0$ .  $\chi_i(t)$  is the time-varying delay, satisfies  $0 \leq \chi_i(t) \leq \chi^*$  and  $\dot{\chi}_i(t) \leq \bar{\chi}_i < 1$ ,  $\bar{\chi} = \max\{\bar{\chi}_1, \bar{\chi}_2, \dots, \bar{\chi}_n\}$ .  $\phi(t)$  is initial continuous function.

*Assumption 1.*  $\psi_i(\cdot)$  is an unknown continuous function, satisfies  $\psi_i^2(\bar{x}_i(t - \chi_i(t))) \leq \sum_{s=1}^i \gamma_s(x_s(t - \chi_s(t)))|x_s(t - \chi_s(t))|$ , where  $\gamma_s(\cdot) > 0$  is an unknown continuous function.

### 2.2 RBFNN and Projection Operator

**RBFNN.** Unknown continuous function  $h(Z) : R^l \rightarrow R$  over  $\Omega_Z \in R^l$  can be approximated as  $\hat{h}(Z) = H^{*T}P(Z)$ , where  $H^*$  is the weight vector,  $P(Z)$  is the vector-valued function, and  $Z \in \Omega_Z \subset R^l$ ,  $H^* = [h_1^*, \dots, h_q^*]^T$ ,  $P(Z) = [p_1(Z), \dots, p_q(Z)]^T \in R^q$  with  $p_i(Z) = \exp\left[\frac{-(Z - \bar{u}_i)^2}{\nu_i^2}\right]$ .

For arbitrary continuous function  $h(Z)$  over  $\Omega_Z$ , we have  $h(Z) = H^T P(Z) + \delta(Z)$ , where  $\delta(Z) \leq \delta'$  is the error function.  $H$  is chosen as  $H := \arg\min_{H^* \in R^q} \{\sup_{Z \in \Omega_Z} |h(Z) - H^{*T} P(Z)|\}$ .

### Projection

**Operator.** For  $\Omega = \{H : \|H\| \leq h_0\}$ , we define  $\text{Proj}(\cdot)$  as  $\dot{\hat{H}}_i = \text{Proj}(\mu_i, \hat{H}_i) = \mu_i - \frac{\eta_{i1}\eta_{i2}}{4(\varepsilon^2 + 2\varepsilon h_0)^{n-i} h_0^2} \nabla p(\hat{H}_i)$ , where  $\eta_{i1} = \begin{cases} p^{n-i}(\hat{H}_i), & \text{if } p(\hat{H}_i) > 0 \\ 0, & \text{otherwise} \end{cases}$ ,  $\eta_{i2} = \|\hat{H}_i\|^2 \|\mu_i\|^2 + \frac{1}{4}$ ,  $p(\hat{H}_i) = \|\hat{H}_i\|^2 - h_0^2$ ,  $\hat{H}_i$  is the estimate function of  $H_i$  with  $\tilde{H}_i = H_i - \hat{H}_i$ , and  $\mu_i$  is  $(n - i)$  times continuously differentiable function.

### 3 Controller Design and Stability Analysis

#### 3.1 Reduced-Order Observer

First, we rewrite system (1) as

$$\begin{cases} \dot{x}_1 = Ax_1 + Bm' + \varphi_1(x_1) + \psi_1, \\ \dot{m}' = Cx_1 + Dm' + Eu + \bar{\varphi} + \bar{\psi}, \\ y(t) = x_1, \end{cases} \quad (2)$$

where  $m' = [x_2, \dots, x_n]^T$  denotes the unmeasured system state vector,  $A, B, C, D, E, \bar{\varphi}, \bar{\psi}$  are matrices and vectors of corresponding dimensions.

To construct a reduced-order observer, we introduce  $v = Cx_1 + Eu$ ,  $z = \dot{x}_1 - Ax_1$ . The observer for the transformed system is designed as  $\dot{\hat{m}}' = A'\hat{m}' + v + Lz$ , where  $L = [l_2, l_3, \dots, l_n]^T$  is a designed constant vector such that  $A' = D - LB$  is a Hurwitz matrix. Furthermore,  $A'^T P + PA' = -M$  holds by appropriate positive-definite matrix  $P, M$ . By defining  $m = \hat{m}' - Ly$  and  $B' = (C - LA) + A'L \triangleq [b'_2, b'_3, \dots, b'_n]$ , one has

$$\dot{m} = A'm + B'y + Eu. \quad (3)$$

The estimation error is  $\varepsilon = m' - \hat{m}' = [\varepsilon_2, \varepsilon_3, \dots, \varepsilon_n]^T$ , and its dynamic is  $\dot{\varepsilon} = A'\varepsilon - L(\varphi_1(x_1) + \psi_1) + \bar{\varphi} + \bar{\psi}$ .

We chose the Lyapunov function as  $V_0 = \varepsilon^T P \varepsilon + M_0$ , with  $M_0 = \frac{e^{\varsigma\chi^*}}{\iota(1-\bar{\chi})} \sum_{i=1}^n \int_{t-\chi_i(t)}^t e^{-\varsigma(t-s)} (n-i+1) \left( \frac{1}{\bar{\iota}_i} x_i^2(s) + \bar{\iota}_i \gamma_i^2(x_i(s)) \right) ds$ , where  $\varsigma, \iota, \bar{\iota}_i$  are positive parameters. Using Young's inequality and Assumption 1, one has

$$\begin{aligned} \dot{V}_0 \leq & -\varepsilon^T M \varepsilon + \iota \varepsilon^T P L L^T P \varepsilon + \iota \varepsilon^T P P \varepsilon + \frac{2}{\iota} \varphi_1^2(x_1(t)) - \varsigma M_0 + \frac{2}{\iota} \|\bar{\varphi}\|^2 \\ & + \frac{e^{\varsigma\chi^*}}{\iota(1-\bar{\chi})} \sum_{i=1}^n \bar{\iota}_i (n-i+1) \gamma_i^2(x_i(t)) + \frac{e^{\varsigma\chi^*}}{\iota(1-\bar{\chi})} \sum_{i=1}^n \frac{(n-i+1)}{\bar{\iota}_i} x_i^2(t). \end{aligned} \quad (4)$$

#### 3.2 Controller Design

The backstepping technique is chosen for the controller design, the change of coordinate is designed as

$$z_1 = y, z_i = m_i - \alpha_{i-1}, i = 2, 3, \dots, n. \quad (5)$$

Define  $\Gamma_i$  as symmetric positive definite matrix,  $\varsigma_i, v_i, \kappa_i, \vartheta_i, \xi_i, \theta_i, \varpi, \zeta'$  as positive constants, and the recursive steps are as follows.

**Step 1.** From (2) and (5), one has  $\dot{z}_1(t) = \varepsilon_2 + z_2 + l_2 z_1 + \alpha_1 + \varphi_1(x_1) + \psi_1$ . The Lyapunov function is defined as  $\dot{V}_1 = \frac{1}{2} z_1^2 + \frac{1}{2} \tilde{H}_1^T \Gamma_1^{-1} \tilde{H}_1 + M_1$ , where  $M_1 = \frac{e^{\varsigma_1\chi^*} \xi_1}{1-\bar{\chi}} \int_{t-\chi_1(t)}^t e^{-\varsigma_1(t-s)} \left( \gamma_1^2(x_1(s)) + x_1^2(s) \right) ds$ , By defining  $h_1(z_1) = \varphi_1(z_1)$ ,

$h_1(z_1) = H_1^T P_1(z_1) + \delta_1(z_1)$ . From the inequality  $z_1 \delta_1(z_1) \leq \frac{v_1}{2} z_1^2 + \frac{1}{2v_1} \delta_1'^2$ , and Assumption 1, we have

$$\begin{aligned} \dot{V}_1 \leq & z_1 \left[ \left( \frac{1}{2\kappa_1} + \frac{1}{2\vartheta_1} + \frac{1}{2\xi_1} + \frac{v_1}{2} + \frac{\xi_1 e^{\varsigma_1 \chi^*}}{(1-\bar{\chi})} \right) z_1 + l_2 z_1 + \alpha_1 + \hat{H}_1^T P_1(z_1) \right] - \varsigma_1 M_1 \\ & + \tilde{H}_1^T (z_1 P_1(z_1) - \Gamma_1^{-1} \dot{\hat{H}}_1) + \frac{\vartheta_1}{2} \|\varepsilon\|^2 + \frac{\kappa_1}{2} z_2^2 + \frac{\xi_1 e^{\varsigma_1 \chi^*}}{(1-\bar{\chi})} \gamma_1^2(x_1(t)) + \frac{1}{2v_1} \delta_1'^2. \end{aligned} \quad (6)$$

The first stabilizing function and the adaptive law are designed as

$$\begin{aligned} \alpha_1 = & - \left( \theta_1 + \frac{1}{2\kappa_1} + \frac{1}{2\vartheta_1} + \frac{1}{2\xi_1} + \frac{v_1}{2} + \frac{\xi_1 e^{\varsigma_1 \chi^*}}{(1-\bar{\chi})} \right) z_1 - l_2 z_1 - \hat{H}_1^T P_1(z_1) \\ & - \frac{n e^{\varsigma \chi^*}}{\bar{\iota} \bar{\iota}_1 (1-\bar{\chi})} z_1 - \sum_{i=2}^n \frac{\xi_i e^{\varsigma_i \chi^*}}{(1-\bar{\chi})} z_1 - \sum_{i=2}^n \frac{4(n-i+1) e^{\varsigma \chi^*}}{\bar{\iota} \bar{\iota}_i (1-\bar{\chi})} l_i^2 z_1, \end{aligned} \quad (7)$$

$$\mu_1 = \Gamma_1 z_1 P_1(z_1), \dot{\hat{H}}_1 = \text{Proj}[\mu_1, \dot{\hat{H}}_1]. \quad (8)$$

The Lyapunov functional candidate is defined as

$$V_1 = V_0 + \dot{V}_1, \quad (9)$$

together with (6)–(8), it's time derivative satisfies

$$\begin{aligned} \dot{V}_1 \leq & -\Theta_1 V_1 + \frac{e^{\varsigma \chi^*}}{\bar{\iota} (1-\bar{\chi})} \sum_{i=2}^n \frac{(n-i+1)}{\bar{\iota}_i} x_i^2(t) - (\lambda - \lambda_1) \|\varepsilon\|^2 + \frac{\kappa_1}{2} z_2^2 \\ & - \sum_{i=2}^n \frac{\xi_i e^{\varsigma_i \chi^*}}{(1-\bar{\chi})} z_1^2 - \sum_{i=2}^n \frac{4(n-i+1) e^{\varsigma \chi^*}}{\bar{\iota} \bar{\iota}_i (1-\bar{\chi})} l_i^2 z_1^2 + \aleph_1 + g_1(t), \end{aligned} \quad (10)$$

where  $M - \iota P L L^T P^T + \iota P^T P - \lambda I = Q$ ,  $Q$  is a positive-defined matrix.  $\Theta_1 = \min\{\Theta'_1, \frac{\lambda_{\min}(Q)}{\lambda_{\max}(P)}, \varsigma\}$ ,  $g_1(t)$  is a combination of unknown functions,  $\lambda_1 = \frac{\vartheta_1}{2}$ ,  $\aleph_1 = \frac{1}{2v_1} \delta_1'^2 + \frac{1}{2} H_m^2$ .

From (7), we have  $|\alpha_1| \leq \ell_{11} |z_1| + \ell_{10}$ ,  $\left| \frac{\partial \alpha_1}{\partial z_1} \right| \leq \ell_{z11}$ ,  $\left\| \frac{\partial \alpha_1}{\partial \hat{H}_1^T} \right\| \leq \|P_1(z_1)\| \triangleq \ell_{h11}$ , where  $\ell_{11}$ ,  $\ell_{10}$ ,  $\ell_{z11}$ ,  $\ell_{h11}$  are positive constants. We define the static gain as

$$\gamma_1 = \max \left\{ \frac{16(n-1) e^{\varsigma \chi^*} \ell_{11}^2}{\bar{\iota} \bar{\iota}_2 (1-\bar{\chi})}, 1 \right\}. \quad (11)$$

**Step  $q$ ,**  $q = 2, \dots, n-1$ . According to the principle of recursive process, and (2), (3), (5), we have

$$\begin{aligned} \dot{z}_q = & -l_q m_2 + z_{q+1} + \alpha_q + b'_q y - \frac{\partial \alpha_{q-1}}{\partial z_1} (\varepsilon_2 + m_2 + l_2 y + \varphi_1(x_1(t)) + \psi_1) \\ & - \sum_{i=2}^{q-1} \frac{\partial \alpha_{q-1}}{\partial m_i} \dot{m}_i - \sum_{i=1}^{q-1} \frac{\partial \alpha_{q-1}}{\partial \hat{H}_i^T} \dot{\hat{H}}_i. \end{aligned} \quad (12)$$

The Lyapunov function in step  $q$  is defined as  $\dot{V}_q = \frac{1}{2}z_q^2 + \frac{1}{2}\tilde{H}_q^T \Gamma_q^{-1} \tilde{H}_q + M_q$ , where  $M_q = \frac{e^{\varsigma_q \chi^*} \xi_q}{1-\bar{\chi}} \int_{t-\chi_1(t)}^t e^{-\varsigma_q(t-s)} \gamma_1^2(x_1(s)) + x_1^2(s) ds$ .

Defining  $h_q(z_1) = \varphi_1(x_1(t))$ , we have  $h_q(z_1) = H_q^T P_q(z_1) + \delta_q(z_1)$ . Defining the virtual control law and the adaptive law as

$$\begin{aligned} \alpha_q = & -\theta_q z_q + l_q m_2 - b'_q y - \frac{4(n-q+1)e^{\varsigma_q \chi^*}}{\bar{\iota}_q(1-\bar{\chi})} z_q - \frac{v_q}{2} z_q - \frac{\kappa_{q-1}}{2} z_q - \frac{\Upsilon_{q-1} \kappa_{q-1}}{2} z_q \\ & - \frac{1}{2\kappa_q} z_q + \frac{\partial \alpha_{q-1}}{\partial z_1} (m_2 + l_2 y) + \frac{\partial \alpha_{q-1}}{\partial z_1} \hat{H}_q^T P_q(z_1) - \frac{1}{2\vartheta_q} \left( \frac{\partial \alpha_{q-1}}{\partial z_1} \right)^2 z_q \\ & - \frac{1}{2\xi_q} \left( \frac{\partial \alpha_{q-1}}{\partial z_1} \right)^2 z_q + \sum_{i=2}^{q-1} \frac{\partial \alpha_{q-1}}{\partial m_i} \dot{m}_i + \sum_{i=1}^{q-1} \frac{\partial \alpha_{q-1}}{\partial \hat{H}_i^T} \dot{\hat{H}}_i, \end{aligned} \quad (13)$$

$$\mu_q = z_q \Gamma_q \frac{\partial \alpha_{q-1}}{\partial z_1} P_q(z_1), \dot{\hat{H}}_q = \text{Proj}[\mu_q, \dot{\hat{H}}_q]. \quad (14)$$

Then, we have

$$\begin{aligned} \dot{V}_q \leq & -\theta_q z_q^2 - \frac{\kappa_{q-1}}{2} z_q^2 - \frac{\Upsilon_{q-1} \kappa_{q-1}}{2} z_q^2 - \frac{4(n-q+1)e^{\varsigma_q \chi^*}}{\bar{\iota}_q(1-\bar{\chi})} z_q^2 + \frac{\kappa_q}{2} z_{q+1}^2 + \frac{\vartheta_q}{2} \|\varepsilon\|^2 \\ & - \varsigma_q M_q + \frac{e^{\varsigma_q \chi^*} \xi_q}{1-\bar{\chi}} \gamma_1^2(x_1(t)) + \frac{e^{\varsigma_q \chi^*} \xi_q}{1-\bar{\chi}} x_1^2(t) + \frac{1}{2v_q} \delta_q'^2. \end{aligned} \quad (15)$$

The static gain is defined as

$$\Upsilon_q = \max_{j=1, \dots, q} \left\{ \frac{8(q+1)(n-q)e^{\varsigma_q \chi^*}}{\bar{\iota}_{q+1}(1-\bar{\chi})} \frac{\ell_{qj}^2}{\prod_{j=1}^{q-1} \Upsilon_j}, \frac{1}{\prod_{j=1}^{q-1} \Upsilon_j} \right\}, \quad (16)$$

where  $\prod_{j=1}^{q-1} \Upsilon_j \triangleq 1$ .

The Lyapunov candidate is defined as  $V_q = \dot{V}_q + V_{q-1}$ , where  $\dot{V}_q = \dot{V}_q + \Upsilon_{q-1} \dot{V}_{q-1}$ . The time derivative of  $V_q$  satisfies

$$\begin{aligned} \dot{V}_q \leq & -\Theta_q V_q - (\lambda - \lambda_q) \|\varepsilon\|^2 + \frac{\kappa_q}{2} z_{q+1}^2 + \frac{e^{\varsigma_q \chi^*}}{\bar{\iota}(1-\bar{\chi})} \sum_{i=q+1}^n \frac{(n-i+1)}{\bar{\iota}_i} x_i^2(t) \\ & - \sum_{i=q+1}^n \frac{\xi_i e^{\varsigma_i \chi^*}}{(1-\bar{\chi})} z_1^2 + g_q(t) + \aleph_q - \sum_{i=q+1}^n \frac{4(n-i+1)e^{\varsigma_q \chi^*}}{\bar{\iota}_i(1-\bar{\chi})} l_i^2 z_1^2, \end{aligned} \quad (17)$$

where  $g_q(t)$  is a combination of unknown functions,  $\Theta_q$ ,  $\lambda_q$  and  $\aleph_q$  are positive constants.

**Step  $n$ .** The Lyapunov function is  $\dot{V}_n = \frac{1}{2}z_n^2 + \frac{1}{2}\tilde{H}_n^T \Gamma_n^{-1} \tilde{H}_n + \frac{1}{2}\tilde{H}^T \Gamma^{-1} \tilde{H} + M_n$ , where  $M_n = \frac{e^{\varsigma_n \chi^*} \xi_n}{1-\bar{\chi}} \int_{t-\chi_1(t)}^t e^{-\varsigma_n(t-s)} \gamma_1^2(x_1(s)) + x_1^2(s) ds$ ,  $\tilde{H} = H - \hat{H}$ . Defining  $h_n(z_1) = \varphi_1(x_1(t))$ , one has  $h_n(z_1) = H_n^T P_n(z_1) + \delta_n(z_1)$ . Together with  $z_n \delta_n(z_1) \leq \frac{v_n}{2} z_n^2 + \frac{1}{2v_n} \delta_n'^2$ , we have

$$\begin{aligned}
\dot{V}_n \leq & z_n \left[ -l_n m_2 + u + b'_n y + \frac{v_n}{2} z_n - \frac{\partial \alpha_{n-1}}{\partial z_1} (m_2 + l_2 y) - \frac{\partial \alpha_{n-1}}{\partial z_1} \hat{H}_n^T P_n(z_1) \right. \\
& + \frac{1}{2\vartheta_n} \left( \frac{\partial \alpha_{n-1}}{\partial z_1} \right)^2 z_n + \frac{1}{2\xi_n} \left( \frac{\partial \alpha_{n-1}}{\partial z_1} \right)^2 z_n - \sum_{i=2}^{n-1} \frac{\partial \alpha_{n-1}}{\partial m_i} \dot{m}_i - \sum_{i=1}^{n-1} \frac{\partial \alpha_{n-1}}{\partial \hat{H}_i^T} \dot{\hat{H}}_i \Big] \\
& + \tilde{H}_n^T \left( -\frac{\partial \alpha_{n-1}}{\partial z_1} P_n(z_1) z_n - \Gamma_n^{-1} \dot{\hat{H}}_n \right) + \frac{\vartheta_n}{2} \|\varepsilon\|^2 + \dot{M}_n + \xi_n \gamma_1^2 (x_1(t) - \chi_1(t)) \\
& + \xi_n x_1^2(t - \chi_1(t)) + \frac{1}{2v_n} \delta_n'^2 - \tilde{H}^T \Gamma^{-1} \dot{\hat{H}}. \tag{18}
\end{aligned}$$

The actual controller and adaption law are given as

$$\begin{aligned}
u = & -\theta_n z_n + l_n m_2 - b'_n y - \frac{v_n}{2} z_n - \frac{4e^{\varsigma\chi^*}}{\bar{\iota}_n(1-\bar{\chi})} z_n - \frac{\kappa_{n-1}}{2} z_n - \frac{\Upsilon_{n-1}\kappa_{n-1}}{2} z_n \\
& + \frac{\partial \alpha_{n-1}}{\partial z_1} (m_2 + l_2 y) + \frac{\partial \alpha_{n-1}}{\partial z_1} \hat{H}_n^T P_n(z_1) - \frac{1}{2\vartheta_n} \left( \frac{\partial \alpha_{n-1}}{\partial z_1} \right)^2 z_n \\
& - \frac{1}{2\xi_n} \left( \frac{\partial \alpha_{n-1}}{\partial z_1} \right)^2 z_n - \frac{z_n}{z_n^2 + \varpi} \hat{H}^T P(\hat{X}) + \sum_{i=2}^{n-1} \frac{\partial \alpha_{n-1}}{\partial m_i} \dot{m}_i + \sum_{i=1}^{n-1} \frac{\partial \alpha_{n-1}}{\partial \hat{H}_i^T} \dot{\hat{H}}_i, \tag{19}
\end{aligned}$$

$$\mu_n = z_n \Gamma_n \frac{\partial \alpha_{n-1}}{\partial z_1} P_n(z_1), \dot{\hat{H}}_n = \text{Proj}[\mu_n, \dot{\hat{H}}_n]. \tag{20}$$

Chose the Lyapunov-Krasovskii function as

$$V_n = \dot{V}_n + V_{n-1}, \tag{21}$$

where  $\dot{V}_n = \dot{V}_n + \Upsilon_{n-1} \dot{V}_{n-1}$ . Thus,  $\dot{V}_n$  satisfies

$$\begin{aligned}
\dot{V}_n \leq & -\theta_n z_n^2 - \frac{1}{2} \tilde{H}_n^T \tilde{H}_n - \varsigma_n M_n - \Upsilon_{n-1} \Theta'_{n-1} \dot{V}_{n-1} - \Theta_{n-1} V_{n-1} - (\lambda - \lambda_n) \|\varepsilon\|^2 \\
& + \Upsilon_{n-1} \dot{V}_{n-1} - \tilde{H}^T \Gamma^{-1} \dot{\hat{H}} + g_n - \frac{z_n^2}{z_n^2 + \varpi} \hat{H}^T P(\hat{X}) + \aleph_{n-1} + \frac{4ne^{\varsigma\chi^*}}{\iota(1-\bar{\chi})} \ell_{n-1,0}^2 \\
& + \sum_{j=1}^n \left( \prod_{s=j}^{n-1} \Upsilon_s \right) \left( \frac{1}{2} H_m^2 + \frac{1}{2v_j} \delta_j'^2 \right), \tag{22}
\end{aligned}$$

where  $\lambda_n = \lambda_{n-1} + \frac{1}{2} \left( \sum_{j=1}^n \left( \prod_{s=j}^{n-1} \Upsilon_s \right) \vartheta_j \right) + \frac{4e^{\varsigma\chi^*}}{\bar{\iota}_n(1-\bar{\chi})}$ ,  $g_n(t)$  is an unknown function. Then, there exists vector variables  $X = [x_1, \dots, x_n]$ ,  $\hat{X} = [x_1, m'_2, \dots, m'_n]$  such that  $g_n(t) \leq H^T P(\hat{X}) + \zeta'$ .



The adaptive law is designed as

$$\dot{\hat{H}} = \Gamma \frac{z_n^2}{z_n^2 + \varpi} P(\hat{X}) - \Gamma \zeta \hat{H}. \quad (23)$$

Therefore, we have

$$\dot{V}_n \leq -\Theta_n V_n + \aleph_n, \quad (24)$$

where  $\Theta_n, \aleph_n$  are positive constants.

**Theorem 1.** If system (1) with time delay satisfies Assumption 1, all the closed loop signals can be kept bonded by utilizing static gains (11), (16), virtual control laws (7), (13), adaptive laws (8), (14), (20), and actual controller (19).

*Proof.* By integrating (24) over  $[0, t)$ , we have  $V_n(t) \leq V_n(0)e^{-\Theta_n t} + \frac{\aleph_n}{\Theta_n}(1 - e^{-\Theta_n t})$ . we have  $V_n(t) \geq \varepsilon^T P \varepsilon + \frac{1}{2} \sum_{i=1}^n z_i^2 + \frac{1}{2} \sum_{i=1}^n \tilde{H}_i^T \Gamma_i^{-1} \tilde{H}_i + \frac{1}{2} \tilde{H}^T \Gamma^{-1} \tilde{H} + \sum_{i=1}^n M_i$ , which implies the boundedness of  $z_i, \varepsilon_i, \hat{H}$ . Furthermore, we have  $\hat{H}_i, i = 1, 2, \dots, n$  are bounded. Therefore, we have  $\alpha_1$  is bounded from (7). According to the recursive analysis in this way, we can get the boundedness of  $m_i, m'_i$ . Finally, the boundedness of actual control law can be obtained from (19). Thus, all the closed-loop signals are bonded.

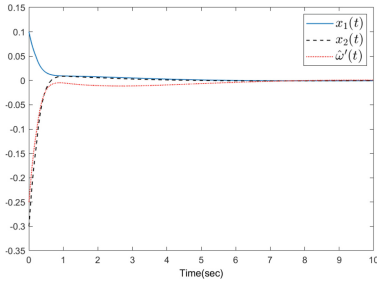
## 4 Simulation Example

$$\begin{cases} \dot{x}_1(t) = x_2 - 0.5x_1 - 0.5x_1(t - \chi_1(t)), \\ \dot{x}_2 = u + 0.5\sin(x_1) + 0.5\sin(x_1(t - \chi_2(t))), \\ y = x_1, \end{cases}$$

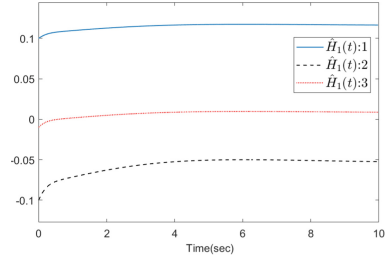
where  $\chi_1(t) = \chi_2(t) = 0.2(\sin^2(t) + 1)$ .

The designed parameters and initial values are chosen as  $l_2 = 1, P = 1, \lambda = 1.9, \theta_1 = \theta_2 = 0.2, \kappa_1 = 2, \vartheta_1 = 1, \vartheta_2 = 1, \xi_1 = 0.5, \xi_2 = 0.1, \varsigma = \varsigma_1 = \varsigma_2 = 0.1, v_1 = 0.3, v_2 = 0.5, \bar{l}_1 = 10, \bar{l}_2 = 20, \iota = 15, \Gamma_1 = I, \Gamma_2 = 1.5I, \Gamma = 1.5I, \zeta = 1, \Upsilon_1 = 1.622, x_1(t) = 0.1, x_2(t) = -0.3, m(t) = -0.35$ .

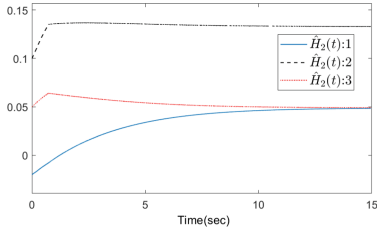
Figures 1, 2, 3, 4 and 5 are the simulation results. Figure 1 gives the responses of  $x_1(t), x_2(t)$  and  $\hat{m}'(t)$ . The responses of  $\hat{H}_1(t), \hat{H}_2(t), \hat{H}(t)$  are shown in Figs. 2, 3 and 4, respectively. Figure 5 shows the responses of actual control law  $u(t)$ . From this simulation results, we verified the effectiveness of the control scheme.



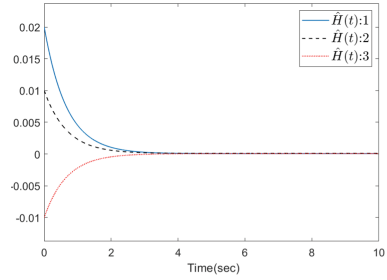
**Fig. 1.** Trajectories of  $x_1(t)$ ,  $x_2(t)$  and  $\hat{n}'(t)$



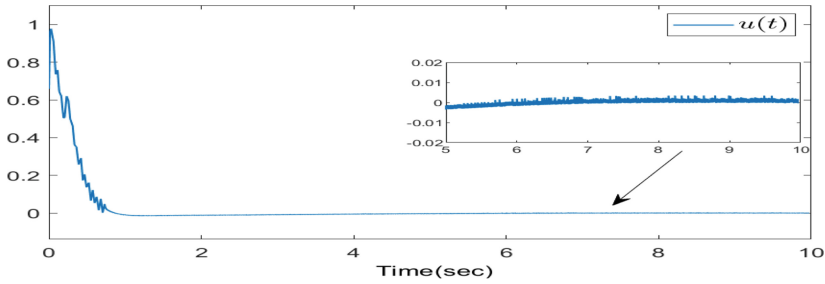
**Fig. 2.** Trajectories of  $\hat{H}_1(t)$



**Fig. 3.** Trajectories of  $\hat{H}_2(t)$



**Fig. 4.** Trajectories of  $\hat{H}(t)$



**Fig. 5.** Trajectory of  $u(t)$

## 5 Conclusion

In this paper, an output feedback control scheme is designed for a class of nonlinear systems with unknown time delays by using reduced-order observer, static gain, neural network, backstepping technique and adaptive control technology. New control gains are designed to compensate for the addition terms generated

by Lyapunov candidate functional. All the closed loop signals can be kept bonded by utilizing the designed control scheme.

## References

1. Isidori, A.: *Nonlinear Control Systems*, 3rd edn. Springer, Berlin (1995)
2. Li, Y.M., Tong, S.C.: Adaptive backstepping control for uncertain nonlinear strict-feedback systems with full state triggering. *Automatica* **163** (2024)
3. Wen, L.Y., Jiang, B., Chen, M., Ma, Y.J.: Adaptive control of noncanonical nonlinear systems with time-varying dynamics. *IEEE Trans. Autom. Control* **69**(1), 590–597 (2024)
4. Gu, K., Kharitonov, V.L., Chen, J.: *Stability of Time-Delay Systems*. Birkhauser, Boston (2003)
5. Sheng, Y., Zeng, Z.G., Huang, T.W.: Finite-time stabilization of competitive neural networks with time-varying delays. *IEEE Trans. Cybern.* **52**(11), 11325–11334 (2022)
6. Cheng, H., Huang, X.C., Cao, H.W.: Asymptotic tracking control for uncertain nonlinear strict-feedback systems with unknown time-varying delays. *IEEE Trans. Neural Netw. Learn. Syst.* **34**12, 9821–9831 (2023)
7. Dong, W.Q., Zhao, Y.B., Cong, Y.H.: Reduced-order observer-based controller design for quasi-one-sided Lipschitz nonlinear systems with time-delay. *Int. J. Robust Nonlinear Control* **31**(3), 817–831 (2021)
8. Hua, C.C., Liu, P.X., Guan, X.P.: Backstepping control for nonlinear systems with time delays and applications to chemical reactor systems. *IEEE Trans. Industr. Electron.* **56**(9), 3723–3732 (2009)
9. Zhang, Z.Q., Xu, S.Y., Zhang, B.Y.: Exact tracking control of nonlinear systems with time delays and dead-zone input. *Automatica* **52**, 272–276 (2015)



# Automatic Ground Collision Avoidance for Aircraft in Uncertain Environments

Rui Li and Rui Zhang<sup>(✉)</sup>

Northwestern Polytechnical University, Xi'an 710072, China  
ruizh@nwpu.edu.cn

**Abstract.** This paper proposes an automatic collision avoidance method to address safety challenges in near-ground flight for combat aircraft in uncertain environments. An adaptive controller is designed to address the escape maneuver control problem under gust disturbances and model uncertainties, combining tracking error feedback with disturbance estimation feedforward using a radial basis function neural network. The stability of the closed-loop system is proven through the Lyapunov method. To address the decision-making problem in aircraft escape maneuvers, a simplified six-degree-of-freedom model is established to reduce the computational load for trajectory prediction. Additionally, a multi-escape trajectory collision detection mechanism is designed to minimize maneuvering demands. Experiments show that the proposed controller achieves high accuracy and robustness, while the collision detection algorithm ensures fast computational efficiency and high safety. The algorithm also determines warning levels based on prediction time, enabling safe obstacle avoidance.

**Keywords:** Auto-GCAS · Neural Network Control · Trajectory Prediction · Escape Decision

## 1 Introduction

Aircraft face heightened risks of terrain collision during near-ground flight due to pilot disorientation, vertical position misjudgment, or delayed reactions. These challenges often lead to accidents, highlighting the necessity of an automatic ground collision avoidance system (Auto-GCAS) [1]. A typical Auto-GCAS consists of four main components: trajectory prediction, terrain scanning, collision assessment, and flight control.

Trajectory prediction algorithms are broadly classified into kinematics-based methods, dynamics-based integration methods, and neural network-based models. Kinematics-based methods rely on simplified aerodynamic principles, which lead to reduced accuracy in complex flight conditions [2]. Examples include constant velocity, constant acceleration, coordinated turn, and Singer models, which segment flight states and consider sub-model interactions to track motion [3]. However, these methods fail to account for variable forces experienced during flight effectively. Dynamics-based methods use aircraft dynamics models to calculate trajectories. For instance, Suplisson proposed an optimal control method for military aircraft, which calculates maneuver trajectories using dynamic models [4]. Bai applied a numerical integration method

based on six-degree-of-freedom equations to compute helicopter escape trajectories [5]. However, environmental factors, such as weather, add complexities that constrain the accuracy of these methods. Neural network-based methods address these limitations by utilizing historical data. Recurrent neural networks (RNNs) and their variants, such as long short-term memory (LSTM) networks, excel at sequential prediction tasks. For example, an LSTM-based method integrated with a fuzzy mechanism within a battlefield spatial grid was applied to address boundary mutation issues [6]. Another study utilized a sequence-to-sequence RNN model combined with Mixture Density Networks (MDNs) to estimate future trajectory probability distributions [7]. CNN-LSTM hybrid models have also been applied for 4-D trajectory prediction in civil aviation [8]. Despite their advantages, neural network-based methods face challenges such as high computational demands, substantial data feature loss, and dependence on large datasets, limiting their applicability in complex, dynamic scenarios.

Collision detection plays a critical role in ensuring aircraft safety. Multi-model trajectory prediction methods [9], probability estimation techniques for wind uncertainty [10], and spatiotemporal reachability analyses [11] have been proposed for conflict detection. Additionally, approaches like chance-constrained optimization [12] and weather-intent-based multi-model methods [13] have been applied to improve detection accuracy and computational efficiency. However, most studies focus on mid-air collision detection, with limited research addressing near-ground collision detection for aircraft. Although advancements in trajectory prediction and collision detection have been achieved, existing methods for near-ground flight conditions remain inadequate. Future research should prioritize developing accurate and computationally efficient systems capable of addressing the complexities of near-ground flight, thereby enhancing safety and operational reliability.

This paper proposes an intelligent control and multi-trajectory escape decision method for uncertain environmental disturbances and model uncertainties, providing comprehensive collision avoidance decisions for the crew. The primary contributions of this paper are summarized as follows.

1. A control method for escape maneuvers under environmental uncertainty is developed. By leveraging a radial basis function neural network, this approach enables parallel estimation of environmental uncertainties and disturbances, enhancing the nonlinear handling capabilities and command tracking precision of the control system during the escape phase.
2. The six-degree-of-freedom simplified model of an aircraft is incorporated into the automatic collision avoidance system. By integrating the low-order equivalent model with the conventional three-degree-of-freedom particle model, the computational complexity of trajectory prediction is significantly reduced.
3. A multi-trajectory escape decision-making mechanism is proposed. By employing a time-varying cylindrical envelope to define collision boundaries, this mechanism provides a flexible decision space for aircraft escape maneuvers, effectively reducing the demand for excessive maneuvering.

This paper is organized with the design methodology of the intelligent escape maneuver controller for composite disturbances being introduced in Sect. 2. The automatic collision avoidance method based on escape trajectory prediction is given in

Sect. 3. Section 4 demonstrates the simulation results of the proposed method. Section 5 concludes the paper.

## 2 Intelligent Control Method for Escape Maneuver Under Compound Disturbances

### 2.1 Aircraft Attitude Dynamics Model

We denote the attitude control states as  $\boldsymbol{\Omega} = [\phi, \beta, \theta]^T$ , where  $\phi$  is the roll angle,  $\beta$  is the sideslip angle, and  $\theta$  is the pitch angle. Then the inputs of the system are  $\boldsymbol{\delta} = [\delta_a, \delta_e, \delta_r]^T$ , which represent the deflections of the aircraft control surfaces (aileron, elevator, and rudder). According to [14], an affine nonlinear form for control can be established as

$$\begin{cases} \dot{\boldsymbol{\Omega}} = \boldsymbol{R} \\ \dot{\boldsymbol{R}} = \boldsymbol{f}_\omega + \boldsymbol{g}_\omega \boldsymbol{\delta} + \boldsymbol{\Delta}_R \end{cases} \quad (1)$$

where  $\boldsymbol{R}$  is the process state,  $\boldsymbol{f}_\omega$  is the nonlinear part,  $\boldsymbol{g}_\omega$  is the expression for the control gain matrix, and  $\boldsymbol{\Delta}_R$  is the lumped disturbance in the angular rate loop.  $\boldsymbol{R}$ ,  $\boldsymbol{f}_\omega$ ,  $\boldsymbol{g}_\omega$  are given as

$$\begin{cases} \boldsymbol{R} = \begin{bmatrix} 0 & \cos \phi & -\sin \phi \\ \sin \alpha & -\cos \alpha & 0 \\ 1 & \tan \theta \sin \phi & \tan \theta \cos \phi \end{bmatrix} \begin{bmatrix} p \\ q \\ r \end{bmatrix} + \boldsymbol{\Delta}_\Omega \\ \boldsymbol{\Delta}_\Omega = [0 \ (Y \cos \beta + mg \cos \gamma \sin \mu - T \sin \beta \cos \alpha) / (mV) \ 0]^T \end{cases} \quad (2)$$

$$\boldsymbol{f}_\omega = [f_{\omega,1}(p, q, r, \phi, \theta) \ f_{\omega,2}(\alpha, \beta, p, q, r, \phi) \ f_{\omega,3}(p, q, r, \phi, \theta)]^T \quad (3)$$

$$\boldsymbol{g}_\omega = \frac{1}{2} \rho V^2 S \begin{bmatrix} I_{zz}/(I_{xx}I_{zz} - I_{xz}^2) & 0 & I_{xz}/(I_{xx}I_{zz} - I_{xz}^2) \\ 0 & 1/I_{yy} & 0 \\ I_{xz}/(I_{xx}I_{zz} - I_{xz}^2) & 0 & I_{xx}/(I_{xx}I_{zz} - I_{xz}^2) \end{bmatrix} \cdot \boldsymbol{B} \quad (4)$$

where  $\alpha$  is the angle of attack,  $p$  is the roll rate,  $q$  is the pitch rate,  $r$  is the yaw rate,  $\boldsymbol{\Delta}_\Omega$  is the nonlinear component of the attitude angle loop,  $Y$  is the aerodynamic side force,  $T$  is the engine thrust,  $\gamma$  is the flight path inclination angle,  $\mu$  is roll angle,  $m$  is the mass of the aircraft,  $g$  is the gravitational constant,  $V$  is the airspeed of the aircraft,  $f_{\omega,i}$ , ( $i = 1, 2, 3$ ) is the  $i$ th element of  $\boldsymbol{f}_\omega$ ,  $I_{xx}$ ,  $I_{yy}$ ,  $I_{zz}$ ,  $I_{xz}$  are the aircraft's moment of inertia,  $S$  is the reference area of the aircraft,  $\rho$  is the air density, and  $\boldsymbol{B}$  is the aerodynamic moment coefficient matrix associated with control surface deflections.

*Assumption 1:* In the near-ground flight of the aircraft, crosswind and wind shear interfere with the aircraft control system by enhancing nonlinearities and coupling effects between control channels, altering aerodynamic coefficients that lead to changes in  $\boldsymbol{B}$ , and generating asymmetric aerodynamic moments that directly affect the aircraft's angular velocities. We encapsulate these effects as  $\boldsymbol{\Delta}_R$ , which satisfies  $\|\boldsymbol{\Delta}_R\| \leq \overline{\Delta}_R$ , where  $\overline{\Delta}_R$  is an unknown but existent positive constant.

## 2.2 Controller Design

Considering aerodynamic uncertainties and wind disturbances for attitude control, the objective is to design a neural adaptive controller to enable the system to achieve the required attitude command tracking performance during the escape maneuver phase. The controller design method is as follows

*Step 1:* Define the attitude angle tracking error  $\mathbf{e}_\Omega = \boldsymbol{\Omega} - \boldsymbol{\Omega}_r$ , where  $\boldsymbol{\Omega}_r$  represents the reference command for the attitude angle. Design the virtual control  $\mathbf{R}_d = [R_{d1} \ R_{d2} \ R_{d3}]^T$  as

$$R_{d_i} = -k_{\Omega_i} e_{\Omega_i} + \dot{\Omega}_{r_i}, i = 1, 2, 3 \quad (5)$$

where  $k_{\Omega_i} (i = 1, 2, 3)$  is the gain parameter, and  $\Omega_{r_i} (i = 1, 2, 3)$  is the  $i$  th element of  $\boldsymbol{\Omega}_r$ .

To prevent noise caused by instruction differentiation, we design a first-order filter to smooth the virtual control signal and obtain the reference instruction signal  $\mathbf{R}_c = [R_{c1} \ R_{c2} \ R_{c3}]^T$  as

$$\alpha_{R_i} \dot{R}_{c_i} + R_{c_i} = R_{d_i}, R_{c_i}(0) = R_{d_i}(0), i = 1, 2, 3 \quad (6)$$

where  $\alpha_{R_i} (i = 1, 2, 3)$  is the filter parameter.

Define the angular rate tracking error  $\mathbf{e}_R = \mathbf{R} - \mathbf{R}_c = [e_{R1} \ e_{R2} \ e_{R3}]^T$ , the derivative of  $\mathbf{e}_\Omega$  can be expressed as

$$\dot{e}_{\Omega_i} = -k_{\Omega_i} e_{\Omega_i} + e_{R_i} + R_{c_i} - R_{d_i}, i = 1, 2, 3 \quad (7)$$

To avoid drastic changes caused by direct dependence on  $\mathbf{e}_R$  and reduce tracking errors through compensation, we define the compensating signal  $\boldsymbol{\chi}_\Omega = [\chi_{\Omega_1} \ \chi_{\Omega_2} \ \chi_{\Omega_3}]^T$  of the attitude filter as

$$\dot{\chi}_{\Omega_i} = -k_{\Omega_i} \chi_{\Omega_i} + \chi_{R_i} + R_{c_i} - R_{d_i}, \chi_{\Omega_i}(0) = 0, i = 1, 2, 3 \quad (8)$$

where  $\boldsymbol{\chi}_R = [\chi_{R1} \ \chi_{R2} \ \chi_{R3}]^T$  will be defined in (14).

Define the compensated attitude tracking error as

$$\boldsymbol{\kappa}_\Omega = \mathbf{e}_\Omega - \boldsymbol{\chi}_\Omega \quad (9)$$

*Step 2:* For the uncertainties in the angular rate loop system, we define a function  $\mathbf{F}$  to represent the sum of the system's nonlinear component  $\mathbf{f}_\omega$  and the lumped disturbance  $\boldsymbol{\Delta}_R$ .

$$\mathbf{F} = [F_1(p, q, r, \phi, \theta) \ F_2(\alpha, \beta, p, q, r, \phi) \ F_3(p, q, r, \phi, \theta)]^T \quad (10)$$

$$F_i = \Gamma_{f_i}(f_{\omega_i} + \Delta_{R_i}), i = 1, 2, 3$$

where  $\Gamma_{f_i} (i = 1, 2, 3)$  is the design parameter,  $f_{\omega_i}$  is the  $i$  th element of  $\mathbf{f}_\omega$ , and  $\Delta_{R_i}$  is the  $i$  th element of  $\boldsymbol{\Delta}_R$

Using the radial basis function neural network (RBFNN) to approximate  $F, \dot{R}$  can be given as

$$\begin{aligned}\dot{R}_i &= \Gamma_{f_i}^{-1} F_i^* + \varepsilon_{R_i} + g_{\omega_i} \delta, i = 1, 2, 3 \\ F^* &= [\varpi_1^{*T} \theta_1(p, q, r, \phi, \theta) \varpi_2^{*T} \theta_2(\alpha, \beta, p, q, r, \phi) \varpi_3^{*T} \theta_3(p, q, r, \phi, \theta)]^T\end{aligned}\quad (11)$$

where  $F^*$  is the optimal neural network approximation,  $\varpi_i^*$  ( $i = 1, 2, 3$ ) are the optimal weight matrices of the neural network,  $\theta_i$  ( $i = 1, 2, 3$ ) are the radial basis function vectors,  $\varepsilon_R \in \mathbb{R}^3$  is the neural network estimation error satisfying  $\|\varepsilon_R\| \leq \bar{\varepsilon}_R$ , with  $\bar{\varepsilon}_R$  is an unknown but existent positive constant, and  $g_{\omega_i}$  ( $i = 1, 2, 3$ ) is the  $i$  th row of  $g_\omega$ .

The final control signal  $\delta$  is designed as

$$\delta = g_\omega^{-1} \left( - \begin{bmatrix} \Gamma_{f_1}^{-1} \hat{\varpi}_1^T \theta_1 \\ \Gamma_{f_2}^{-1} \hat{\varpi}_2^T \theta_2 \\ \Gamma_{f_3}^{-1} \hat{\varpi}_3^T \theta_3 \end{bmatrix} - \begin{bmatrix} k_{R_1} e_{R_1} \\ k_{R_2} e_{R_2} \\ k_{R_3} e_{R_3} \end{bmatrix} - e_\Omega + \dot{R}_c \right) \quad (12)$$

where  $k_{R_i}$  ( $i = 1, 2, 3$ ) is the gain parameter, and  $\hat{\varpi}_i$  ( $i = 1, 2, 3$ ) are the estimates of the optimal weight matrices. Then the derivative of the angular rate tracking error  $e_R$  is obtained as:

$$\dot{e}_R = \begin{bmatrix} \Gamma_{f_1}^{-1} \hat{\varpi}_1^T \theta_1 \\ \Gamma_{f_2}^{-1} \hat{\varpi}_2^T \theta_2 \\ \Gamma_{f_3}^{-1} \hat{\varpi}_3^T \theta_3 \end{bmatrix} + \varepsilon_R - \begin{bmatrix} k_{R_1} e_{R_1} \\ k_{R_2} e_{R_2} \\ k_{R_3} e_{R_3} \end{bmatrix} - e_\Omega \quad (13)$$

where  $\tilde{\varpi}_i = \varpi_i^* - \hat{\varpi}_i$  ( $i = 1, 2, 3$ ).

Define the compensating signal  $\chi_R = [\chi_{R_1} \chi_{R_2} \chi_{R_3}]^T$  of the angular rate command filter as

$$\dot{\chi}_{R_i} = -k_{R_i} \chi_{R_i} - \chi_{\Omega_i}, \chi_{R_i}(0) = 0, i = 1, 2, 3 \quad (14)$$

The compensated angular rate tracking error is defined as

$$\kappa_R = e_R - \chi_R \quad (15)$$

Define the prediction error as

$$\chi_{RN} = [\chi_{RN_1} \chi_{RN_2} \chi_{RN_3}]^T = R - \hat{R} \quad (16)$$

where  $\hat{R}$  is constructed through the serial-parallel estimation model as

$$\hat{R} = \begin{bmatrix} \Gamma_{f_1}^{-1} \hat{\varpi}_1^T \theta_1 & \Gamma_{f_2}^{-1} \hat{\varpi}_2^T \theta_2 & \Gamma_{f_3}^{-1} \hat{\varpi}_3^T \theta_3 \end{bmatrix}^T + g_\omega \delta + [v_1 \chi_{RN_1} \ v_2 \chi_{RN_2} \ v_3 \chi_{RN_3}]^T \quad (17)$$

with  $\hat{R}(0) = R(0)$ ,  $v_i$  ( $i = 1, 2, 3$ ) is the design parameter.

The weight update rule for  $\hat{\varpi}_i$  ( $i = 1, 2, 3$ ) is designed as:

$$\dot{\hat{\varpi}}_i = \lambda_{R_i} [\Gamma_{f_i}^{-1} (\kappa_{R_i} + \lambda_{RN_i} \chi_{RN_i}) \theta_i - \varsigma_i \hat{\varpi}_i], i = 1, 2, 3 \quad (18)$$

where  $\lambda_{R_i}, \lambda_{RN_i}, \varsigma_i$  ( $i = 1, 2, 3$ ) are design parameters.



### 2.3 Stability Analysis

Define the Lyapunov function as

$$V_L = \frac{1}{2} \sum_{i=1,2,3} V_{L_i} = \frac{1}{2} \sum_{i=1,2,3} (\kappa_{\Omega_i}^2 + \kappa_{R_i}^2 + \lambda_{RN_i} \chi_{RN_i}^2 + \tilde{\mathbf{w}}_i^T \lambda_{R_i}^{-1} \tilde{\mathbf{w}}_i) \quad (19)$$

where  $\kappa_{\Omega_i}$  and  $\kappa_{R_i}$  represent the  $i$  th element of  $\kappa_{\Omega}$  and  $\kappa_R$ , respectively.

From (7)–(9), (13)–(15), and (16)–(18), the derivative of  $V_{L_i}$  can be calculated as

$$\begin{aligned} \dot{V}_{L_i} &= \kappa_{\Omega_i} \dot{\kappa}_{\Omega_i} + \kappa_{R_i} \dot{\kappa}_{R_i} + \lambda_{RN_i} \chi_{RN_i} \dot{\chi}_{RN_i} + \tilde{\mathbf{w}}_i^T \lambda_{R_i}^{-1} \dot{\tilde{\mathbf{w}}}_i \\ &= -k_{\Omega_i} \kappa_{\Omega_i}^2 - k_{R_i} \kappa_{R_i}^2 + \kappa_{R_i} \varepsilon_{R_i} - \lambda_{RN_i} v_i \chi_{RN_i}^2 + \lambda_{RN_i} \chi_{RN_i} \varepsilon_{R_i} + \varsigma_i \tilde{\mathbf{w}}_i^T (-\tilde{\mathbf{w}}_i + \mathbf{w}_i^*) \end{aligned} \quad (20)$$

Consider the following inequality

$$\begin{aligned} \kappa_{R_i} \varepsilon_{R_i} &\leq \frac{1}{2} \kappa_{R_i}^2 + \frac{1}{2} \varepsilon_{R_i}^2 \\ \chi_{RN_i} \varepsilon_{R_i} &\leq \frac{1}{2} \chi_{RN_i}^2 + \frac{1}{2} \varepsilon_{R_i}^2 \\ \tilde{\mathbf{w}}_i^T (-\tilde{\mathbf{w}}_i + \mathbf{w}_i^*) &\leq -\frac{1}{2} \tilde{\mathbf{w}}_i^T \tilde{\mathbf{w}}_i + \frac{1}{2} \|\mathbf{w}_i^*\|^2 \end{aligned} \quad (21)$$

Proceeding further  $\dot{V}_{L_i}$ , we obtain

$$\dot{V}_{L_i} \leq -k_{\Omega_i} \kappa_{\Omega_i}^2 - (k_{R_i} - \frac{1}{2}) \kappa_{R_i}^2 - (v_i - \frac{1}{2}) \lambda_{RN_i} \chi_{RN_i}^2 - \frac{1}{2} \varsigma_i \tilde{\mathbf{w}}_i^T \tilde{\mathbf{w}}_i + C_i \quad (22)$$

where

$$C_i = \frac{1}{2} \varsigma_i \|\mathbf{w}_i^*\|^2 + (\lambda_{RN_i} + 1) \varepsilon_{R_i}^2, i = 1, 2, 3 \quad (23)$$

For control parameters  $k_{R_i} > \frac{1}{2}$ ,  $v_i > \frac{1}{2}$ , we have

$$\dot{V}_{L_i} \leq -K_i V_{L_i} + C_i \quad (24)$$

where

$$K_i = \min\{2k_{\Omega_i}, 2k_{R_i} - 1, 2v_i - 1, \varsigma_i \lambda_{R_i}\}, i = 1, 2, 3 \quad (25)$$

Proceeding further, we obtain

$$0 \leq \frac{1}{2} \sum_{i=1,2,3} V_{L_i} \leq \frac{1}{2} \sum_{i=1,2,3} \frac{C_i}{K_i} + (V_{L_i}(0) - \frac{C_i}{K_i}) e^{-K_i t} \quad (26)$$

Thus, it is proved that the terms  $\kappa_{\Omega_i}$ ,  $\kappa_{R_i}$ ,  $\chi_{RN_i}$ ,  $\tilde{\mathbf{w}}_i$  ( $i = 1, 2, 3$ ) in the Lyapunov function are uniformly bounded.

### 3 Automatic Ground Collision Avoidance Based on Escape Trajectory Prediction

#### 3.1 Simplified Six-Degree-of-Freedom Model for Aircraft

Considering that the 3-DOF model neglects the control system and the 6-DOF model is computationally complex, we use a simplified the 6-DOF model like [15] by applying a lower-order equivalent transformation to the control system and combining it with a simplified point-mass model to predict aircraft flight trajectory.

Traditional PID controllers struggle to track commands accurately under wind disturbances, resulting in imprecise modeling of attitude angle responses. In contrast, the proposed intelligent control method ensures precise tracking of continuous commands, allowing the attitude angle response transfer function to be simplified as

$$\begin{aligned}\frac{\phi}{\phi_r} &= \frac{\omega_{n\phi_1} \omega_{n\phi_2}^2}{(s + \omega_{n\phi_1})(s^2 + 2\xi_\phi \omega_{n\phi_2} + \omega_{n\phi_2}^2)} \\ \frac{\theta}{\theta_r} &= \frac{\omega_{n\theta_1} \omega_{n\theta_2}^2}{(s + \omega_{n\theta_1})(s^2 + 2\xi_\theta \omega_{n\theta_2} + \omega_{n\theta_2}^2)}\end{aligned}\quad (27)$$

where  $\phi_r$  and  $\theta_r$  refer to step command inputs,  $\omega_{n\phi_1}$  and  $\omega_{n\theta_1}$  are the first-order pole frequency,  $\omega_{n\phi_2}$  and  $\omega_{n\theta_2}$  are the second-order natural frequency,  $\xi_\phi$  and  $\xi_\theta$  are the damping ratio.

By transforming it into state-space form, the equivalent model of the pitch angle and roll angle control system can be established as

$$\begin{cases} \dot{\mathbf{x}}_\phi = \mathbf{A}_\phi \mathbf{x}_\phi + \mathbf{B}_\phi \phi_r \\ \dot{\phi} = \mathbf{C}_\phi \dot{\mathbf{x}}_\phi \end{cases} \quad \text{and} \quad \begin{cases} \dot{\mathbf{x}}_\theta = \mathbf{A}_\theta \mathbf{x}_\theta + \mathbf{B}_\theta \theta_r \\ \dot{\theta} = \mathbf{C}_\theta \dot{\mathbf{x}}_\theta \end{cases}\quad (28)$$

where  $\mathbf{x}_\phi, \mathbf{x}_\theta \in \mathbb{R}^3$  are intermediate state variables,  $\mathbf{A}_\phi, \mathbf{B}_\phi, \mathbf{C}_\phi, \mathbf{A}_\theta, \mathbf{B}_\theta$ , and  $\mathbf{C}_\theta$  are the state matrices that can be obtained as

$$\begin{aligned}\mathbf{A}_\phi &= \begin{bmatrix} -\omega_{n\phi_1} - 2\omega_{n\phi_2} \xi_\phi & -\omega_{n\phi_2}^2 & -2\omega_{n\phi_1} \xi_\phi \omega_{n\phi_2} & -\omega_{n\phi_1} \omega_{n\phi_2}^2 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix}, \mathbf{B}_\phi = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \mathbf{C}_\phi = \begin{bmatrix} 0 \\ 0 \\ \omega_{n\phi_1} \omega_{n\phi_2}^2 \end{bmatrix}^T \\ \mathbf{A}_\theta &= \begin{bmatrix} -\omega_{n\theta_1} - 2\omega_{n\theta_2} \xi_\theta & -\omega_{n\theta_2}^2 & -2\omega_{n\theta_1} \xi_\theta \omega_{n\theta_2} & -\omega_{n\theta_1} \omega_{n\theta_2}^2 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix}, \mathbf{B}_\theta = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \mathbf{C}_\theta = \begin{bmatrix} 0 \\ 0 \\ \omega_{n\theta_1} \omega_{n\theta_2}^2 \end{bmatrix}^T\end{aligned}$$

Then, during the steady flight phase of the aircraft, where  $\mu \approx \phi$ ,  $\beta \approx 0$ ,  $Y \approx 0$ ,  $p \approx 0$ ,  $q \approx 0$ ,  $r \approx 0$ , the simplified point-mass dynamics model of the aircraft can be expressed as follows

$$\begin{cases} \dot{x} = V \cos \gamma \cos \chi \\ \dot{y} = V \cos \gamma \sin \chi \\ \dot{z} = V \sin \gamma \\ \dot{V} = (T \cos \alpha - D - mg \sin \gamma)/m \\ \dot{\gamma} = (L \cos \mu - mg \cos \gamma + T \sin \alpha \cos \mu)/(mV) \\ \dot{\chi} = (L \sin \mu + T \sin \alpha \sin \mu)/(mV \cos \gamma) \end{cases}\quad (29)$$

where  $x$ ,  $y$  and  $z$  are the position of the aircraft in the inertial frame,  $T$  is the engine thrust,  $L$  is the aerodynamic lift force,  $D$  is the aerodynamic drag force, and  $\chi$  is the flight path heading angle.

In (29),  $\alpha$ ,  $L$ ,  $D$  are simplified as

$$\begin{aligned}\alpha &= \theta - \gamma \\ L &= 0.5\rho V^2 S \tilde{C}_L(\alpha) \\ D &= 0.5\rho V^2 S \tilde{C}_D(\alpha)\end{aligned}\quad (30)$$

where  $\tilde{C}_L$  is the estimated value of the lift coefficient  $C_L$ , and  $\tilde{C}_D$  is the estimated value of the drag coefficient  $C_D$ .

Based on actual flight conditions, within a certain range of values,  $\tilde{C}_L$  is modeled as a linear function of the angle of attack, i.e.

$$\tilde{C}_L = k_{C_L}\alpha + b_{C_L} \quad (31)$$

where  $k_{C_L}$  and  $b_{C_L}$  are the fitting parameters of  $\tilde{C}_L$ .

Similarly,  $\tilde{C}_D$  is modeled as a quadratic function of the angle of attack, i.e.

$$\tilde{C}_D = a_{C_D}\alpha^2 + b_{C_D}\alpha + c_{C_D} \quad (32)$$

where  $a_{C_D}$ ,  $b_{C_D}$  and  $c_{C_D}$  are the fitting parameters of  $\tilde{C}_D$ .

From (28) to (32), a simplified six-degree-of-freedom model for the aircraft can be established as

$$\dot{X} = f(X, U) \quad (33)$$

where  $X = [x, y, z, V, \gamma, \chi, \mathbf{x}_\varphi^T, \mathbf{x}_\theta^T]^T$  is the state vector of the model,  $U = [T, \phi_r, \theta_r]^T$  is the input vector of the model.

### 3.2 The Trajectory-Based Collision Detection Method

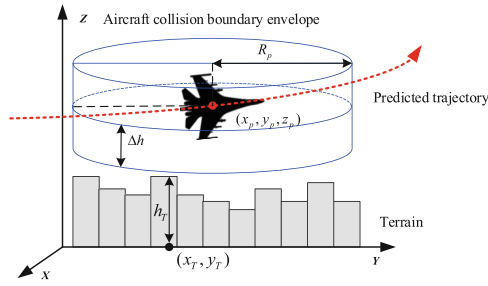
Currently, most escape maneuvers use a standard vertical pull-up maneuver, which involves first rolling the wings to a horizontal position and then pulling up. However, when the aircraft faces threatening terrain, the pull-up maneuver is not always the most effective evasive decision. This paper designs five types of escape maneuvers in the flight space and predicts the escape trajectories as shown in Table 1.

First, based on the simplified six-degree-of-freedom equations above, the Runge-Kutta method is used for integration to calculate the predicted trajectory for the next 30 s. The integration step size is denoted as  $h$  and is selected according to the required prediction accuracy.

Then, the trajectory-based collision detection method assesses collision risk by analyzing the relationship between the aircraft collision boundary volume and terrain threats. Collision detection requires constructing the aircraft collision boundary volume and integrating it with the digital elevation model (DEM) to extract threat information. In this study, we adopt a cylindrical boundary volume as the aircraft's safety representation, as shown in Fig. 1.

**Table 1.** Trajectory Prediction Algorithm Design Parameters

|   | Roll Angle Command (°) | Pitch Angle Command (°) | Throttle Command | Prediction Time (s) |
|---|------------------------|-------------------------|------------------|---------------------|
| 1 | −60                    | 20                      | max              | 30                  |
| 2 | −30                    | 20                      | max              | 30                  |
| 3 | 0                      | 20                      | max              | 30                  |
| 4 | 30                     | 20                      | max              | 30                  |
| 5 | 60                     | 20                      | max              | 30                  |

**Fig. 1.** Aircraft Collision Boundary Envelope

The cylinder is centered at the aircraft's center of mass and oriented perpendicular to the horizontal plane. The radius  $R_p$  describes the lateral collision range of the aircraft, while the half-height  $\Delta h$  represents the longitudinal collision range. Since prediction errors increase with the prediction time, the values of  $R_p$  and  $\Delta h$  are determined as follows

$$\begin{aligned} R_p(t) &= R_N + \sqrt{R_T^2 + R_{pre}(t)^2} \\ \Delta h(t) &= \Delta h_N + \sqrt{\Delta h_T^2 + \Delta h_{pre}(t)^2} \end{aligned} \quad (34)$$

where  $R_N$  is the lateral navigation error,  $\Delta h$  is the vertical navigation error,  $R_T$  is the lateral terrain error,  $\Delta h_T$  is the vertical terrain error,  $R_{pre}(t)$  is the lateral prediction error, and  $\Delta h_{pre}(t)$  is the vertical prediction error.

After obtaining the predicted flight trajectory and equivalent terrain data for the next 30 s, collision detection is performed. Assuming the aircraft's predicted position at a given time is  $(x_p, y_p, z_p)$ . The collision warning condition is defined as

$$\forall (x_T, y_T) \in \left\{ (x_T, y_T) | (x_p - x_T)^2 + (y_p - y_T)^2 \leq R_p(t)^2 \right\}, \exists |z_p - h_T| \leq \Delta h(t) \quad (35)$$

where  $(x_T, y_T)$  are the coordinates in DEM,  $h_T$  is the terrain height.

Finally, the escape maneuver trigger logic of the Auto-GCAS in this paper, as illustrated in Fig. 2. The system initially predicts the aircraft's original task trajectory in

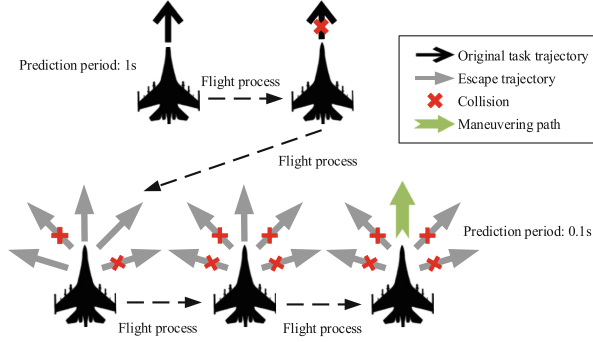


Fig. 2. Collision Avoidance Command Trigger Algorithm Illustration

real-time at 1 s intervals. Upon detecting a potential collision, it switches to predicting five escape trajectories at 0.1 s intervals. If any predicted trajectory indicates a collision, the algorithm does not trigger a maneuver. A maneuver is initiated only when a single escape trajectory remains collision-free, and the aircraft executes the avoidance maneuver based on this final collision-free trajectory.

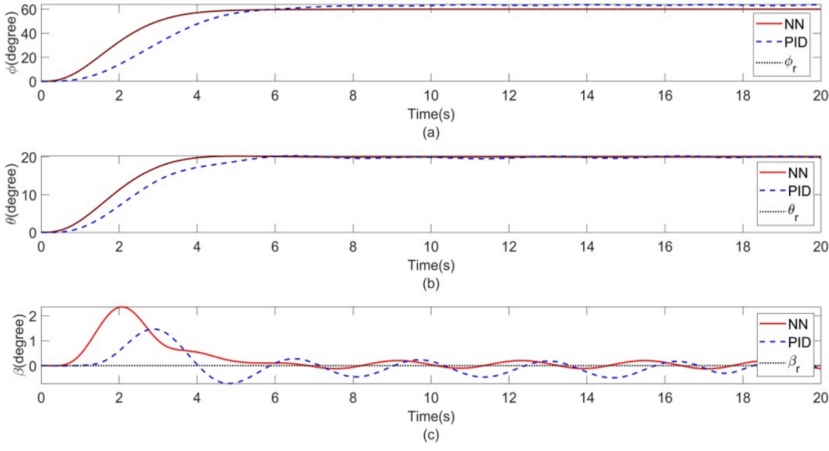
## 4 Simulation Study

### 4.1 Controller and Simplified Model Performance Evaluation

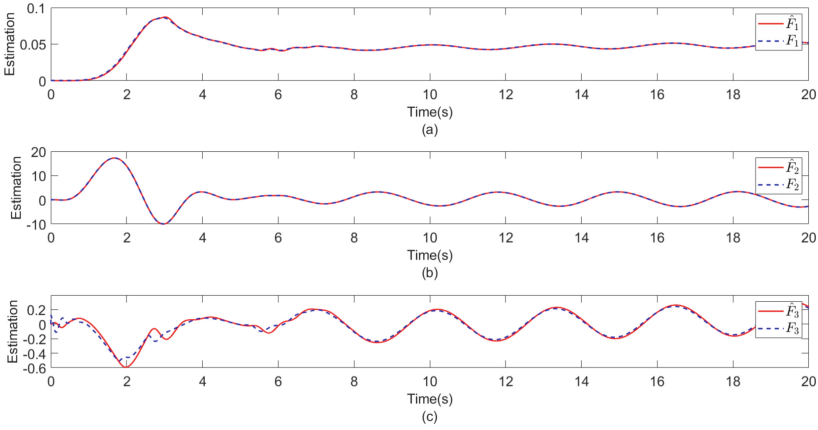
The initial state of the aircraft system is set as  $z = 3800$  m,  $V = 122$  m/s,  $\alpha = 0^\circ$ ,  $\beta = 0^\circ$ ,  $\phi = 0^\circ$ ,  $\theta = 0^\circ$ . During the flight, a constant thrust  $T = 44482.2$  N is maintained. The parameters are set as  $k_{\Omega_1} = 15$ ,  $k_{\Omega_2} = 0.05$ ,  $k_{\Omega_3} = 15$ ,  $\alpha_{R_1} = 0.005$ ,  $\alpha_{R_2} = 0.005$ ,  $\alpha_{R_3} = 0.005$ ,  $\Gamma_{f_1} = 1$ ,  $\Gamma_{f_2} = 1$ ,  $\Gamma_{f_3} = 1$ ,  $k_{R_1} = 15$ ,  $k_{R_2} = 0.51$ ,  $k_{R_3} = 50$ ,  $\lambda_{R_1} = 0.1$ ,  $\lambda_{R_2} = 2.6$ ,  $\lambda_{R_3} = 2$ ,  $\lambda_{RN_1} = 0.1$ ,  $\lambda_{RN_2} = 1$ ,  $\lambda_{RN_3} = 1$ ,  $\varsigma_1 = 0.3$ ,  $\varsigma_2 = 1$ ,  $\varsigma_3 = 3$ ,  $v_1 = 1$ ,  $v_2 = 1.4$ ,  $v_3 = 5$ ,  $\omega_{n\phi_1} = 0.9$ ,  $\omega_{n\phi_2} = 1.5$ ,  $\xi_\phi = 0.7$ ,  $\omega_{n\theta_1} = 1.2$ ,  $\omega_{n\theta_2} = 1.3$ ,  $\xi_\theta = 0.7$ . Considering the impact of external disturbances and aerodynamic parameter perturbations during flight, a simulated disturbance  $[0.2, 0.2, 0.01]^T \sin(2t + 0.1)$  is introduced into the angular rate loop. The control algorithm designed in this paper is denoted as “NN” and is compared with PID control algorithm for performance evaluation.

As shown in Fig. 3, the NN control offers better accuracy in attitude tracking, whereas PID control results in oscillatory responses. It is also reasonable to model the attitude control system as a lower-order system. Additionally, the parallel estimation algorithm employed in this paper effectively estimates model uncertainty in Fig. 4.

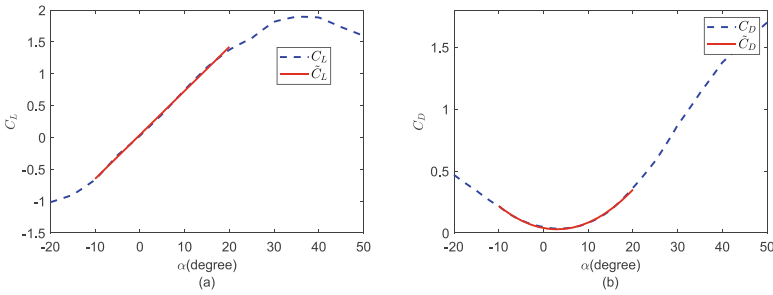
In Fig. 5, it can be observed that the aerodynamic parameters can be accurately approximated within a certain angle of attack range. In the 20 s trajectory prediction, the simplified model achieves an average computation time of 0.32 s, compared to 5.19 s for the 6-DOF model. This indicates that the simplified model offers significantly faster computation. Furthermore, as shown in Fig. 6, the prediction error of the simplified model remains within an acceptable range.



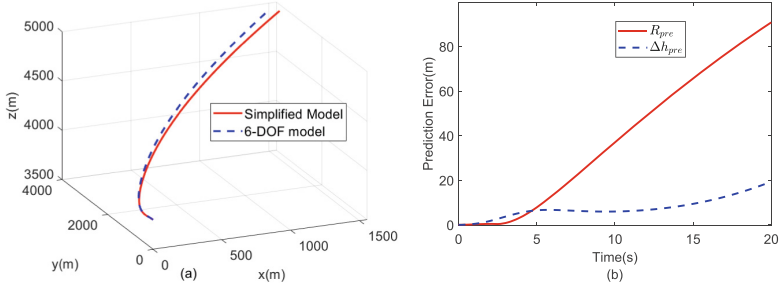
**Fig. 3.** Control Command Tracking. (a)  $\phi$ , (b)  $\theta$ , (c)  $\beta$



**Fig. 4.** Compound Estimation. (a)  $\hat{F}_1$ , (b)  $\hat{F}_2$ , (c)  $\hat{F}_3$



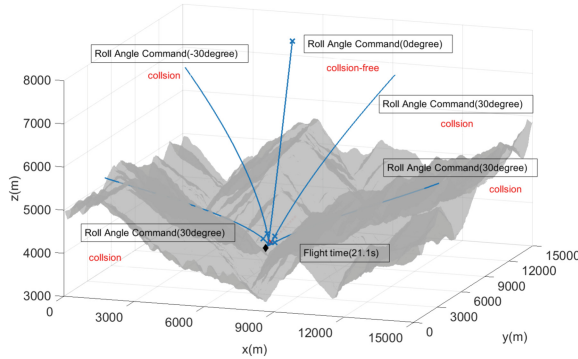
**Fig. 5.** Aerodynamic Parameters Fitting Results. (a)  $C_L$ , (b)  $C_D$



**Fig. 6.** Trajectory Prediction Results. (a) Prediction Trajectory, (b) Prediction Error

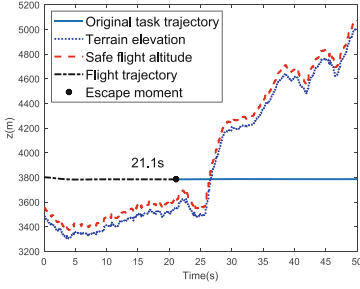
## 4.2 Automatic Collision Avoidance Comprehensive Simulation

As shown in Fig. 7, The initial state of the aircraft is set as  $x = 960$  m,  $y = 7400$  m,  $z = 3800$  m,  $\phi = 0^\circ$ ,  $\theta = 2.92^\circ$ ,  $\psi = -7^\circ$ ,  $\alpha = 2.92^\circ$ ,  $\beta = 0^\circ$ . We assume the aircraft is flying at a steady speed of Ma 0.5 over a mountain. The flight trajectory is predicted in real-time every 1 s. When a collision is detected on the original mission trajectory, a warning is issued to the pilot, five escape trajectories are predicted in real-time every 0.1 s. The trajectory prediction and escape results are shown in Fig. 8.

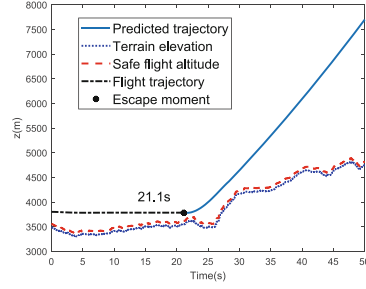


**Fig. 7.** Automatic Collision Avoidance Scenarios

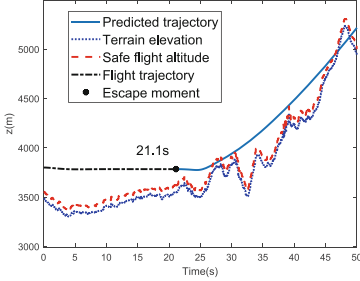
From Fig. 8, the algorithm determined that only the vertical climb trajectory avoided collision, triggering the corresponding maneuver at 21.1 s. The aircraft successfully executed the climb maneuver, maintaining a safe distance from terrain threats. The collision detection algorithm demonstrated high computational efficiency, with an average processing time of 0.9 s per iteration, ensuring timely responses to dynamic flight conditions. The results validate that the collision detection and escape decision mechanisms effectively reduced maneuver demand while ensuring safety. These findings highlight the practical potential of the proposed method for enhancing flight safety in near-ground operations.



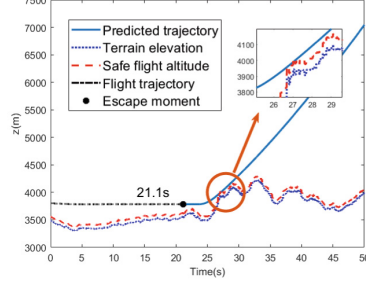
(a) Original Flight Trajectory



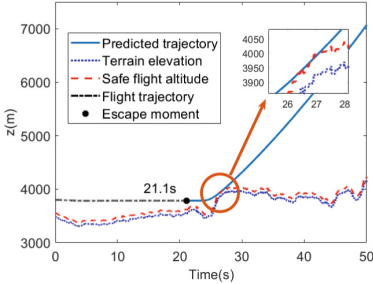
(b) Roll Angle Command(0degree)



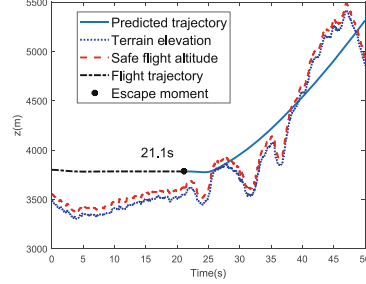
(c) Roll Angle Command(-60degree)



(d) Roll Angle Command(-30degree)



(e) Roll Angle Command(30degree)



(f) Roll Angle Command(60degree)

**Fig. 8.** Escape Trajectory Prediction and Collision Detection Results

## 5 Conclusion

In this work, we propose an intelligent control and collision avoidance method for combat aircraft in uncertain environments. An adaptive control strategy, combining tracking error feedback with radial basis function neural network-based disturbance estimation, addresses gust disturbances and model uncertainties. The stability of the closed-loop system is proven via Lyapunov theory. A simplified six-degree-of-freedom model reduces computational complexity in trajectory prediction, complemented by a multi-trajectory escape decision mechanism to optimize maneuver demands. Experimental results validate the controller's robustness and accuracy under disturbances, while the collision detection algorithm ensures computational efficiency and safety. This



framework enhances near-ground flight safety and operational effectiveness for combat aircraft.

**Acknowledgments.** This work was supported in part by the National Natural Science Foundation of China under Grant 62403382 and 61933010, in part by the Fundamental Research Funds for the Central Universities under Grant G2024KY05106, in part by the Aeronautical Science Foundation Project under Grant 20240008053005.

**Disclosure of Interests.** The authors have no competing interests to declare that are relevant to the content of this article.

## References

1. Swihart, D.E., Barfield, A.F., Griffin, E.M.: Automatic ground collision avoidance system design, integration, & flight test. *IEEE Aerosp. Electron. Syst. Mag.* **26**(5), 4–11 (2011)
2. Ye, X., Zhang, O., Li, X.: Tracking algorithm for cruise missile based on IMM-singer model. In: *AIP Conference Proceedings*. AIP Publishing (2019)
3. Yang, X., Wei, P.: Scalable multi-agent computational guidance with separation assurance for autonomous urban air mobility. *J. Guid. Control. Dyn.* **43**(8), 1473–1486 (2020)
4. Suplisson, A.W.: Optimal recovery trajectories for automatic ground collision avoidance systems. Air Force Institute of Technology (2015)
5. Bai, Y.: Research on aircraft near-ground collision avoidance technology based on multi-trajectory prediction. *Aircraft Des.* **43**(04), 61–65 (2023). (in Chinese)
6. Li, M., Lu, F., Zhang, H.: Predicting future locations of moving objects with deep fuzzy-LSTM networks. *Transportmetrica A. Transp. Sci.* **16**(1), 119–136 (2020)
7. Zyner, A., Worrall, S., Nebot, E.: Naturalistic driver intention and path prediction using recurrent neural networks. *IEEE Trans. Intell. Transp. Syst.* **21**(4), 1584–1594 (2019)
8. Zhang, H.P., Huang, C.Q., Tang, S.Q.: CNN-based real-time prediction method of flight trajectory of unmanned combat aerial vehicle. *Acta Armamentarii.* **41**(9), 1894–1903 (2020)
9. Jilkov, V.P., Li, X.R., Ledet, J.H.: Improved estimation of conflict probability for aircraft collision avoidance. In: *17th International Conference on Information Fusion (FUSION)*, pp. 1–7. IEEE (2014)
10. Hernandez-Romero, E., Valenzuela, A., Rivas, D.: Probabilistic multi-aircraft conflict detection and resolution considering wind forecast uncertainty. *Aerosp. Sci. Technol.* **105**(10), 105973 (2020)
11. Siqi, H., Shaowu, C., Yaping, Z.: A multi-aircraft conflict detection and resolution method for 4-dimensional trajectory-based operation. *Chin. J. Aeronaut.* **31**(7), 1579–1593 (2018)
12. Yang, Y., Zhang, J., Cai, K.Q.: Multi-aircraft conflict detection and resolution based on probabilistic reach sets. *IEEE Trans. Control Syst. Technol.* **25**(1), 309–316 (2016)
13. Jilkov, V.P., Ledet, J.H., Li, X.R.: Multiple model method for aircraft conflict detection and resolution in intent and weather uncertainty. *IEEE Trans. Aerosp. Electron. Syst.* **55**(2), 1004–1020 (2018)
14. Nguyen, L.T.: Simulator study of stall/post-stall characteristics of a fighter airplane with relaxed longitudinal static stability. National Aeronautics and Space Administration, vol. 12854 (1979)
15. Liu, C., Sun, S., Tao, C.: Optimizing evasive maneuvering of planes using a flight quality driven model. *Sci. China Inf. Sci.* **67**(3), 132206 (2024)



# Investigations into the Existing State of Distribution Network Security Using Dispersed Generators

Jiage Zhang, Chen Lei<sup>(✉)</sup>, and Liping He

School of Automation and Electrical Engineering, Linyi University, Linyi 276005, China  
lp-h@163.com

**Abstract.** The access of distributed generators to the network of distribution has changed its internal structure and power flow trend, posed a threat to the traditional distribution network relay protection system, and reduced the operational dependability and safety of the entire power system. In this paper, the impact of distributed generators' access location and capacity on the existing relay protection system of distribution network is analyzed from the current passing through the area of protection while failures with three-phase short circuits happen at several places, and MATLAB/Simulink serves as the foundation for the simulation model to take the three-phase short circuits fault for a special example. In view of the effect of access to the protection system's distribution network via distributed generation, a master protection strategy on the basis of adaptive current protection based on composite sequence components was proposed.

**Keywords:** Composite order components · Distribution grids · Adaptive current protection · Distributed power generation

## 1 Introduction

As society and the continuous growth of energy demand continues to advance, problems of energy shortage have gradually become prominent. For the sustainable development of the country, many scholars have devoted themselves to research in the field of energy. In this grim situation, new and efficient distributed power sources (such as solar, wind, etc.) are developing rapidly. These distributed power sources have the advantages of lower cost, less pollution, and higher reliability, which are of great significance for alleviating the problem of energy shortage. However, the integration of a significant quantity of distributed power sources poses several difficulties, especially the impact on traditional distribution networks.

First of all, the distributed generators' access has changed the topology of the traditional distribution network, which makes the architecture of the network of distribution makes the network of distribution's architecture more intricate and necessitates redesigning and planning. Second, the access to distributed generators affects the distribution of power flows and the magnitude of currents in the network of distribution, which can lead

to oversupply of power in some areas and power shortages in others. In addition, the access of distributed generators has an impact on the correct operation of the network of distribution protection, which may lead to the malfunction or rejection of the protection device. With the aim of the stable operation of the network of distribution and the rapid recovery in the event of a fault, it is essential to carry out comprehensive research on the current protection of the network of distribution with distributed generators.

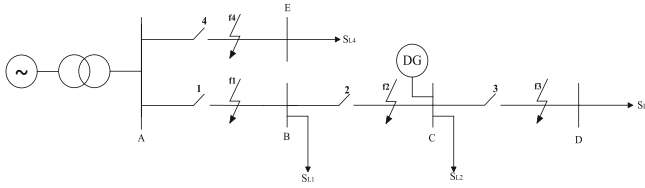
At present, many scholars have studied the effect of distributed generators on the network of distributions after connecting grid, and the influence [1] of the network of distribution's dispersed generators is analyzed from four influencing factors: different locations, different quantities, different capacities, and power factors connected to the network of distribution. In Ref. [2], a simulation model was established to show that the distributed power generation is situated downstream of the faults, downstream of the fault, which will make short circuits currents at the fault smaller, so that the protection might not work. The distributed power supplies are located upstream of the fault point causing a large amount of current to flow through the protection installation point, which can cause protection malfunction. Ref. [3] analyzes the influence of line loss and line operation law from three aspects: the capacity, location, and the manner in which distributed power generation linked to the power grid operates. Ref. [4] improves and optimizes the traditional distribution network protection scheme, adds a Fault Current Limiter (FCL) to limit the fault current, and connects the FCL in series in the access branch of the distributed generator, which not only limits the size of the short circuits currents but also provides the voltage required by direction elements, which greatly improves the reliability of the protection. Using the Thevenin equivalence, combining an adaptive directed current protection technique is suggested in Ref. [5–9] to address the constant-state portion of the fault using the real-time voltage and current measured at the protection installation.

## 2 The Impact of Distributed Power Generation Access on the Network of Distribution's Fault Current

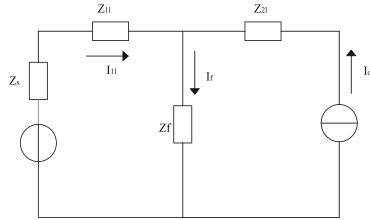
Due to the distributed power sources' access, the conventional radiant distribution network's power flow distribution and current flow direction have evolved, and it is no longer a simple matter of power supply to load [10, 11]. Therefore, when a system fails, the situation of the fault currents become complicated. As the photovoltaic power supply has a connection with the network of distribution, its voltage remains basically unchanged, so it can be considered that its current is changing, so It might be comparable to a fluctuating source of current.

When a fault arises at the upstream point linked to the distributed power generation in the typical distribution network depicted in Fig. 1, it can be simplified to the corresponding circuit diagram in Fig. 2 below.

In the figure above, where  $\dot{U}_s$  is the system power supply and  $Z_s$  is the system impedance; On the upstream side of the fault site,  $\dot{I}_{1l}$  is the short circuit current, and  $\dot{I}_f$  is the short circuits current at fault point; The short circuit current ( $\dot{I}_{dg}$ ) supplied by the photovoltaic cells downstream of the fault points;  $Z_{1l}$  is the resistance between the upstream side line and the fault point in this system,  $Z_{2l}$  is the resistance between the



**Fig. 1.** Topology of distribution network containing DG.



**Fig. 2.** Equivalent circuit diagram during upstream fault.

line on the distributed power side and the fault point, and  $Z_f$  is the short circuits ground resistance.

The comparable circuit diagram indicates that the system power supply's fault current is:

$$I_{l1} = \frac{\dot{U}_s - Z_f \dot{I}_{dg}}{Z_s + Z_{l1} + Z_f} \quad (1)$$

At the actual point of failure, the fault current is:

$$\dot{I}_f = \frac{\dot{U}_s + (Z_s + Z_{l1} + Z_f) \dot{I}_{dg}}{Z_s + Z_{l1} + Z_f} \quad (2)$$

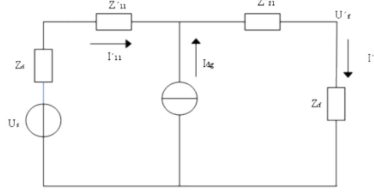
When there is no access to a distributed power source, The system power supply provides the fault current, i.e. by:

$$\dot{I}_f = \dot{I}_{l1} = \frac{\dot{U}_s}{Z_s + Z_{l1} + Z_f} \quad (3)$$

It is evident that fault current is supplied to the fault point  $\dot{I}_{l1}$  by both the system power supply  $\dot{U}_s$  and the distributed power supply  $\dot{I}_{dg}$ . The fault current is connected to both the system and distributed power supplies, as well as to the ground resistance  $Z_f$ .

The analogous design in Fig. 3 can be used to simplify a fault that arises downstream of a distributed power source.

The fault short circuit current  $\dot{I}'_{l1}$  is supplied by the system power supply, and  $\dot{I}'_f$  is the fault short circuit current at the actual fault. In the picture above,  $Z'_{l1}$  represents the fault point's upstream line impedance,  $Z'_{l2}$  represents its downstream line impedance, and so on.



**Fig. 3.** Equivalent circuit diagram during upstream fault.

The system power supply provides the following fault currents:

$$I'_{1l} = \frac{\dot{U}_s - (Z'_{2l} + Z_f)I_{dg}}{Z_s + Z'_{1l} + Z'_{2l} + Z_f} \quad (4)$$

The fault short circuits currents at the actual fault are:

$$I'_f = \frac{\dot{U}_s + (Z'_{1l} + Z_s)I_{dg}}{Z_s + Z'_{1l} + Z'_{2l} + Z_f} \quad (5)$$

If there is no access to a distributed power supplies, and the system power supplies provide all fault short circuits currents, then:

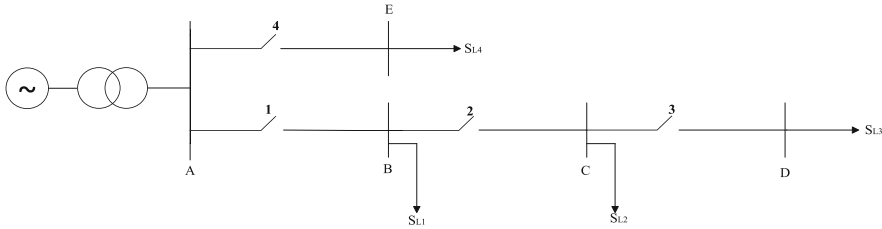
$$I''_{1l} = I''_f = \frac{\dot{U}_s}{Z_s + Z'_{1l} + Z'_{2l} + Z_f} \quad (6)$$

From the above equations, the fault current that the system power supply provides to the fault point in the network containing distributed power generation is often smaller than that of the network of distribution system without distributed power generation, and the actual fault short circuits current size of fault points after the distributed power generation is connected is greater than that of distribution network when there is no distributed power source access, because the distributed power generation's access changes the fault current's distribution and flow direction.

### 3 Simulation Analysis After Distributed Access to the Network of Distribution

#### 3.1 The Network of Distribution System is Operating Normally

According to the transmission network model's Fig. 2, 3 and 4, the system voltage  $ES = 10$  kV,  $X_s = 8 \Omega$  the frequency  $f$  is 50 Hz, the unit impedance of each line is  $(0.2 + j0.4) \Omega/\text{km}$ , the line length  $AB = 2$  km,  $CD = 6$  km,  $BC = 4$  km,  $AE = 8$  km, and  $Z_{AB} = 0.4 + j0.8 \Omega$ ,  $Z_{BC} = 0.8 + j1.6 \Omega$ ,  $Z_{CD} = 1.2 + j2.4 \Omega$ ,  $Z_{AE} = 1.6 + j3.2 \Omega$ . Load capacity at the end of the line  $S_{L1} = S_{L2} = S_{L3} = S_{L4} = 3$  MVA, the reliability coefficient is taken as  $K'_{rel} = 1.2$ ,  $K''_{rel} = 1.2$ ,  $K'''_{rel} = 1.15$ , the self-starting coefficient  $K_{MS} = 1.8$ , the return coefficient  $K_{re} = 0.85$ , and the running time is 1s.



**Fig. 4.** Network of distribution without distributed power sources.

As so to investigate the impact of distributed power generation on the current of the network of distribution, it is decided to connect the distributed power generating to the network of distribution from busbar A, and the network of distribution without photovoltaic power is compared with the network of distribution with the distributed power generation, and the current amplitude obtained by simulation is shown in Table 1.

**Table 1.** Current amplitudes obtained from simulation

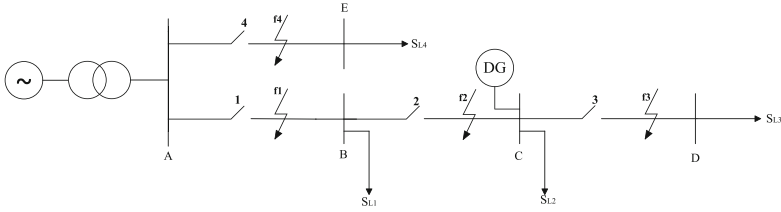
|          | Distribution network current amplitude without PV (A) | Distribution network current amplitude with PV (A) | Increased amplitude |
|----------|-------------------------------------------------------|----------------------------------------------------|---------------------|
| Busbar A | 517                                                   | 536.8                                              | 19.8                |
| Busbar B | 336.9                                                 | 347.1                                              | 10.2                |
| Busbar C | 165.3                                                 | 170.4                                              | 5.1                 |
| Busbar E | 179.2                                                 | 184.9                                              | 5.7                 |

Obviously, the distributed generation’s access from the beginning of the network of distribution will cause the protective area’s current to grow, but the amplitude of the current increase will be smaller as the transmission line becomes farther away from the location of the distributed generation access.

**3.2 Three-Phase Short Circuits Faults Occur at Various Sites Following Their Connection to the Network of Distribution**

In this study, the change of short circuits current caused by the access of distributed generators when the upstream position experiences a three-phase short circuit fault, downstream position and adjacent feeder respectively at the connection busbar C of the distributed generator is analyzed, and the fault occurrence time is 0.4 s–0.5 s for the relay protection device, as shown in Fig. 5.

Let the system supply voltage be and  $E_s$  its equal resistance is  $X_s$ ; The resistance of the transformer is  $X_T$ ; The voltage of the distributed power source is  $E_{dg}$ , and its equivalent resistance is  $X_{dg}$ ;  $X_1, X_2, X_3, X_4$  are the equal resistances of transmission lines L1, L2, L3, L4 respectively;  $\alpha, \beta, \gamma$  and  $\sigma$  the percentage of the fault distance due to faults on transmission lines L1, L2, L3 and L4 to the entire transmission line.



**Fig. 5.** A model of distributed power sources connecting to the network of distribution from busbar C.

### The Distributed Power Generation is Upstream of a Three-Phase Short Circuit Problem

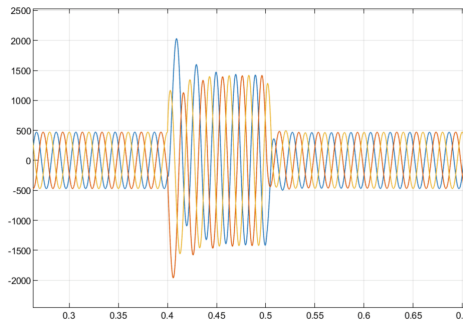
The distributed power supplies are positioned downstream of the fault, taking the fault at  $f_1$  as an example, the fault current at protection 1 is analyzed:

$$I_{k1} = \frac{E_s}{X_s + X_T + \alpha X_1} \quad (7)$$

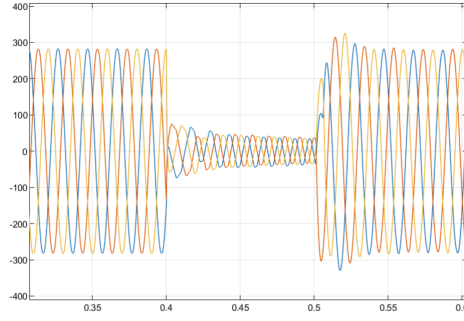
The fault current at protection 2 is:

$$I_{k2} = \frac{E_{dg}}{X_{dg} + X_2 + (1 - \alpha)X_1} \quad (8)$$

From the above analysis, we can deduce that the fault short-circuit current identified by protection 1 is independent of the distributed generator when there is a fault in the upstream of the dispersed generator. Nevertheless, the reverse fault current supplied by the power source that is distributed is the fault current that protection 2 detects, and the direction and impact of the current generated during the fault need to be carefully analyzed on a case-by-case basis. The measured fault short-circuit current passing through busbar A is displayed in Fig. 6, and the fault short-circuit current passing through busbar B is displayed in Fig. 7.



**Fig. 6.** Current at busbar A during fault  $f_1$ .



**Fig. 7.** Current at busbar B during fault  $f1$ .

### The Distributed Power Generation is Downstream of a Three-Phase Short-Circuit Issue

Using the fault at  $f3$  as an illustration, the following shows the fault current at protection 2, and the distributed power supply is situated upstream of the fault occurrence.

$$I_{k2} = \frac{E_s - \frac{E_{sd}}{X_{sd} + \gamma X_3} \cdot \gamma X_3}{X_s + X_1 + X_2 + X_T} \quad (9)$$

The fault currents at protection 3 are:

$$I_{k3} = \frac{E_{sd}}{X_{sd} + \gamma X_3} \quad (10)$$

Among them,  $E_{sd} = \frac{E_s(X_{dg} + X_T) + E_{dg}(X_s + X_1 + X_2)}{X_s + X_1 + X_2 + X_{dg} + X_T}$ ,  $X_{sd} = \frac{(X_{dg} + X_T)(X_s + X_1 + X_2)}{X_s + X_1 + X_2 + X_{dg} + X_T}$ .

From the above analysis, the upstream system side protection 2's short-circuit fault current will be decreased when the distributed power generation's downstream  $f3$  malfunctions because of the connection to the dispersed power source. This could result in the protection device being rejected. Protection 3 will malfunction as a result of the connected distributed power supply increasing the fault short-circuit current at downstream protection 3. The measured fault short-circuit current passing through busbar B is displayed in Fig. 8, and the fault short-circuit current passing through busbar C is displayed in Fig. 9.

### A Three-Phase Short-Circuit Problem Develops on the Feeder Near the Line that Houses the Distributed Power

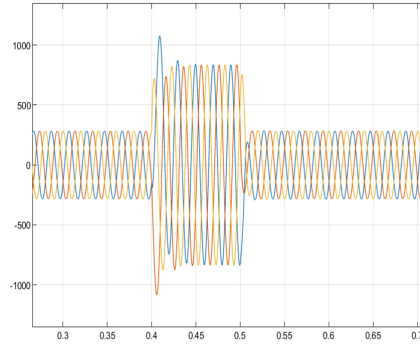
Now the fault at  $f4$  is analyzed, and the fault currents at protection 1 and protection 2 can be obtained as:

$$I_{k1} = I_{k2} = \frac{E_{dg} - \frac{E_{sd}}{X_{sd} + \sigma X_4} \cdot \sigma X_4}{X_{dg} + X_1 + X_2 + X_T} \quad (11)$$

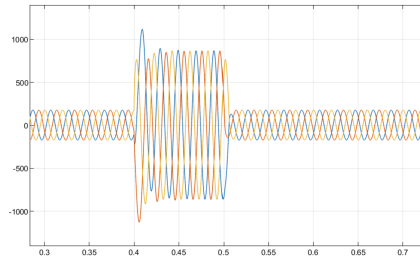
The fault short-circuits currents flowing through protection 4 are:

$$I_{k4} = \frac{E_{sd}}{X_{sd} + \sigma X_4} \quad (12)$$





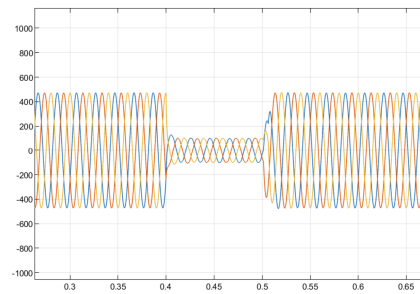
**Fig. 8.** Current at busbar B during fault  $f3$ .



**Fig. 9.** Current at busbar A during fault  $f3$ .

Among them,  $E_{sd} = \frac{E_s(X_{dg} + X_T + X_1 + X_2) + E_{dg}X_s}{X_s + X_1 + X_2 + X_{dg} + X_T}$ ,  $X_{sd} = \frac{X_s(X_{dg} + X_T + X_1 + X_2)}{X_s + X_1 + X_2 + X_{dg} + X_T}$ .

Figure 10 shows the measured fault short-circuit current flowing through busbar A.

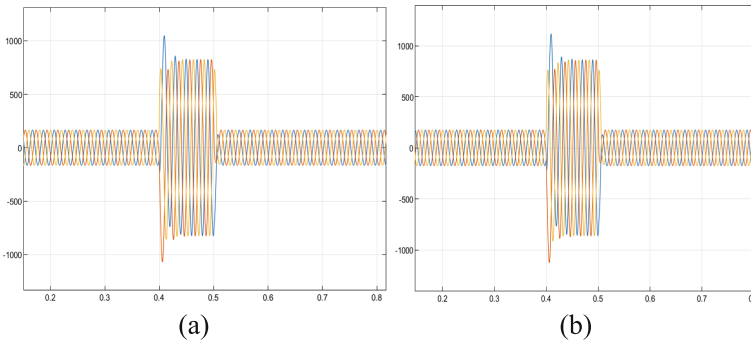


**Fig. 10.** Current at busbar A during fault  $f4$ .

### 3.3 Distributed Power Sources with Varying Capacity Linked to the Distribution Network Might Result in Three-Phase Short-Circuit Failures.

From the analysis in the previous section, we may deduce that the various locations of distributed generators that are affixed to the electrical grid will have an impact on their electrical properties and power grid operations, and when the distributed generators are close to the fault point, the fault impact may be expanded and there will be a significant short-circuit current available. Conversely, short-circuit current is reduced via dispersed power sources that are located further from the site of failure.

For instance, consider the three-phase short-circuit failure at f3, the influence of the short-circuit current caused by the different capacities of the connected distributed generation of power is analyzed. The fault short-circuit current passing through protection 3 is measured as seen in Fig. 11(a) when busbar C's capacity is 0MW. Figure 11(b) displays the measured fault short-circuit current passing through protection 3 when the connected capacity is 4 MW.



**Fig. 11.** Protect the fault current at position3

When the connected distributed power generation capacity is 0MW, the maximum short-circuits current measured through protection 3 is 1049A; When the connected distributed power generation capacity is 4MW, the maximum short-circuits current measured at the flow through protection 3 is 1117A. It can be concluded that when the location of distributed power generation access is located upstream of the fault point, the influence on the protection of the distribution network increases with the growing distributed power source's capability, as does the fault current at the fault point.

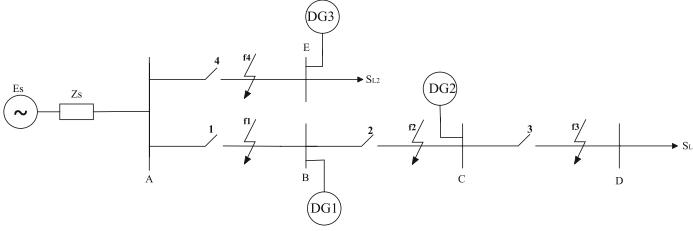
## 4 Present Safeguarding of Electrical Distribution Systems Utilizing Dispersed Electricity Generation

### 4.1 Setting Calculation of Current I Section Protection

The distribution network with dispersed generators can no longer be adequately protected by the conventional three-stage current protection. The inverter distributed generator only produces positive sequence current When the distribution network malfunctions;

zero sequence current and negative sequence current, which may be comparable to a controlled current source, are not produced.

To dissect the composite arrangement network schematic, the schematic layout of the electrical distribution system incorporating several dispersed power sources, as shown in Fig. 12 below, is offered as an instance. When an abrupt three-phase circuit interruption occur, the adaptive current I segment is then modified and computed based on the positive sequence current.



**Fig. 12.** Distribution network structure diagram with multiple distributed power sources added

The positive sequence network diagram is its composite sequence network diagram, and Eq. 13 illustrates the connection between the direct sequence potential and the positive sequence currents at the fault point. If a three-phase short circuits fault occurs, it only contains the positive sequence element, that is, the positive sequence voltage and the positive sequence current; there are no adverse or null phase components.

$$\dot{I}_m = \frac{\dot{U}_m}{Z_k} \quad (13)$$

Among them,  $\dot{U}_m$  the positive sequence voltage measured on the place of fault, the  $\dot{I}_m$  positive sequence current measured at the fault site, and  $Z_k$  the short circuits impedance.

In the following analysis, the  $\alpha$  distance The percentage of the entire line between the protection installation point and the failure point, the  $K'_{rel}$  reliability coefficient of the adaptive current I stage protection, and  $K'''_{rel}$  the reliability factor of the adaptive current II stage protection.

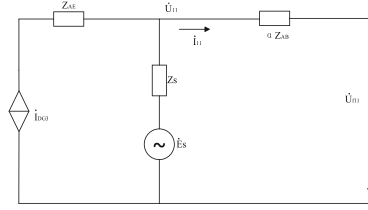
### A Three-Phase Short Circuits Fault Occurs on $fI$

Figure 12 shows the composite sequence network diagram of the power system when a three-symmetrical short-circuit fault occurs at  $fI$ , as shown in Fig. 13.

$$\dot{I}_{11} = \frac{\dot{U}_{11}}{\alpha Z_{AB}} \quad (14)$$

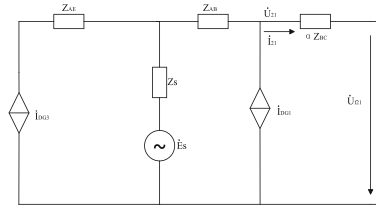
The adaptive current I section of protection 1 is set as follows to assure reliability: the line proportional coefficient is taken to be  $\alpha = 1$ , and the adaptive current I section is set based on the maximum short circuits current that prevents the line's termination short circuit.

$$I'_{1.act(3)} = K'_{rel} \left| \frac{\dot{U}_{11}}{Z_{AB}} \right| \quad (15)$$



**Fig. 13.** Composite sequence network diagram when  $f1$  fault

### *At $f2$ , a Three-Phase Short-Circuit Failure Happens*



**Fig. 14.** Network diagram of a composite sequence when an  $F2$  fault occurs

As a three-symmetrical short circuits faults take place at  $F2$ , the system's combined sequence network diagram is displayed in Fig. 14, and the positive sequence current passing through protection 2 looks like this:

$$\dot{I}_{21} = \frac{\dot{U}_{21}}{\alpha Z_{BC}} \quad (16)$$

The adaptive current I section of protection 2's setting value is therefore:

$$I'_{2.act(3)} = K'_{rel} \left| \frac{\dot{U}_{21}}{Z_{BC}} \right| \quad (17)$$

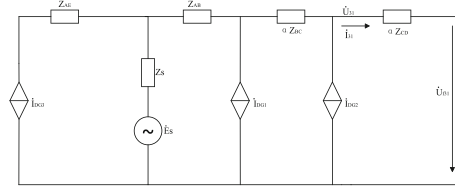
### **At $f3$ , a Three-Phase Short Circuits Fault Happens**

When three symmetrical short circuits faults happen on  $f3$ , the system's combined sequence network diagram is displayed in Fig. 15, and the sequence network diagram is used to determine the positive sequence currents passing through protection 3.

$$\dot{I}_{31} = \frac{\dot{U}_{31}}{\alpha Z_{CD}} \quad (18)$$

At fault time, the adaptive current I part of protection 3's setting value is:

$$I'_{3.act(3)} = K'_{rel} \left| \frac{\dot{U}_{31}}{Z_{CD}} \right| \quad (19)$$



**Fig. 15.** Composite sequence network diagram when  $f_3$  fault

In the same way, when the  $f_4$  position is faulty, The adaptive current I section of protection 4's setting value is:

$$I'_{4.act(3)} = K'_{rel} \left| \frac{\dot{U}_{41}}{Z_{AE}} \right| \quad (20)$$

## 4.2 Calculation of the Current II Section Protection Setting

The adaptive current II stage protection design must guarantee that its protection range is strictly controlled within the protection range of section I of the next line, and the current setting value of section II protection of this section of line must match the protection setting value of section I of the adjacent next line in order to meet the stringent requirements of selectivity and rapidity of power system relay protection. For example, the II-stage adaptive current setting at protection 2 needs to be coordinated with the I-stage protection of protection 3. In order for the positive sequence current passing through protection 3 to equal the setting value of the protection 3's I section,  $f_3$  must first be faulted. The greatest range that the protection 3's I section can cover is as follows.

$$I'_{\max} = \frac{1}{K'_{rel}} \quad (21)$$

Establishing a connection between protections 2 and 3 is essential since it is the setting calculation of protection 2's second section.  $\dot{U}_{21}$  is utilized to show that the positive sequence current passing through protection 3 is:

$$\dot{I}_{31} = \left| \frac{\dot{U}_{21}}{Z_{BC} + \alpha Z_{CD}} \right| \quad (22)$$

By introducing the maximum protection range of section I of protection 3, as determined by Eq. 22, the current setting value of section I of protection 3 is:

$$I'_{3.act(3)} = K'_{rel} \left| \frac{\dot{U}_{21}}{K'_{rel} Z_{BC} + Z_{CD}} \right| \quad (23)$$

The II-stage adaptive current setting value of the available protection 2 is:

$$\begin{cases} I''_{2.act(3)} = K''_{rel} K'_{rel} \left| \frac{\dot{U}_{21}}{K'_{rel} Z_{BC} + Z_{CD}} \right| \\ t_2 = t_1 + \Delta t \end{cases} \quad (24)$$

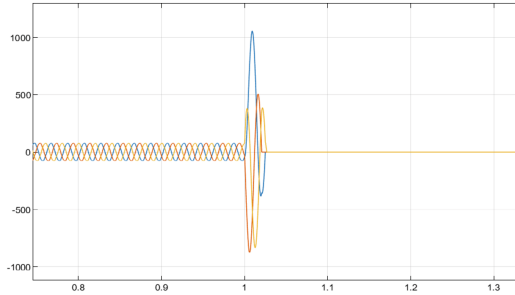
Similarly, the adaptive current II stage of protection 1's setting value is:

$$I''_{1.act(3)} = K''_{rel} K'_{rel} \left| \frac{\dot{U}_{11}}{K'_{rel} Z_{AB} + Z_{BC}} \right| \quad (25)$$

### 4.3 Simulation Verification

To confirm the viability of the protection plan, the distribution network model depicted in Fig. 12 was created using the MATLAB/Simulink simulation platform. The load at the end of the line, assuming that the distributed power generation capacity is 4 MVA, is  $(2 + j0.3)$  MVA, the unit positive impedance of the line is  $(0.03 + j0.27)\Omega/\text{km}$ , and the length of each line segment is 20 km, that is, the line  $LAB = LBC = LCD = LAE = 20$  km, the running time is 5s, and the three-phase short-circuit fault is set to occur in one second.

Using line BC's three-phase short-circuit fault as an example, the short-circuit current at the fault spot and the functioning of protection sections I and II may be determined by simulating and analyzing the scale coefficients  $\alpha = 0.4$  and  $\alpha = 1$  (Fig. 16).

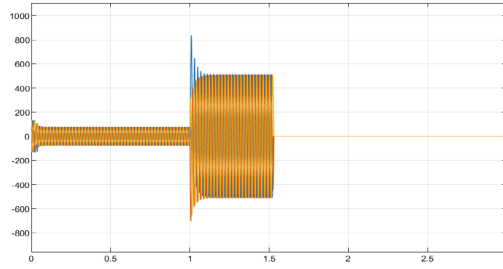


**Fig. 16.** 3–5 Short circuit current at fault  $\alpha = 0.4$  on BC line

According to the simulation results in the above figure, when a three-phase short-circuit fault occurs at 40% of the BC line, section I's current protection and section II's protection delay are both set to 0.5 s, but the circuit breaker trips due to the signal that is sent to it after section I's operation (Fig. 17).

The current protection of sections I and II safeguard the line's whole length because, of course, when the end of the line experiences a three-phase short circuits fault in the distribution network, the current protection of section I does not act when it is 0.5s, and the delay of section II current protection occurs for 0.5s, after which the circuit breaker sends a trip signal, trips, and the fault is cut off at 1s.

When a three-phase short-circuit fault occurs at different locations of the AB line and the BC line, the current settings at protection 1 and protection 2 are displayed in Table 2 and Table 3 below, respectively. The capacity of the distributed power generation added at each location is set at 4MW.



**Fig. 17.** 3–7 Short circuit current at fault  $\alpha = 1$  on BC line

**Table 2.** Current situation at protection 1 in case of fault in section AB

| The location of the fault in the AB segment | Fault Current(A) | Setting value of the first segment (A) | Stage II setting value (A) |
|---------------------------------------------|------------------|----------------------------------------|----------------------------|
| 0.3                                         | 1192             | 431.76                                 | 398.6                      |
| 0.4                                         | 1105             | 534.24                                 | 493.2                      |
| 0.7                                         | 903              | 765.24                                 | 706.5                      |

**Table 3.** Current situation at protection 2 in case of fault in section BC

| The location of the fault in the BC segment | Fault Current(A) | Setting value of the first segment (A) | Stage II setting value (A) |
|---------------------------------------------|------------------|----------------------------------------|----------------------------|
| 0.3                                         | 675              | 250                                    | 230.8                      |
| 0.4                                         | 647              | 319.8                                  | 295.2                      |
| 0.7                                         | 572              | 496.9                                  | 458.7                      |

Tables 2 and 3 demonstrate that in the event of a three-phase symmetrical short circuit in the transmission line, the proportional factor increases with the setting value of the I and II sections of the current protection, while the fault current magnitude decreases as the proportional coefficient of the fault increases.

## 5 Summary

One effective solution to the energy and resource problems is distributed power generation, which is integrated into the power system as a clean, renewable energy source. However, the distribution network's protection has gotten more complicated as a result of distributed generators' access, and the original protection plan and principle have failed, rendering the current protection system unstable. Studying how distributed power generation access affects the distribution network protection system is therefore extremely important. The following are the primary study findings of this paper in order to completely comprehend and address these agreements:

- 1) After creating a model of a solar system, it is connected to the distribution network for research purposes. Taking the simple distributed generator access distribution network model as the research object, the research focuses on analyzing the impact of the capacity, access location and fault location of the distributed generator connected to the power grid on the distribution network. As an example, the three-phase short-circuit fault occurs at the upstream, downstream, and adjacent positions when the distributed generator is connected to the intermediate busbar. The fault current magnitude passing through each protection point is measured, and simulation is used to confirm that the theoretical analysis is accurate.
- 2) The distribution network model with multiple distributed generators is further created based on the detailed analysis of the grid connection of single distributed generators. In the face of the challenges of protection device rejection, maloperation and reliability reduction caused by distributed power access, an innovative adaptive current main protection scheme was proposed, and the simulation platform was used to confirm the protection scheme's efficacy.

**Acknowledgments.** This study was funded by Design Reform of Automation and Electrical Engineering Specialty Based on Engineering Education Professional Certification (No. JG2021Z12).

## References

1. Zhang, X., Zhang, N., Long, F., et al.: Multi-angle study on the influence of distributed generation access and distribution network on its quiescent voltage stability. *Power Syst. Prot. Control* **45**(06), 120–125 (2017)
2. Li, J., Li, X., Ma, Y., et al.: *J. Electr. Power Syst. Autom.* **28**(02), 98–102 (2016)
3. Liu, R., Yang, Y., Chen, H., et al.: Research on the influence index of distribution network line loss considering distributed power access. *Electr. Eng. Electr.* (01), 17–22 (2018)
4. Lin, X., Lu, Y., Wang, L.: *Autom. Electr. Power Syst.* (20), 50–56 (2008)
5. Ma, J., Liu, J., Deng, Z., et al.: An adaptive directional current protection scheme for distribution network with DG integration based on fault steady-state component. *Int. J. Electr. Power Energy Syst.* **102**, 223–234 (2018)
6. Wang, X., Cao, A., Sheng, Y.: Influence of distributed generation on current protection of distribution network. *Electr. Measur. Instrum.* **54**(18), 37–41+78 (2017)
7. Zhang, D., Zhao, H., Liu, C.: *Yunnan Electr. Power Technol.* **40**(06), 27–29 (2012)
8. Wang, Q.: Research on current protection applicable to multiple distributed generator distribution networks. Chongqing University of Technology (2023)
9. Qi, Y.: Research on distribution network protection optimization scheme with distributed generators. Southwest Jiao tong University (2022)
10. Xie, M., Wang, T., Xu, J., et al.: Influence of distributed generation on relay protection of distribution network and comprehensive improvement protection scheme. *Power Syst. Prot. Control* **47**(19), 78–84 (2019)
11. Wu, H.: Influence of distributed generation on current protection of distribution network. *Electr. Technol.* (04), 72–75 (2012)





# Parkinglot Obstacle Detection System Using Infrastructure-Based Cameras

Yuesheng He<sup>1</sup>✉, Fei Wang<sup>1</sup>, Zihan Zong<sup>2</sup>, and Ming Yang<sup>1</sup>

<sup>1</sup> Department of Automation, Shanghai Jiao Tong University, Shanghai, China  
heyuesh@sjtu.edu.cn

<sup>2</sup> University of Michigan - Shanghai Jiao Tong University Joint Institute,  
Shanghai, China

**Abstract.** This paper presents a parking lot obstacle detection system that utilizes infrastructure-based cameras to detect and track obstacles within a parking lot. The system begins by capturing images from ceiling-mounted cameras and employs object detection algorithms to identify 2D detection boxes and estimate 3D bounding boxes' ground points. By leveraging camera intrinsics and extrinsics, along with the assumption that all objects are on a plane, we transform these points into the world coordinate system. The system then fuses detection results from multiple cameras using spatial and temporal fusion techniques, including the mean shift algorithm and the Expanded Kalman Filter, to optimize and refine the detection results. Our approach effectively addresses the challenge of depth estimation in monocular camera setups and provides a robust solution for parking lot obstacle detection.

**Keywords:** Obstacle Detection · Monocular Camera · 3D Bounding Box · Spatial Fusion · Temporal Fusion

## 1 Introduction

Parking lots are intricate environments where vehicles and pedestrians frequently navigate in close proximity, necessitating reliable obstacle detection systems to ensure safety and efficiency. Traditional object detection methods, such as YOLO and DETR, have been widely employed for various applications but encounter challenges when applied to parking lot scenarios due to the lack of depth information in monocular camera images. This paper introduces a novel system that addresses these challenges by reformulating the problem into keypoint estimation and employing both spatial and temporal fusion techniques to enhance detection accuracy. Our system not only detects obstacles but also estimates their 3D bounding boxes, providing a comprehensive solution for parking lot monitoring.

The primary contributions of this paper are twofold:

1. **Efficient and Accurate Detection of Ground Bounding Boxes:** We reformulate the problem of detecting object bounding boxes in the parking lot by estimating the distribution of key points of the objects. This innovative approach

allows us to efficiently and precisely detect the ground-bound boxes of targets within the parking lot. By assuming all targets are on the ground, we focus on detecting the four points of the 3D bounding box on the ground and estimate their distribution through a network. This method effectively addresses the challenge of depth estimation in monocular camera setups, which is a significant advancement over traditional methods that struggle with the lack of depth information.

2. **Real-time Obstacle Detection System Based on Edge-side Monocular Cameras:** We have developed a real-time obstacle detection system based on edge-side monocular cameras for underground parking lots. This system captures images from ceiling-mounted cameras and employs object detection algorithms to identify 2D detection boxes and estimate 3D bounding boxes' ground points. By leveraging camera intrinsics and extrinsics, along with the assumption that all objects are on a plane, we transform these points into the world coordinate system. The system then fuses detection results from multiple cameras using spatial and temporal fusion techniques, including the mean shift algorithm and the Expanded Kalman Filter, to optimize and refine the detection results. This system is designed to operate in real time, providing a robust solution for parking lot obstacle detection.

In summary, this paper presents a cutting-edge approach to parking lot obstacle detection that leverages keypoint estimation and fusion techniques to achieve high accuracy and real-time performance. Our system's ability to estimate 3D bounding boxes and its real-time capabilities make it a valuable tool for enhancing safety and efficiency in parking lot environments. The following sections will detail our methodology, experimental setup, and results, demonstrating the effectiveness of our approach in various scenarios.

## 2 Related Works

The field of parking lot obstacle detection has seen significant advancements with the integration of intelligent vision systems and advanced algorithms. This section reviews the latest research pertinent to our parking lot obstacle detection system.

**Deep Learning in Object Detection** – The application of deep learning, particularly convolutional neural networks (CNNs), has revolutionized object detection. He et al. provide a comprehensive survey on the use of transformers in vision tasks, highlighting their potential in object detection [14]. Dosovitskiy et al. demonstrate the effectiveness of transformers for image recognition at scale, offering a new paradigm for detecting objects in images [10].

**3D Object Detection Techniques** – The transition from 2D to 3D object detection is crucial for accurate obstacle detection in parking lots. Qi et al. discuss deep learning methodologies for 3D object detection and segmentation, which are essential for understanding the spatial context of objects [27]. Chen et al. present a multi-view 3D object detection network tailored for autonomous driving, emphasizing the importance of multiple perspectives in detecting objects [7].

**Monocular Camera Depth Estimation** – One of the challenges in monocular camera setups is the lack of depth information. Zhou et al. survey the field of monocular 3D object detection for autonomous driving, addressing the depth estimation problem [8]. Wang et al. (2022) provide an overview of depth estimation techniques for monocular object detection, which is vital for our system’s accuracy [13].

**Fusion Techniques in Obstacle Detection** – The integration of data from multiple sensors is key to improving detection reliability. Geiger et al. introduce the KITTI vision benchmark suite, which includes multi-sensor fusion for 3D object detection [12]. Li et al. review multi-sensor fusion techniques for 3D object detection, underlining the significance of combining data for enhanced accuracy [24].

**Extended Kalman Filter in Object Tracking** – The Extended Kalman Filter (EKF) is a popular method for fusing detection results over time. Arulampalam et al. offer a tutorial on particle filters for online nonlinear/nongaussian Bayesian tracking, which includes the application of EKF [4]. Simon discusses optimal state estimation techniques, including EKF, which is integral to our system’s object tracking capabilities [18].

**Parking Lot Monitoring Systems** – Specific to parking lots, Zhang et al. explore the detection and occupancy estimation of parking spaces using aerial images, providing insights into monitoring parking lots [2]. Li et al. survey deep learning applications for parking space detection and status estimation, which is directly relevant to our system [20].

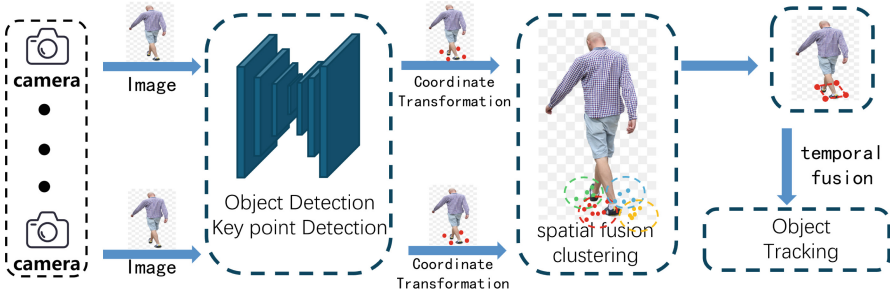
**Camera-Based Obstacle Detection Systems** – The effectiveness of camera-based systems in obstacle detection is demonstrated by Bochkovskiy et al. with YOLOv4, an optimal speed and accuracy object detection framework [5]. Redmon et al. further improve upon this with YOLOv5, an incremental improvement over previous versions [3].

**Multi-View Object Detection and Fusion** – Wang et al. survey multi-view fusion for 3D object detection, emphasizing the importance of integrating information from different viewpoints [1]. Li et al. introduce a graph-based message passing and attention mechanism for multi-view 3D object detection, which could apply to our system [25].

**Mean Shift Algorithm in Clustering** – For spatial fusion, the mean shift algorithm is a robust clustering method. Comaniciu et al. discuss the mean shift approach to feature space analysis, which applies to our system’s spatial fusion of key points [9]. Chen et al. review mean shift clustering techniques, providing further insights into its application [26].

**Real-Time Obstacle Detection Systems** – Real-time detection is crucial for practical applications. Wang et al. present YOLOv7, which sets a new state-of-the-art for real-time object detectors [22]. Bochkovskiy et al. introduce YOLOv8, focusing on scalability and performance for real-time detection [19].

These works provide a solid foundation for the design and implementation of our parking lot obstacle detection system. By building upon these studies, we



**Fig. 1.** System Overview.

aim to enhance the safety and efficiency of parking lot monitoring through the integration of advanced vision systems and algorithms.

### 3 Method

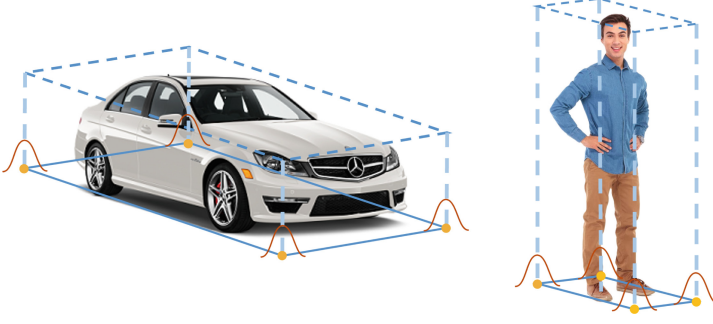
#### 3.1 System Overview

An overview of the parking lot obstacle detection system is shown in Fig. 1. The system starts by receiving images from the ceiling-mounted cameras in the parking lot and then performs object detection to obtain the 2D detection box of the object and its four ground points of the 3d bounding box, i.e., front-left, front-right, back-left, and back-right points. After obtaining the detection results, first, since these four points are on the ground, we project them from the camera coordinate system to the world coordinate system using the camera's intrinsic and extrinsic, obtaining a coarse bounding box of the target in the world coordinates. Then, the spatial fusion of the detection results is required as multiple cameras may detect the same object. After obtaining the spatial fusion results, we perform a temporal fusion of the targets using Expanded Kalman filtering to optimize the target detection results and obtain the final parking lot obstacle detection results.

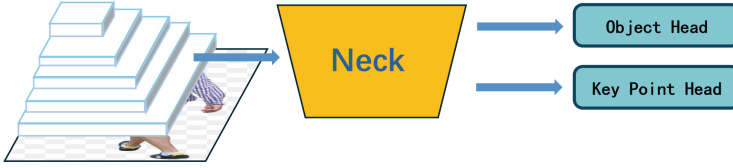
#### 3.2 Obstacle Keypoint Detection

There are already many mature methods for object detection in images, such as YOLO [5, 11, 15–17, 21–23] and DETR [6], which have been applied in numerous real-world tasks. However, for cameras mounted on parking garage ceilings, due to the lack of depth information in the images, the 2D bounding box of an object detected directly in the image cannot be readily converted into a 3D bounding box of the object in the world coordinate system.

Many works are also dedicated to 3D object detection with monocular cameras. However, due to the lack of depth information, this problem often reduces to monocular depth estimation of the object.



**Fig. 2.** The illustration of keypoint detection result



**Fig. 3.** The network structure

We note that in the parking lot, the objects are all on a plane, which is a useful constraint. Therefore, we reformulate the problem of detecting object bounding boxes in the parking lot as estimating the distribution of keypoints of the objects. As shown in Fig. 2, each object has a 3D bounding box that defines its boundaries. However, we assume that all targets in the parking lot are on the ground, so we only detect the four points of the 3D bounding box on the ground (marked as orange points in Fig. 2) and estimate the distribution of these four points through the network.

**Network.** For object and projection point detection, we propose a key point detection method. The network architecture we designed is shown in Fig. 3. The network uses a feature extraction network to extract multi-level features and then fuses these features using a feature pyramid. Finally, the network has two heads to obtain the obstacle detection results and ground key point estimation results separately.

Through the network, we can obtain the 2D bounding box of the object in the image, its category, and the four key points on the ground. And these four points also have corresponding classes, i.e., front-left, front-right, back-left, and back-right points.

**Transformation.** After detecting the object and its four points of the object's 3D bounding box on the ground, and then using the camera's intrinsic and extrinsic along with the ground plane constraint, we obtain the distribution of

these four points in the world coordinate system. The projection formula is as follows:

First, transform the ground plane to the camera coordinate system.

$$\begin{bmatrix} \hat{a} \\ \hat{b} \\ \hat{c} \\ \hat{d} \end{bmatrix} = T \begin{bmatrix} a \\ b \\ c \\ d \end{bmatrix} \quad (1)$$

where  $a, b, c, d$  are the plane parameters( $ax + by + cz = d$ ).

Then, for each point

$$depth = \frac{-\hat{d}}{\hat{a}(x - c_x)/f_x + \hat{b}(y - c_y)/f_y + \hat{c}} \quad (2)$$

where  $c_x, c_y, f_x, f_y$  is camera's intrinsic.

Thus, the point is

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = T^{-1} \begin{bmatrix} \frac{depth(x - c_x)}{f_x} \\ \frac{depth(y - c_y)}{f_y} \\ depth \end{bmatrix} \quad (3)$$

### 3.3 Result Fusion

Because multiple cameras from different perspectives may simultaneously detect the same object, and because the estimated ground key points are assumed to follow a normal distribution, fusing results from multiple views can provide a better estimation of the object's ground bounding box.

Since the estimated ground key points have class information, we only need to cluster the key points of the same class for all detection results. Considering that the detection results will have outliers, the mean shift algorithm is a suitable algorithm to compute the clustering algorithm that removes outliers.

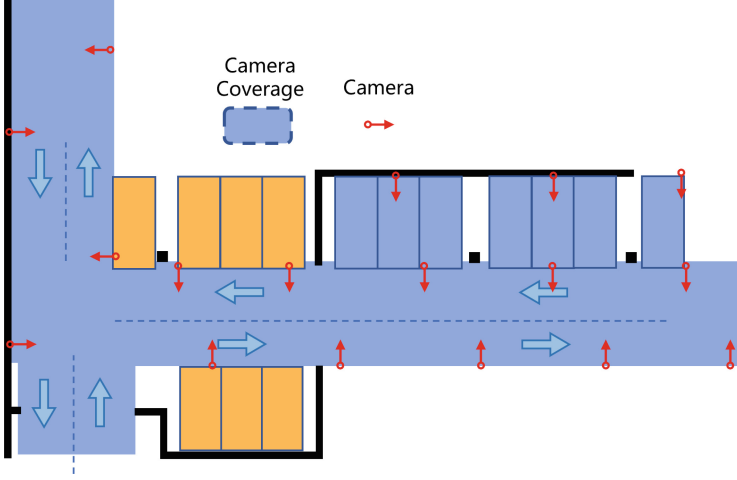
The mean shift algorithm finds stable outlier-free keypoint means by repeatedly calculating the direction of shift of the mean of points within the Euclidean distance threshold to achieve spatial fusion of the same keypoint class.

### 3.4 Object Tracking

An Extended Kalman Filter (EKF) is adopted for tracking and fusing the obstacle detection results of object 3D bounding box in coordinate system of the world. The EKF is an ideal filter for removing noise between the detection noise using temporal fusion.

The obstacle state aims to estimate is

$$x = \begin{bmatrix} x \\ y \\ \theta \\ w \\ h \end{bmatrix}, \quad (4)$$



**Fig. 4.** The sensor network.

representing obstacle's position and its rotation angle, the bounding box's width and height. Then, the standard procedure of EKF propagation procedure is applied to update the obstacle's state.

First, use the prediction equation to predict the state of the obstacle,

$$\hat{x}_{\bar{k}} = A\hat{x}_{k-1} + Bu_{k-1} \quad (5)$$

$$P_{\bar{k}} = AP_{k-1}A^T + Q \quad (6)$$

Then after obtaining the observations  $z_k$ , the state of the Kalman filter is updated to obtain the spatial fusion results.

$$K_k = \frac{P_{\bar{k}}H^T}{HP_{\bar{k}}H^T + R} \quad (7)$$

$$\hat{x}_k = \hat{x}_{\bar{k}} + K_k(z_k - H\hat{x}_{\bar{k}}) \quad (8)$$

$$P_k = (I - K_kH)P_{\bar{k}} \quad (9)$$

## 4 Experiments

### 4.1 Sensor Network and Data Prepare

We deployed the system proposed in Sect. 3 in an indoor parking lot. We installed a total of 17 cameras on the roadside and connected these cameras with network cables to form a sensor network. As shown in Fig. 4, the blue area represents the coverage of the camera, containing roughly 50 m of the driveway and 7 parking spaces, the red arrows represent the camera's mounting position and orientation,

with a field of view of  $120^\circ$  the camera was selected to be mounted at about  $45^\circ$  downwards.

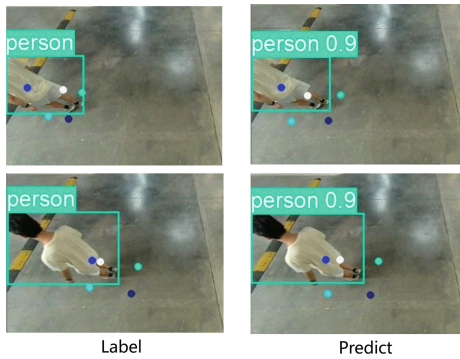
To prepare data for training the detection network, we also installed LiDARs in the parking lot to assist in accurately annotating 3D detection boxes for objects in the images. After calibrating the LiDAR and camera, we only need to annotate the 3D detection boxes for objects in the LiDAR point cloud. Then, we can convert these detection boxes into the camera’s perspective to obtain the 3D detection boxes of objects in the images.

## 4.2 Detection Experiment

**Table 1.** Model Detection Result

| Method | Bounding Box |              | Key Point |              |
|--------|--------------|--------------|-----------|--------------|
|        | $AP_{50}$    | $AP_{50-95}$ | $AP_{50}$ | $AP_{50-95}$ |
| Ours   | 0.98         | 0.938        | 0.97      | 0.955        |

First, we labeled approximately 800 images for training the model and we trained and validated the proposed network using the self-collected dataset. After training, our model achieved the accuracy shown in Table 1. The evaluation metrics of the model primarily consist of two parts: object detection and detection of object’s ground points, i.e. bounding box and key points. For evaluation metrics, Intersection over Union (IoU) is used for bounding boxes. For key points, we used the object keypoint similarity(OKS) from COCO. Therefore, the  $AP_{50}$  for key point represents the average precision when keypoint similarity is greater than 0.5, considering those as true positives.



**Fig. 5.** Object Detection and Keypoint Results

Due to the small training set, we achieved high detection accuracy and key-point estimation accuracy on this dataset. The detection results are shown in



the Fig. 5. The results indicate that our proposed network can detect objects in the parking lot and estimates the distribution of key points for their ground bounding boxes.

### 4.3 System Experiment

In the previous section, we demonstrated that the proposed network can accurately detect obstacles in the parking lot. Our proposed obstacle detection system also needs to evaluate the accuracy of detected obstacles in the world coordinate system.

Therefore, we deployed the trained model into the proposed system and used ground truth bounding boxes in the world coordinate system to evaluate the accuracy of our system. The experiment results are shown in the Table 2. We mainly used two metrics to evaluate our proposed system: the Root Mean Square Error (RMSE) of the object trajectory and the mIoU of the object ground bounding box. These metrics assess the accuracy of the detected objects positions and bounding boxes, respectively. Our method achieves a lower RMSE and a higher IoU for ground bounding boxes compared to the direct 2D projection detection method.

**Table 2.** Detection System Result

| Method | Trajectory(m) | Ground mIoU  |
|--------|---------------|--------------|
| 2D Box | 1.908         | 0.046        |
| Ours   | <b>0.162</b>  | <b>0.561</b> |

At the same time, since the obstacle detection system is a real-time system, we also evaluated the inference time of our proposed system. Given that our system includes multiple cameras, we perform inference with a batch size of 8 on RTX3090. The inference time is shown in the Table 3.

**Table 3.** Inference Time

| Method | Per Batch(ms) | Per Image(ms) |
|--------|---------------|---------------|
| Ours   | 37.6          | 4.7           |

## 5 Conclusion

In this study, we have successfully developed and evaluated a parking lot obstacle detection system that utilizes infrastructure-based cameras, addressing the

critical need for accurate and real-time monitoring in complex parking environments. Our system's innovative approach of reframing the ground bounding box detection problem as a keypoint distribution estimation issue has resulted in a significant advancement in the field of computer vision for parking lot applications.

The primary contributions of this research are the effective transformation of the detection problem into a keypoint estimation challenge and the implementation of a real-time, edge-based monocular camera system for underground parking lot obstacle detection. Our system demonstrates the ability to not only detect obstacles but also to estimate their 3D bounding boxes with high precision, which is a substantial improvement over traditional monocular camera setups that struggle with depth estimation.

The employment of spatial and temporal fusion techniques, including the mean shift algorithm and the Expanded Kalman Filter, has proven to be effective in refining detection results and enhancing the robustness of our system. The experiments conducted within an indoor parking lot, equipped with a network of 17 cameras, have yielded promising results. The system achieved high accuracy in detecting objects and estimating key points, with an AP50 score of 0.98 for bounding boxes and 0.97 for key points. Furthermore, the system's real-time performance, with an inference time of 4.7 milliseconds per image, underscores its practical applicability.

The results of our system experiment, which used ground truth bounding boxes in the world coordinate system for evaluation, showed a significant reduction in the Root Mean Square Error (RMSE) of object trajectory and an increase in the mean Intersection over Union (mIoU) of object ground bounding boxes compared to traditional 2D box methods. These improvements highlight the effectiveness of our approach in providing accurate spatial of obstacles within the parking lot.

In conclusion, our parking lot obstacle detection system offers a comprehensive solution for monitoring and ensuring safety in parking lots. The system's ability to estimate 3D bounding boxes in real time, coupled with its high accuracy and robustness, makes it a valuable asset for intelligent transportation systems and autonomous vehicle technologies. Future work will focus on expanding the system's capabilities, incorporating additional sensors, and testing the system in diverse parking lot scenarios to further enhance its performance and reliability.

**Disclosure of Interests.** The authors have no competing interests to declare that are relevant to the content of this article.

## References

1. Review of multi-view 3D object recognition methods based on deep learning. *Displays* **69**, 102053 (2021)
2. Acharya, D., Yan, W., Khoshelham, K.: Real-time image-based parking occupancy detection using deep learning. *Research@ Locate* **4**, 33–40 (2018)

3. Jocher, G., et al.: ultralytics/YOLOv5: v6.0 - YOLOv5n 'Nano' Models, Roboflow Integration, TensorFlow Export, OpenCV DNN support (2021)
4. Arulampalam, M.S., Maskell, S., Gordon, N., Clapp, T.: A tutorial on particle filters for online nonlinear/non-gaussian Bayesian tracking. *IEEE Trans. Signal Process.* **50**(2), 174–188 (2002)
5. Bochkovskiy, A., Wang, C.Y., Liao, H.Y.M.: YOLOv4: optimal speed and accuracy of object detection. *arXiv preprint [arXiv:2004.10934](https://arxiv.org/abs/2004.10934)* (2020)
6. Carion, N., Massa, F., Synnaeve, G., Usunier, N., Kirillov, A., Zagoruyko, S.: End-to-end object detection with transformers. In: *European Conference on Computer Vision*, pp. 213–229. Springer, Cham (2020)
7. Chen, X., et al.: Multi-view 3D object detection network for autonomous driving. *IEEE Trans. Pattern Anal. Mach. Intell.* (2022)
8. Chen, X., Kundu, K., Zhang, Z., Ma, H., Fidler, S., Urtasun, R.: Monocular 3D object detection for autonomous driving. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 2147–2156 (2016)
9. Comaniciu, D., Meer, P.: Mean shift: a robust approach toward feature space analysis. *IEEE Trans. Pattern Anal. Mach. Intell.* **24**(5), 603–619 (2002)
10. Dosovitskiy, A., et al.: An image is worth 16 × 16 words: transformers for image recognition at scale. *arXiv preprint [arXiv:2010.11929](https://arxiv.org/abs/2010.11929)* (2021)
11. Farhadi, A., Redmon, J.: YOLOv3: an incremental improvement. In: *Computer Vision and Pattern Recognition*, vol. 1804, pp. 1–6. Springer, Heidelberg (2018)
12. Geiger, A., Lenz, P., Stiller, C.: Are we ready for autonomous driving? The KITTI vision benchmark suite. In: *2012 IEEE Conference on Computer Vision and Pattern Recognition*, pp. 3365–3372. IEEE (2012)
13. Huang, K.C., Wu, T.H., Su, H.T., Hsu, W.H.: MonoDTR: monocular 3D object detection with depth-aware transformer. In: *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 4012–4021 (2022)
14. Khan, S., Naseer, M., Hayat, M., Zamir, S.W., Khan, F.S., Shah, M.: Transformers in vision: a survey. *54*(10s) (2022)
15. Li, C., et al.: YOLOv6: a single-stage object detection framework for industrial applications. *arXiv preprint [arXiv:2209.02976](https://arxiv.org/abs/2209.02976)* (2022)
16. Redmon, J.: You only look once: unified, real-time object detection. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (2016)
17. Redmon, J., Farhadi, A.: YOLO9000: better, faster, stronger. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 7263–7271 (2017)
18. Simon, D.: *Optimal State Estimation: Kalman, H Infinity, and Nonlinear Approaches*. Wiley, Hoboken (2019)
19. Varghese, R., Sambath, M.: YOLOv8: a novel object detection algorithm with enhanced performance and robustness. In: *2024 International Conference on Advances in Data Engineering and Intelligent Computing Systems (ADICS)*, pp. 1–6 (2024)
20. Vu, H.T., Huang, C.C.: Parking space status inference upon a deep CNN and multi-task contrastive network with spatial transform. *IEEE Trans. Circuits Syst. Video Technol.* **29**(4), 1194–1208 (2019)
21. Wang, A., et al.: YOLOv10: real-time end-to-end object detection. *arXiv preprint [arXiv:2405.14458](https://arxiv.org/abs/2405.14458)* (2024)
22. Wang, C.Y., Bochkovskiy, A., Liao, H.Y.M.: YOLOv7: trainable bag-of-freebies sets new state-of-the-art for real-time object detectors. In: *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 7464–7475 (2023)

23. Wang, C.Y., Yeh, I.H., Liao, H.Y.M.: YOLOv9: learning what you want to learn using programmable gradient information. arXiv preprint [arXiv:2402.13616](https://arxiv.org/abs/2402.13616) (2024)
24. Wang, X., Li, K., Chehri, A.: Multi-sensor fusion technology for 3D object detection in autonomous driving: a review. *IEEE Trans. Intell. Transp. Syst.* **25**(2), 1148–1165 (2024)
25. Wu, S., et al.: Graph-based 3D multi-person pose estimation using multi-view images, pp. 11128–11137 (2021)
26. Yang, J., Rahardja, S., Fränti, P.: Mean-shift outlier detection and filtering. *Pattern Recogn.* **115**, 107874 (2021)
27. Yi, F., Jeong, O., Moon, I., Javidi, B.: Deep learning integral imaging for three-dimensional visualization, object detection, and segmentation. *Opt. Lasers Eng.* **146**, 106695 (2021)



# Predefined-Time Consensus for First-Order Heterogeneous Nonlinear Multi-agent Systems

Shufen Liu, Jianhua Liu()<sup>✉</sup>, Feng Zhao, and Xiangyong Chen

School of Automation and Electrical Engineering, Linyi University, 27600 Linyi,  
China

fzhao1986163@163.com

**Abstract.** This paper studies the predefined-time consensus of first-order heterogeneous nonlinear multi-agent systems. In a leaderless communication topology, a novel control protocol is devised based on the states of the agents neighbors. By employing graph theory, predefined-time control techniques, and constructing relevant Lyapunov functions, the predefined-time consensus can be actualized for the multi-agent system. Eventually, the correctness of the results is tested through a simulation example.

**Keywords:** Predefined-time consensus · Multi-agent system · Leaderless case · Nonlinear system · Heterogeneous system

## 1 Introduction

A multi-agent system(MAS) is consisted of lots of agents that it can mutually communicate and exchange information. Due to its independent perception, decision-making, and communication capabilities [1], it has drawn the interest and research of many savants in other domains, and has been widely applied in physics, biology, computer research, mathematics, management science research, robotics technology, and communication technology [2]. The study of consensus issues is the foundation of multi-agent systems(MASs) research. Only when the MASs achieves consensus can further research on multi-agent systems be conducted at a deeper level. In the process of studying the uniformity of multi-agent systems, from studying finite-time consensus [5–7] to studying fixed-time consensus [8,9], and then to studying predefined-time consensus [10–12], numerous researchers have achieved many research results in each part, laying the foundation for studying other parts of MASs, for example, formation control, inclusion control, clustering and swarming, distributed optimization and decision-making, controllability, etc. [3,4].

Predefined-time control(PTC) is a special form of fixed-time control(FTC) method. Unlike FTC, predefined-time control is not relying on the initial state of the system and is not affected by the setting of controller parameters. Therefore,

many scholars have conducted in-depth research on prespecified-time control. For example, reference [13] proposed a new distributed controller to achieve PTC in high-order dynamic MASs. Reference [14] presented a PTC control algorithm based on dynamic event-triggering mechanisms applicable to nonlinear MASs. Additionally, reference [15] came up with a distributed observer and constructed a new sliding surface by introducing a terminal sliding mode consensus protocol, achieving PTC tracking in second-order MASs. Reference [16] also proposed a new stochastic prespecified-time consensus protocol for PTC in stochastic MASs based on connected undirected graphs. There have been many methods and achievements in the study of predefined-time consensus. In this paper, a new protocol framework will be constructed to achieve PTC in first-order nonlinear multi-agent systems.

*Notations:*  $\mathbb{R}^{m \times n}$  represents a  $m \times n$  real matrix.  $w = [w_1, w_2, \dots, w_n]^T \in \mathbb{R}^n$  is a vector, denotes  $\text{sig}^\alpha(w) = [|w_1|^\alpha \text{sign}(w_1), \dots, |w_n|^\alpha \text{sign}(w_n)]^T$ , the  $\text{sign}(\cdot)$  represents the sign function.  $\|\cdot\|_p$  represents the  $p$  norm. For any  $t \geq 0$ , abbreviate  $r(t)$  as  $r$ , where  $r$  represent symbols used in this article.

## 2 Preliminaries

### 2.1 Graph Theory

In the study of MASs, graph theory is commonly employed to depict the relationships among agents, utilizing its vertices and edges to signify the connections between agents. A graph  $G = \{M, \Xi, A\}$  is utilized to represent the communication topology of intelligent agents, where  $M = \{m_i, i \in \mathcal{M}, \mathcal{M} = 1, 2, \dots, \sigma\}$  is the vertices set and represents the collection of agents;  $\Xi \subseteq M \times M$  is the edges set and indicates a communication channel between two agents; the adjacency matrix  $A$  represents the connectivity of the graph  $G$ , where  $a_{ij} > 0$  denotes the existence of a communication channel between two agents, and  $a_{ij} = 0$  signifies its absence.  $L = [l_{ij}]_{\sigma \times \sigma}$  as the Laplacian matrix of graph  $G$ , when  $i \neq j$ ,  $l_{ij} = -a_{ij}$ , and  $l_{ii} = \sum_{j=1, j \neq i}^{\sigma} a_{ij}$ .

### 2.2 Related Lemmas

Consider the following system,

$$\dot{h}(t) = g_i(h(t)) \quad (1)$$

where  $h \in \mathbb{R}^n$ ,  $g_i : \mathbb{R}^+ \times \mathbb{R}^n \rightarrow \mathbb{R}^n$  represents a non-linear function.

*Definition 1* [20]: For system (1), if its origin is fixed-time stable, it is said to be predefined-time stable and the stable time function is  $T : \mathbb{R}^n \rightarrow \mathbb{R}$ , when satisfied  $\forall h_0 \in \mathbb{R}^n : T(h_0) \leq T_c$ ,  $T_c$  is defined as the predefined-time.

*Lemma 1* [20]: If there exists an unbounded Lyapunov function  $V(h)$  for the system (1), the following formula is satisfied:

$$\dot{V} \leq -\frac{1}{T_c} \frac{2}{\varphi} (2V + V^{1-\frac{\varphi}{2}} + V^{1+\frac{\varphi}{2}}) \quad (2)$$

where  $T_c > 0$  is a constant,  $\varphi \in (0, 1)$ , and system (1) is said to be predefined-time stable.

*Lemma 2* [18]: If  $L$  is the Laplacian matrix of a graph that is a connected graph with  $\sigma$  vertices: (1) For any  $z = [z_1, z_2, \dots, z_\sigma]^T \in \mathbb{R}^n$ , have  $z^T L z = \frac{1}{2} \sum_{i=1}^{\sigma} \sum_{j=1}^{\sigma} a_{ij} (z_j - z_i)^2$ . (2) Only one eigenvalue of  $L$  is zero, while all other eigenvalues are positive.

*Lemma 3* [19]:  $\forall m \in \mathbb{R}^n, h > q \geq 1$ , it holds that  $\|m\|_h \leq \|m\|_q \leq n^{\frac{1}{h}-\frac{1}{q}} \|m\|_h$ .

### 2.3 System Description

Research a multi-agent system composed of  $\sigma$  agents. Define  $\chi = [\chi_1, \chi_2, \dots, \chi_\sigma]^T$ ,  $\chi_i \in \mathbb{R}$ , where  $\chi_i$  is the state of the  $i$ th intelligent agent. The dynamic model of each agent is as follows,

$$\dot{\chi}_i(t) = \chi_i(t) + u_i(t) + g_i(\chi_i(t)) \quad (3)$$

where  $u_i(t) \in \mathbb{R}$  is the control input of the  $i$ th agent,  $g_i : \mathbb{R}^+ \times \mathbb{R}^n \rightarrow \mathbb{R}^n$  is an unknown non-linear function.

*Assumption 1* [19]:  $\bar{g}_i(\chi_i(t))$  is a known function and satisfy a inequality as follows,

$$|g_i(\chi_i(t))| \leq \bar{g}_i(\chi_i(t)) \quad (4)$$

*Assumption 2* [17]: The topology structure of intelligent agents is connected and undirected.

## 3 Main Results

Design a new control protocol as follows,

$$\begin{aligned} \dot{\chi}_i^s &= p_{i,1}(\chi_j - \chi_i, j \in \xi_i) \\ u_i &= p_{i,2}(\chi_j - \chi_i, j \in \xi_i, \chi_i - \chi_i^s), \forall i \in \mathcal{M} \end{aligned} \quad (5)$$

where,  $\chi_i^s$  represents the target trajectory of the  $i$ th agent;  $p_{i,1}(\cdot)$  and  $p_{i,2}(\cdot)$  represent operators that has not been designed and will be designed later;  $\xi_i$  represents the neighbor set of the  $i$ th agent.

Let  $\rho_i = \chi_i - \chi_i^s$ , construct  $p_{i,1}(\cdot), p_{i,2}(\cdot)$  as follows,

$$\begin{aligned} & p_{i,1}(\chi_j - \chi_i, j \in \xi_i) \\ &= \sum_{j=1}^{\sigma} a_{ij}(K_1 \text{sig}^{1-\beta}(\chi_j - \chi_i) + K_2 \text{sig}^{1+\beta}(\chi_j - \chi_i) + K_3 \text{sig}(\chi_j - \chi_i)) \end{aligned} \quad (6)$$

$$\begin{aligned} & p_{i,2}(\chi_j - \chi_i, j \in \xi_i, \chi_i - \chi_i^s) \\ &= p_{i,1}(\chi_j - \chi_i, j \in \xi_i) - C \text{sig}(\rho_i) - C_1 \text{sig}^{1-\alpha}(\rho_i) - C_2 \text{sig}^{1+\alpha}(\rho_i) - \chi_i \\ & \quad - \bar{g}_i(\chi_i) \text{sign}(\rho_i) \end{aligned} \quad (7)$$

where  $K_1, K_2, K_3 > 0, C, C_1, C_2 > 0, \beta(0, 1), \alpha(0, 1)$ .

*Theorem 1* : Under Assumption 1–2, the multi-agent system (3) can achieve PTC through protocol (5), (6), (7).

*Proof* : Construct Lyapunov function  $V_1 = \frac{1}{2}\rho_i^2, \forall i \in \mathcal{M}$ , its derivative as follows,

$$\begin{aligned} \dot{V}_1 &= \rho_i \dot{\rho}_i = \rho_i(\dot{\chi}_i - \dot{\chi}_i^s) \\ &= \rho_i(\chi_i + u_i + j_i(\chi_i) - p_{i,1}(\chi_j - \chi_i, j \in \xi_i)) \\ &= \rho_i(\chi_i + u_i + j_i(\chi_i) - p_{i,2}(\chi_j - \chi_i, j \in \xi_i, \chi_i - \chi_i^s) - C \text{sig}(\rho_i) \\ & \quad - C_1 \text{sig}^{1-\alpha}(\rho_i) - C_2 \text{sig}^{1+\alpha}(\rho_i) - \bar{g}_i(\chi_i) \text{sign}(\rho_i) - \chi_i) \\ &= \rho_i(g_i(\chi_i) - C \text{sig}(\rho_i) - C_1 \text{sig}^{1-\alpha}(\rho_i) - C_2 \text{sig}^{1+\alpha}(\rho_i) - \bar{g}_i(\chi_i) \text{sign}(\rho_i)) \end{aligned} \quad (8)$$

according to (4), (8), we have

$$\begin{aligned} \dot{V}_1 &\leq \rho_i(-C \text{sig}(\rho_i) - C_1 \text{sig}^{1-\alpha}(\rho_i) - C_2 \text{sig}^{1+\alpha}(\rho_i)) \\ &\leq \rho_i(-C|\rho_i| - C_1|\rho_i|^{1-\alpha} - C_2|\rho_i|^{1+\alpha}) \\ &\leq -C|\rho_i|^2 - C_1|\rho_i|^{2-\alpha} - C_2|\rho_i|^{2+\alpha} \end{aligned} \quad (9)$$

because of  $V_1 = \frac{1}{2}\rho_i^2$ , we have

$$\begin{aligned} \dot{V}_1 &\leq -2CV_1 - C_1 2^{\frac{2-\alpha}{2}} V_1^{1-\frac{\alpha}{2}} - C_2 2^{\frac{2+\alpha}{2}} V_1^{1+\frac{\alpha}{2}} \\ &\leq -\frac{1}{T_{c1}} \frac{2}{\alpha} (2V_1 + V_1^{1-\frac{\alpha}{2}} + V_1^{1+\frac{\alpha}{2}}) \end{aligned} \quad (10)$$

by Lemma 1, the actual trajectory of each agent is consistent with the target trajectory at a predefined time, that is  $\chi_i = \chi_i^s$ .

In order to make all agents follow a same target, take the average trajectory of the agent's target, denoted as  $\bar{\chi}^s(t) = \frac{1}{\sigma} \sum_{i=1}^{\sigma} \chi_i^s(t)$ , and let  $\Delta p_i = p_{i,1}(\chi_j - \chi_i, j \in \xi_i) - p_{i,1}(\chi_j^s - \chi_i^s, j \in \xi_i)$ . Differentiating  $\bar{\chi}^s(t) = \frac{1}{\sigma} \sum_{i=1}^{\sigma} \chi_i^s(t)$  yields that

$$\begin{aligned} \dot{\bar{\chi}}^s(t) &= \frac{1}{\sigma} \sum_{i=1}^{\sigma} \dot{\chi}_i^s(t) \\ &= \frac{1}{\sigma} \sum_{i=1}^{\sigma} (p_{i,1}(\chi_j^s - \chi_i^s, j \in \xi_i) + \Delta p_i) \end{aligned} \quad (11)$$



because of

$$\begin{aligned} & \sum_{i=1}^{\sigma} \sum_{j=1}^{\sigma} a_{ij} (K_1 \text{sig}^{1-\beta}(\chi_j - \chi_i) + K_2 \text{sig}^{1+\beta}(\chi_j - \chi_i) + K_3 \text{sig}(\chi_j - \chi_i)) \\ &= - \sum_{i=1}^{\sigma} \sum_{j=1}^{\sigma} a_{ij} (K_1 \text{sig}^{1-\beta}(\chi_j - \chi_i) + K_2 \text{sig}^{1+\beta}(\chi_j - \chi_i) + K_3 \text{sig}(\chi_j - \chi_i)) \end{aligned}$$

so we can obtain  $\dot{\bar{\chi}}^s(t) = \frac{1}{\sigma} \sum_{i=1}^{\sigma} \Delta p_i$  and  $\dot{\bar{\chi}}^s(t) = 0$ , which indicates that the average trajectory of the agent's target is a constant when  $T > T_{c1}$ .

Let  $\rho_i^s = \chi_i^s - \bar{\chi}^s$ , select  $V_2$  as the Lyapunov function, where  $V_2$  is expressed as follows,

$$V_2(t) = \frac{1}{2} \sum_{i=1}^{\sigma} \rho_i^s(t)^2 \quad (12)$$

then, differentiating  $V_2(t)$  yields that

$$\begin{aligned} \dot{V}_2(t) &= \sum_{i=1}^{\sigma} \rho_i^s \dot{\rho}_i^s \\ &= \sum_{i=1}^{\sigma} \rho_i^s (p_{i,1}(\chi_j^s - \chi_i^s, j \in \xi_i)) = \sum_{i=1}^{\sigma} \rho_i^s (p_{i,1}(\rho_j^s - \rho_i^s, j \in \xi_i)) \\ &= \sum_{i=1}^{\sigma} \rho_i^s \left( \sum_{j=1}^n a_{ij} (K_1 \text{sig}^{1-\beta}(\rho_j^s - \rho_i^s) + K_2 \text{sig}^{1+\beta}(\rho_j^s - \rho_i^s) + K_3 \text{sig}(\rho_j^s - \rho_i^s)) \right) \\ &= \sum_{i=1}^{\sigma} \sum_{j=1}^{\sigma} a_{ij} \rho_i^s (K_1 |\rho_j^s - \rho_i^s|^{1-\beta} + K_2 |\rho_j^s - \rho_i^s|^{1+\beta} + K_3 |\rho_j^s - \rho_i^s|) \\ &= \frac{1}{2} \sum_{i=1}^{\sigma} \sum_{j=1}^{\sigma} a_{ij} (K_1 |\rho_j^s - \rho_i^s|^{2-\beta} + K_2 |\rho_j^s - \rho_i^s|^{2+\beta} + K_3 |\rho_j^s - \rho_i^s|^2) \\ &= \frac{K_1}{2} \sum_{i=1}^{\sigma} \sum_{j=1}^{\sigma} (a_{ij}^{\frac{2}{2-\beta}} |\rho_i^s - \rho_j^s|^2)^{\frac{2-\beta}{2}} + \frac{K_2}{2} \sum_{i=1}^{\sigma} \sum_{j=1}^{\sigma} (a_{ij}^{\frac{2}{2+\beta}} |\rho_i^s - \rho_j^s|^2)^{\frac{2+\beta}{2}} \\ &\quad + \frac{K_3}{2} \sum_{i=1}^{\sigma} \sum_{j=1}^{\sigma} (a_{ij} |\rho_i^s - \rho_j^s|^2) \end{aligned} \quad (13)$$

from Lemma 3, have

$$\begin{aligned} \dot{V}_2(t) &\leq -\frac{K_1}{2} \left( \sum_{i=1}^{\sigma} \sum_{j=1}^{\sigma} a_{ij}^{\frac{2}{2-\beta}} |\rho_i^s - \rho_j^s|^2 \right)^{\frac{2-\beta}{2}} \\ &\quad - \frac{K_2}{2} (2\sigma)^{-\frac{\beta}{2}} \left( \sum_{i=1}^{\sigma} \sum_{j=1}^{\sigma} a_{ij}^{\frac{2}{2+\beta}} |\rho_i^s - \rho_j^s|^2 \right)^{\frac{2+\beta}{2}} - \frac{K_3}{2} \left( \sum_{i=1}^{\sigma} \sum_{j=1}^{\sigma} a_{ij} |\rho_i^s - \rho_j^s|^2 \right) \end{aligned} \quad (14)$$

by Lemma 2, have

$$\begin{aligned}
\sum_{i=1}^{\sigma} \sum_{j=1}^{\sigma} a_{ij}^{\frac{2}{2-\beta}} |\rho_i^s - \rho_j^s|^2 &= 2\rho^{s^T} L_{\beta 1} \rho^s \geq 4\lambda_2(L_{\beta 1})V_2 \\
\sum_{i=1}^{\sigma} \sum_{j=1}^{\sigma} a_{ij}^{\frac{2}{2+\beta}} |\rho_i^s - \rho_j^s|^2 &= 2\rho^{s^T} L_{\beta 2} \rho^s \geq 4\lambda_2(L_{\beta 2})V_2 \\
\sum_{i=1}^{\sigma} \sum_{j=1}^{\sigma} a_{ij} |\rho_i^s - \rho_j^s|^2 &= 2\rho^{s^T} L_{\beta 3} \rho^s \geq 4\lambda_2(L_{\beta 3})V_2
\end{aligned} \tag{15}$$

according to (14), (15), we have

$$\begin{aligned}
\dot{V}_2(t) &\leq -K_1 2^{1-\beta} \lambda_2^{1-\frac{\beta}{2}}(L_{\beta 1}) V_2^{1-\frac{\beta}{2}} \\
&\quad - K_2 (2\sigma)^{-\frac{\beta}{2}} 2^{1+\beta} \lambda_2^{1+\frac{\beta}{2}}(L_{\beta 2}) V_2^{1+\frac{\beta}{2}} - 2K_3 \lambda_2(L_{\beta 3}) V_2 \\
&\leq -\frac{1}{T_{c2}} \frac{2}{\beta} (2V_2 + V_2^{1-\frac{\beta}{2}} + V_2^{1+\frac{\beta}{2}})
\end{aligned} \tag{16}$$

by Lemma 1, it holds that  $\chi_i^s = \chi_j^s$ ,  $\forall t \geq T_c$ ,  $T_c = T_{c1} + T_{c2}$ . In the formula (15),  $L_{\beta 1}, L_{\beta 2}, L_{\beta 3}$  represent three Laplacian matrices corresponding to graphs  $G(A^{\frac{2}{2-\beta}})$ ,  $G(A^{\frac{2}{2+\beta}})$  and  $G(A)$ .

Finally, it can be concluded that the system (3) can attain the PTC within the given time  $T_c$ .

## 4 Simulation

Consider six intelligent agents with the dynamic characteristics described in (3), and the nonlinear functions are as follows:

$$\begin{aligned}
g_1(\chi_1) &= 0.2\chi_1 + 0.3 \sin \chi_1, g_2(\chi_2) = 0.6 \sin \chi_2 \\
g_3(\chi_3) &= 0.5 \sin \chi_3, g_4(\chi_4) = |\chi_4| \\
g_5(\chi_5) &= 0.1 \cos \chi_5, g_6(\chi_6) = 0.2e^{-|\chi_6|}
\end{aligned} \tag{17}$$

where can learn  $\forall i \in \{1, 2, 3, 4, 5, 6\}$ , and Fig. 1 represents the topology of these six intelligent agents. Give initial state of intelligent agents:  $\chi_1(0) = -2$ ,  $\chi_2(0) = 2.5$ ,  $\chi_3(0) = -1.4$ ,  $\chi_4(0) = 2.5$ ,  $\chi_5(0) = 1.5$ ,  $\chi_6(0) = -2.5$ , and protocol parameters:  $\alpha = 0.5$ ,  $\beta = 0.5$ ,  $K_1 = 2$ ,  $K_2 = 3$ ,  $K_3 = 2$ ,  $C = 2$ ,  $C_1 = 1$ ,  $C_2 = 3$ . Incorporate  $|f_i(\chi_i)| \leq |\chi_i| + 1$  into the designed protocol.

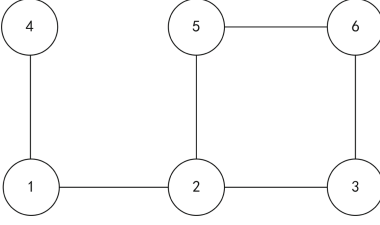
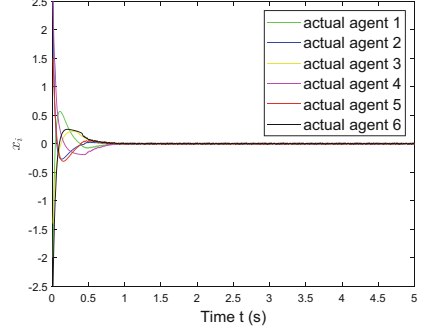
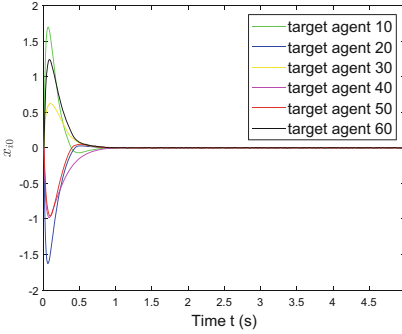
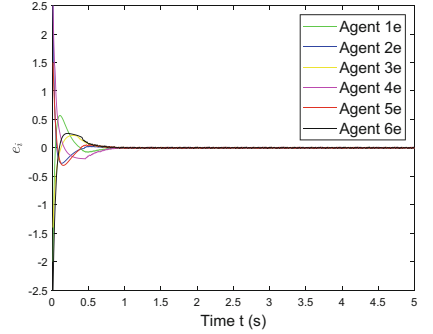
**Fig. 1.** Topology of multi-agent system**Fig. 2.** Agents' actual state**Fig. 3.** Agents' target state**Fig. 4.** Actual trajectory and average target trajectory error

Figure 2 represents the actual trajectory of the agents, Fig. 3 represents the target trajectory of the agents, and Fig. 4 represents the difference between the actual trajectory and the target trajectory. By analyzing Figs. 2, 3, and 4, it is clear that the system (3) can achieve PTC and the predefined-time is 1 s.

## 5 Conclusion

This study addresses the PTC problem in first-order nonlinear heterogeneous multi-agent systems. To achieve this, a protocol that relies solely on state transmission is designed. Within this protocol framework, the system can achieve consensus within a predetermined time, thereby realizing predefined-time consensus control. In future work, we will consider applying this method to multi-agent systems with leader following.

**Acknowledgments.** This research was supported in part by the National Natural Science Foundation of China under Grants 62173175, 61877033 and 62406132, in part by the Natural Science Foundation of Shandong Province under Grants ZR2024MF032 and ZR2024QF039.

**Disclosure of Interests.** The authors declare that they have no conflicts of interest regarding the publication of this paper.

## References

1. Zhu, Y.R., Li, S.L., Ma, J.Y., Zheng, Y.S.: Bipartite consensus in networks of agents with antagonistic interactions and quantization. *IEEE Trans. Circuits Syst. II Express Briefs* **65**, 2012–2016 (2018)
2. Wen, G., Wang, H., Yu, X., Yu, W.: Bipartite tracking consensus of linear multi-agent systems with a dynamic leader. *IEEE Trans. Circuits Syst* **65**, 1204–1208 (2018)
3. Wen, G.H., Zheng, W.X.: On constructing multiple lyapunov functions for tracking control of multiple agents with switching topologies. *IEEE Trans. Autom. Control* **64**, 3796–3803 (2019)
4. Zhang, W.B., Ho, D., Tang, Y., Liu, Y.R.: Quasi-consensus of heterogeneous-switched nonlinear multiagent systems. *IEEE Trans. Cybern.* **50**, 3136–3146 (2020)
5. Song, J., Niu, Y.G., Zou, Y.Y.: Finite-time stabilization via sliding mode control. *IEEE Trans. Autom. Control* **62**, 1478–1483 (2017)
6. Wang, L.M., Dong, T.D., Ge, M.F.: Finite-time synchronization of memristor chaotic systems and its application in image encryption. *Appl. Math. Comput.* **347**, 293–305 (2019)
7. Xu, C., Yang, X., Lu, J., Feng, J., Alsaadi, F.E., Hayat, T.: Finite-time synchronization of networks via quantized intermittent pinning control. *IEEE Trans. Cybern.* **48**(10), 3021–3027 (2017)
8. Shirkavand, M., di Pourgholi, M.: Robust fixed-time synchronization of fractional order chaotic using free chattering nonsingular adaptive fractional sliding mode controller design. *Chaos Solitons Fractals* **113**, 135–147 (2018)
9. Ni, J., Liu, L., Liu, C., Hu, X.: Fractional order fixed-time nonsingular terminal sliding mode synchronization and control of fractional order chaotic systems. *Nonlinear Dyn.* **89**(3), 2065–2083 (2017). <https://doi.org/10.1007/s11071-017-3570-6>
10. Anguiano-Gijón, C.A., Muñoz-Vázquez, A.J., Sánchez-Torres, J.D., Romero-Galván, G., Martínez-Reyes, F.: On predefined-time synchronisation of chaotic systems. *Chaos Solitons Fractals* **122**, 172–178 (2019)
11. Assali, E.: Predefined-time synchronization of chaotic systems with different dimensions and applications. *Chaos Solitons Fractals* **147**, 11 (2021)
12. Li, Q.P., Yue, C.: Predefined-time polynomial-function-based synchronization of chaotic systems via a novel sliding mode control. *IEEE Access* **8**, 162149–62 (2020)
13. Yang, P., Wang, X.: Predefined-time consensus tracking of uncertain leader-following multiagent systems with high-order dynamics via event-triggered control. In: *China Automation Congress (CAC) 2023*, pp. 7342–7347 (2023)
14. Ma, X., Tan, Y., Mei, H.: Predefined-time consensus of nonlinear multi-agent input delay/dynamic event-triggered under switching topology. *IEEE Access* **11**, 29883–29895 (2023)
15. Ni, J., Liu, L., Tang, Y., Liu, C.: Predefined-time consensus tracking of second-order multiagent systems. *IEEE Trans. Syst. Man Cybern. Syst.* **51**, 2550–2560 (2021)
16. Zhang, L., Zhang, T., Jiang, X., Sun, Y., Zhang, W.: Predefined-time consensus of stochastic nonlinear multi-agent systems. In: *2024 14th Asian Control Conference (ASCC)*, pp. 492–497 (di)

17. Fu, J.J., Wen, G.H., Yu, W.W., Huang, T.W., Cao, J.D.: Exponential consensus of multiagent systems with lipschitz nonlinearities using sampled-data information. *IEEE Trans. Circ. Syst.* **I(65)**, 4363–4375 (2018)
18. Jiang, F., Wang, L.: Finite-time weighted average consensus with respect to a monotonic function and its application. *Syst. Control Lett.* **60**, 718–725 (2011)
19. Zou, W.C., Qian, K.W., Xiang, Z.G.: Fixed-time consensus for a class of heterogeneous nonlinear multiagent systems. *IEEE Trans. Circuits Syst. II Express Briefs* **67**, 1279–1283 (2020)
20. Zhang, M.J., Zang, H.Y., Bai, L.Y.: A new predefined-time sliding mode control scheme for synchronizing chaotic systems. *Chaos Solitons Fractals* **164**, 112745 (2022)



# Research on Automatic Throttle Speed Stability System Based on Magic Carpet Landing Technology

Letian Zhao, Hang Chen<sup>(✉)</sup>, Yangyang Zhou, Xiaolei Ma, and Yidi Lei

The First Aircraft Institute of AVIC, Beijing 710089, China  
18291033099@163.com

**Abstract.** Carrier based aircraft are in a low dynamic pressure state during landing, with static and unstable anti zone characteristics. In order to improve their flight quality, a power compensation system needs to be added. This article analyzes the three commonly used types of automatic throttle in power compensation systems and obtains the automatic throttle for magic carpet landing. Using a non-linear dynamic inverse development method to design the incremental mode of the flight path and comparing with PID control. Simulation verification shows that NDI control has higher efficiency, better tracking effect and the designed automatic throttle can achieve good speed maintenance in two states: correcting back to the reference path at different trajectory angles and changing the landing trajectory.

**Keywords:** Orbital transfer · Auto throttle · Dynamic inversion · Incremental mode of flight path

## 1 Introduction

Carrier based aircraft are in a low dynamic pressure state during landing, exhibiting static instability and anti zone characteristics [1]. The aircraft's maneuverability deteriorates, and relying solely on elevators cannot accurately control the flight trajectory. To ensure the low-speed stability of carrier based aircraft during approach and landing, it is necessary to solve the coordinated control problem of multiple control mechanisms such as throttle and aerodynamic control surfaces. At the same time, during the final stage of landing, rapid response of the trajectory angle is required, and the aircraft needs to track the movement of the hull and achieve it when it is very close to the ship [2]. This special requirement requires the use of a power compensation system to ensure the speed of aircraft trajectory control. The power compensation system is based on the input angle of attack and overload, forming quasi velocity feedback to achieve speed control [3]. In an ideal situation, if the angle of attack can be controlled to a constant zero deviation during the dynamic response process, the control of the aircraft attitude angle can achieve fast and accurate response of the trajectory angle to the attitude angle. However, in practical situations, it is necessary to use a power compensation system to control the speed vector

of the aircraft, so that the speed axis of the aircraft can quickly track the changes in the body axis when the aircraft attitude changes, and the angle of attack of the aircraft can be quickly controlled to zero deviation, achieving rapid response of the aircraft trajectory angle. The control command is applied to the automatic throttle lever system of the engine, which controls the speed of the aircraft by changing the power of the engine. The approach power compensation of carrier based aircraft requires automatic control of the engine, which means that the traditional manual throttle control system needs to be replaced with an automatic throttle control system (ATCS). ATCS controls the throttle according to the corresponding control law, thereby achieving control of the engine thrust [4].

In order to improve the safety of ship landing and the success rate of carrier based aircraft landing, the US military has launched the “MAGIC CARPET” project [5]. The magic carpet technology reduces the number of maneuvers for carrier based aircraft before docking from hundreds to over ten, greatly simplifying the pilot’s control pressure [6]. The “magic carpet” control technology includes two trajectory control strategies: trajectory angular velocity control mode (FPAH mode) and flight trajectory incremental control mode (DP mode) [7]. The core of FPAH mode lies in decoupling the attitude and trajectory of the aircraft. Pilots can quickly correct the angular velocity of the trajectory through longitudinal stick commands, thereby adjusting the aircraft trajectory and achieving precise trajectory tracking. Real time monitoring of trajectory angle deviation in DP mode, where pilots directly correct the aircraft’s trajectory angle through longitudinal instructions to achieve precise landing.

During the descent and landing period, by controlling the aircraft’s attitude, the aircraft can quickly track the angle of the descent trajectory, ensuring that the aircraft slides down according to the prescribed descent trajectory. Due to the need for rapid response of trajectory angle during aircraft landing, especially during the final period of landing, the aircraft needs to track the movement of the hull and achieve it when it is very close to the ship. This special requirement requires the use of a power compensation system to ensure the speed of aircraft trajectory control. The power compensation system forms quasi velocity feedback through the input of angle of attack and overload, achieving speed control. The rapid control of the aircraft speed vector is actually to enable the angle of attack of the aircraft to be quickly controlled, so that the control of the aircraft attitude angle can indirectly achieve rapid control of the trajectory angle. In an ideal situation, if the angle of attack can be controlled to keep the deviation constant at zero during the dynamic response process, then by controlling the aircraft attitude angle, the trajectory angle can achieve fast and accurate response to the attitude angle. In practical situations, in order to quickly control the angle of attack of the aircraft, a power compensation system is used to control the speed vector of the aircraft, so that the speed axis of the aircraft can quickly track the changes in the body axis after the aircraft attitude changes. In this way, the angle of attack of the aircraft is also quickly controlled to zero deviation, achieving a fast and accurate response of the aircraft attitude angle to the trajectory angle. The control command is applied to the automatic throttle lever system of the engine, which controls the speed of the aircraft by changing the power of the engine.

Based on the above analysis, this article conducts research on the design technology of an automatic throttle speed stability system that adapts to the needs of track changing maneuvers. By analyzing and clarifying the type of automatic throttle for magic carpet landing, and based on nonlinear dynamic inverse design of DP modal control law, the speed stability of the designed control law is verified through simulation.

## 2 Automatic Throttle Control System

### 2.1 Power Compensation

Approach power compensation refers to the control of engine thrust while the control surface deflects, so that the velocity  $\Delta V = 0$ . At this time, the elevator deflection can change the flight attitude, and the carrier based aircraft lift increment change rate is used to make the trajectory angle consistent with the attitude angle, thereby eliminating the long-period trajectory response deviation caused by gravity affecting the rudder surface change. If the pitch attitude is not maintained, the single throttle system cannot achieve the expected control purpose. This indicates that the control of attitude and trajectory of carrier based aircraft requires the joint action of elevators and throttle levers.

In the 1950s, when pilots of carrier based aircraft entered the aircraft, they simultaneously controlled the throttle and elevator to ensure the low-speed performance and trajectory tracking performance of the aircraft, greatly increasing the pilot's control load. The proposal of a power compensation system effectively solves this problem, typically by designing an effective automatic throttle control scheme that can choose angle of attack control, speed control, trajectory control, etc.

### 2.2 Automatic Throttle Control Scheme

#### (1) Automatic throttle scheme for angle of attack maintenance

The angle of attack control scheme is to use the angle of attack variation  $\Delta\alpha$  and its integral to adjust the engine thrust  $\Delta T$ , so that the aircraft can always maintain the reference angle of attack  $\alpha_0$  during landing. Due to  $\Delta\gamma = \Delta\theta - \Delta\alpha$ , the trajectory angle  $\Delta\gamma$  has a good tracking effect on pitch angle  $\Delta\theta$ , enhancing the glide trajectory response and ensuring the controllability of the angle of attack. This compensation system that maintains the reference angle of attack  $\alpha_0$  ( $\Delta\alpha = 0$ ) is equivalent to a trajectory response enhancer.

#### (2) Automatic throttle scheme for speed maintenance

When  $\Delta\theta$  changes during landing, the velocity change caused by gravity cannot be compensated for, making it difficult for  $\Delta\gamma$  to track the changes in  $\Delta\theta$ , and even the steady-state  $\Delta\gamma$  may have negative values. The speed maintenance control scheme is to suppress the speed changes caused by attitude changes by increasing the reciprocal of the stable speed of the carrier based aircraft. That is, the speed change signal is sent to the throttle, causing the thrust  $T$  to change and  $\Delta V = 0$  to improve the damping of the carrier based aircraft's long-term motion.

#### (3) Automatic throttle scheme for trajectory control

The automatic throttle of trajectory control is a control strategy for Backside manipulation in conventional landing control. Pilots control the inclination angle of the



flight path by controlling the throttle thrust on the one hand, and control the flight attitude on the other hand to ensure that the angle of attack remains unchanged and accelerate the dynamic response [8].

Considering the delayed throttle response of the engine and the slow dynamic characteristics of trajectory response to thrust, the automatic throttle maneuverability of trajectory control is not strong. Pilots need to manipulate the control stick to quickly correct the angle of attack changes caused by thrust control while using thrust to control the aircraft trajectory, in order to maintain the angle of attack constant and quickly correct the trajectory.

### **2.3 The Selection of Automatic Throttle for Magic Carpet Landing**

The conventional landing control mainly includes Frontside control and Backside control [9]: Frontside control completes the glide by controlling the pitch angle through the longitudinal rod, and uses a power compensation system that maintains the angle of attack to restore controllability under low dynamic pressure. The angle of attack signal is introduced into the transmitter to assist the longitudinal rod in controlling the attitude; Backside operates a trajectory control based power compensation system, controlling the downward trajectory through throttle and restoring controllability under low dynamic pressure by controlling the angle of attack with a longitudinal rod. The magic carpet landing control adopts automatic throttle with speed maintenance, mainly for the following reasons:

- (1) Automatic throttle that maintains speed can decouple normal and tangential control, reducing the burden on pilots;
- (2) When the magic carpet lands and slides down for correction, the gravitational potential energy is converted into kinetic energy. The speed maintenance scheme can better maintain the speed, and the comprehensive direct lift control can effectively maintain the speed;
- (3) The elevator in the magic carpet landing control can effectively control the angle of attack without the need for an engine to maintain the angle of attack [10];
- (4) The magic carpet landing control directly controls the trajectory by controlling the lift generated by the direct lift control surface, making the trajectory angle a fast variable. It has a faster response and better effect than the engine controlled glide trajectory in Backside control, so there is no need to use an automatic throttle scheme based on trajectory control.

## **3 Design of Automatic Throttle Speed Stability Control System Based on Dynamic Inverse**

In the final stage of landing, the DP mode is generally used for landing control, and according to the above analysis, the automatic throttle scheme of speed maintenance is preferred during the magic carpet landing control. In this paper, the nonlinear dynamic inverse control theory is used to design a speed stability control system based on DP mode.

### 3.1 Angle of Attack Holding Circuit

Simple direct lift requires the angle of attack to remain constant, the pitch angle to change the same as the trajectory angle, and trajectory correction does not rely on angle of attack changes. The main function of the elevator is to control the angle of attack and maintain its stability, ensuring the achievement of simple direct lift. The angle of attack holding loop refers to the dynamic inverse control law of the dynamic inverse airflow angle loop, and the differential equation  $\alpha, \beta, \mu$  is written in affine nonlinear form:

$$\begin{bmatrix} \dot{\alpha} \\ \dot{\beta} \\ \dot{\mu} \end{bmatrix} = \begin{bmatrix} f_{\alpha}(x_1) \\ f_{\beta}(x_1) \\ f_{\mu}(x_1) \end{bmatrix} + g_{s1}(x_1) \begin{bmatrix} p \\ q \\ r \end{bmatrix} + g_{s2}(x_1) \begin{bmatrix} \delta_{el} \\ \delta_{er} \\ \delta_{al} \\ \delta_{ar} \\ \delta_r \\ \delta_f \end{bmatrix} \quad (1)$$

Among them:  $(\alpha, \beta, \mu)^T$  is the attitude vector in the airflow system,  $(p, q, r)^T$  is the angular velocity vector under the machine system, and  $\delta_x$  is the control surface input.  $f_{\alpha}(x_1), f_{\beta}(x_1), f_{\mu}(x_1)$  represent the combined external forces acting on the aircraft after removing the forces generated by the deflection of the control surface and the rotational angular velocity;  $g_{s1}(x_1)$  represents the kinematic relationship between  $[p, q, r]^T$  and  $[\dot{\alpha}, \dot{\beta}, \dot{\mu}]^T$ ;  $g_{s2}(x_1)$  represents the control force generated by the rudder surface.

The deflection of the aircraft control surface can generate corresponding control force and control torque. For general aircraft, the control surface is mainly used to generate large torque and relatively small direct force. The influence of the deflection of the control surface on fast state variables is much greater than that on slow state variables. At the same time, using the angular velocity and control surface as the airflow angle slow loop will produce the minimum phase zero point, leading to system instability. Under the support of singular perturbation theory, ignoring the influence of rudder deflection on slow loops, there are:

$$\begin{bmatrix} \dot{\alpha} \\ \dot{\beta} \\ \dot{\mu} \end{bmatrix} \approx \begin{bmatrix} f_{\alpha}(x_1) \\ f_{\beta}(x_1) \\ f_{\mu}(x_1) \end{bmatrix} + g_{s1}(x_1) \begin{bmatrix} p \\ q \\ r \end{bmatrix} \quad (2)$$

Among them,  $\omega_{\alpha} = \omega_{\beta} = \omega_{\mu} = 2.3 \text{ rad/s}$ .

The control law for the angle of attack holding loop is obtained by applying nonlinear dynamic inversion as follows:

$$\begin{bmatrix} p_c \\ q_c \\ r_c \end{bmatrix} = g_{s1}^{-1}(x_1) \left( \begin{bmatrix} \dot{\alpha}_d \\ \dot{\beta}_d \\ \dot{\mu}_d \end{bmatrix} - \begin{bmatrix} f_{\alpha}(x_1) \\ f_{\beta}(x_1) \\ f_{\mu}(x_1) \end{bmatrix} \right) \quad (3)$$

$\dot{\alpha}_d, \dot{\beta}_d$  and  $\dot{\mu}_d$  represent the expected slow state variable dynamic response.

The control law for the angular rate control loop obtained by dynamic inversion is as follows:

$$\begin{bmatrix} \delta_e \\ \delta_a \\ \delta_r \\ \delta_f \end{bmatrix} = g^{-1}(x) \left( \begin{bmatrix} \dot{p}_c \\ \dot{q}_c \\ \dot{r}_c \end{bmatrix} - \begin{bmatrix} f_p(x) \\ f_q(x) \\ f_r(x) \end{bmatrix} \right) \quad (4)$$

Among them:

$$\begin{bmatrix} \dot{p}_d \\ \dot{q}_d \\ \dot{r}_d \end{bmatrix} = \begin{bmatrix} \omega_p & 0 & 0 \\ 0 & \omega_q & 0 \\ 0 & 0 & \omega_r \end{bmatrix} \begin{bmatrix} p_c - p \\ q_c - q \\ r_c - r \end{bmatrix} \quad (5)$$

$\omega_p = \omega_q = \omega_r = 10 \text{ rad/s}$ ,  $f_p(x)$ ,  $f_q(x)$ ,  $f_r(x)$  is the nonlinear control torque,  $g(x)$  is the control surface control torque, which are:

$$\begin{aligned} f_p(x) &= \frac{I_z \hat{l} + I_{xz} \hat{n} + pq(I_x - I_y + I_z)I_{xz} + rq(I_y I_z - I_z^2 - I_{xz}^2)}{I_x I_z - I_{xz}^2}, \\ f_q(x) &= \frac{\hat{m} + (r^2 - p^2)I_{xz} + pr(I_z - I_x)}{I_y}, \\ f_r(x) &= \frac{I_{xz} \hat{l} + I_x \hat{n} + pq(I_x^2 - I_x I_y + I_{xz}^2) - rq(I_x - I_y + I_z)I_{xz}}{I_x I_z - I_{xz}^2} \end{aligned}$$

Among them,  $\hat{l}$ ,  $\hat{m}$ ,  $\hat{n}$  respectively represent roll, pitch, and yaw moments other than the aircraft control surface torque.

$$g(x) = \begin{bmatrix} g_{p\delta_e}(x) & g_{p\delta_a}(x) & g_{p\delta_r}(x) & g_{p\delta_f}(x) \\ g_{q\delta_e}(x) & 0 & 0 & g_{q\delta_f}(x) \\ g_{r\delta_e}(x) & g_{r\delta_a}(x) & g_{r\delta_r}(x) & g_{r\delta_f}(x) \end{bmatrix}$$

Among them,  $g_{pi}(x) = \frac{I_x C_{li} + I_{xz} C_{ni}}{I_x I_z - I_{xz}^2} \bar{q} S b$ ,  $i = \delta_e, \delta_a, \delta_r, \delta_f$ ,

$g_{qi}(x) = \frac{C_{mi}}{I_y} \bar{q} S c$ ,  $i = \delta_e, \delta_f$ ,  $g_{ri}(x) = \frac{I_x C_{mi} + I_{xz} C_{li}}{I_x I_z - I_{xz}^2} \bar{q} S b$ ,  $i = \delta_e, \delta_a, \delta_r, \delta_f$ .

### 3.2 Track Angle Control Circuit

The direct lift directly acts on the trajectory control loop. Unlike conventional NDI designs, this paper designs a two-layer parallel dynamic inverse control loop to decouple the trajectory direct lift control from the angle of attack holding loop. For the trajectory angle direct lift control loop, first construct the trajectory loop direct lift NDI affine nonlinear system equation:

$$\dot{X}_F = F_F(X_F) + G_F(X_F)u \quad (6)$$

In the formula:  $F_F$  represents the dynamic relationship of the trajectory control state itself;  $G_F$  represents the control effectiveness relationship of the direct lift control surface;  $u$  represents the control input control surface  $\delta_f$ . The differential equation of speed and trajectory angle is as follows:

$$\dot{V} = \frac{1}{m}(F_T \cos \alpha \cos \beta - D - mg \sin \gamma) \quad (7)$$

$$\dot{\gamma} = \frac{1}{mV}[F_T(\sin \mu \cos \alpha \sin \beta + \cos \mu \sin \alpha) - Y \sin \mu + L \cos \mu - mg \cos \gamma] \quad (8)$$

For channel  $\gamma$ , an analytical method is used to solve the multivariate nonlinear problem by transforming it into a univariate problem.

$$F_{Tc} = \frac{m\dot{V} + D + mg \sin \gamma}{\cos \alpha_c} \quad (9)$$

According to variables  $\dot{V}$ ,  $\gamma$ ,  $\alpha_c$ ,  $F_{Tc}$  can be obtained analytically.

For channel  $\gamma$ , the affine nonlinear system equation can be written as:

$$\dot{\gamma}_F = F_F(X_F) + G_F(X_F)\delta_f \quad (10)$$

Let  $L_F$  be the lift of the aircraft after removing the direct force control surface ( $\delta_f$ ), that is:

$$F_F(X_F) = \frac{1}{mV}[F_T(\sin \mu \cos \alpha \sin \beta + \cos \mu \sin \alpha) - Y \sin \mu + L_F \cos \mu - mg \cos \gamma]$$

$$G_F(X_F) = \frac{1}{mV}(C_{Ldf} \delta_f \cos \mu)$$

Let the expected closed-loop dynamic system of the airflow angle state subsystem be a first-order link, that is:

$$\begin{bmatrix} \dot{V}_d \\ \dot{\gamma}_d \end{bmatrix} = \begin{bmatrix} \omega_V & 0 \\ 0 & \omega_\gamma \end{bmatrix} \begin{bmatrix} V_c - V \\ \gamma_c - \gamma \end{bmatrix} \quad (11)$$

Among them,  $\omega_V = 1 \text{ rad/s}$ ,  $\omega_\gamma = 2 \text{ rad/s}$ .

To test the control performance of the DP mode, a height control loop is also constructed in the outer loop. The altitude error is calculated by the guidance law to obtain the trajectory angle command signal, which is output to the flight control system. The guiding law is as follows:

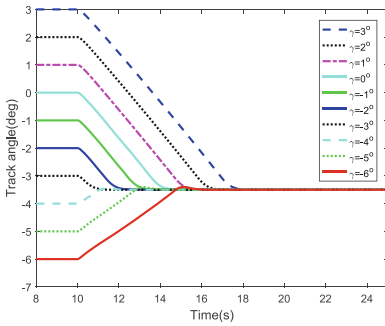
$$\gamma_c = K_p(H_c - H) + \gamma_{ref} \quad (12)$$

Among them,  $K_p = 0.325$ ,  $\gamma_{ref} = -3.5^\circ$ .

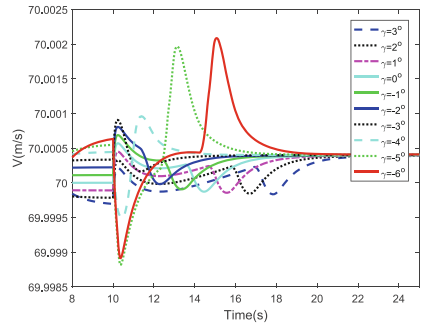
## 4 Simulation Verification

### 4.1 Simulation Analysis of Velocity Stability Under Trajectory Angle Correction

The purpose of this section is to verify the trajectory angle control performance and automatic throttle speed stability performance of the designed incremental flight path mode. During the landing process of the magic carpet, it is necessary to glide down the ship at track angle  $\gamma = -3.5^\circ$ , which means that the designed system needs to be tested for its ability to correct back to  $\gamma = -3.5^\circ$  and the stability of its speed at different track angles. Figure 1, Fig. 2 and Table 1 respectively show the correction of carrier based aircraft from different trajectory angles back to  $-3.5^\circ$ , the stability of speed, as well as the time response and dynamic performance under each correction condition.



**Fig. 1.** Correction situation under different trajectory angle states.



**Fig. 2.** Correction capability under different trajectory angle states.

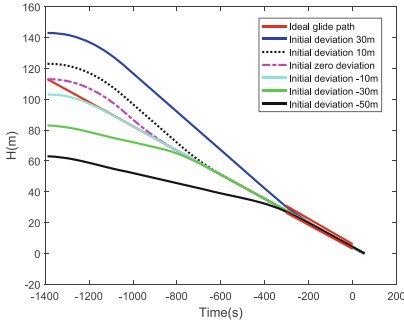
From the simulation results, it is evident that carrier based aircraft can accurately correct to  $\gamma = -3.5^\circ$  at any trajectory angle, with almost no overshoot or steady-state error. The simulation results have verified that the designed control system has good trajectory control and correction capabilities, meeting the control performance requirements for magic carpet landing. The change in trajectory angle will affect the speed, but the maximum change in speed does not exceed  $0.002m/s$ , which can be basically ignored, indicating that the speed is well maintained under this control.

**Table 1.** Correction capability under different trajectory angle states.

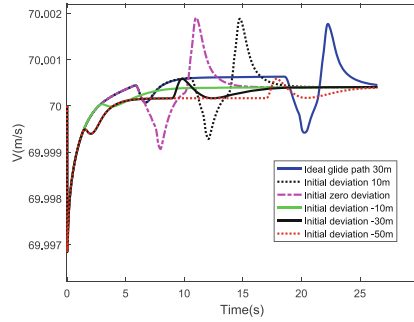
| $\gamma_c(^{\circ})$  | -6    | -5    | -4   | -3   | -2   | -1   | 0   | 1    | 2    | 3   |
|-----------------------|-------|-------|------|------|------|------|-----|------|------|-----|
| $t_s(s)$              | 3.94  | 2.85  | 1.82 | 0.86 | 1.81 | 2.84 | 3.9 | 4.94 | 6.02 | 7.1 |
| $\sigma(100\%)$       | 0.04  | 0.06  | 0.19 | 0    | 0    | 0    | 0   | 0    | 0    | 0   |
| $e$                   | 0     | 0     | 0    | 0    | 0    | 0    | 0   | 0    | 0    | 0   |
| $\max \Delta V (m/s)$ | 0.002 | 0.001 | 0    | 0    | 0    | 0    | 0   | 0    | 0    | 0   |

## 4.2 Simulation Analysis of Velocity Stability During Ship Landing and Orbit Change

In order to verify the trajectory correction capability of the incremental mode landing system based on dynamic inverse control, i.e. the maximum correction capability boundary, different initial approach altitude deviations were set for simulation analysis. The results are shown in Figs. 3, 4, 5 and 6.



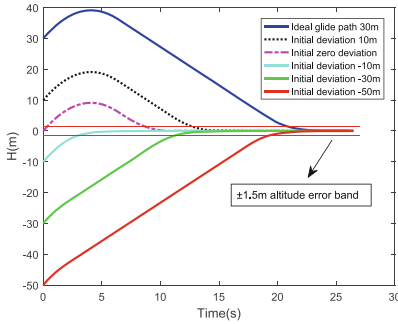
**Fig. 3.** Track correction of initial deviation controlled by NDI.



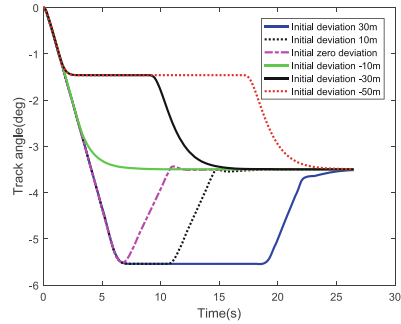
**Fig. 4.** Stability of NDI control speed.

From the chart, it is easy to see that the designed NDI control method has good correction ability when there is a certain initial deviation in the approach of the carrier based aircraft, which can effectively eliminate the deviation, accurately track the ideal glide trajectory, and ultimately glide down to the ship with a constant  $-3.5^\circ$  trajectory angle. At the same time, when the initial deviation reaches 30 m and  $-55$  m, the carrier based aircraft only tracks the ideal glide path when it is almost touching the ship, and descends to the ship at the  $-3.5^\circ$  trajectory angle. At this point, the boundary for safe landing of the carrier based aircraft has been reached. Therefore, when the initial deviation of the carrier based aircraft entering the site is outside the  $-55 \text{ m} \leq \Delta H \leq 30 \text{ m}$  range, it is not allowed for the carrier based aircraft to enter and land on the ship, and a go around should be carried out.

From the stable speed situation in Fig. 4, it can be seen that the speed variation is mainly affected by the trajectory angle and can be well maintained around  $70 \text{ m/s}$ . During the height correction process, the maximum change in speed is shown in Table 2, which is less than  $|\Delta V| \leq 0.004 \text{ m/s}$ . This change can be basically ignored, and the speed can be well maintained.



**Fig. 5.** Tracking error under NDI control initial deviation.

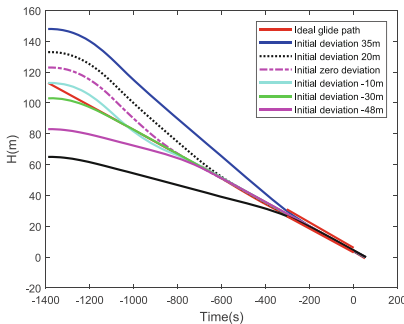


**Fig. 6.** Track angle response of initial deviation controlled by NDI.

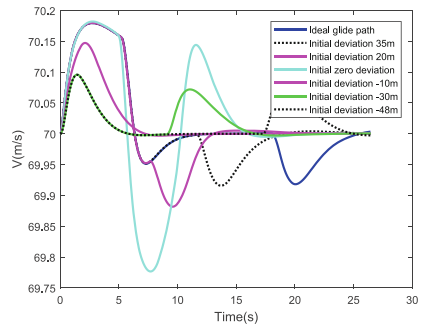
**Table 2.** Speed variation during height correction

| Initial deviation $\Delta H(m)$ | 30    | 15    | 0     | -15   | -30   | -45   | -55   |
|---------------------------------|-------|-------|-------|-------|-------|-------|-------|
| $\max \Delta V (m/s)$           | 0.002 | 0.002 | 0.002 | 0.004 | 0.004 | 0.004 | 0.004 |

In order to verify the trajectory correction capability and maximum correction capability boundary of DP mode based on PID control, simulation verification analysis was conducted by setting different initial approach altitude deviations. The simulation results are shown in Figs. 7, 8, 9 and 10.

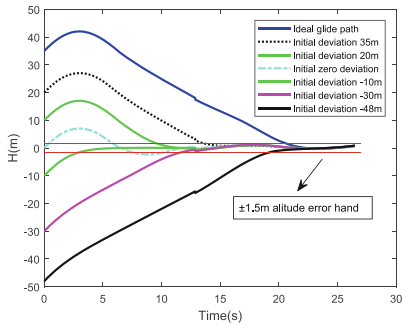


**Fig. 7.** PID control for trajectory correction of initial deviation.

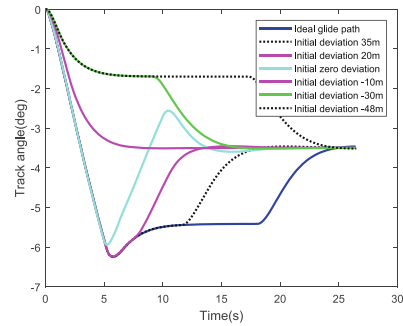


**Fig. 8.** Stability of PID control speed.

From the simulation results in the above figure, it can be seen that the PID controlled trajectory angle control landing system has the ability to correct the trajectory when there is a certain initial deviation during the approach of the carrier based aircraft. It can eliminate the deviation, track the reference glide trajectory, and glide down the ship at a constant trajectory angle of  $-3.5^\circ$ . Meanwhile, as shown in the above figure, when the initial deviation of PID control reaches 25 m and  $-40$  m, the carrier based aircraft almost



**Fig. 9.** Tracking error under initial deviation of PID control



**Fig. 10.** PID control of trajectory angle response with initial deviation

touches the ship before tracking the reference glide path. At this point, it has reached the boundary of safe landing of the carrier based aircraft, which is far less efficient than NDI control.

## 5 Conclusion

This article conducts research on the design technology of an automatic throttle speed stability system that adapts to the needs of orbit maneuvering. The main research conclusions are as follows:

- (1) Analyzed three types of automatic throttle: angle of attack maintenance, speed maintenance, and trajectory control, and selected the automatic throttle scheme for the magic carpet landing;
- (2) Designed a DP mode and speed maintaining automatic throttle based on nonlinear dynamic inverse control method, and conducted simulation verification analysis. Compared with PID control, the results showed that NDI control had higher efficiency and better tracking effect;
- (3) The simulation results show that the designed control law achieves good speed maintenance in two states: correcting the reference trajectory at different trajectory angles and changing the landing orbit.

**Disclosure of Interests.** The authors have no competing interests to declare that are relevant to the content of this article.

## References

1. Ke, T., Wei, W., Yue, M., et al.: Flight control and airwake suppression algorithm for carrier landing based on model predictive control. *Trans. Inst. Meas. Control* **41**(8), 014233121878096 (2018)



2. Zhen, Z., Wang, X., Ju, J., et al.: Research progress in guidance and control of automatic carrier landing of carrier-based aircraft. *Acta Aeronautica et Astronautica Sinica* **38**(2), 127–148 (2017)
3. Luo, F., Zhang, J., Wang, B., et al.: Direct lift trajectory control for carrier aircraft based on NDI. *Flight Dyn.* **39**(01), 40–45+53(2021)
4. Urnes, J.M., Hess, R.K.: Development of the F/A-18A automatic carrier landing system. *J. Guid. Control. Dyn.* **8**(3), 289–295 (2012)
5. Wenhai, W., Jie, W., Li, G., et al.: Analysis on MAGIC CARPET carrier landing technology. *Syst. Eng. Electr.* **40**(9), 2079–2091 (2018)
6. Yulian, Z., Ziyang, Z., Yuxuan, J., et al.: Direct Lift control for auto-landing of shipboard aircraft. *Electron. Opt. Control.* **27**(11), 1–5 (2020)
7. Duan, Z., Wang, W., Geng, J., et al.: Precision trajectory manual control technologies for carrier-based aircraft approaching and landing. *Acta Aeronautica et Astronautica Sinica* **40**(4), 6–20 (2019)
8. Zhu, H., Liu, X., Zhang, Y., et al.: Design of carrier-based aircraft landing control law based on direct force, Singapore, pp. 1153–1163 (2022)
9. Wang Jie, Wu Wenhai, Zhang Yang, et al.: Study on control strategy characteristics and physical mechanism of the aircraft in backside area. *Flight Dynamics* **35**(5), 44–47+52(2017)
10. Zhu, Q., Shi, J., Huang, S., et al.: Carrier-based aircraft landing based on deck motion prediction, Singapore, pp. 265–274 (2023)



# The Decentralized Trusted Identity Management in the Collaboration of Supply Chain and Logistics Operations

Zhangliang Li , Zhengyuan Yue , Zhongli Qiao , Keqing Wang , Ziqi Liu ,  
Jianlong Qiu , and Deqian Fu

Linyi University, Linyi, China  
fudeqian@lyu.edu.cn

**Abstract.** In the era of digital transformation in the logistics industry, the integration of processes such as transportation, warehousing, handling, processing, distribution, and information management has become essential for establishing a seamless supply chain. However, traditional identity verification methods within logistics systems are often centralized, leading to vulnerabilities and hindering the independence, autonomy, and interoperability of collaborative supply chain processes. To address these challenges, we have developed the LDICM framework, a distributed identity verification and authorization system. LDICM enables different roles in supply chain logistics collaboration to maintain control over their data while accessing specific information based on business collaboration requirements. This framework is built on the Hyperledger Fabric blockchain and leverages Decentralized Identifiers (DIDs) for entity identity verification and Verifiable Credentials (VCs) for consent management in information access. Our proposed solution significantly enhances data security, privacy, immutability, interoperability, and user autonomy by incorporating selective disclosure, zero-knowledge proofs, decentralized identifiers, verifiable credentials, and ontology-based access control. The outcomes of this study have practical implications for decentralized identity management and authentication within collaborative supply chain logistics.

**Keywords:** Blockchain · Decentralized Identity · Supply Chain Logistics · Ontology-Based Access Control

## 1 Introduction

Supply chain logistics refers to all logistics activities and processes involved in supply chain management. It encompasses the entire process from the procurement of raw materials, production, warehousing, and transportation to the distribution of final products to end customers [1, 2]. In the supply chain, logistics involves not only the physical movement of goods but also a large amount of information exchange related to inventory, orders, packaging, handling, and cash flow. The supply chain logistics process involves multiple parties and complex business scenarios, often leading to information silos and inefficient information sharing, which negatively impacts the efficiency and quality of supply chain logistics. Ensuring data security and trust among participants is critical in the entire supply chain logistics process.

Currently, most participants in supply chain logistics use traditional identity management systems (IDMS). These systems rely on centralized institutions to act as identity providers, and in some business integration and coordination activities, manual operations are still used. As a result, identity owners are unable to control their own data and cannot prevent the misuse of their identity and privacy [3]. More importantly, achieving interoperability between different services remains a challenge. Blockchain-based decentralized identity, as a form of Self-Sovereign Identity (SSI), grants individuals or organizations full control over their identity information, allowing them to selectively share it with third parties while maintaining control over their privacy. The decentralized identity system is built on the trust triangle of issuers, identity owners (or holders), and verifiers, with Decentralized Identifiers (DIDs) and Verifiable Credentials (VCs) as its two main pillars, together constructing a verifiable digital identity infrastructure on distributed ledgers, which is entirely self-managed by the identity owner [4–6].

Decentralized identity management has become a highly researched and promising field, offering many benefits for supply chain logistics business scenarios. Hongzhi Li [7] proposed a blockchain-based logistics data security storage solution that ensures secure storage and auditing of data through smart contracts and efficient consensus mechanisms, while also analyzing its performance and privacy protection. Humayun M [8] proposed a layered framework called BCTLF, which integrates the Internet of Things (IoT) with blockchain to build intelligent logistics and transportation systems. Ahmad R W [9] explored the potential role of blockchain technology in transforming port logistics operations and services, proposing a framework for applying blockchain to port logistics and adopting a permissioned chain architecture to map port logistics services, demonstrating the deployment of system components, participants, and automated operations. Their studies on decentralized identity management in logistics scenarios were limited to certain aspects and did not consider interoperability between different identity management systems. Ghosh B C [10] proposed a distributed identity management architecture and protocol based on Hyperledger Indy and Aries, aiming to ensure secure and dynamic data sharing and interoperability between permissioned chain networks through self-sovereign decentralized identity (DID) and verifiable membership credentials. They demonstrated its feasibility and security in trade finance and logistics networks. Dang R [11] addressed traceability, transparency, and authenticity challenges in high-value goods supply chains by integrating consortium blockchain technology, Non-Fungible Tokens (NFTs), and decentralized identity (DIDs). Sun N [12] proposed a self-sovereign identity privacy protection scheme designed for logistics transportation, which ensures the secure transmission and privacy protection of user identities and data by combining one-time use tokens and decentralized identity with obfuscated identity encryption techniques.

In this paper, we analyze the various entities involved in supply chain logistics collaboration and propose an innovative solution called Logistics Distributed Identity and Credential Management (LDICM), which aims to address trust issues in supply chain logistics collaboration. This solution leverages distributed technologies to extract and manage the identity information of all entities in the supply chain, providing users with sustainable, reliable, available, and highly secure identity services, thereby effectively enhancing trust and collaboration among all parties. Additionally, the solution introduces

an Ontology-Based Access Control (OBAC) mechanism for strict access control during data exchange, further ensuring data confidentiality and privacy protection.

The rest of this paper is organized as follows. In Sect. 2, we introduce the theoretical foundations and related terminology of this study. Section 3 describes the process of supply chain logistics collaboration. In Sect. 4, we propose an authentication and ontology-based access control solution based on decentralized identifiers. Section 5 presents case testing and performance analysis of the proposed solution. Finally, Sect. 6 concludes the paper.

## 2 Related Technical Concepts

### 2.1 Decentralized Identifier

The core building block of decentralized identity is the Decentralized Identifier (DID). A DID is a unique identifier that enables verifiable, decentralized digital identity by linking individuals or organizations to objects or other entities they control [9]. A DID consists of a short, unique string composed of three parts: the first part is the DID URI scheme identifier, the second part specifies the DID method that identifies the ledger from which the DID originates, and the third part is a unique identifier specific to that method [13]. A DID is designed to resolve into a DID Document. Each DID corresponds to a DID Document, which contains descriptive information about the DID.

### 2.2 Verifiable Credential

A Verifiable Credential (VC) is an open standard established by the W3C, designed to provide secure, privacy-preserving, and machine-verifiable credentials on the internet [14]. VCs can carry the same data as physical credentials, but through cryptographic techniques, such as digital signatures, they are more secure and tamper-resistant than physical credentials. Using the VC model, applications can be built on a strong foundation of digital trust, with its built-in cryptographic mechanisms ensuring the authenticity and integrity of the credentials [15].

### 2.3 Zero-Knowledge Proof and Selective Disclosure

Zero-Knowledge Proof (ZKP) is a cryptographic technique used to verify the validity of personal attributes without revealing the actual values of those attributes. It allows the prover to demonstrate the truth of specific information, enabling the verifier to trust the accuracy of the claims made [16]. The key advantage of ZKP is that it validates the claim without disclosing any personal information, thereby reducing the risks of identity theft and privacy breaches. In decentralized identity verification, ZKP is used to prove that the holder possesses a verifiable credential without providing any identifiable personal data to the verifier. The verifier receives the necessary information, but the holder's unique identity remains concealed during the verification process. An important privacy feature of Verifiable Credentials (VCs) is their support for Selective Disclosure. Selective Disclosure allows users to reveal only specific claims within a VC, without having to

disclose all the attributes contained in the credential. When sharing information, users can choose which data to share and generate a Verifiable Presentation (VP). This approach gives users greater control over how their data is used and shared, enhancing privacy and data management.

## 2.4 Hyperledger Fabric

Hyperledger Fabric [17] is a modular and scalable enterprise-grade blockchain platform designed to support distributed ledger technology for complex business scenarios involving multi-party collaboration. Fabric adopts a permissioned network architecture, where participant identities are verified to ensure system security and privacy. Unlike other blockchain platforms, Fabric offers pluggable components, allowing users to customize the consensus mechanism, identity management, and data privacy protection measures according to their needs. Fabric enables the definition and execution of business logic through chaincode, with transactions being executed on multiple endorsement nodes and achieving consensus before being packaged and sent to the ledger by ordering nodes. Its decentralized architecture records all transactions on a distributed ledger, ensuring immutability and transparency of the transactions.

## 2.5 Ontology-Based Access Control

Ontology-Based Access Control (OBAC) is a method that utilizes ontology technology to define and manage access control policies [18]. Ontology formalizes the concepts and relationships within a domain, providing a flexible and scalable knowledge model for access control. Compared to traditional Role-Based Access Control (RBAC), OBAC is capable of handling more complex contextual information and dynamic rules. In an OBDA system, users query data distributed across different data sources via ontologies, but access to this data is subject to authorization controls. OBAC can be integrated with OBDA to infer, through ontology reasoning, whether users are authorized to access specific query results [19].

# 3 Supply Chain Logistics Business Collaboration

Supply chain logistics encompasses the comprehensive management of the entire process from raw material procurement, product manufacturing, storage, transportation, and distribution, to the final delivery to the consumer. It involves the coordinated flow of goods, information, and finances, ensuring that products are efficiently distributed in the right quantities, at the right time, and to the right location. However, without sufficient trust between the parties involved at various stages of the supply chain, achieving smooth business collaboration can be challenging [20]. By introducing a reliable online collaboration mechanism, trust can be established among the different entities, thereby ensuring the smooth execution of business processes. Based on actual business processes (as shown in Fig. 1), we can conduct a detailed analysis of each stage of supply chain logistics.

The core method of information exchange in supply chain logistics is through the transmission of waybills. Different stages of the logistics process use distinct waybill information. The starting point of the supply chain typically involves interaction between the shipper and the consignee through an ERP (Enterprise Resource Planning) or OMS (Order Management System). Once a collaboration agreement is reached, the shipper prepares to dispatch the goods. During the dispatch phase, the shipper can choose between two main methods to deliver the goods to the distribution center: one option is through urban distribution, where a third-party logistics company handles short-distance transportation, and the other is self-pickup, where the shipper arranges transportation to deliver the goods to the distribution center.

Once the goods arrive at the distribution center, the center is responsible for receiving and confirming the delivery, as well as coordinating with logistics companies to arrange the sorting and dispatching of the goods. Upon receiving the distribution center's information, the logistics company utilizes its transportation resources (such as trucks, freight trains, etc.) to transport the goods to various destinations. In this process, due to differences in the scale and coverage of logistics companies, goods often undergo multiple transfers and redistribution, especially for long-distance or cross-regional transportation. This requires collaboration between different logistics companies to fully leverage their transportation resources, ensuring that goods reach their final destinations efficiently.

In the terminal logistics stage, the end-point logistics company is responsible for delivering the goods to the consignee. At this stage, the logistics company can choose different delivery methods depending on the consignee's preferences and specific circumstances. These may include self-pickup by the consignee, urban delivery, or the logistics company's own delivery services. Regardless of the method, the goal of terminal logistics is to ensure that the goods are delivered to the consignee in a timely and secure manner, thereby closing the loop in the entire supply chain.

Each stage of the supply chain logistics involves different participants and business processes. Through the coordinated flow of information, finances, and goods, the efficient operation of the entire supply chain is ensured. This complex multi-party collaboration mechanism guarantees the smooth progression of goods from production to consumption, while also providing reliable support for companies' production processes and meeting customers' needs.

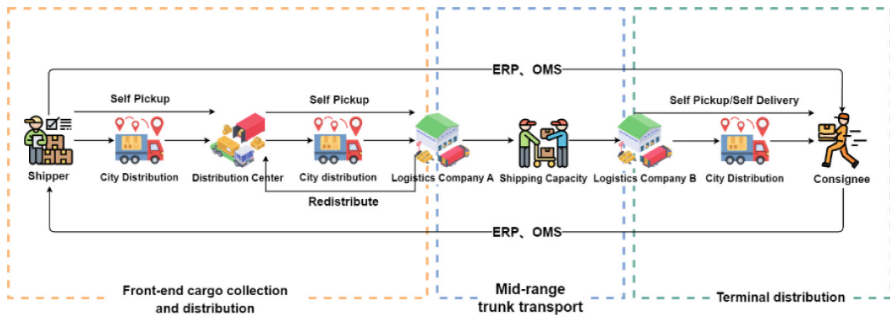


Fig. 1. Supply chain logistics business processes.

## 4 Distributed Trusted Identity Management Solution for Supply Chain Logistics

To achieve trusted identity verification and authorization in the supply chain logistics scenario, we propose a distributed identity verification and authorization framework for supply chain logistics (LDICM). This framework aims to address the complex issue of multi-party participation within the supply chain, ensuring that each participant can perform identity authentication and data access in a trusted and secure environment. Based on the analysis in Fig. 1, we categorize the seven main participants in supply chain logistics as shippers, urban distribution providers, distribution centers, logistics companies, carriers, consignees, and government entities. These participants play different roles in various stages of the supply chain and have different requirements for identity verification and data access. Therefore, the design goal of LDICM is to ensure that these diverse participants can effectively collaborate by using a distributed identity authentication mechanism and ontology-based access control policies to securely complete their respective tasks.

We first define the functionality, architecture, and workflow of the system. The core of LDICM lies in achieving distributed trusted identity verification through the collaborative operation of various components, ensuring that data access rights are strictly controlled during business collaboration. The framework's functionalities include identity verification, data access control, privacy protection, and a trusted environment for data exchange. The architecture is based on blockchain technology, which, through decentralization, ensures trust between participants without relying on a central authority and provides a highly secure data storage and transmission mechanism. The workflow encompasses identity registration, verification, authorization, and data exchange, ensuring that each step, from identity creation to data access, follows established rules and procedures.

LDICM includes several key components that work together to achieve overall functionality. First, the Hyperledger Fabric node (Fabric Node) is the core part of the framework, responsible for managing all distributed identity identifiers (DIDs). In the LDICM framework, Fabric nodes are used not only to store DIDs but also to ensure that each network participant can securely read blockchain content, with credential (VC) verification procedures in place. Through this mechanism, all participants can interact based on verified identity information, ensuring the credibility of identities and the integrity of data.

The ontology and rule base are fundamental for implementing access control. According to the different responsibilities of roles in supply chain logistics and the privacy protection requirements associated with waybills at different stages, we established an ontology to define the relationships and data access rules among the various participants. Additionally, an industry-level rule base imposes strict limitations on the data accessible by each role, ensuring that each participant can only view waybill data relevant to their permissions. For example, shippers can only view the details of their related waybills, while government agencies have broader access to waybill data for regulatory purposes.

The basic connector serves as a key component connecting participants' systems to the LDICM framework. It is a gateway device that can be installed locally, on servers,

or elsewhere. It connects to the identity authentication connector, other users' connectors, and the blockchain. In addition to bridging data and services, the basic connector enforces ontology-based access control based on the requester's VC (credentials) and the predefined rule base. Through this mechanism, the connector ensures that data exchange between participants takes place in a trusted environment, and all data access complies with the security and privacy requirements established by the framework.

The identity authentication connector is an extension of the basic connector, specifically designed to manage identity creation, verification, and revocation. It generates cryptographic keys and issues identity credentials to holders. When participants submit requests, the identity authentication connector can issue related certificates to holders according to the rule base and perform identity revocation when needed. Furthermore, the identity authentication connector has an industry-standard rule base and ontology, ensuring that identity management across the entire supply chain is consistent and reliable. Given the critical importance of identity security, only authorized government agencies can use this component, ensuring that identity management and credential issuance maintain the highest level of trustworthiness.

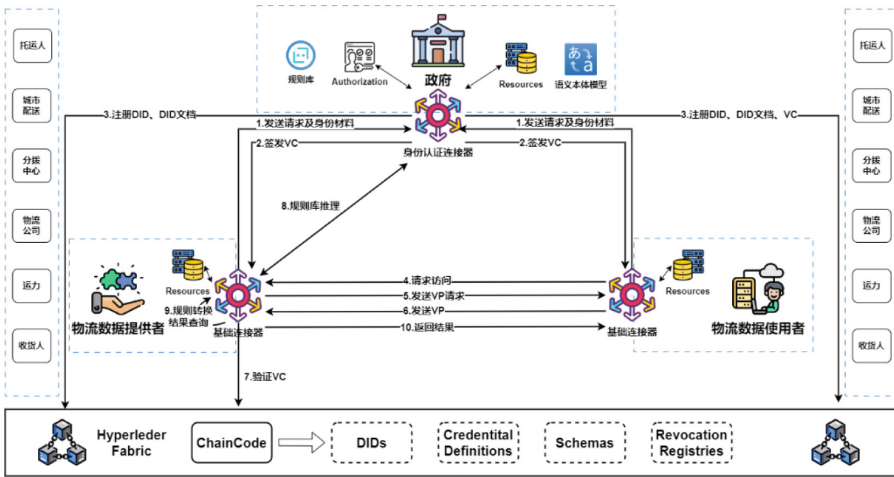


Fig. 2. LDICM Overall Architecture.

## 4.1 Functionality

### 4.1.1 Registration and Authentication

All functional modules need to be developed in chaincode and deployed on the Hyperledger Fabric platform. Participants submit an application to the identity authentication connector through the basic connector, as shown in Formula (1). They provide personal basic information and company qualifications (such as business licenses, ID cards, etc.) and select the type of Verifiable Credential (VC) they wish to apply for (such as Shipper,



City Delivery, Distribution Center, Logistics Company, Carrier, Consignee, or Government). Users can choose from different VC types to determine their role within the supply chain.

$$Request = \{Personal\ Info, Company\ Documents, VC\ Type\} \quad (1)$$

A public-private key pair based on elliptic curve cryptography is employed. The public key information and DID description are added to the DID document and uploaded to the blockchain via chaincode. The system issues the VC according to the type requested by the participant and stores the private key along with the VC in the participant's local basic connector.

#### 4.1.2 VC Generation and Verification

The government authority confirms the type and necessary attributes of the Verifiable Credential (VC) and interacts with the blockchain through a smart contract, storing relevant schema records, signature types, revocation information, and issuer details of the credential. The government authority logs into the system via the identity authentication connector, reviews, and approves the application information. Once approved, the system creates a valid identity within the supply chain logistics collaboration, as shown in Formula 2, that complies with the W3C DID standard using Fabric chaincode. Here, *sba* represents the DID method identifier, and the DID Method-Specific Identifier is an 18-character string.

$$did : sba : \{DID\ Method - Specific\ Identifier\} \quad (2)$$

The system generates a public-private key pair based on the Elliptic Curve Digital Signature Algorithm (ECDSA) for each user. The public key information and DID description are added to the DID document and uploaded to the blockchain via chaincode. The Verifiable Credential (VC) requested by the participant is issued, and the private key is stored along with the VC in the participant's local basic connector. Participants can generate a Verifiable Presentation (VP) within the connector based on the attributes in the VC. The VP contains the necessary attribute information for the current scenario, which can be used by the verifier.

$$DIDDocument = \{PublicKey, DIDDescription, Signature\} \quad (3)$$

$$VC = \{Issuer, Subject, Claims, Proof, Context\} \quad (4)$$

The verification of the Verifiable Credential (VC) is initiated by the verifier, who sends a request to the holder. Data users such as city delivery services, logistics companies, and others submit the VC to the data holder. The data holder verifies the identity of the issuer, the user, and the digital signature of the VC through the identity authentication connector, leveraging the blockchain to ensure the validity of the credential.

### 4.1.3 Ontology Mapping Establishment

Based on the order information exchange and participant roles in supply chain collaboration, a supply chain ontology is constructed and placed within the identity authentication connector. The supply chain ontology includes classes and properties that define various roles (such as Shipper, Logistics Company, etc.) and their associated attributes, represented using the Web Ontology Language (OWL). Data providers establish mappings between their local data and the ontology through the basic connector, constructing a virtual knowledge graph to ensure interoperability across different business systems.

### 4.1.4 Rule Base Reasoning

The rule base defines the unified industry standards mandated by the government and regulates the exchange of supply chain logistics information. Different identities must adhere to the corresponding rules. After verifying the requester's Verifiable Credential (VC), the data provider extracts the relevant attributes and sends them to the identity authentication connector, where rule-based reasoning is performed using Jena. Based on the reasoning outcomes, the core connector utilizes the virtual knowledge graph to extract the corresponding data and returns the results to the requester.

## 4.2 Workflow

This paper discusses the workflow among the shipper, City Distribution, and government. Figure 2 shows the overall architecture, and the workflow is the same for the other participants.

### 4.2.1 Decentralized Identity Authentication

Step 1. Preparatory Work and Issuing Invitation:

- The government, as the issuer, creates the DID, DID document, and VC format, and deploys them on the server through a smart contract, generating its own certificate. The basic connector and identity authentication connector are also established.
- The government issues an invitation to the public through a website or a QR code.
- Participants in the supply chain logistics can accept or reject the invitation from the government.
- Step 2. Government Issues VC to Shipper:
  - After accepting the invitation, the shipper downloads the basic connector, registers, and submits personal information to the government. Upon first login, the user selects a role; currently, the shipper selects the "Shipper" role. After approval, additional roles can be applied for by submitting information through the basic connector to obtain other VCs.
  - The government views the registration request through the identity authentication connector, reviews the submitted information, and, upon approval, generates a pair of elliptic curve cryptography (ECC) keys.
  - A DID and DID document are generated, with the public key, government signature, and other details written into the DID document, which is then uploaded to the blockchain.

- The requested VC is generated, containing the user’s attributes and personal information. The VC and user’s private key are returned to the applicant’s basic connector (Fig. 3).

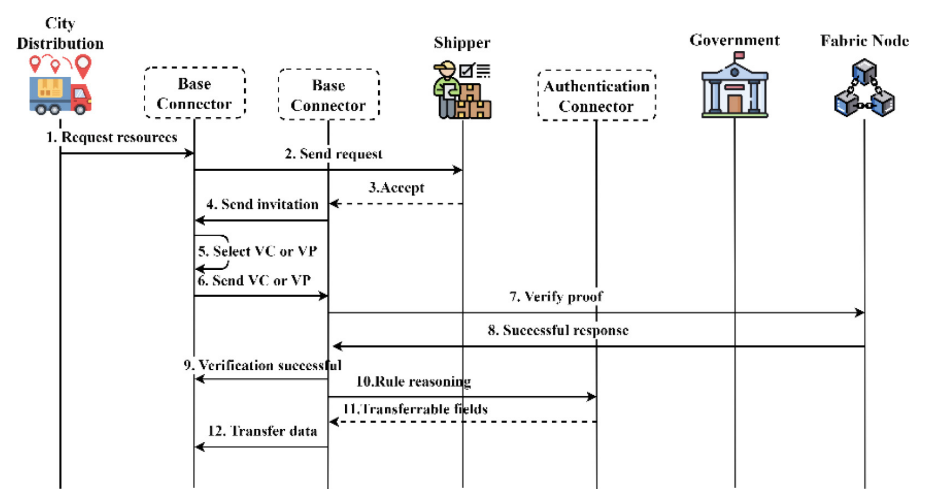


Fig. 3. DID Generation and VC Issuance Process.

Step 3. City Distribution Requests Data from Shipper:

- City Distribution requests access to a set of resources from the shipper.
- The shipper creates an invitation, specifying the requirements for identity verification.
- City Distribution selects its own VC or generates a verifiable presentation (VP) by disclosing partial attributes from the VC, then sends it to the shipper.
- The shipper connects to the smart contract through the basic connector, searches the blockchain for the government’s DID document, and uses the public key in it to verify the issuer’s authenticity. The shipper also verifies the authenticity of City Distribution’s identity using the public key from City Distribution’s DID document.

Tables are “float elements” which should be inserted after their first text reference and have specific styles for identification. Do not use images to present tables, or they will be inaccessible to readers using assistive technologies (Fig. 4).

4.2.2 Ontology-Based Access Control.

Step 4. Reasoning Based on Ontology Rules:

- After verifying the VC, the shipper sends the attributes contained in the requester’s VC to the identity authentication connector.
- The identity authentication connector performs reasoning based on the rules defined for the entire supply chain logistics industry, determining which data the requester is allowed to access, and returns the result to the shipper.

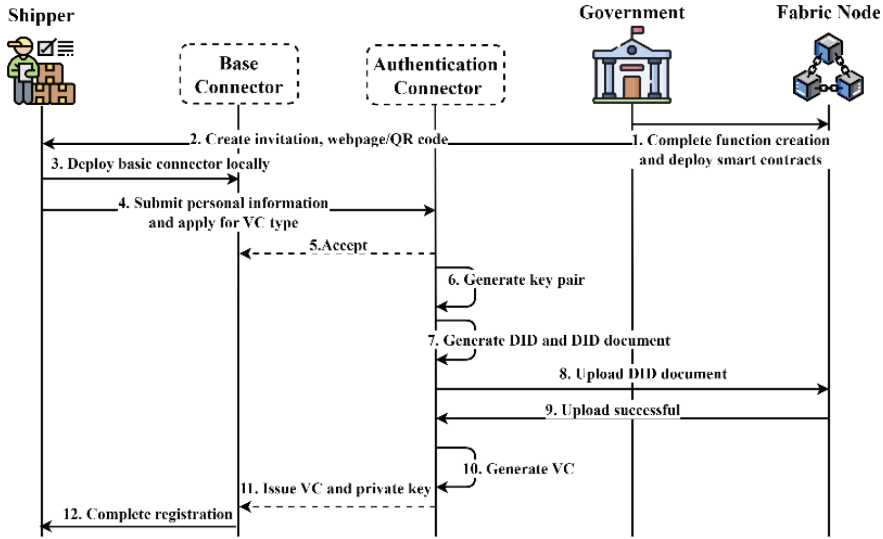


Fig. 4. DID Generation and VC Issuance Process.

- The shipper retrieves the specified fields of data via Ontology-Based Data Access (OBDA) technology through the basic connector, and transmits it to the City Distribution.

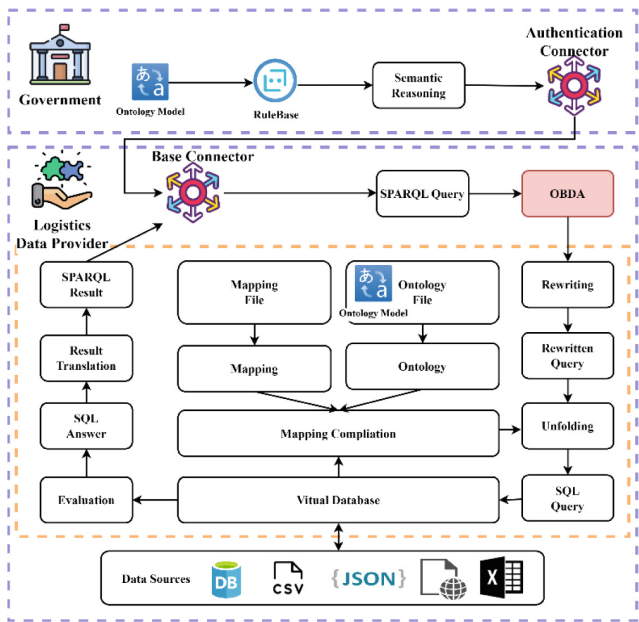


Fig. 5. OBDA Process.

The operational principle of OBDA is illustrated in Fig. 5. Firstly, a mapping file and an ontology file are required, which are the two key elements in this process. The mapping file maps the tables and fields from the underlying database to the concepts and properties of the ontology. The ontology file defines the concepts and relationships in the knowledge graph, which form the semantic-level data model. The ontology file used here is the same as the one in the identity authentication connector. The mapping compilation process combines the mapping file with the ontology file to ensure the correspondence between queries and data sources. It defines how information related to the ontology is retrieved from the database.

Based on the reasoning results from the rule base, the basic connector generates SPARQL queries and sends a request to the virtual knowledge graph endpoint to retrieve certain information. OBDA rewrites the user's SPARQL query. By analyzing the query, the system converts it into a structured query that can be mapped to the relational database. OBDA provides a virtual database through which users can query and access information from the underlying data sources without copying the data. The virtual database is linked to the data sources through predefined mapping files. After mapping and compilation, the SQL query is sent to the underlying data source, and the database returns the corresponding results. These results are provided in SQL format. The results of the SQL query are then retranslated into RDF triples as required by the SPARQL query during a transformation process. Finally, the transformed RDF data is returned to the user as a response, allowing them to view the results relevant to their query.

## 5 Results and Analysis

### 5.1 Overview of Experimental Environment

The experimental environment utilized four virtual machines, each configured with different hardware resources and purposes (see Table 1 for specific configurations and uses). To achieve efficient service management and deployment, all services were encapsulated in Docker containers and run on the virtual machine servers. This container-based deployment approach not only improves the portability of the system but also facilitates service scalability and maintenance.

### 5.2 Experimental Results

Based on the constructed test environment, we provide a detailed description of the experimental results of LDICM and analyze its practical application. In the experiment, the supplier and City Distribution roles connect to the identity authentication connector via the basic connector for identity registration. Figure 6 shows the decentralized identifier (DID) and its associated DID document generated for the supplier after approval by the government authority, and Fig. 7 shows the verifiable credential (VC) issued to the supplier. The registration process for the City Distribution role is similar to that of the supplier.

During the data request phase, the City Distribution role initiates a data request to the supplier and sends its VC to the supplier. The supplier verifies the authenticity



```
[rule1:(?x :role ?a)(?a rdf:type :CityDistribution)(?y rdf:type :Delivery)->( ?x :allow ?y)]
[rule1:(?x :role ?a)(?a rdf:type :CityDistribution)(?y rdf:type :Pickup)->( ?x :allow ?y)]
[rule1:(?x :role ?a)(?a rdf:type :CityDistribution)(?y rdf:type :Name)->( ?x :allow ?y)]
[rule1:(?x :role ?a)(?a rdf:type :CityDistribution)(?y rdf:type :CarrierInformation)->( ?x :allow ?y)]
[rule1:(?x :role ?a)(?a rdf:type :CityDistribution)(?y rdf:type :Weight)->( ?x :allow ?y)]
[rule1:(?x :role ?a)(?a rdf:type :CityDistribution)(?y rdf:type :Volume)->( ?x :allow ?y)]
[rule1:(?x :role ?a)(?a rdf:type :CityDistribution)(?y rdf:type :Short-haulFreightFee)->( ?x :allow ?y)]
```

Fig. 8. Partial Rule File.

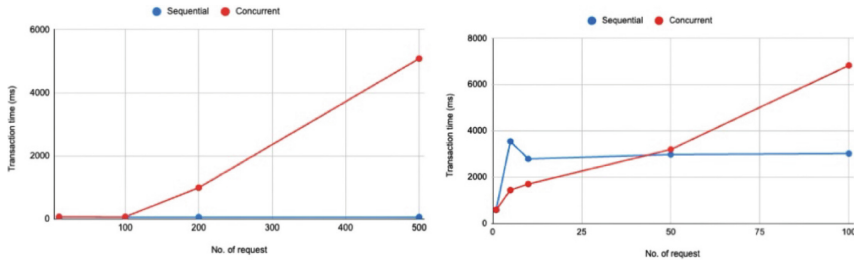
Table 2. Response data for query based on waybill “233089423”.

| Description            | Value                  |
|------------------------|------------------------|
| Waybi Number           | 233089423              |
| Delivery               | 2024-10-01 15:30       |
| Pickup                 | 2024-09-29 09:00       |
| Name                   | Eletronic Components   |
| Carrier Information    | XYZ Logis-tics Co.,Ltd |
| Weight                 | 500 kg                 |
| Volume                 | 2.5 m <sup>3</sup>     |
| Short-haul Freight Fee | \$150                  |

5.3 Performance Analysis

We conducted a comprehensive performance evaluation by decomposing the model into various modules and testing their individual functionalities. The tests covered several key components, ranging from the generation and uploading of DID documents to the creation and verification of VC/CP. Specifically, for cryptographic operations, we utilized the elliptic curve encryption algorithm to assess its performance within the model. To evaluate the model’s performance in terms of semantic reasoning and knowledge graph querying, we analyzed the query rewriting mechanism in the virtual knowledge graph and transformed SPARQL queries into SQL for processing. Additionally, we conducted thorough testing on data access, data processing, and overall response time.

The test results indicate that, under typical conditions, the model maintains solid performance and response speed. This is particularly evident in single-task scenarios and situations with lower concurrency, where all modules exhibit excellent response times. However, when the number of concurrent requests surges, the system’s transaction processing time increases linearly with the growing request volume. Despite this, the system demonstrates a high degree of scalability, capable of handling a large number of requests while maintaining efficiency within a certain range. Figure 9 clearly illustrate the variation in transaction processing times under different request loads, further proving the model’s potential to handle high concurrency scenarios.



**Fig. 9.** Average Transaction Time for Connection Invitations and VC Issuance.

## 6 Results and Analysis

This paper introduces LDICM, a novel solution designed to address the challenges of distributed identity authentication and access control within collaborative supply chain logistics. LDICM integrates W3C standards for verifiable credentials (VCs) and decentralized identifiers (DIDs) to establish a decentralized trust mechanism among identity providers, users, and verifiers. By reducing reliance on centralized authorities, LDICM enhances system robustness and resistance to attacks. Leveraging the immutability, decentralization, and transparency of the Fabric blockchain, LDICM ensures the security, transparency, and trustworthiness of the identity management process. Additionally, LDICM employs an ontology-based access control model to enhance security and privacy protection during data exchange through semantic rule definition and reasoning. This mechanism meets the permission needs of different roles and ensures data confidentiality and controllability in a dynamic collaborative environment.

The paper provides a comprehensive overview of the framework architecture, technical implementation details, and performance analysis of LDICM, showcasing its significant role in enhancing supply chain logistics collaboration. Experimental results demonstrate that LDICM efficiently handles identity authentication and access control, while also offering strong security and privacy protection capabilities in a distributed environment. The application of LDICM underscores the potential of blockchain technology and decentralized identity management in complex business scenarios, providing valuable insights for future research and practice.

## References

1. Zenden, R.: The identity mangement solution that improves data sharing in logistics. University of Twente (2023)
2. Bowersox, D.J., Closs, D.J., Cooper, M.B., et al.: Supply Chain Logistics Management. Mcgraw-hill, Boston (2020)
3. Javed, I.T., Alharbi, F., Bellaj, B., et al.: Health-ID: A blockchain-based decentralized identity management for remote healthcare. *Healthcare MDPI* **9**(6), 712 (2021)
4. Naik, N., Jenkins, P.: Sovrin network for decentralized digital identity: analysing a self-sovereign identity system based on distributed ledger technology. In: 2021 IEEE International Symposium on Systems Engineering (ISSE), pp. 1–7. IEEE (2021)



5. Sun, G., Dai, M., Zhang, F., et al.: Blockchain-enhanced high-confidence energy sharing in internet of electric vehicles. *IEEE Internet Things J.* **7**(9), 7868–7882 (2020)
6. Son, S., Lee, J., Park, Y., et al.: Design of blockchain-based lightweight V2I handover authentication protocol for VANET. *IEEE Trans. Netw. Sci. Eng.* **9**(3), 1346–1358 (2022)
7. Li, H., Han, D., Tang, M.: A privacy-preserving storage scheme for logistics data with assistance of blockchain. *IEEE Internet Things J.* **9**(6), 4704–4720 (2021)
8. Humayun, M., Jhanjhi, N.Z., Hamid, B., et al.: Emerging smart logistics and transportation using IoT and blockchain. *IEEE Internet Things Maga.* **3**(2), 58–62 (2020)
9. Ahmad, R.W., Hasan, H., Jayaraman, R., et al.: Blockchain applications and architectures for port operations and logistics management. *Res. Transp. Bus. Manag.* **41**, 100620 (2021)
10. Ghosh, B.C., Ramakrishna, V., Govindarajan, C., et al.: Decentralized cross-network identity management for blockchain interoperation. In: 2021 IEEE International Conference on Blockchain and Cryptocurrency (ICBC), pp. 1–9. IEEE (2021)
11. Dang, R., Li, H., Yang, A., et al.: Decentralized identity system for high-value commodity supply chain based on consortium blockchain. In: Proceedings of the 2023 4th International Conference on Big Data Economy and Information Management, pp. 337–343 (2023)
12. Sun, N., Zhu, C., Liu, Y.: A self-sovereign identity privacy-preserving scheme for logistics transportation based on one-time-use tokens. *Electronics* **13**(14), 2799 (2024)
13. W3C. Decentralized identifiers (dids) v1.0 (2022). <https://www.w3.org/TR/did-core/>
14. Brunner, C., Gellersdörfer, U., Knirsch, F., et al.: Did and vc: untangling decentralized identifiers and verifiable credentials for the web of trust. In: Proceedings of the 2020 3rd International Conference on Blockchain Technology and Applications, pp. 61–66 (2020)
15. Sedlmeir, J., Smethurst, R., Rieger, A., et al.: Digital identities and verifiable credentials. *Bus. Inf. Syst. Eng.* **63**(5), 603–613 (2021)
16. Yang, X., Li, W.: A zero-knowledge-proof-based digital identity management scheme in blockchain. *Comput. Secur.* **99**, 102050 (2020)
17. Androulaki, E., Barger, A., Bortnikov, V., et al.: Hyperledger fabric: a distributed operating system for permissioned blockchains. In: Proceedings of the Thirteenth EuroSys Conference, pp. 1–15 (2018)
18. Choi, C., Choi, J., Kim, P.: Ontology-based access control model for security policy reasoning in cloud computing. *J. Supercomput.* **67**, 711–722 (2014)
19. Xiao, G., Calvanese, D., Kontchakov, R., et al.: Ontology-based data access: a survey. In: International Joint Conferences on Artificial Intelligence (2018)
20. Tian, X., Zhang, M.: Research on spatial correlations and influencing factors of logistics industry development level. *Sustainability* **11**(5), 1356 (2019)



# Unify the Research and Design of Power Quality Regulators

Yaoyao Qin<sup>1</sup>, Chen Lei<sup>1</sup>, Liping He<sup>1</sup>(✉), and Longjing Liu<sup>2</sup>

<sup>1</sup> School of Automation and Electrical Engineering, Linyi University, Linyi 276005, China  
lp-h@163.com

<sup>2</sup> Shandong Electric Power Group Linyi Power Supply Company, Linyi 276005, China

**Abstract.** In order to ensure a stable power supply and meet the development needs of the power market, the Unified Power Quality Controller (UPQC) device has quickly attracted the attention of experts in the field of electric power. In this paper, the topology of UPQC, the model of the series and parallel parts of UPQC, and its power flow are introduced. The command signal detection method based on the Resonant Attenuation Principle (RAC) and the ip-iq detection method based on the instantaneous reactive power theory are discussed and analyzed. The two controllers were used for analysis and the selection of suitable controllers for the series-parallel portions. Finally, simulations are carried out to simulate voltage drops, rises, current distortions and other working conditions to verify the effectiveness of the UPQC model for improving power quality.

**Keywords:** unified power quality regulator · ip-iq assay · resonant attenuation control · PR control

## 1 Introduction

UPQC stands for Unified Power Quality Conditioner (UPQC), which is a device with comprehensive power quality regulation function, and is an important representative of the Active Power Filter (APF) family. It combines series active filters and parallel active filters, so that it can give full play to the advantages of series and parallel active filters in power system applications, and has a comprehensive power quality adjustment function [1].

In 1991, the United States Westinghouse Electric Engineer Laszlo Gyugyi proposed a unified power flow controller (Unified Power FlowController, UPFC) [2], but it is mainly used in the transmission system, and the distribution side of the power quality problems are also increasing, so the Japanese scholar Hirofumi Akagi was inspired by the UPFC for the first time at the conference proposed the concept of UPQC. Therefore, Japanese scholar Hirofumi Akagi [3], inspired by UPFC, first proposed the concept of UPQC in the conference, which can simultaneously manage grid voltage and nonlinear load current. For UPQC, the current research focuses on topology improvement and control strategies [4, 5]. For topology improvement, [6] applied a modular multilevel converter to UPQC to make its output level closer to sinusoidal waveform, which to

some extent can eliminate the output-side filter, but due to the need for a large number of switching devices, the inconsistency in the characteristics of the switching devices may easily cause uneven flow, and the uneven voltage and circulating current suppression between modules need to be taken into account, which increases the control difficulty. In addition, it is also necessary to consider the voltage equalization and loop current suppression between the modules, which increases the control difficulty. For the control strategy of UPQC, most of the studies are based on the indirect control strategy [7, 8] and the direct control strategy [9–11] for extended derivation. Although the two control strategies differ in control objects and objectives, the final control effects are basically the same [12], which realize the harmonic compensation and reactive power compensation of grid-side input currents, regulate the load voltage as a stable and adjustable sinusoidal voltage, and ensure that the dc bus voltage does not fluctuate significantly. In addition, novel approaches based on the two control strategies derived from the literature [13–16] have been applied to UPQC control.

With the closeness of artificial intelligence to life, new technologies such as fuzzy control combined with new disciplines such as neural networks, genetic algorithms and chaos theory, artificial neural network control, and intelligent control of sliding mode variable structure are emerging, which greatly improves UPQC's ability to improve power quality.

## 2 Topology of UPQC

According to the connection position of the series-parallel APF, UPQC can be divided into the following two structures: (1) the left side is connected in series and the right side is connected in parallel. The series converter is partially close to the power supply, while the parallel converter is close to the load; When this topology is applied, the problems of voltage rise, drop, and distortion are compensated by the series APF, and the parallel APF mainly compensates for the current quality problems, such as the compensation of reactive power and the elimination of harmonics for the load. The left side is connected in parallel and the right side is connected in series. The series converter part is close to the load, and the parallel converter part is close to the power supply side; Parallel APF maintains DC side voltage stability; The series APF mainly isolates the load to generate harmonics into the power grid. However, due to the proximity of the parallel APF to the power supply side, when the system voltage is distorted, the working performance of the parallel APF part will be affected.

In this paper, the UPQC of the left-series-right-parallel three-phase three-wire system shown in is studied, which is composed of a power supply, a series transformer, a capacitor, an inductor, a series APF, a parallel APF, and a nonlinear load, and the series-parallel part becomes a back-to-back structure through DC capacitors (Fig. 1).

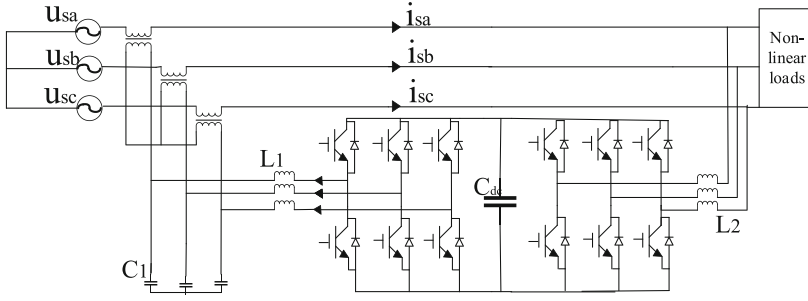


Fig. 1. Overall topology construction of UPQC

### 3 Detection of UPQC Command Signal

In this paper, the ip-iq detection method based on the basic instantaneous reactive power theory and the signal detection method based on the RAC resonance attenuation controller (RAC) are compared and analyzed.

The ip-iq detection method first obtains the phase of the power supply voltage, uses the sine and cosine signal generation circuit to generate the sine and cosine signal, obtains the transformation matrix  $C$  from the signal, and then converts the three-phase current into the two-phase coordinate system, and multiplies the current in the two-phase coordinate system with  $C$  to obtain  $i_p$  and  $i_q$ :

$$\begin{pmatrix} i_p \\ i_q \end{pmatrix} = C \begin{pmatrix} i_a \\ i_b \\ i_c \end{pmatrix}, \begin{pmatrix} i_a \\ i_b \\ i_c \end{pmatrix} = C_{32} \begin{pmatrix} i_p \\ i_q \end{pmatrix}, C = \begin{pmatrix} \sin \omega t & -\cos \omega t \\ -\cos \omega t & -\sin \omega t \end{pmatrix} \quad (1)$$

Signal detection method based on RAC Resonance Attenuation Controller (RAC): Firstly, the Clarke transform of the three-phase current in the abc coordinate system is performed to obtain  $\alpha\beta$  the two current signal components in the stationary coordinate system. Secondly, the RAC controller is used to attenuate the signal of the non-resonant frequency to obtain the  $\alpha\beta$  fundamental positive sequence component in the coordinate system. Finally, the Clarke inverse transform is performed to obtain the three-phase fundamental current component, which is then differentiated with the load current to obtain the command current. Through the complete process of the above three parts, the three-phase load current can be controlled, and the fundamental components can be extracted, and finally the required command current can be obtained.

The resonant attenuation control method includes a positive sequence controller and a negative sequence controller, which are used to control the positive and negative sequence components of the unbalanced current, respectively. The transfer function is used to implement this method by controlling the system to effectively suppress resonance phenomena. The transfer function is shown in Eq. (2):

$$\begin{cases} GPRAC = \frac{K_{iwc}}{s-jw+wc} \\ GNRAC = \frac{K_{iwc}}{s+jw+wc} \end{cases} \quad (2)$$

From the transfer function of the principle of resonant attenuation, it can be seen that:

$$\begin{cases} i_{\alpha\beta}^+ = i\alpha\beta \frac{kiwc}{s-jw+wc} \\ i_{\alpha\beta}^- = i\alpha\beta \frac{kiwc}{s+jw+wc} \end{cases} \quad (3)$$

In the above equation,  $K_i$  is the gain coefficient, which  $i\alpha\beta$  is  $\alpha\beta$  the component of the load current in the coordinate system, and the positive and negative signs represent the current components of the positive and negative sequences, which can be obtained by further simplification

$$\begin{cases} i_{\alpha\beta}^+ = \frac{jwi_{\alpha\beta}^+ + i\alpha\beta wc - wci_{\alpha\beta}^+}{s} \\ i_{\alpha\beta}^- = \frac{-jwi_{\alpha\beta}^- + i\alpha\beta wc - wci_{\alpha\beta}^-}{s} \end{cases} \quad (4)$$

In the  $\alpha\beta$  coordinate system, the positive and negative order components can be transformed as shown in Eq. (5).

$$\begin{cases} ji_{\alpha}^+ = -i_{\beta}^+ \\ ji_{\beta}^+ = i_{\alpha}^+ \\ ji_{\alpha}^- = i_{\beta}^- \\ ji_{\beta}^- = -i_{\alpha}^- \end{cases} \quad (5)$$

Sorting out the above two formulas, we can get Eq. (6):

$$\begin{cases} i\alpha^+ = \frac{\omega ci\alpha - \omega ci\alpha^+ - \omega i\beta^+}{s} \\ i\beta^+ = \frac{\omega ci\beta - \omega ci\beta^+ - \omega i\alpha^+}{s} \\ i\alpha^- = \frac{\omega ci\alpha - \omega ci\alpha^- - \omega i\beta^-}{s} \\ i\beta^- = \frac{\omega ci\beta - \omega ci\beta^- - \omega i\alpha^-}{s} \end{cases} \quad (6)$$

From this, the basic structure of the resonant attenuation controller can be completed, as shown in Fig. 2 below:

When the resonant frequency of the model is consistent with the fundamental angular frequency  $w$  (usually  $100 \pi \text{ rad/s}$ ) in the power grid,  $K_i$  take 1 and then set the appropriate  $wc$  to extract the fundamental component.

In this paper, the signal detection method based on the resonant attenuation controller and the  $i_p$ - $i_q$  detection method are studied, and the fundamental components and command signals extracted from them are compared and analyzed.

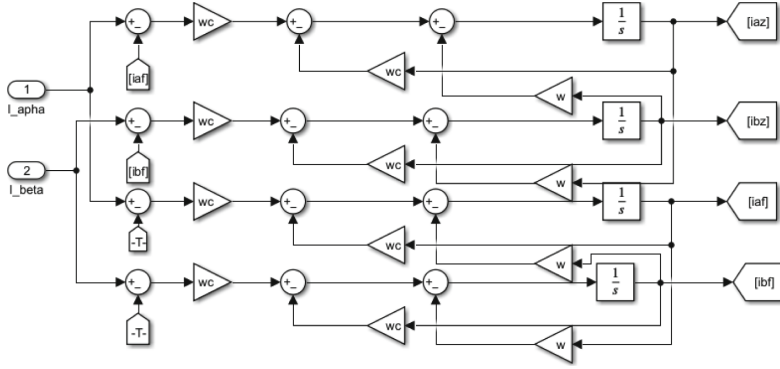


Fig. 2. Basic structure of RAC

## 4 Simulation Study of UPQC Overall System Based on Three-Phase Three-Wire Topology

### 4.1 UPQC Series Part Control Design

The composition structure of the PI controller is relatively simple and easy to implement. In the signal tracking system, the setting of the proportional link directly affects the response speed of the signal tracking, while the setting of the integration link affects the accuracy of the tracking. Therefore, when designing the PI controller, it is necessary to comprehensively consider the parameter settings of these two links to ensure that the system can achieve a good performance state.

Transfer Functions of the PI Controller:

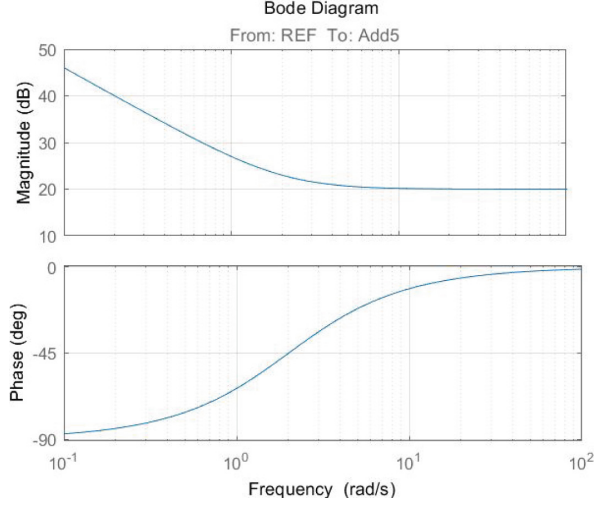
$$G(s) = k_p + \frac{k_i}{s} \quad (7)$$

where:  $k_p$  is the scale coefficient and  $k_i$  is the integration coefficient.

Figure 3 is the Bode diagram at  $k_p = 10$  and  $k_i = 20$ , and it can be seen that the gain of the PI controller is relatively high in the low frequency region. As the frequency increases, the gain gradually decreases and eventually stabilizes around a value.

In an AC circuit, voltage and current typically contain AC components. Therefore, when dealing with three-phase circuits, the stationary three-phase coordinates need to be converted, which is easier to analyze and calculate when converted into the dq coordinate system. At the same time, in order to avoid partial coupling in the dq coordinate system and make the system independent of each other in the dq coordinate axis, it is necessary to further decouple the system control. In the second chapter of this paper, the mathematical expression for the loop equation with the filtered inductance in series is (8) in the dq coordinate system (Fig. 4):

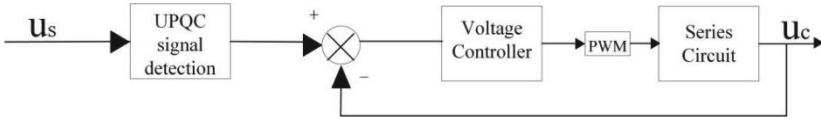
$$\begin{aligned} S1dU_{dc} &= -(L1 \frac{di1d}{dt} + R1i1d) - \omega L1i1q + u1d \\ S1qU_{dc} &= -(L1 \frac{di1q}{dt} + R1i1q) - \omega L1i1d + u1q \end{aligned} \quad (8)$$



**Fig. 3.** PI controller Bode diagram

$$\begin{aligned} i1d &= C1 \frac{duc d}{dt} - \omega C1 ucq + icd \\ i1q &= C1 \frac{duc q}{dt} + \omega C1 ucd + icq \end{aligned} \quad (9)$$

The harmonic content of the voltage in the power system is relatively low, and the sudden drop or rise in the system voltage often occurs. The PI controller is selected to realize the compensation control work of the series part to ensure that the tracking performance of the compensation is guaranteed. This method can effectively reduce the voltage fluctuation in the system and improve the stability and reliability of the system.



**Fig. 4.** Flowchart of series compensation control

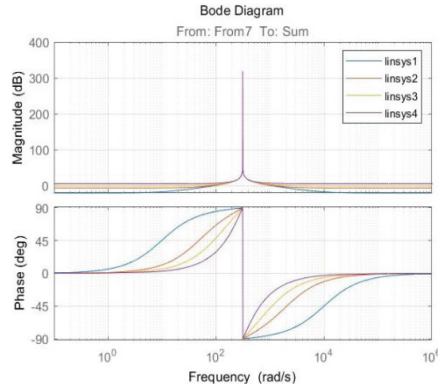
## 4.2 UPQC Parallel Part Control Design

The transfer function of the PR controller is:

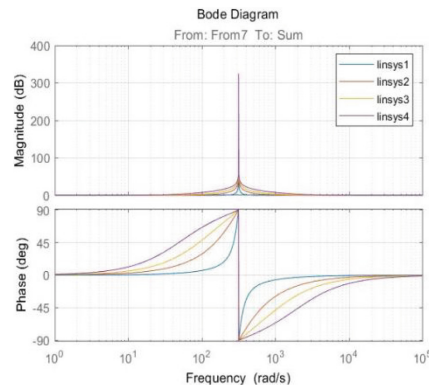
$$GPR(s) = kp + \frac{k_r s}{s^2 + \omega_0^2} \quad (10)$$

where is the  $k_p$  scale coefficient and is the  $k_r$  resonance coefficient.  $\omega_0 = 2\pi f_0$

As shown in Fig. 5, the PR controller has a higher gain at the resonant frequency, which can effectively compensate for the signal at the resonant frequency and suppress the signal at the non-resonant frequency. Therefore, it can be seen that the PR controller has better performance in compensating and suppressing the signal.



**Fig. 5.** PR controller bode diagram



**Fig. 6.** PR controller bode diagram

Analyze  $k_p$  the  $k_r$  influence of two parameters on the performance of the PR controller. If the parameter  $k_r = 1000$  is unchanged, it  $k_p$  changes from 0.1, 0.5, 1, and 5 in turn. As can be seen in Fig. 5, the gain of the PR controller at the resonant frequency is significantly higher than at other frequencies, so the signal at the resonant frequency can be effectively compensated. In addition, the simulation results show that  $k_p$  although the change of parameters will affect the harmonic content, it will not change the gain at the resonant frequency.



In the case that the parameter  $k_p$  does not change and is equal to 1, it  $k_r$  changes from 100, 500, 1000, and 2000 in turn. As can be seen from Fig. 6, the parameter changes of the PR controller  $k_r$  is related to the gain of the frequency, but does not affect the bandwidth. These findings provide an important reference for further optimization of controller parameters.

From the parallel model introduced in Sect. 2, the mathematical expression in the dq coordinate system is:

$$\begin{aligned} S2dU_{dc} &= -(L2 \frac{di_{2d}}{dt} + R2i_{2d}) - \omega L2i_{2q} + u_{1d} \\ S2qU_{dc} &= -(L2 \frac{di_{2q}}{dt} + R2i_{2q}) - \omega L2i_{2d} + u_{1q} \end{aligned} \quad (11)$$

Similarly, from Eq. (11), it can be seen that the voltage and current of the D-axis of the system interact with the Q-axis; In order to reduce this effect, the parallel part adopts a cross-decoupling control method. The parallel part is controlled by the PR controller, and the control strategy adopted makes  $U_{dc}$  and  $i_{L2}$  as the control quantity of the outer loop voltage and the feedback quantity of the inner loop respectively.

The current command signal  $i_{L2d}^*$ ,  $i_{L2q}^*$  the actual value of the inductor current  $i_{L2d}$ , and  $i_{L2q}$  subtract to obtain the error signal, and then send it to the current regulator for processing, and after processing, the result is sent to the PWM module, and the parallel side converter is controlled to achieve the purpose of real-time tracking compensation, as shown in Fig. 7.

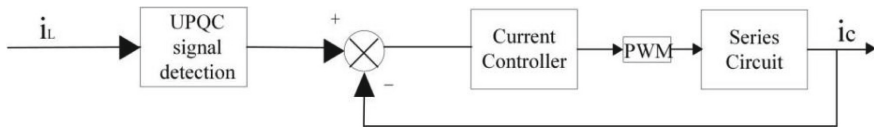
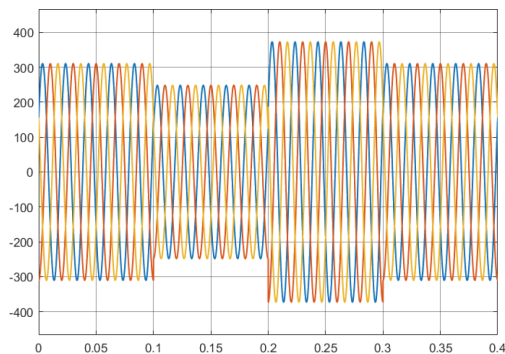


Fig. 7. Working flowsheet of based on PR controller

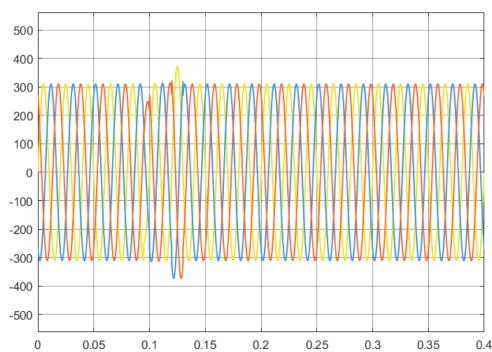
### 4.3 Simulation Analysis of PQC System Model

The UPQC model was built on the Simulink simulation platform, as shown in Fig. 8. The size of the parameter setting is as follows: the power supply is three-phase sinusoidal alternating current, and the size of the voltage is 220V; The filter inductance of the series part is 2 mH; Filter capacitance 50.5  $\mu F$ ; The filter inductance of the parallel unit is 2.5 mH, the capacitor voltage reference value is set to 800 V, the DC unit is set to the size of 3000  $\mu F$  energy storage capacitor, and the nonlinear load is composed of an uncontrollable rectifier bridge and  $20\sqrt{3} \Omega$  resistor, and a 10 mH inductor is connected in parallel.

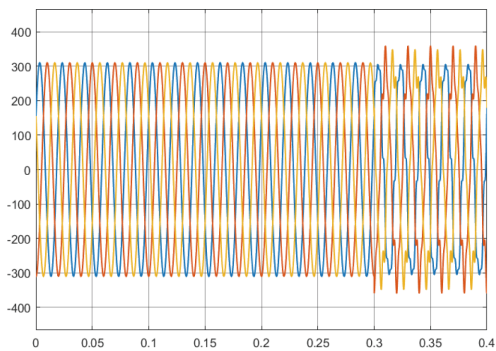




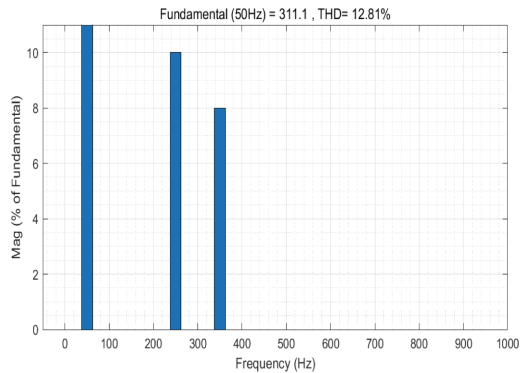
**Fig. 9.** Waveform of temporary rise and fall of grid voltage



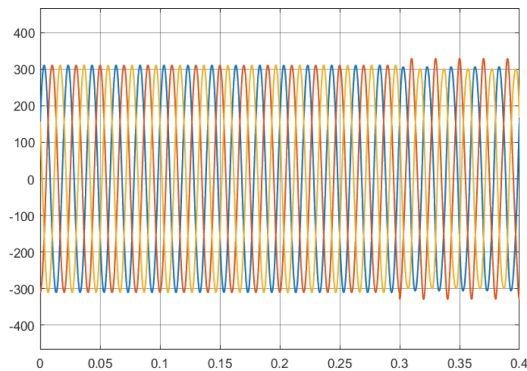
**Fig. 10.** Compensated grid voltage



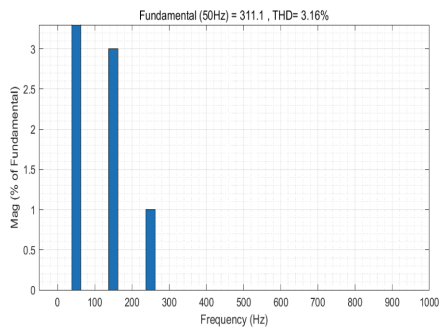
**Fig. 11.** Grid voltage containing harmonics



**Fig. 12.** Analysis of voltage distortion rate



**Fig. 13.** Compensated grid voltage

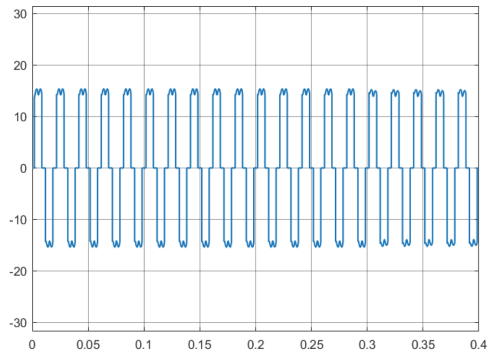


**Fig. 14.** Compensated analysis of voltage distortion rate

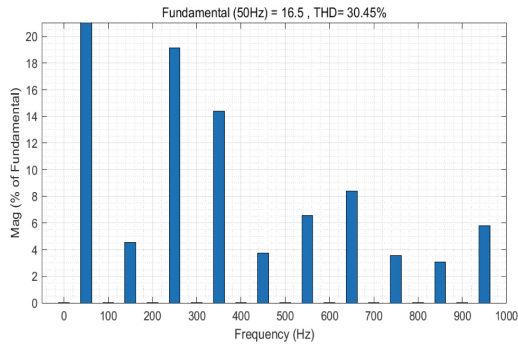
**4.3.2 Parallel Simulation Analysis**

The harmonic current generated by the nonlinear load is processed by the parallel converter, which is the main reason for the pollution current in the odd harmonic of the

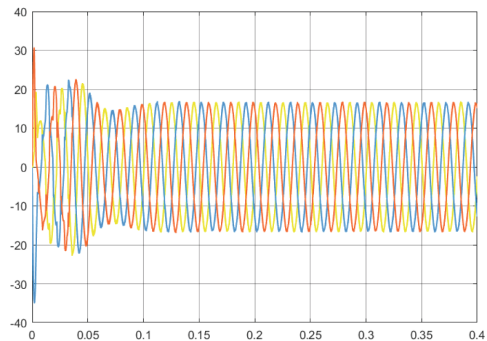
5th and 7th order. The current waveform of the front load is shown in Fig. 15, and the waveform of the load current after the parallel unit compensation is shown in Fig. 17.



**Fig. 15.** Load current waveform

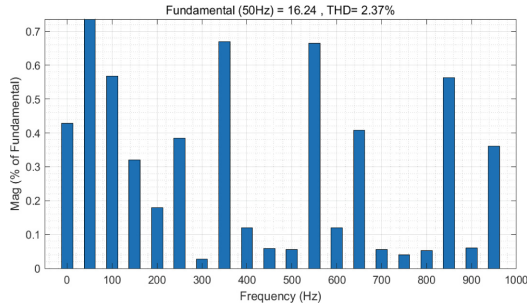


**Fig. 16.** Analysis of load current distortion rate



**Fig. 17.** Compensated load current

As can be seen from Figs. 15 and 17, the grid current has been greatly improved after the parallel compensator work compensation. It can be concluded that the control



**Fig. 18.** Compensated analysis of load current distortion rate

compensation strategy selected in this paper is effective, and the UPQC reduces the content of the third and fifth harmonics. As can be seen from Figs. 16 and 18, the current distortion rate is reduced from 30.45% to 2.37%, and the current quality is improved to a certain extent. As a result, the UPQC system can suppress power quality problems in the grid.

## 5 Summary

This paper presents the  $i_p$ - $i_q$  and resonant-attenuation controller methods. Both methods were used to extract instruction signals, with their results compared. The RAC method effectively reduced harmonic components but had inaccuracies in phase and amplitude of the fundamental wave, indicating the need for further model refinement. In contrast, the  $i_p$ - $i_q$  method yielded ideal extraction results, leading to the selection of its detection method, based on instantaneous reactive power theory, for the design of the subsequent compensation module.

Before the compensating device is built, the model needs to be decoupled to reduce the influence of system coupling on the simulation results. After a brief analysis of the characteristics of the PI controller, it is found that it meets the requirements of the series part and the DC part control compensation strategy, that is, the PI controller is determined. When selecting the parallel part of the controller, the working characteristics of the PR controller were analyzed in detail, and the appropriate parameters were selected by comprehensive consideration.

Finally, the complete simulation model was built and simulated, and the verification and analysis were carried out for several different working conditions from the perspective of grid-side voltage and load current. The simulation results show that UPQC can keep the voltage stability on the grid side and reduce the harmonic content of the load current, and it is concluded that the UPQC device can effectively suppress a variety of power quality problems.

**Acknowledgments.** This study was funded by Design Reform of Automation and Electrical Engineering Specialty Based on Engineering Education Professional Certification (No. JG2021Z12).

## References

1. Sawadogo, W., Abiodun, B.J., Okogbue, E.C.: Impacts of global warming on photovoltaic power generation over West Africa. *Renew. Energy* **151**, 263–277 (2020)
2. Gyugyi, L.: A unified flow control concept for flexible AC transmission systems. In: *International Conference on Ac & Dc Power Transmission*. IET (1991)
3. Fujita, H., Akagi, H.: The unified power quality conditioner: the integration of series active-filters and shunt active filters. In: *PESC Record. 27th Annual IEEE Power Electronics Specialists Conference* (1998)
4. Geury, T., Pinto, S., Gyselinck, J.: Direct control method for a PV system integrated in an Indirect Matrix Converter-based UPQC. In: *The 4TH Ieee International Conference on Smart Energy Grid Engineering (SEGE)* (2016)
5. Zhong, Y., Xia, M., Chiang, H.D.: Electric vehicle charging station micro-grid providing unified power quality conditioner support to local power distribution networks. *Int. Trans. Electr. Energy Syst.* **27**(3), 2262 (2017)
6. Garikapati, R., Kumar, S.R., Karthik, N.: ANFIS controlled MMC-UPQC to mitigate power quality problems in solar PV integrated power system. *J. Adv. Res. Appl. Sci. Eng. Technol.* **36**(1), 102–130 (2023)
7. Karanki, S.B., Geddada, N., Mishra, M., et al.: A modified three-phase four-wire UPQC topology with reduced DC-link voltage rating. *IEEE Trans. Ind. Electron.* **60**(9), 3555–3566 (2013)
8. Raphael, J., dos Santos, M., Carlo, J., da Cunha, M., Mezaroba, A.: A simplified control technique for a dual unified power quality conditioner. *IEEE Trans. Ind. Electron.* **61**(11), 5851–5860 (2014). <https://doi.org/10.1109/TIE.2014.2314055>
9. Kesler, M., Ozdemir, E.: Synchronous-reference-frame-based control method for UPQC under unbalanced and distorted load conditions. *IEEE Trans. Ind. Electron.* **58**(9), 3967–3975 (2011). <https://doi.org/10.1109/TIE.2010.2100330>
10. Modesto, R.A., Silva, S.A.O., Oliveira, A.A., et al.: A versatile unified power quality conditioner applied to three-phase four-wire distribution systems using a dual control strategy. *IEEE Trans. Power Electron.* **31**(8), 5503–5514 (2016)
11. Patjoshi, R.K., Mahapatra, K.: High-performance unified power quality conditioner using command generator tracker-based direct adaptive control strategy. *IET Power Electron.* **9**(6), 1267–1278 (2016)
12. Trinh, Q.N., Lee, H.H.: Improvement of power quality under distorted source and nonlinear load conditions. In: *Proceedings of the 7th International Power Electronics and Motion Control Conference*. IEEE (2012)
13. Krishna, D., Sasikala, M., Kiranmayi, R.: FOPI and FOFL controller based UPQC for mitigation of power quality problems in distribution power system. *J. Electr. Eng. Technol.* **17**(3), 1543–1554 (2022)
14. Meng, L., Ma, L., Zhu, W., et al.: Control strategy of single-phase UPQC for suppressing the influences of low-frequency DC-link voltage ripple. *IEEE Trans. Power Electron.* **37**(2), 2113–2124 (2021)
15. Mansor, M.A., Hasan, K., Othman, M.M., et al.: Construction and performance investigation of three-phase solar PV and battery energy storage system integrated UPQC. *IEEE Access* **8**, 103511–103538 (2020)
16. Çelik, D., Ahmed, H.: Enhanced control of superconducting magnetic energy storage integrated UPQC for power quality improvement in EV charging station. *J. Energy Storage* **62**, 106843 (2023)

# **Learning and Systems**





# A Zero Trust Continuous Authentication Scheme Based on Keystroke Dynamics

Ping Gong<sup>(✉)</sup>

School of Artificial Intelligence and Big Data, Zibo Vocational Institute, Zibo, Shandong, China  
990500701@qq.com

**Abstract.** The zero trust model enhances the security of the Internet of Things, but frequent identity verification affects user efficiency, especially the inconvenience caused by multi system switching, complex processes, and remote login. Moreover, multi-step authentication exacerbates learning costs. Keystroke dynamics, a biometric technology for analyzing typing habits, adapted to a zero trust framework, achieves low interference verification and improves user experience by continuously monitoring user behavior. This solution optimizes the authentication process without compromising security, reduces the burden of frequent verification, and is particularly beneficial for users who frequently travel and need to quickly switch devices, creating a smoother workflow.

**Keywords:** Zero Trust · Continuous Authentication · Keystroke Dynamics

## 1 Introduction

Under a zero trust architecture, continuous authentication of user identity is crucial for ensuring network security. Traditional static authentication methods (such as passwords or tokens) or multi factor authentication are difficult to continuously verify user identity throughout the entire session [1]. The keystroke dynamics provides a dynamic and continuous user authentication method that can compensate for the shortcomings of traditional security measures.

Introducing keystroke dynamics into zero trust continuous authentication can provide a dynamic and seamless authentication method. In this scheme, first of all, when the user initially uses the system, the system will guide them to type for a period of time to collect their keystroke dynamics data, such as keystroke duration, keystroke interval, keystroke speed, etc. These data will be used to train a machine learning model that can learn and recognize each user's unique keystroke patterns. In daily use, the system backend program continuously monitors the user's keystroke behavior, collects relevant data in real time, and regularly uses this data to verify the user's identity, especially when the user accesses sensitive resources. This verification process is almost imperceptible to users, ensuring that it does not interfere with normal workflow. If the system detects keystrokes that do not match the user's normal mode, it will trigger additional authentication steps or security warnings. Users can provide feedback on incorrect authentication results to help the system better adjust and optimize the keystroke dynamics model. In order to

ensure user privacy and data security, all keystroke data is encrypted during the collection and transmission process.

Integrating keystroke dynamics into a zero trust security framework not only enhances the security of the Internet of Things, but also provides a continuous and dynamic verification mechanism. Attackers often find it difficult to imitate this habit or need to pay a higher price, greatly reducing potential security threats. This method is particularly suitable for scenarios with high security requirements, such as sensitive data processing and access control of critical infrastructure. At the same time, this method not only maintains high security standards, but also avoids the inconvenience caused to users in traditional security measures. The application of keystroke dynamics improves the concealment of security measures, making it difficult for potential attackers to identify and exploit vulnerabilities in security systems.

## **2 Relevant Technical Background**

### **2.1 Zero Trust Architecture**

With the increasing scale and complexity of network systems in the Internet of Things environment, traditional boundary based network security models are no longer able to meet the security needs of rapidly developing network systems. To address this issue, the concept of zero trust was first proposed in 2010 [2] and applied through Google's BeyondCorp project [3]. The innovation of zero trust architecture lies in abandoning the traditional boundary based security model, deploying all applications on a public network, and no longer dividing the network into internal and external. The zero trust architecture accepts the fact that threats are everywhere in the network. Assuming that the network environment is always insecure, authorization no longer depends on network location, and all devices, users, and network traffic need to be authenticated and authorized according to dynamic security policies. Even within the network, continuous identity authentication is required for authorization. This approach enables users to access securely from any location and using any device, meeting the current demand for remote work.

The essence of zero trust is identity centric dynamic access control [4], whose core idea is to distrust devices, users, and network traffic inside and outside the network. In a zero trust architecture, permission allocation is no longer based on static rules, but requires continuous identity authentication, real-time evaluation of trust based on user behavior during resource access applications, and dynamic allocation of permissions based on trust scores.

### **2.2 Keystroke Dynamics**

Keystroke dynamics is a behavioral biometric technology that measures and evaluates the rhythm of human typing on digital devices, verifying user identity based on their keystroke patterns [5]. The emergence of keystroke dynamic biometric technology can be traced back to the late 19th century, when the telegraph revolution was at its peak and it was the main long-distance communication tool of that era. Telegraphers can

seamlessly distinguish each other by simply listening to the tapping rhythm of dots and strokes. With the development of computer and sensing technology, computer keyboards, mobile phone keyboards, and touch screens have become common input devices in the 21st century.

This technology identifies unique user behavior characteristics by measuring and evaluating the intensity, rhythm, and time interval of individual keystrokes. The study of keystroke dynamics is based on the premise that everyone has their own unique gestures and habits when using a keyboard, just like biological characteristics. A typical keystroke dynamic authentication system consists of several parts, including data collection, feature extraction, classification/matching, decision-making, and retraining. Data collection is the fundamental stage of collecting raw keystroke data through various input devices. Feature extraction is the process of processing raw keystroke data and storing it as a reference template for future use. Prior to feature extraction, some pre-processing processes can be applied to ensure or improve the quality of feature data. Classification/matching is the process of classifying and distinguishing feature data for making decisions during identity verification. Decision making is based on the results of classification or matching algorithms to determine whether a user is legitimate.

Further applications include disease diagnosis and human-computer interaction research. For example, in the field of disease diagnosis, by analyzing keystroke dynamics data, subtle hand movement changes caused by muscle diseases or neurodegenerative diseases such as Parkinson's disease can be discovered [6]. In human-computer interaction research, keystroke dynamics can help design keyboards that are more ergonomic and reduce fatigue or damage caused by typing.

Overall, keystroke dynamics, as a user behavior analysis technique, has broad application prospects, especially in the fields of security verification and health monitoring. With the development of data processing and artificial intelligence technology, keystroke dynamics can enable more accurate identity verification.

### 2.3 Particle Swarm Optimization

Particle Swarm Optimization (PSO) is a heuristic optimization algorithm inspired by observing the behavior of groups such as birds or fish. In PSO, candidate solutions in the solution space are called particles, and each particle adjusts its position and velocity based on its own experience and the experience of its neighbors to find the optimal solution.

In PSO, each particle has two attributes: position and velocity. It independently searches for the optimal solution in the solution space, and the individual optimal solution found is called the individual extremum. The optimal solution among the individual extremum of all particles is called the global optimal solution. Particles adjust their position and velocity based on individual extremum and global optimum, in order to find the global optimum during the search process. The specific formula is as follows:

$$v_{ij}(t+1) = wv_{ij}(t) + c_1r_1(t)[p_{ij}(t) - x_{ij}(t)] + c_2r_2(t)[g_{ij}(t) - x_{ij}(t)] \quad (1)$$

$$x_{ij}(t+1) = x_{ij}(t) + v_{ij}(t+1) \quad (2)$$



results to help the system better adjust and optimize the keystroke dynamics model. In order to ensure user privacy and data security, all keystroke data is encrypted during the collection and transmission process.

Integrating keystroke dynamics into a zero trust security framework not only enhances the security of the Internet of Things, but also provides a continuous and dynamic verification mechanism. Attackers often find it difficult to imitate this habit or need to pay a higher price, greatly reducing potential security threats. This method is particularly suitable for scenarios with high security requirements, such as sensitive data processing and access control of critical infrastructure. At the same time, this method not only maintains high security standards, but also avoids the inconvenience caused to users in traditional security measures [8]. The application of keystroke dynamics improves the concealment of security measures, making it difficult for potential attackers to identify and exploit vulnerabilities in security systems.

### 3.2 Key Steps of the Plan

The specific scheme of using keystroke dynamics for zero trust continuous authentication can be divided into five key steps, as shown in Fig. 2.

#### 1) Data collection and processing

Data collection of keystroke dynamics: Collect data on the user's keystroke behavior during normal use, including keystroke duration, keystroke interval, keystroke speed, etc.

Preprocessing: Preprocessing the collected data, such as removing noise, standardizing data, and handling missing values.

#### 2) User behavior modeling

Feature extraction: Extracting key features from processed data, such as button delay, flight time, etc.

Machine learning model training: Use extracted features to train machine learning models (such as support vector machines, random forests, neural networks, etc.) to recognize each user's unique button pressing habits.

#### 3) Real time monitoring and verification

Real time data capture: Capture users' keystroke behavior in real-time while using the system.

Continuous verification: Apply trained models to analyze real-time captured keystroke data to verify user identity. This can continue after the user logs in, ensuring that the user's identity remains unchanged during the session.

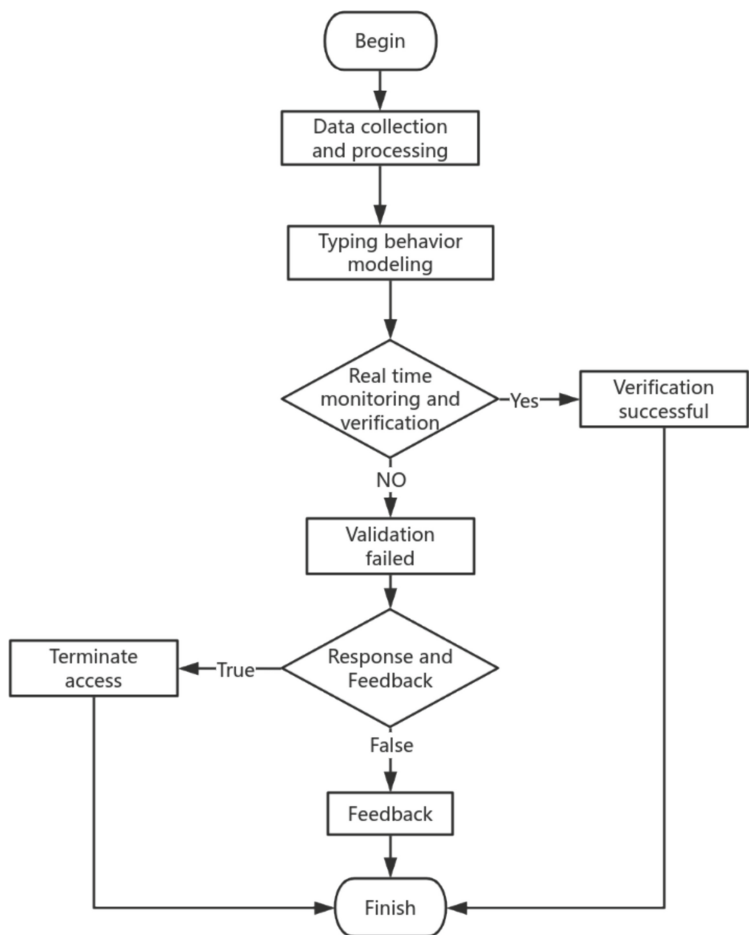
#### 4) Response and Feedback

Authentication feedback: If a keystroke pattern is detected that does not match a known user pattern, the system can take appropriate measures, such as requesting additional authentication, issuing a security warning, or terminating the session directly.

User feedback mechanism: Allow users to provide feedback on incorrect authentication results to improve the accuracy of verification.

#### 5) Model update and optimization

Continuous learning: Over time, users' keystroke habits may change. Therefore, it is necessary to regularly update and retrain the model to adapt to these changes.



**Fig. 2.** Key steps for continuous authentication

Due to the fact that the authentication process is based on the user’s natural behavior rather than mandatory security steps, the user experience is greatly improved [9]. No longer requiring frequent password input or multi factor authentication, users can seamlessly continue their work while the system remains silently alert in the background [10]. By integrating keystroke dynamics into a zero trust security framework, a more seamless and dynamic authentication process can be achieved, ensuring security while maintaining a comfortable user experience.

## 4 Conclusions

This article introduces keystroke dynamics into the zero trust framework to achieve continuous identity authentication, providing a dynamic and seamless security solution for the frequent authentication mechanism of zero trust. Compared to traditional authentication methods, this approach can continuously verify identity without the user's awareness, enhancing both security and user experience. Especially suitable for high security environments that handle sensitive data and protect critical infrastructure, it not only reduces potential security threats but also minimizes interference with users. Meanwhile, by combining the improved particle swarm optimization algorithm to optimize the random forest model, the accuracy of identity verification has been further improved, achieving a balance between security and convenience.

**Data Availability.** The data used to support the findings of this study are included within the article.

**Conflicts of Interest.** The authors declares that there is no conflict of interest regarding the publication of this paper.

## References

1. Ahmad, M.O., Tripathi, G., Siddiqui, F., et al.: BAAuth-ZKP—a blockchain-based multi-factor authentication mechanism for securing smart cities. *Sensors* **23**(5), 2757 (2023)
2. Osborn, B., McWilliams, J., Beyer, B., et al.: Beyondcorp: design to deployment at google. In: *Usenix*, vol. 41, p. 28 (2016)
3. Ward, R., Beyer, B.: Beyondcorp: a new approach to enterprise security. *Login* **39**(6), 6–11 (2014)
4. Yan, X., Wang, H.: Survey on zero-trust network security. In: *Artificial Intelligence and Security: 6th International Conference, ICAIS 2020, Hohhot, China, 17–20 July 2020, Proceedings, Part I 6*, pp. 50–60. Springer Singapore (2020)
5. Teh, P.S., Teoh, A.B.J., Yue, S.: A survey of keystroke dynamics biometrics. *Sci. World J.* **2013**, 24 (2013)
6. Tripathi, S., Arroyo-Gallego, T., Giancardo, L.: Keystroke-dynamics for Parkinson's disease signs detection in an at-home uncontrolled population: a new benchmark and method. *IEEE Trans. Biomed. Eng.* **70**(1), 182–192 (2022)
7. Peška, L., Veselý, P., Skopal, T., et al.: Person authentication using visual representations of keyboard typing dynamics. In: *2022 Ninth International Conference on Social Networks Analysis, Management and Security (SNAMS)*, pp. 1–6. IEEE (2022)
8. Naz, M., Shah, J.H., Khan, M.A., et al.: From ECG signals to images: a transformation based approach for deep learning. *PeerJ Comput. Sci.* **7**, e386 (2021)
9. Lis, K., Niewiadomska-Szynkiewicz, E., Dziewulska, K.: Siamese neural network for keystroke dynamics-based authentication on partial passwords. *Sensors* **23**(15), 6685 (2023)
10. Tufail, S., Parvez, I., Batool, S., et al.: A survey on cybersecurity challenges, detection, and mitigation techniques for the smart grid. *Energies* **14**(18), 5894 (2021)



# Cognitive-Based Autonomous Orbit of Non-Cooperative Targets

Jianliang Ma<sup>1</sup>, Lele Zhang<sup>1,2(✉)</sup>, Jian Yang<sup>3</sup>, and Fang Deng<sup>1,2</sup>

<sup>1</sup> Greater Bay Area Innovation Research Institute, Beijing Institute of Technology, Zhuhai 519088, China

zhanglele@bit.edu.cn

<sup>2</sup> School of Automation, Beijing Institute of Technology, Beijing 100081, China

<sup>3</sup> School of Optics and Photonics, Beijing Institute of Technology, Beijing 100081, China

**Abstract.** Unmanned aerial vehicle (UAV) has received extensive attention in recent years. It is significant to achieve proximal observation of non-cooperative, highly dynamic, and large solidity targets by using of UAV. In this paper, a three-stage control method is proposed to solve this problem, which includes bearing-only tracking, obstacle-avoidance tracking, and proximity-based observation. Specifically, radar and depth camera, are used for observation, and the three-stage control method is designed by refinement of the traditional path planning method Rapidly-exploring Random Tree (RRT) and Model Predictive Control (MPC) technologies. A quadcopter model was developed in Gazebo, where its feasibility in validating the control method was demonstrated, providing essential technical support for the approach and observation of non-cooperative targets.

**Keywords:** UAV · RRT · Obstacle avoidance · MPC

## 1 Introduction

UAV has been widely used in academic, military and industrial fields. Due to the flexibility and miniaturization, real-time information such as location, speed, and status can be obtained to support mission execution by using UAV for tracking targets. The information can enhance situational awareness and monitoring, support decision-making, to enhance mission accuracy and response time. Unmanned aerial vehicles (UAVs) are extensively employed in the areas such as reconnaissance, command and control, and environmental modeling.

The current research primarily focuses on improving Artificial Intelligence (AI) algorithms to achieve tracking and detection independent of target size [1, 2]; additionally, it includes target tracking and detection in different environments (e.g., indoor environments) [3, 4]; multi-angle tracking using UAV clusters [5]; target tracking and detection under limited sensor conditions (e.g., using only direction angle) [6]. There is limited research on proximity observation for large solidity non-cooperative targets, and the obstacle avoidance problem must be addressed in proximity observation, complicating the tracking task. In the field of autonomous vehicles, a path prediction method [7] using



obstacle orientation and road information is proposed, a nonlinear model predictive control[8] is proposed, kinematic model of obstacles is established, position and velocity error feedback is used to optimize the trajectory prediction. However, in practice, when approaching and observing large solidity targets, it is challenging to determine the target's orientation. Additionally, non-cooperative targets complicate the understanding of the obstacle's kinematic model, making the methods proposed in [7] and [8] insufficient for solving this issue. To solve these problems, we can divide approach observation into several steps, and adopt different control algorithms in different stages to approach gradually. The following are the steps and some challenges required to achieve proximity observation.

- (1) The first step is to detect the target according to the onboard sensor. For airborne sensors, vision is one of the most suitable solutions. Some deep learning algorithms are capable of target detection [9, 10] pointing out that YOLOv8 is considered to be the most advanced and robust target detection platform. Once the object is detected, we can obtain bearing measurements from the image based on the internal parameters of the camera [11]. However, in many cases, UAV is far away from the target, and the depth camera or radar sensor cannot detect the specific position and velocity of the target. Therefore, the guidance algorithm that only relies on the bearing is often an effective tracking strategy. The challenge is that when the sensor's detection range is short and neither position nor velocity of the target can be accurately obtained, traditional guidance methods like PNG [12] struggle to control speed near the target, often failing to decelerate in time if approaching too quickly. In [6], it is verified that circular or elliptical motion of the observer will often improve the observability, and a pseudo-linear Kalman filter is designed to estimate the target position and velocity, but this method ignores energy consumption and often increases the difficulty of control when moving at high speed.
- (2) The second step is to control the aircraft to approach the target when the sensor can detect the target position and velocity information. Once the depth camera detects the exact position, we need to approach the target while avoiding its obstacles. Widely adopted obstacle avoidance strategies include graph-based search algorithms such as Improved A\* [13], sample-based search algorithms such as Anytime Rapidly exploring Random Tree algorithm(Anytime RRT) [14], and machine learning-based algorithms such as [15]. Algorithms based on virtual potential field, such as artificial potential field(APF) method [16], etc. The challenge arises from the simultaneous motion of both the obstacle and the target, imposing higher real-time performance demands on the algorithm compared to scenarios with a dynamic obstacle and a stationary target. The real-time performance of existing algorithms remains a significant challenge [17].
- (3) The third step is to achieve circling observation after approaching the target. We assume that the vision algorithm can capture key points of different parts of the large solid target. Once in the designated observation position, circling observation becomes crucial. For a moving target, several strategies have been proposed based solely on bearing information, enabling multiple observers to circumnavigate the target [18]. Additionally, a control method has been developed to allow the aircraft to orbit the target along any smooth, closed trajectory [19]. The challenge is that

their control method is suitable for circling fixed targets, and designing effective circling motions for moving targets is the key to improving observability.

This paper aims to overcome the above challenges, and the following solutions are proposed to address them.

- (1) The static target encirclement guidance [18] is enhanced in consideration of allowing dynamic adjustment of the encircling radius, enabling end-velocity control while rapidly approaching the target based on bearing-only information.
- (2) For dynamic obstacles and targets, this paper proposes a real-time RRT planning method for dynamic root node movement. It achieves on-the-fly UAV trajectory planning, and obtains better real-time performance than RRT.
- (3) For the circling problem, this paper combines MPC control to achieve spiral motion trajectory and complete the circling observation of the target.

## 2 Methodology

### 2.1 Bearing-Only Target Following

To address the problem of end-speed control when approaching a target, we can improve the target encirclement guidance proposed in [18], which relies solely on the bearing. The flight control can be divided into two components: the z-axis and the x-y-plane. For the analysis of the x-y plane, we assume that the velocity heading angle of the UAV is denoted by  $\theta$  and the robot's position is represented by  $(x, y)$  with respect to a global coordinate frame, the target position in the robot's world coordinate frame  $(x_t, y_t)$ , the distance between the aircraft and the target point  $\rho$ . We define

$$\alpha := \arctan2(x_t, y_t) \quad (1)$$

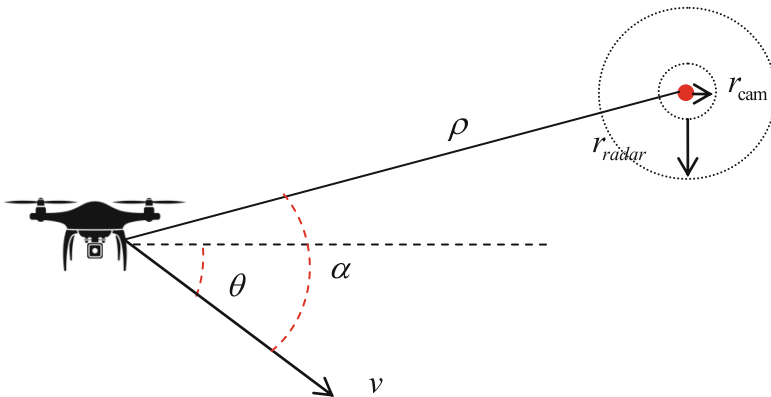


Fig. 1. An illustration of defined parameters.

To complete obstacle avoidance and tracking detection, we assume that the UAV is equipped with a depth camera and Lidar. It is assumed that the detection distance of the depth camera is  $r_{cam}$ , and the detection distance of the lidar is  $r_{radar}$ . (see Fig. 1).

[18] Using Lyapunov stability theory, a control law that can move in a circle around a stationary target is proposed:

$$\begin{cases} v = k_v v_{\max} \cos(\alpha + \phi), \\ w = k_w w_{\max} \sin(\alpha + \phi), \end{cases} \quad (2)$$

where  $k_v, k_w \in (0, 1]$  are control gains for forward speed and angular speed, respectively. Let  $v_{\max}$  and  $w_{\max}$  represent the maximum linear velocity and maximum angular velocity, respectively. Suppose we aim to move in a circular path with a radius  $r$  around the target. The control will be determined by the following formula:

$$\phi = -\arctan \frac{r \cdot k_w w_{\max}}{k_v v_{\max}} \quad (3)$$

This method is often very effective in tracking the target when it is stationary, but it is often difficult to achieve the ideal effect for moving objects, especially when  $r$  is large, the aircraft typically begins preparing for circular motion at a considerable distance, which reduces its tracking ability and increases energy consumption. Conversely, when  $r$  is small, while the tracking performance at a distance is improved, the terminal velocity after approach tends to be low, and the speed variation is significant, making it challenging to complete the maneuver effectively. Therefore, a dynamic control law must be designed. A large value of  $r$  can be selected in the initial phase, and the radius  $r$  can be progressively reduced once the range is detectable by radar. This paper proposes a strategy that leverages the relative stability of deceleration when  $r$  is large and the strong tracking capability when  $r$  is small.

$$\begin{cases} r_{end} = \frac{v_{tar} + \Delta v}{k_w w_{\max} \sin(\phi)}, \\ r = r_{end} \frac{(r_{radar} - (d - r_{cam}))}{r_{radar}} + r_{start} \frac{(d - r_{cam})}{r_{radar}}, \end{cases} \quad (4)$$

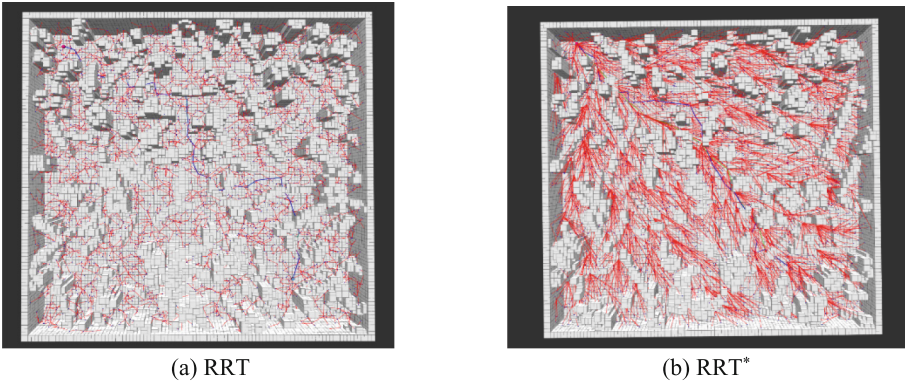
$d$  is the distance between the aircraft and the target,  $r_{end}$  is defined according to the terminal velocity  $v_{tar}$ ,  $\Delta v$  is greater than the target velocity,  $r_{start}$  needs to be defined in terms of actual vehicle and sensor performance. When  $d = r_{radar} + r_{cam}$ , it indicates that the system is far from the target, with  $r = r_{start}$ . When  $d = r_{cam}$ , it signifies that the system has reached the vicinity of the target, with  $r = r_{end}$ . It is worth noting that (4) is not the only guidance law. The actual attenuation law can adopt other methods, such as exponential attenuation, or it can be based on time attenuation, which can be independent of radar observation data. However, why not choose other parameters to adjust? Firstly, to meet the terminal condition, it is essential to ensure that (3) is satisfied. Therefore, altering  $k_v$  and  $k_w$  simultaneously changes at least two parameters, which increases the control complexity. Additionally, changes in  $k_v$  and  $k_w$  will inevitably affect  $v$  and  $w$ , thereby influencing the UAV's maneuvering performance and reducing its tracking capability.

For the z-axis, the PNG control law proposed in [12] can be applied to achieve control.

## 2.2 Obstacle Avoidance and Target Tracking for UAV

In this paper, the classical RRT path planning method is improved, and an NMPC [25] control strategy is applied to achieve trajectory planning that satisfies UAV kinematics, enabling the control of the UAV to perform the desired maneuvers. The following provides a brief introduction to the RRT algorithm.

The principle of RRT is similar to a path search tree. Starting from the root node, random sampling is performed in the space to identify the nearest branch that can be connected without obstacles, continuing until the goal range is reached, thereby forming a feasible path (Fig. 2(a)). This method is not optimal, which led to the development of the RRT\* algorithm. In RRT\*, the newly added sampling points serve as the center of a circle, from which better parent nodes are identified and rewiring operations are performed. As iterations increase, this method converges to the optimal path (Fig. 2(b)).

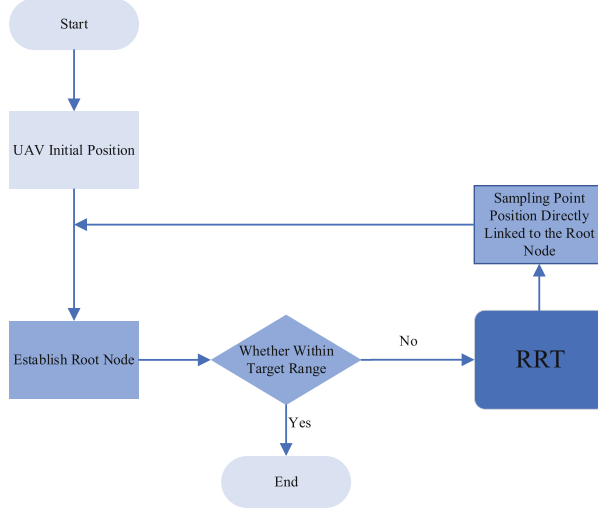


**Fig. 2.** RRT and RRT\* in Gazebo.

In fact, several algorithms have been proposed to enhance the optimality and exploration speed of the RRT algorithm. For example, RRT# utilizes lazy optimization, Informed-RRT\* [20] focuses sampling on regions with potential optimal paths, and Anytime RRT gradually refines the path within a limited time frame [21]. For dynamic obstacles, DRRT\* [22] has been proposed, which monitors environmental changes in real time during path planning, particularly the movement of dynamic obstacles, allowing for replanning and path adjustment. However, real-time challenges remain when obstacles or targets move rapidly [23]. In this paper, the real-time requirements become more stringent when both obstacles and targets are moving simultaneously.

Therefore, the RRT algorithm is improved in this paper. After an effective path is formed each time, we take the first sampling point, that is, the sampling point directly connected to the root node, as the effective path point and take this sampling point as the new root node. The next exploration starts from the new root node to explore the new path. This process continues until the new root node reaches the target range. This process will generate a pursuit path, where the root node pursues the target, as illustrated in Fig. 3. It is evident from this process that it imposes high demands on the real-time exploration capabilities of the RRT algorithm. Therefore, the number of RRT explorations should

be limited, and meeting these requirements is often challenging using traditional RRT methods. Once RRT fails to find a suitable path within a limited number of explorations, the UAV will get trapped. Therefore, improving the exploration speed of RRT is an issue that needs to be considered.



**Fig. 3.** Path Planning Process.

Due to the movement of the target points and obstacles, it is difficult to find a global path and then locally optimize it. However, we can adopt the efficient approach of Informed-RRT\* to identify the potential exploration range, with the added requirement of dynamically adjusting the range. Although the obstacles are moving, in the context of large entity pursuit, the actual obstacles consistently move in close to the target point, in conjunction with its motion. Consequently, no obstacles will be present at a distance, allowing the obstacle distance threshold to be set based on the radar's detection range. It is important to note that the distance to the obstacle here refers to the expanded boundary of the obstacle. Let us assume that  $\text{Consider\_Obstacle} = 1$  means the obstacle is considered, and  $\text{Consider\_Obstacle} = 0$  means the obstacle is not considered. The logical preference expression is as follows:

$$\text{Consider\_Obstacle} = \begin{cases} 1, & d \leq d_{obs} \\ 0, & d > d_{obs} \end{cases} \quad (5)$$

This implies that when the distance to the obstacle is relatively large, the obstacle can be ignored as the system moves toward the target. Ignoring obstacles can effectively increase the speed of exploration. Besides,  $\text{Consider\_Obstacle} = 0$  means that the UAV is relatively far from the obstacle. In this case, the exploration step size and exploration range can be set relatively large to improve the exploration speed. When the UAV is close to the obstacle, the exploration step size and exploration range can be smaller. The

exploration range can be represented by polar coordinates. The pitch  $\theta \in (0, \pi)$ , and yaw  $\psi \in (0, 2\pi)$  are set, and the control law of the range ( $P_x, P_y, P_z$ ) is as follows:

$$\begin{cases} P_x = x_{start} + r_{std} \sin(\theta) \cos(\psi) \\ P_y = y_{start} + r_{std} \sin(\theta) \sin(\psi) \\ P_z = z_{start} + r_{std} \cos(\theta) \end{cases} \quad (6)$$

where  $r_{std} \propto r$ ,  $r$  is the distance between the UAV and the target point, and  $(x_{start}, y_{start}, z_{start})$  denotes the position of the UAV. As seen from the expression, the range is always constrained by the starting point and the distance with the UAV. The range can be adjusted based on the actual situation. Similarly, the step size can also be modified according to the actual conditions. The key idea is that when entering the obstacle detection area ( $d \leq d_{obs}$ ), the step size should be gradually reduced. The following law is adopted in this paper:

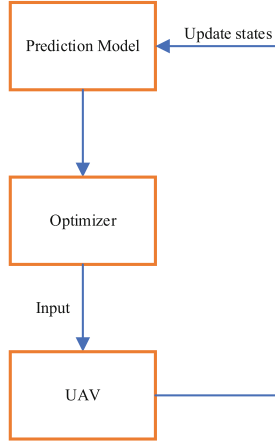
$$step = \begin{cases} step_1, d \geq d_{obs} \\ \frac{step_2}{k_1 - k_2 d}, d < d_{obs} \end{cases} \quad (7)$$

where  $step_1$  and  $step_2$  are designed with step sizes based on practical considerations, while  $k_1$  and  $k_2$  are gains that determine the rate of attenuation. After the RRT path is planned, the MPC is used to control the UAV for path tracking. The following is a brief introduction to MPC control methods.

MPC control can be divided into three steps. First, the kinematic model of the UAV is established and discretized, and the predictive model is derived. Then the optimized objective function and related kinematic constraints are set up. Finally, the optimization solution is carried out according to the constraints and objective functions. The following takes quadrotor UAV as an example, Fig. 4 shows this process. We define  $x = [p^T v^T \eta^T w^T]$ , where  $p$  is the UAV position in world frame,  $v$  is the UAV velocity in world frame,  $\eta = [\varphi, \theta, \psi]^T$  is Z-X-Y euler angle representation of the UAV orientation, and  $w$  is the angular velocity of UAV. [25] define input as the  $u = [\dot{f}_1, \dot{f}_2, \dot{f}_3, \dot{f}_4]$ . The quadrotor UAV is generally taken as  $f = c_f w^2$ , where  $w$  is the angular velocity of the rotor and  $c_f$  is the thrust coefficient.

$$\min_{u_0 \dots u_N} \sum_{k=0}^N \|x_k - x_{ref,k}\|_Q^2 + \sum_{k=0}^{N-1} \|u_k - u_{ref,k}\|_R^2 \quad (8)$$

In the objective function,  $x_k$  denotes the system state at time step  $k$ , while  $x_{ref,k}$ , represents the desired reference state. Similarly,  $u_k$  is the control input, and  $u_{ref,k}$  is the reference control input at time step  $k$ . The matrices  $Q$  and  $R$  are the weighting matrices for the state and control deviations, respectively, which determine their relative importance in the optimization. The parameter  $N$  defines the prediction horizon for the optimization. This paper utilizes the qpOASES solver [26], an efficient open-source library for solving quadratic programming (QP) problems. In practical applications, nonlinear MPC combined with the qpOASES solver is commonly used in control scenarios that require real-time handling of nonlinear constraints and dynamic systems. Given the nonlinearities and constraints of flight dynamics, employing nonlinear MPC to optimize the flight path and solving it in real time using qpOASES is an effective approach.



**Fig. 4.** MPC Process.

### 2.3 Orbit and Target Tracking

The ideal circling observation trajectory is usually a closed trajectory around the target, such as a circle or spiral shape, to maintain the distance and Angle change between the drone and the target, so that the drone can continue to observe the target from multiple angles. This trajectory can achieve all-round observation of the target, and is suitable for target monitoring and three-dimensional reconstruction tasks in various environments. The actual circling trajectory needs to be calculated according to the actual requirements and actively explore the unobserved area. For example, [24] design a sampling-based active reconstruction strategy to explore the unobserved areas for 3D Gaussians splatting. In this paper, spiral circling trajectory is designed for approaching the target. After obtaining the target pose by using the depth camera, we can design the spiral circling path for this, and then directly use the MPC to control the UAV to complete the path tracking. For example, suppose the vision sensor detects a decision to make a spiral loop around the target in the Z-axis direction. We define  $(x_{goal}, y_{goal}, z_{goal})$  and  $(x_{UAV}, y_{UAV}, z_{UAV})$  is the position of goal and UAV respectively. Path control is defined as follows:

$$\begin{cases} x = r \cos(\theta) + x_{goal} \\ y = r \sin(\theta) + y_{goal} \\ z = k \cdot \theta + z_{goal} \end{cases} \quad (9)$$

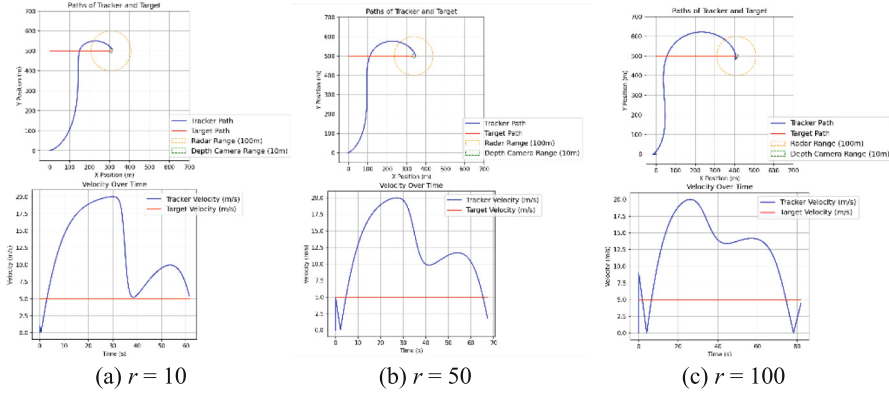
where  $r$  determines the radius of the orbit.  $\theta$  is the angle that increases over time, with the initial value of  $\theta$  ensuring  $(x, y, z) = (x_{UAV}, y_{UAV}, z_{UAV})$ .  $k$  is the growth coefficient along the z-axis. To improve the observability, we can adjust different trajectories according to the actual scene:

$$z = k \cdot \theta + z_{goal} + A \sin(wt) \quad (10)$$

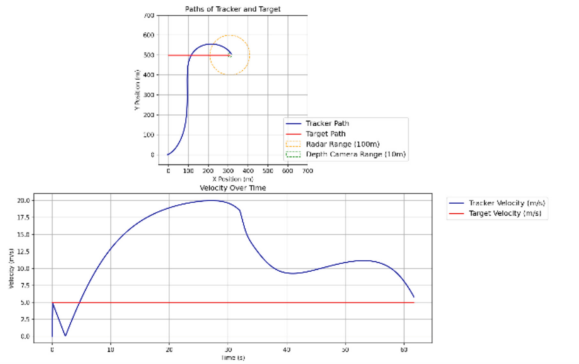
where  $A$  and  $w$  represent the amplitude and angular frequency of the sine function, respectively.

### 3 Result

#### 3.1 Bearing-Only Target Following



**Fig. 5.** Presents the simulation results of control guidance (2) against a constant-velocity target.



**Fig. 6.** Presents the simulation results of control guidance (4).

We define  $v_{max} = 20$  m/s,  $v_{tar} = 20$  m/s,  $w_{max} = 1$  rad/s,  $k_v = 1$ ,  $k_w = 1$ . As shown in Fig. 5, the original control law does not achieve dynamic tracking. When  $r = 10$ , there is a rapid speed change, which presents a challenge for UAV control. For  $r = 50$  and  $r = 100$ , since the radius is greater than the detection range of the depth camera, the UAV tends to decelerate as it approaches the target, causing the terminal speed to be lower than the actual speed, which is unfavorable for close-range observation.

In this paper, we set  $r_{start} = 50$ ,  $\Delta v = 1$  m/s, and as shown in Fig. 6, the improved tracking control law avoids the rapid speed changes caused by a smaller radius while ensuring that the terminal speed approaches the target speed, providing an appropriate initial speed for subsequent close-range observations.



### 3.2 Obstacle Avoidance and Target Tracking for UAV

In this paper, the UAV kinematics model will be built under ROS and Gazebo simulation environment to demonstrate the effectiveness of the above obstacle avoidance algorithm. It will be divided into three parts, static obstacle avoidance, dynamic tracking, dynamic obstacle avoidance and tracking.

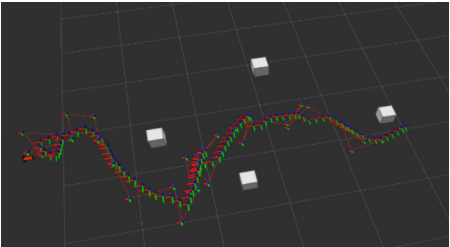
- (1) **Static obstacle avoidance.** Static obstacle avoidance algorithms are very mature. We use the above improved algorithms to continuously push the root node to move towards the target to verify the effectiveness of obstacle avoidance. As can be seen from Fig. 7, this algorithm can achieve the obstacle avoidance functionality of traditional RRT, but the number of updates of the algorithm in this paper is equal to the number of nodes in the figure, and the relative time is longer than that of traditional algorithms. However, the algorithm in this paper enables simultaneous tracking and path planning, achieving effective results.
- (2) **Dynamic Tracking:** Traditional RRT algorithms are typically obstacle avoidance methods and cannot pursue targets effectively. The improved RRT algorithm in this paper demonstrates good tracking ability, with the target moving in a circular motion at a constant speed, as shown in Fig. 8. Compared to the traditional PNG algorithm, this algorithm has a non-optimal path, lower pursuit ability, and higher energy consumption. However, it offers better obstacle avoidance, a feature not present in the traditional PNG algorithm.
- (3) **Dynamic Obstacle Avoidance and Tracking:** The second stage of this paper focuses on pursuing the target while avoiding obstacles. The target moves in a constant circular motion, with obstacles moving along with it. As shown in Fig. 9, after adjusting the logical preferences, step size, and obstacle search range, the desired results have been achieved.

### 3.3 Orbit and Target Tracking

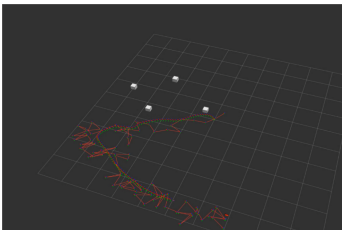
This paper validates a feasible circling trajectory in ROS and Gazebo. As shown in Fig. 10, the circling of the target is achieved, with the specific circling strategy needing to be designed based on different requirements, such as actual modeling or measurement.

## 4 Limitations

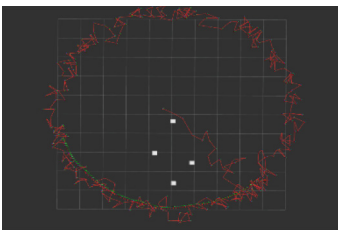
Although the three-stage control can initially achieve tasks such as fast tracking, obstacle avoidance, and circular observation, it assumes that the target's motion velocity is not too high. Especially for obstacle avoidance, rapidly changing obstacles impose higher real-time requirements on RRT. This paper does not incorporate the velocities of obstacles and target point into the RRT, meaning the algorithm still treats the target as stationary during each path update, which hinders fast obstacle avoidance. Additionally, obstacle consideration is based solely on distance, without integrating the UAV's and obstacle's speeds. The tracking strategy divides control into the x-y plane and z-plane, addressing control robustness and speed reduction. The circling path is controlled solely by the target's position, and in cases of high target speed, collisions are likely, as target velocity is not considered. Future work will address these limitations.



**Fig. 7.** Static obstacle avoidance.

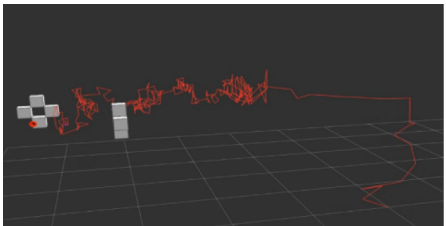


(a) Path planning at time = 12s

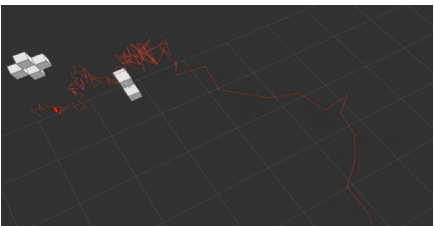


(b) Path planning at time = 40s

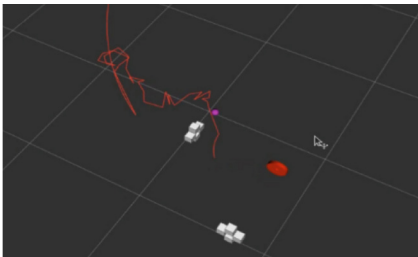
**Fig. 8.** Different time of dynamic tracking.



(a) The first execution

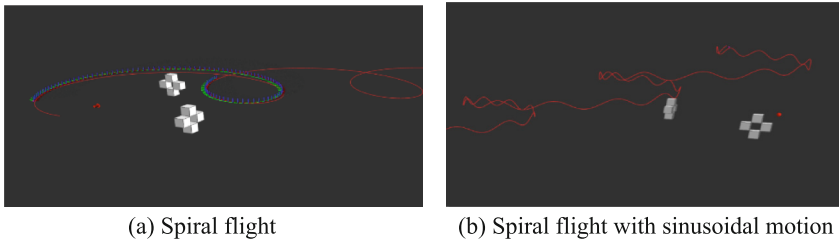


(b) The second execution



(c) The third execution

**Fig. 9.** Different scenarios of dynamic obstacle avoidance and tracking.



**Fig. 10.** Presents the simulation results of control guidance (9) and (10).

## 5 Conclusion

This paper presents a three-stage control law for fast tracking, obstacle avoidance, and circular flight from a distance. It implements a bearing-only guidance law to control terminal velocity, improves RRT for dynamic obstacle avoidance and fast tracking, and integrates MPC for spiral circling control. The proposed control law enhances UAV tracking and observation capabilities for large moving entities and demonstrates potential for dynamic target tracking in complex environments.

**Acknowledgments.** This work was supported in part by the National Natural Science Foundation of China National Science Fund for Distinguished Young Scholars under Grant 62025301, and the National Natural Science Foundation of China General Program under Grant 62473054, and the National Natural Science Foundation of China Basic Science Center Program under Grant 62088101, and the Young Elite Scientist Sponsorship Program by CAST under Grant YESS20220509.

**Disclosure of Interests.** The authors declare that they have no competing interests.

## References

1. Fan, H., et al.: Lasot: a high-quality benchmark for large-scale single object tracking. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Long Beach, USA, pp. 5374–5383 (2019)
2. Duan, R., Fu, C., Alexis, K., Kayacan, E.: Online recommendation-based convolutional features for scale-aware visual tracking. In: 2021 IEEE International Conference on Robotics and Automation (ICRA), Xi'an, China, pp. 14206–14212 (2021)
3. Khosiawan, Y., Nielsen, I.: A system of UAV application in indoor environment. *Prod. Manuf. Res.* **4**(1), 2–22 (2016)
4. Sampedro, C., Rodriguez-Ramos, A., Bavlé, H., Carrio, A., De la Puente, P., Campoy, P.: A fully-autonomous aerial robot for search and rescue applications in indoor environments using learning-based techniques. *J. Intell. Rob. Syst.* **95**, 601–627 (2019)
5. Azam, M.A., Dey, S., Mittelman, H.D., Ragi, S.: Decentralized UAV swarm control for multitarget tracking using approximate dynamic programming. In: 2021 IEEE World AI IoT Congress (AIIoT), Online, pp. 0457–0461 (2021)
6. Li, J., Ning, Z., He, S., Lee, C.H., Zhao, S.: Three-dimensional bearing-only target following via observability-enhanced helical guidance. *IEEE Trans. Rob.* **39**(2), 1509–1526 (2022)

7. Flores, M.J., Armingol, J.M., de la Escalera, A.: Real-time drowsiness detection system for an intelligent vehicle. In: 2008 IEEE Intelligent Vehicles Symposium, pp. 637–642. IEEE (2008)
8. Lindqvist, B., Mansouri, S.S., Agha-mohammadi, A.A., Nikolakopoulos, G.: Nonlinear MPC for collision avoidance and control of UAVs with dynamic obstacles. *IEEE Rob. Autom. Lett.* **5**(4), 6001–6008 (2020)
9. Zheng, Y., Chen, Z., Lv, D., Li, Z., Lan, Z., Zhao, S.: Air-to-air visual detection of micro-uavs: an experimental evaluation of deep learning. *IEEE Rob. Autom. Lett.* **6**(2), 1020–1027 (2021)
10. Alhafnawi, M., et al.: A survey of indoor and outdoor uav-based target tracking systems: current status, challenges, technologies, and future directions. *IEEE Access* **11**, 68324–68339 (2023)
11. Ma, Y., Soatto, S., Košecká, J., Sastry, S.: An Invitation to 3-d Vision: From Images to Geometric Models. Springer, New York (2004)
12. Roelofs, G.: PNG: The Definitive Guide. O'Reilly & Associates, Inc. (1999)
13. Ju, C., Luo, Q., Yan, X.: Path planning using an improved a-star algorithm. In: 2020 11th International Conference on Prognostics and System Health Management (PHM-2020 Jinan), Jinan, China, pp. 23–26 (2020)
14. Karaman, S., Walter, M.R., Perez, A., Frazzoli, E., Teller, S.: Anytime motion planning using the RRT\*. In: 2011 IEEE International Conference on Robotics and Automation, Shanghai, China, pp. 1478–1483 (2011)
15. Dong, Y., Zou, X.: Mobile robot path planning based on improved DDPG reinforcement learning algorithm. In: 2020 IEEE 11th International Conference on Software Engineering and Service Science (ICSESS), Beijing, China, pp. 52–56 (2020)
16. Agirrebeitia, J., Avilés, R., De Bustos, I.F., Ajuria, G.: A new APF strategy for path planning in environments with obstacles. *Mech. Mach. Theory* **40**(6), 645–658 (2005)
17. Yang, Y., Xiong, X., Yan, Y.: UAV formation trajectory planning algorithms: a review. *Drones* **7**(1), 62 (2023)
18. Zheng, R., Liu, Y., Sun, D.: Enclosing a target by nonholonomic mobile robots with bearing-only measurements. *Automatica* **53**, 400–407 (2015)
19. Sinha, A., Cao, Y.: Three-dimensional autonomous guidance for enclosing a stationary target within arbitrary smooth geometrical shapes. *IEEE Trans. Aerosp. Electron. Syst.* **59**(6), 9247–9256 (2023)
20. Gammell, J.D., Srinivasa, S.S., Barfoot, T.D.: Informed RRT\*: optimal sampling-based path planning focused via direct sampling of an admissible ellipsoidal heuristic. In: 2014 IEEE/RSJ International Conference on Intelligent Robots and Systems, Chicago, IL, USA, pp. 2997–3004 (2014)
21. Ferguson, D., Stentz, A.: Anytime rrt\*. In: 2006 IEEE/RSJ International Conference on Intelligent Robots and Systems, pp. 5369–5375. IEEE (2006)
22. Connell, D., La, H.M.: Dynamic path planning and replanning for mobile robots using RRT. In: 2017 IEEE International Conference on Systems, Man, and Cybernetics (SMC), pp. 1429–1434. IEEE (2017)
23. Muhsen, D.K., Raheem, F.A., Sadiq, A.T.: A systematic review of rapidly exploring random tree RRT algorithm for single and multiple robots. *Cybern. Inf. Technol.* **24**(3), 78–101 (2024)
24. Jin, R., Gao, Y., Wang, Y., Lu, H., Gao, F.: GS-planner: a gaussian-splatting-based planning framework for active high-fidelity reconstruction. *arXiv preprint [arXiv:2405.10142](https://arxiv.org/abs/2405.10142)* (2024)
25. Jacquet, M., Corsini, G., Bicego, D., Franchi, A.: Perception-constrained and motor-level nonlinear MPC for both underactuated and tilted-propeller UAVS. In: 2020 IEEE International Conference on Robotics and Automation (ICRA), Paris, France, pp. 4301–4306 (2020)
26. Ferreau, H.J., Kirches, C., Potschka, A., Bock, H.G., Diehl, M.: QpOASES: a parametric active-set algorithm for quadratic programming. *Math. Program. Comput.* **6**, 327–363 (2014)



# Finite-Time Stabilizing Control of Furuta Pendulum System Based on Disturbance Observer and Integral Sliding Mode Method

Zhujun Wang<sup>1</sup>, Shuli Gong<sup>2</sup>(✉), Ancai Zhang<sup>1</sup>, and Junyao Yu<sup>1</sup>

<sup>1</sup> School of Automation and Electrical Engineering, Linyi University,  
Shandong 276000, China

<sup>2</sup> School of Mathematics and Statistics, Linyi University, Shandong 276000, China  
gongshuli@lyu.edu.cn

**Abstract.** This paper studies the stabilizing control of an underactuated Furuta pendulum system subjected to an unknown matched disturbance. Firstly, the system's dynamic model is derived using Lagrange modeling method. Through approximate linearization and a coordinate transformation, this model is converted into a standard chained form system around the zero equilibrium point. Subsequently, the disturbance is estimated in finite time by a Levant differentiator disturbance observer. And then, a robust stabilization controller is designed using an integral sliding mode technique and the estimated disturbance value. It enables the Furuta pendulum system to be robustly stabilized at the zero equilibrium point in a finite time. The effectiveness of our presented method is verified from numerical examples.

**Keywords:** Furuta pendulum system · Disturbance observer · Integral sliding mode · Finite-time stabilization

## 1 Introduction

Mechanical systems can be categorized as fully actuated systems, underactuated systems, or redundant systems depending on the number of degrees of freedom (DOF) and control inputs [1]. The underactuated systems are often lightweight, energy-efficient, and capable of performing flexible movements [2, 3]. Their motion control is very useful research topic. However, this control problem is difficult to solve because the underactuated systems often have high complexity and nonholonomic constraints [4].

To facilitate the study on motion control of underactuated systems, researchers have developed some simplified benchmark models in laboratory. Furuta pendulum model is such an example. It is a classic underactuated system introduced by Professor Katsuhisa Furuta in the 1990s [5]. And it has been widely applied in the teaching and research of control theory [6].

Nowadays, researchers have conducted extensive study on the stabilizing control of the Furuta pendulum. And a number of control strategies have been developed, e.g., a linear quadratic regulator (LQR) approach [7], a neural networks method [8]. In [9], a sliding mode control (SMC) strategy was introduced, which has fast response and high robustness [10]. For the SMC control method, the linear sliding mode is a commonly used form. However, the linear sliding mode method often results in chattering, which fails to meet the performance requirements of many high-demand systems [11]. Additionally, the terminal sliding mode method can rapidly drive system to sliding surface. However, it may cause the control input to approach infinity, which is not feasible in practical applications [12]. In contrast, the integral sliding mode control (ISMC) method addresses these issues by introducing an integral term. It can effectively reduce steady-state errors [13,14]. Nevertheless, this method currently only achieves asymptotic stabilization. This limits the ability of control system to operate safely and efficiently.

On the basis of research background above, this paper discusses the robust finite-time stabilizing control problem for Furuta pendulum system with an unknown matched disturbance. First, a finite-time disturbance observer (DO) based on Levant differentiator is constructed to accurately estimate unknown external disturbance. A stabilizing controller is then designed by combining the DO with the nested ISMC method. It can guarantee the Furuta pendulum to be robustly stabilized at the origin within a finite time.

## 2 Definition and Lemmas

**Definition 1.** [15] *Consider the nonlinear system*

$$\dot{\chi}(t) = \kappa(t, \chi(t)), \quad \chi(0) = \chi_0 \quad (1)$$

with  $\kappa(t, 0) = 0, \forall t > 0$ . If the system (1) is globally asymptotically stable and its any solution can reach the origin within finite time, then we call (1) globally finite-time stable, where  $T(\chi_0)$  is a settling time constant related to the initial value  $\chi_0$ .

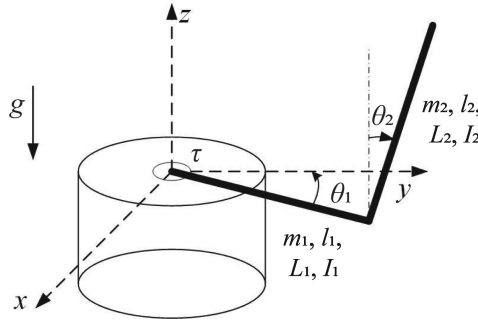
**Lemma 1.** [16] *The normal chain system*

$$\begin{cases} \dot{x}_1 = x_2, \\ \dot{x}_2 = x_3, \\ \vdots \\ \dot{x}_{n-1} = x_n, \\ \dot{x}_n = -\sum_{i=1}^n \alpha_i |x_i|^{\beta_i} \text{sgn}(x_i) \end{cases} \quad (2)$$

converges to zero within a finite time, where the constant  $\alpha_i$  ensures that the polynomial  $p^n + \alpha_n p^{n-1} + \dots + \alpha_2 p + \alpha_1$  is Hurwitz, the constant  $\beta_i$  satisfies

$$\beta_{i-1} = \frac{\beta_i \beta_{i+1}}{2\beta_{i+1} - \beta_i}, \quad \beta_n = \beta, \quad \beta_{n+1} = 1, \quad i = 2, \dots, n, \quad (3)$$

and  $\beta \in (1 - \varepsilon, 1)$ ,  $\varepsilon \in (0, 1)$ .



**Fig. 1.** Model of Furuta pendulum.

**Lemma 2.** [17, 18] Consider the nonlinear system given below,

$$\phi(t) + \mu h(\phi(t)) = 0, \quad (4)$$

where  $\mu > 0$ ,  $h(\phi(t))$  is a nonlinear function and the condition  $h(0) = 0$  needs to be satisfied. If

$$\dot{h}(\phi(t)) = \text{sgn}(\phi(t))|\phi(t)|^\nu, \quad (5)$$

where  $\nu \in (0, 1)$ , in that case, the system (4) is capable of reaching the origin within a finite time  $T(\phi_0) = \frac{|\phi(0)|^{1-\nu}}{\mu(1-\nu)}$ .

**Lemma 3.** [19] If a function  $V(x)$  satisfies that  $V(x) \geq 0$  and  $\dot{V}(x) \leq -kV^q$ , then the system (1) is finite-time stable and the settling time is

$$T(x_0) = \frac{1}{k(1-q)} V^{1-q}(x(0)), \quad (6)$$

where  $k > 0$  and  $0 < q < 1$ .

**Table 1.** Parameters and their meanings for the Furuta pendulum system

| Parameters ( $i = 1, 2$ )              | Meanings                                                                |
|----------------------------------------|-------------------------------------------------------------------------|
| $m_i$ (kg)                             | Masses of rotating arm and pendulum                                     |
| $L_i$ (m)                              | Lengths of rotating arm and pendulum                                    |
| $l_1$ (m)                              | The distance from the origin to the rotating arm's center of mass (COM) |
| $l_2$ (m)                              | The distance from the rod's attachment point to the pendulum's COM      |
| $I_i$ ( $\text{kg} \cdot \text{m}^2$ ) | Moments of inertia                                                      |
| $\theta_i$ (rad)                       | Deflection angles                                                       |
| $\tau$ (Nm)                            | Control torque                                                          |
| $g$ ( $\text{m/s}^2$ )                 | Gravitational acceleration                                              |

### 3 Dynamic Equations of Furuta Pendulum System

Figure 1 presents the physical model of Furuta pendulum system, which comprises rotating arm and pendulum rod. The parameters illustrated in Fig. 1 are defined in Table 1. If the *xoy* plane is chosen as the surface of zero potential energy, it is straightforward to derive the system's kinetic and potential energies as

$$\begin{cases} \mathcal{T} = 0.5 \dot{\theta}_1^2 (m_1 l_1^2 + I_1 + m_2 L_1^2 + m_2 l_2^2 \sin^2 \theta_2) + 0.5 \dot{\theta}_2^2 (m_2 l_2^2 + I_2) \\ \quad - \dot{\theta}_1 \dot{\theta}_2 m_2 L_1 l_2 \cos \theta_2, \\ \mathcal{V} = m_2 g l_2 \cos \theta_2. \end{cases} \quad (7)$$

We choose  $\mathcal{L} = \mathcal{T} - \mathcal{V}$  to be the Lagrangian function. It gives the system's Euler-Lagrange dynamic equations as

$$\begin{cases} J_2 \dot{\theta}_1 \dot{\theta}_2 \sin 2\theta_2 + (J_1 + J_2 \sin^2 \theta_2) \ddot{\theta}_1 + J_3 \dot{\theta}_2^2 \sin \theta_2 - J_3 \ddot{\theta}_2 \cos \theta_2 = \tau + d, \\ -0.5 J_2 \dot{\theta}_1^2 \sin^2 \theta_2 - J_3 \ddot{\theta}_1 \cos(\theta_2) + \dot{\theta}_2 (J_2 + I_2) - J_4 \sin \theta_2 = 0, \end{cases} \quad (8)$$

where  $d$  is an unknown differentiable external disturbance, which satisfies  $\max \{|d|, |\dot{d}|, |\ddot{d}|\} < k_d$ , and  $k_d > 0$ . Moreover, the system (8) can be represented as

$$\mathbf{T}(\boldsymbol{\theta}) \ddot{\boldsymbol{\theta}} + \mathbf{N}(\boldsymbol{\theta}, \dot{\boldsymbol{\theta}}) + \mathbf{G}(\boldsymbol{\theta}) = \mathbf{F}, \quad (9)$$

where  $\boldsymbol{\theta} = [\theta_1, \theta_2]^\top$ , and

$$\mathbf{T}(\boldsymbol{\theta}) = \begin{bmatrix} J_1 + J_2 \sin^2 \theta_2 & -J_3 \cos \theta_2 \\ -J_3 \cos \theta_2 & J_2 + I_2 \end{bmatrix}, \quad \mathbf{N}(\boldsymbol{\theta}, \dot{\boldsymbol{\theta}}) = \begin{bmatrix} N_1 \\ N_2 \end{bmatrix},$$

$$\mathbf{G}(\boldsymbol{\theta}) = \begin{bmatrix} 0 \\ -J_4 \sin \theta_2 \end{bmatrix}, \quad \mathbf{F}(\boldsymbol{\theta}) = \begin{bmatrix} \tau + d \\ 0 \end{bmatrix}, \quad \begin{cases} N_1 = \dot{\theta}_1 \dot{\theta}_2 J_2 \sin 2\theta_2 + \dot{\theta}_2^2 J_3 \sin \theta_2 \\ N_2 = -\frac{1}{2} \dot{\theta}_1^2 J_2 \sin 2\theta_2 \end{cases},$$

$$J_1 = m_1 l_1^2 + I_1 + m_2 L_1^2, \quad J_2 = m_2 l_2^2, \quad J_3 = m_2 L_1 l_2, \quad J_4 = m_2 g l_2.$$

We choose the state space variables of (9) as

$$\boldsymbol{\delta} = [\delta_1, \delta_2, \delta_3, \delta_4]^\top, \quad \delta_1 = \theta_1, \quad \delta_2 = \theta_2, \quad \delta_3 = \dot{\theta}_1, \quad \delta_4 = \dot{\theta}_2.$$

Then, the state-space form of (9) is

$$\begin{cases} \dot{\delta}_1 = \delta_3, \\ \dot{\delta}_2 = \delta_4, \\ \dot{\delta}_3 = f_1(\boldsymbol{\delta}) + g_1(\boldsymbol{\delta})(\tau + d), \\ \dot{\delta}_4 = f_2(\boldsymbol{\delta}) + g_2(\boldsymbol{\delta})(\tau + d), \end{cases} \quad (10)$$

where

$$\begin{bmatrix} f_1(\boldsymbol{\delta}) \\ f_2(\boldsymbol{\delta}) \end{bmatrix} = -\mathbf{T}^{-1} [\mathbf{N} + \mathbf{G}], \quad \begin{bmatrix} g_1(\boldsymbol{\delta}) \\ g_2(\boldsymbol{\delta}) \end{bmatrix} = \mathbf{T}^{-1} \begin{bmatrix} 1 \\ 0 \end{bmatrix}.$$



For the system (10),  $\delta = \mathbf{0}$  represents an open-loop equilibrium point. Our concerned issue is to design  $\tau$  to robustly stabilize (10) at this equilibrium point.

In a small neighborhood of  $\delta = \mathbf{0}$ , the nonlinear system (10) can be approximately linearized into a system of the following form

$$\dot{\delta} = \mathbf{A}\delta + \mathbf{B}(\tau + d), \quad (11)$$

where

$$\mathbf{A} = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & a_1 & 0 & 0 \\ 0 & a_2 & 0 & 0 \end{bmatrix}, \mathbf{B} = \begin{bmatrix} 0 \\ 0 \\ b_1 \\ b_2 \end{bmatrix}, a_1 = \frac{J_3 J_4}{J_1(J_2 + I_2) - J_3^2},$$

$$a_2 = \frac{J_1 J_4}{J_1(J_2 + I_2) - J_3^2}, b_1 = \frac{J_2 + I_2}{J_1(J_2 + I_2) - J_3^2}, b_2 = \frac{J_3}{J_1(J_2 + I_2) - J_3^2}.$$

Based on the characteristic polynomial of  $\mathbf{A}$ , we can construct the following two matrices.

$$\mathbf{\Xi} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ -A_2 & 0 & 1 & 0 \\ 0 & -A_2 & 0 & 1 \end{bmatrix}, \quad \mathbf{\Gamma} = [\mathbf{A}^3 \mathbf{B}, \mathbf{A}^2 \mathbf{B}, \mathbf{A} \mathbf{B}, \mathbf{B}] \mathbf{\Xi}.$$

A simple calculation yields

$$\mathbf{\Psi}_1 = \mathbf{\Gamma}^{-1} \mathbf{A} \mathbf{\Gamma} = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & A_2 & 0 \end{bmatrix}, \quad \mathbf{\Psi}_2 = \mathbf{\Gamma}^{-1} \mathbf{B} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix}. \quad (12)$$

We apply a coordinate transformation to the system (11) defined by  $\mathbf{\Pi} = \mathbf{\Gamma}^{-1} \delta$ , where  $\mathbf{\Pi} = [\Pi_1, \Pi_2, \Pi_3, \Pi_4]^\top$ . Combining (11) with (12) yields

$$\dot{\mathbf{\Pi}} = \mathbf{\Psi}_1 \mathbf{\Pi} + \mathbf{\Psi}_2(\tau + d), \quad (13)$$

that is,

$$\begin{cases} \dot{\Pi}_1 = \Pi_2, \\ \dot{\Pi}_2 = \Pi_3, \\ \dot{\Pi}_3 = \Pi_4, \\ \dot{\Pi}_4 = A_2 \Pi_3 + \tau + d. \end{cases} \quad (14)$$

It is clear that  $\delta = \mathbf{0}$  and  $\mathbf{\Pi} = \mathbf{0}$  are equivalent. Therefore, we just need to design a controller  $\tau$  that robustly stabilizes (14) at  $\mathbf{\Pi} = \mathbf{0}$  now.

## 4 Design of Disturbance Observer and Robust Controller

### 4.1 Finite-Time Disturbance Observer

Note that the system (14) has an unknown external disturbance  $d$ . To eliminate its effect on the control system, we use the fourth equation of (14) and the Levant differentiator method to design a disturbance observer as

$$\begin{cases} \dot{\hat{\Pi}}_4 = A_2 \Pi_3 + \tau + v_0, \\ \dot{\hat{d}} = v_1, \\ \dot{\hat{d}} = v_2, \end{cases} \quad (15)$$

where

$$\begin{cases} v_0 = -\lambda_0 k_d^{\frac{1}{3}} |\hat{\Pi}_4 - \Pi_4|^{\frac{2}{3}} \operatorname{sgn}(\hat{\Pi}_4 - \Pi_4) + \hat{d}, \\ v_1 = -\lambda_1 k_d^{\frac{1}{2}} |\hat{d} - v_0|^{\frac{1}{2}} \operatorname{sgn}(\hat{d} - v_0) + \hat{d}, \\ v_2 = -\lambda_2 k_d \operatorname{sgn}(\hat{d} - v_1), \end{cases} \quad (16)$$

and  $\lambda_i$  ( $i = 0, 1, 2$ ) are positive constants. The estimated values of  $\Pi_4$ ,  $d$ , and  $\dot{d}$  are  $\hat{\Pi}_4$ ,  $\hat{d}$ , and  $\dot{\hat{d}}$  in (15), respectively. By choosing the appropriate parameters  $\lambda_i$ , the disturbance error  $\tilde{d} = d - \hat{d}$  will approach to the origin within a finite time from the analytical results in [20]. Furthermore, Levant differentiator observer shows better robustness and a faster rate of convergence than other disturbance observer design methods [21]. Consequently, an estimated value of  $d$  can be quickly obtained.

### 4.2 Finite-Time Robust Controller

A terminal sliding mode function is designed to be

$$\xi = \dot{\Pi}_4 + \sum_{i=1}^4 \bar{\alpha}_i |\Pi_i|^{\bar{\beta}_i} \operatorname{sgn}(\Pi_i), \quad (17)$$

where  $\bar{\alpha}_i$  and  $\bar{\beta}_i$  ( $i = 1, \dots, 4$ ) satisfy the requirements in Lemma 1 with  $n = 4$ . And an integral sliding function is constructed to be

$$S = \xi + \rho \xi_I, \quad \dot{\xi}_I = |\xi|^\nu \operatorname{sgn}(\xi), \quad (18)$$

where  $\rho$  is a constant greater than zero, and  $0 < \nu < 1$ . Moreover, the initial value of  $\xi_I$  is set to be

$$\xi_I(0) = -\rho^{-1} \xi(0). \quad (19)$$

Combining (18) and (19) gives  $S(0) = 0$ . This denotes that the system (14) arrives at the sliding surface at  $t = 0$ .

**Theorem 1.** *For the chained-form system (14) with matched disturbances  $d$ , if the controller is designed as*

$$\begin{cases} \tau = \tau_{eq} + \tau_{sw}, \\ \tau_{eq} = -A_2 \Pi_3 - \sum_{i=1}^4 \bar{\alpha}_i |\Pi_i|^{\bar{\beta}_i} \text{sgn}(\Pi_i) - \rho \xi_I - \dot{d}, \\ \dot{\tau}_{sw} = -k_d \text{sgn}(S), \end{cases} \quad (20)$$

then it can be stabilized at zero point within a finite time.

*Proof.* Combining (14) and (17) gives

$$\dot{\xi} = A_2 \dot{\Pi}_3 + \dot{\tau} + \dot{d} + \sum_{i=1}^4 \bar{\alpha}_i \bar{\beta}_i \dot{\Pi}_i |\Pi_i|^{\bar{\beta}_i-1} \text{sgn}(\Pi_i). \quad (21)$$

Thus, we have

$$\begin{aligned} \dot{S} &= A_2 \dot{\Pi}_3 + \dot{\tau} + \dot{d} + \sum_{i=1}^4 \bar{\alpha}_i \bar{\beta}_i \dot{\Pi}_i |\Pi_i|^{\bar{\beta}_i-1} \text{sgn}(\Pi_i) + \rho \dot{\xi}_I \\ &= A_2 \dot{\Pi}_3 + \dot{\tau}_{eq} + \dot{\tau}_{sw} + \dot{d} + \sum_{i=1}^4 \bar{\alpha}_i \bar{\beta}_i \dot{\Pi}_i |\Pi_i|^{\bar{\beta}_i-1} \text{sgn}(\Pi_i) + \rho \dot{\xi}_I \\ &= \dot{\tau}_{sw} + \dot{d}. \end{aligned} \quad (22)$$

The Lyapunov function is constructed as  $V = S^2$ . From equation (22), we conclude that

$$\dot{V} = 2S\dot{S} = 2S(\dot{\tau}_{sw} + \dot{d}) = -2k_d|S| + S\dot{d} \leq -k_d V^{\frac{1}{2}} + S\dot{d}. \quad (23)$$

Based on (23) and the above analysis, we can conclude that  $\dot{V} \leq -k_d V^{\frac{1}{2}}$  is achievable within a finite time. From Lemma 3, we get that  $S = 0$  can also be achieved within a finite time. By combining (18) and Lemma 2, it is easy to obtain that  $\xi$  converges to zero quickly. When  $\xi = 0$ , the system (14) becomes

$$\begin{cases} \dot{\Pi}_1 = \Pi_2, \\ \dot{\Pi}_2 = \Pi_3, \\ \dot{\Pi}_3 = \Pi_4, \\ \dot{\Pi}_4 = -\sum_{i=1}^4 \bar{\alpha}_i |\Pi_i|^{\bar{\beta}_i} \text{sgn}(\Pi_i). \end{cases} \quad (24)$$

As a result, Lemma 1 guarantees the finite-time stabilization of (24) at  $\Pi = \mathbf{0}$ . The finite-time robust stabilization of Furuta pendulum system is realized since  $\Pi = \mathbf{0}$  and  $\delta = \mathbf{0}$  are equivalent.

# 5 Numerical Example

**Table 2.** Physical parameters of Furuta pendulum.

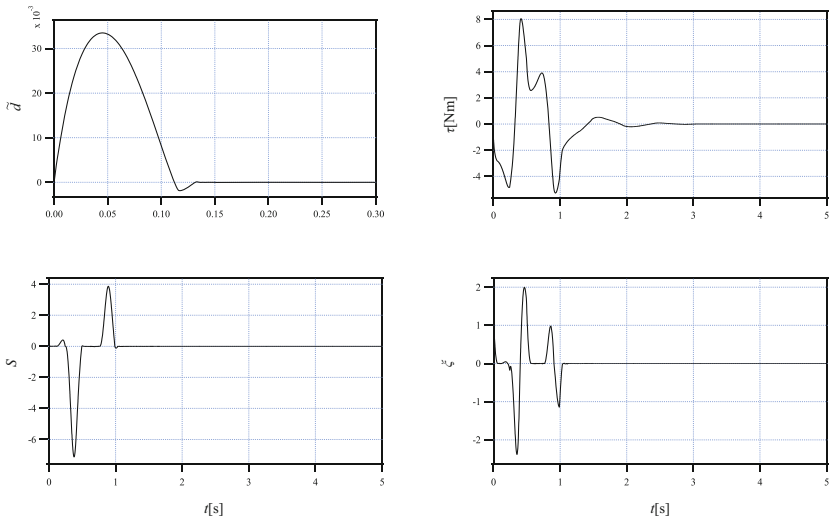
| $M_1$ | $M_2$ | $l_1$ | $l_2$ | $L_1$ | $L_2$ | $I_1$                  | $I_2$                  |
|-------|-------|-------|-------|-------|-------|------------------------|------------------------|
| 0.080 | 0.098 | 0.100 | 0.148 | 0.150 | 0.215 | $3.127 \times 10^{-2}$ | $2.619 \times 10^{-3}$ |

In numerical simulation experiments, the physical and designed parameters we used are shown in the Table 2 and Table 3 [22]. The external disturbance is set to be  $d = \sin(2t)$ . And the initial state of Furuta pendulum is chosen as  $\delta_0 = [\pi/3, \pi/6, 0, 0]^\top$ .

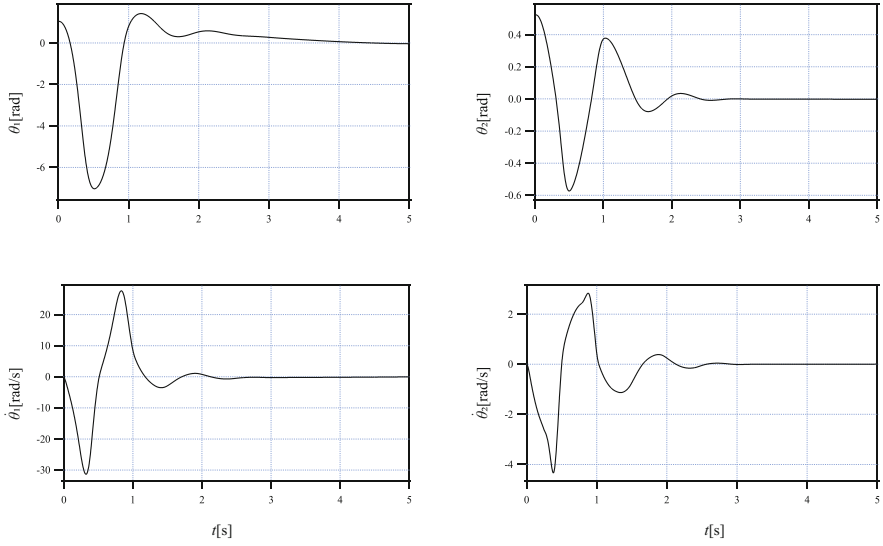
By the conditions given above, the simulation results of  $\tilde{d}$ ,  $\tau$ ,  $S$  and  $\xi$  are shown in the Fig. 2. It is obvious that the disturbance  $d(t)$  can be quickly estimated by the disturbance observer (15). And the estimated error convergence time is less than 0.15 s. Moreover, the simulation show that  $S = 0$  is achieved in 1 s. After that, the terminal sliding mode function  $\xi$  rapidly converges to the origin. In addition, the system's simulation results are given in Fig. 3. Note that the Furuta pendulum system can be stabilized within 5 s. The convergence time is faster than [23]. This proves the superiority of the method in this study.

**Table 3.** Designed parameters in (15) and (20).

| $\lambda_0$ | $\lambda_1$ | $\lambda_2$ | $k_d$ | $\rho$ | $\nu$ | $\bar{\alpha}_1$ | $\bar{\beta}_1$ | $\bar{\alpha}_2$ | $\bar{\beta}_2$ | $\bar{\alpha}_3$ | $\bar{\beta}_3$ | $\bar{\alpha}_4$ | $\bar{\beta}_4$ |
|-------------|-------------|-------------|-------|--------|-------|------------------|-----------------|------------------|-----------------|------------------|-----------------|------------------|-----------------|
| 3           | 1.5         | 1           | 19    | 40     | 0.6   | 4                | $\frac{13}{21}$ | 20               | $\frac{13}{19}$ | 15               | $\frac{13}{17}$ | 4                | $\frac{13}{15}$ |



**Fig. 2.** The simulation results of  $\tilde{d}$ ,  $\tau$ ,  $S$  and  $\xi$  with  $\delta_0 = [\pi/3, \pi/6, 0, 0]^\top$ .



**Fig. 3.** The simulation results of  $\theta_1$ ,  $\theta_2$ ,  $\dot{\theta}_1$ ,  $\dot{\theta}_2$  with  $\delta_0 = [\pi/3, \pi/6, 0, 0]^\top$ .

## 6 Conclusions

To solve the robust stabilization of Furuta pendulum system with unknown matched disturbance, a new control method based on disturbance observer and integral sliding mode approach was developed. Through approximate linearization and coordinate transformation, the origin model of Furuta pendulum was firstly converted into a standard chained form system around the zero equilibrium point. Next, the unknown disturbance was estimated within finite time using a Levant differentiator disturbance observer. Following this, a robust stabilizing controller was designed by utilizing an integral sliding mode approach and the estimated value of disturbance. It robustly stabilized Furuta pendulum system at the zero equilibrium point within a finite time.

**Acknowledgments.** This study was partially supported by the Taishan Scholar Project (No. tsqn202211240) and the Natural Science Foundation Program of Shandong Province (No. ZR2024MF041).

**Disclosure of Interests.** The authors explicitly declared that they had nopotential conflicts of interest.

## References

1. He, B., Wang, S., Liu, Y.: Underactuated robotics: a review. *Int. J. Adv. Rob. Syst.* **16**(4), 1729881419862164 (2019)
2. Lu, B., Fang, Y., Sun, N.: Continuous sliding mode control strategy for a class of nonlinear underactuated systems. *IEEE Trans. Autom. Control* **63**(10), 3471–3478 (2018)
3. Zhang, A., Lai, X., Wu, M., She, J.: Nonlinear stabilizing control for a class of underactuated mechanical systems with multi degree of freedoms. *Nonlinear Dyn.* **89**(3), 2241–2253 (2017). <https://doi.org/10.1007/s11071-017-3582-2>
4. Liu, P., Huda, M.N., Sun, L., Yu, H.: A survey on underactuated robotic systems: bio-inspiration, trajectory planning and control. *Mechatronics* **72**, 102443 (2020)
5. Furuta, K., Yamakita, M., Kobayashi, S., Nishimura, M.: A new inverted pendulum apparatus for education. *IFAC Proc. Vol.* **25**(12), 133–138 (1991)
6. Song, H.: Research on control method of rotating inverted pendulum. *Acad. J. Eng. Technol. Sci.* **5**(9), 37–43 (2022)
7. Nguyen, C.X., Tran, T.T., Pham, T.X., Le, K.M.: Design of control laws for rotary inverted pendulum based on LQR and Lyapunov function. *IOP Conf. Ser. Mater. Sci. Eng.* **1029**(1), 012021 (2021)
8. Kim, D., Nguyen, N.P., Moon, J.: Neural-network based swing-up and stabilization control of rotary inverted pendulum systems. *IFAC PapersOnLine* **56**(2), 6304–6307 (2023)
9. Furuta, K., Pan, Y.: Variable structure control of continuous-time system with sliding sector. *IFAC Proc. Vol.* **32**(2), 2564–2569 (1999)
10. Edwards, C., Spurgeon, S.K.: *Sliding Mode Control: Theory And Applications*. CRC Press (1998)
11. Gambhire, S.J., Kishore, D.R., Londhe, P.S., Pawar, S.N.: Review of sliding mode based control techniques for control system applications. *Int. J. Dyn. Control* **9**(1), 363–378 (2020). <https://doi.org/10.1007/s40435-020-00638-7>
12. Feng, Y., Yu, X., Han, F.: On nonsingular terminal sliding-mode control of nonlinear systems. *Automatica* **49**(6), 1715–1722 (2013)
13. Xu, J., Niu, Y., Lim, C.C., Shi, P.: Memory output-feedback integral sliding mode control for Furuta pendulum systems. *IEEE Trans. Circuits Syst. I Regul. Pap.* **67**(6), 2042–2052 (2020)
14. Rubagotti, M., Estrada, A., Castaños, F., Ferrara, A., Fridman, L.: Integral sliding mode control for nonlinear systems with matched and unmatched perturbations. *IEEE Trans. Autom. Control* **56**(11), 2699–2704 (2011)
15. Bhat, S.P., Bernstein, D.S.: Finite-time stability of continuous autonomous systems. *SIAM J. Control. Optim.* **38**(3), 751–766 (2000)
16. Feng, Y., Han, F., Yu, X.: Chattering free full-order sliding-mode control. *Automatica* **50**(4), 1310–1314 (2014)
17. Chiu, C.S.: Derivative and integral terminal sliding mode control for a class of MIMO nonlinear systems. *Automatica* **48**(2), 316–326 (2012)
18. Shao, K., Zheng, J., Huang, K., Wang, H., Man, Z., Fu, M.: Finite-time control of a linear motor positioner using adaptive recursive terminal sliding mode. *IEEE Trans. Industr. Electron.* **67**(8), 6659–6668 (2020)
19. Huang, X., Lin, W., Yang, B.: Global finite-time stabilization of a class of uncertain nonlinear systems. *Automatica* **41**(5), 881–888 (2005)
20. Borkar, A., Patil, P.M.: Non-linear state feedback control for uncertain systems using a finite time disturbance observer. *Int. J. Dyn. Control* **10**(2), 427–434 (2022)

21. Shtessel, Y.B., Shkolnikov, I.A., Levant, A.: Smooth second-order sliding modes: missile guidance application. *Automatica* **43**(8), 1470–1476 (2007)
22. Zhao, L., Gong, S., Zhang, A., Cong, L.: Nonlinear stabilization control of Furuta pendulum only using angle position measurements. *J. Math. Comput. Sci.* **16**(4), 452–460 (2016)
23. Rudra, S., et al.: Stabilization of Furuta Pendulum: a backstepping based hierarchical sliding Mode approach with disturbance estimation. In: 2013 7th International Conference on Intelligent Systems and Control (ISCO), pp. 99–105. IEEE (2013)



# Method of Heading Estimation Based on Underwater Acoustic Sensors

Qinglin Wen, Chun Jia<sup>(✉)</sup>, and Jiachang Zhang

College of Intelligent System Science and Engineering, Harbin Engineering University,  
Harbin 150001, China  
jiachun@hrbeu.edu.cn

**Abstract.** Heading information is key navigation information for underwater vehicles to perform missions successfully. Conventional underwater heading sensors, such as inertial measurement unit can measure the vehicle's heading information, but its heading error will be accumulated over time; magnetic compass is susceptible to interference from external magnetic fields, resulting in a lack of effective underwater heading estimation methods. An underwater heading estimation method based on dual-beacon acoustic sensors is proposed in this paper. The dual beacons are deployed on the underwater vehicle with a fixed baseline, and the heading is estimated using the geometric relationship of the dual-beacon design. Considering that many factors may impact heading accuracy in terms of positioning accuracy, depth, and baseline, a multi-dimensional simulation test is designed to quantitatively analyze these effects. The results show that different positioning accuracy, depths, and baselines influence heading accuracy to varying degrees. When the baseline length is 10 m, a change in positioning accuracy from 0.1%D to 2% D leads to nearly a tenfold decrease in heading accuracy.

**Keywords:** Underwater heading estimation · Acoustic sensor · Ultra-short baseline positioning

## 1 Introduction

Heading information is crucial for underwater vehicles to navigate safely and accurately to designated operational locations, directly impacting mission success [1]. Existing conventional underwater heading estimation methods are mainly based on the following sensors. The underwater vehicles are typically equipped with an Inertial Measurement Unit (IMU) to provide heading information. IMU costs directly influence its accuracy. High-precision IMUs are costly, and standalone IMUs suffer from cumulative errors over time, limiting their ability to provide accurate long-duration heading information. They often need to be integrated with other heading sensors to maintain heading accuracy, which increases both cost and system complexity [2]. Magnetic compasses are also widely used for underwater heading estimation, as they obtain heading information based on the geomagnetic field. However, magnetic compasses are susceptible to geomagnetic instability and external magnetic interference, necessitating error modeling and compensation before use. In high-latitude regions, significant variations in magnetic field



direction can lead to decreased heading accuracy, requiring additional magnetic calibration [3]. Doppler Velocity Logs (DVLs) determine the velocity of underwater vehicles by measuring the Doppler shift between transmitted and received signals [4]. By integrating the vehicle's motion state, DVLs can accurately estimate heading. However, DVL speed accuracy is affected by various factors, such as water currents, temperature, depth, and salinity, which in turn impacts heading accuracy. In strong current environments, DVLs require error compensation for water flow [5]. As ocean exploration intensifies, underwater scenarios are becoming increasingly complex and variable. High-accuracy, long-duration heading estimation capabilities are needed for underwater heading, a need yet to be fully met.

Common acoustic navigation sensors include Long Baseline (LBL), Short Baseline (SBL), and Ultra-Short Baseline (USBL) systems, all of which provide accurate positioning information for underwater vehicles [6]. As a real-time underwater sensor, USBL offers the advantages of compact size and easy deployment, and methods to improve USBL positioning accuracy have been widely studied [7]. This paper proposes an underwater heading estimation method based on a dual-beacon USBL acoustic positioning system. By utilizing the geometric relationship between the two USBL beacons, heading information can be determined for the underwater vehicle. Furthermore, underwater vehicles typically utilize a USBL positioning system to obtain positioning information, which typically provides only positional information, the heading estimation method proposed in this paper requires only the addition of a low-cost beacon to provide extra heading estimation capabilities.

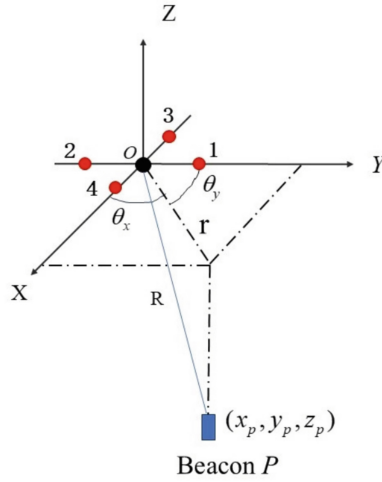
The subsequent arrangement of this paper is as follows: First, the underwater USBL dual-beacon heading estimation method is presented in detail. Then, the error characteristics of USBL dual-beacon heading estimation are analyzed. A simulation model is used to comprehensively analyze the effects of USBL positioning accuracy, depth, and baseline length on heading accuracy. Finally, conclusions are drawn.

## 2 Heading Estimation Method

Taking an acoustic array containing four acoustic elements (1, 2, 3, 4) as an example, the principle of the USBL positioning based on slant range and the azimuth is given in Fig. 1. For Fig. 1,  $\theta_x$  denotes the angle between the direction of sound wave propagation and the X-axis, and  $\theta_y$  denotes the angle between the direction of sound wave propagation and the Y-axis. Considering that the acoustic signals are shot into the acoustic elements in parallel,  $\cos \theta_x = \varphi_{3,4}\lambda/2\pi d$  and  $\cos \theta_y = \varphi_{1,2}\lambda/2\pi d$  can be obtained from the geometric relationship. The baseline between acoustic elements 1 and 2 is  $d$ .  $\lambda$  is the wavelength of the acoustic signals.  $\varphi_{1,2}$  denotes the phase difference of the underwater acoustic signal received by two acoustic elements on the X-axis.  $\varphi_{3,4}$  denotes the phase difference of the underwater acoustic signal received by two acoustic elements on the Y-axis. Then, the position of beacon  $P$  is calculated:

$$X_P = \begin{bmatrix} x_p \\ y_p \\ z_p \end{bmatrix} = \begin{bmatrix} R \cos \theta_x \\ R \cos \theta_y \\ R \sqrt{1 - \cos^2 \theta_x - \cos^2 \theta_y} \end{bmatrix} \quad (1)$$

where  $R$  represents the distance from the array center to the beacon, also known as the slant range.



**Fig. 1.** Diagram of the USBL positioning based on slant range and azimuth.

As shown in Fig. 2, using USBL positioning technology, the specific positions of the dual beacons in the array coordinate system can be calculated. This allows for the estimation of the baseline vector between the USBL dual beacons, and consequently, the heading information of the underwater vehicle in the navigation coordinate system. The positions of beacons 1 and 2 are denoted as  $(x_1, y_1)$  and  $(x_2, y_2)$ , respectively. By the geometric projection, the heading information of the underwater vehicle can be solved in the two-dimensional plane XOY coordinate system, which is obtained by the following equation:

$$\theta = \arctan \frac{y_2 - y_1}{x_2 - x_1} \quad (2)$$

where  $\theta$  is the heading of the underwater vehicle.

Combined with Fig. 2 and Eq. (2), the heading accuracy of the proposed method is determined by the horizontal positioning accuracy. In actual measurements, phase difference ambiguities cannot typically be ignored, making azimuth estimation error one of the primary factors affecting USBL positioning accuracy. The operating depth of the underwater vehicle (hereafter referred to as depth) also influences heading accuracy. As the depth between the acoustic array and the underwater beacon increases, the slant range grows, which impacts horizontal positioning error due to the increased slant range, ultimately affecting heading accuracy. In summary, USBL positioning accuracy, depth, and baseline are key factors influencing heading accuracy.

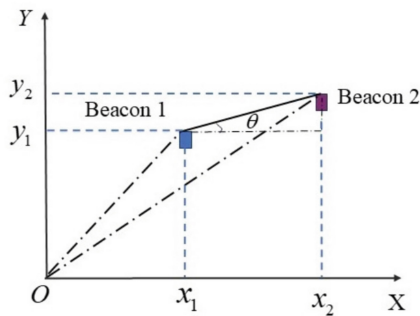


Fig. 2. Dual beacon heading estimation principle.

3 Experiment and Analysis

In this study, numerical simulations were conducted to verify the feasibility of the proposed method. To reliably assess the impact of USBL positioning accuracy, depth, and baseline length on heading accuracy, we used simulated data with parameters consistent with actual underwater navigation sensors [8]. In the simulation experiments, a fixed baseline of 10 m was set for the dual-beacon design. Additionally, abnormal noise was introduced into the simulated data to mimic interference effects on acoustic signal propagation in complex underwater environments.

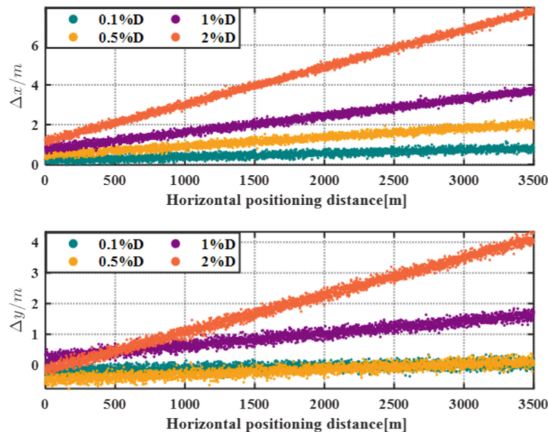
3.1 Analysis of USBL Positioning Accuracy

To analyze the impact of USBL positioning accuracy on heading accuracy, this study maintains a fixed depth for the underwater vehicle, with the depths of the acoustic array and beacon set to 0 m and 2000 m, respectively. Different USBL systems have varying sound velocity probes (SVP) accuracy and azimuth measurement accuracy. In this study, SVP accuracy and azimuth measurement accuracy are set based on different USBL positioning accuracy to simulate the effect of USBL positioning accuracy on heading accuracy [9], as shown in Table 1. It is noted that the root mean square (RMS) is used as the error statistical standard.

Table 1. SVP accuracy and azimuth accuracy corresponding to different USBL positioning accuracy

| USBL positioning accuracy | SVP errors | Azimuth errors |
|---------------------------|------------|----------------|
| 0.1%D                     | 0.2 m/s    | 0.02°          |
| 0.5%D                     | 0.5 m/s    | 0.05°          |
| 1%D                       | 0.8 m/s    | 0.1°           |
| 2%D                       | 1.2 m/s    | 0.25°          |

Figure 3 and Fig. 4 show the relationship between different USBL positioning accuracy and positioning accuracy as well as heading accuracy. Table 2 summarizes the positioning and heading accuracy for different USBL positioning accuracy. The analysis based on Fig. 4 and Table 2 is as follows:



**Fig. 3.** The impact of different USBL positioning accuracy on positioning errors

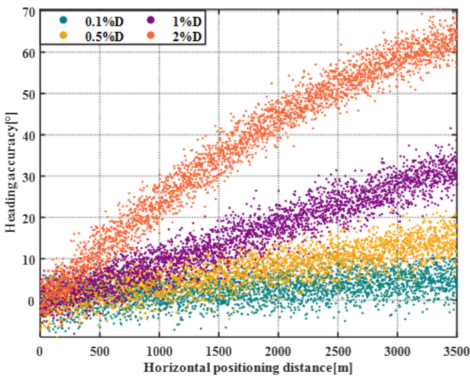
**Table 2.** The impact of different USBL positioning accuracy on heading accuracy

| USBL accuracy | $\Delta x/m$ | $\Delta y/m$ | $\Delta r/m$ | Heading Accuracy/ $^{\circ}$ |
|---------------|--------------|--------------|--------------|------------------------------|
| 0.1%D         | 0.56         | 0.17         | 0.59         | 4.35                         |
| 0.5%D         | 1.34         | 0.29         | 1.37         | 8.94                         |
| 1%D           | 2.39         | 0.43         | 2.43         | 18.57                        |
| 2%D           | 4.82         | 1.57         | 5.07         | 41.44                        |

Based on Fig. 3 and Table 2. The impact of different USBL positioning accuracy on heading accuracy, it can be observed that as the USBL positioning accuracy decreases, the horizontal positioning accuracy also decreases. When the USBL positioning accuracy increases from 0.1%D to 2%D, the positioning error in the X-axis direction increases by a factor of 8, while the positioning error in the Y-axis direction increases from 0.17 m to 1.57 m, a growth of over 800%. The horizontal positioning accuracy decreases nearly ninefold.

Figure 4 shows that as the horizontal positioning distance increases, the heading error gradually increases. The scatter points of heading errors for USBL systems with different accuracy show clear stratification, with a USBL positioning accuracy of 0.1%D offering a tenfold improvement in heading accuracy compared to 2%D. This is because when the baseline length is 10 m, the positioning error in the Y-axis significantly increases as the

USBL positioning accuracy changes from 0.1%D to 2%D. Combining this with Eq. (2), the heading accuracy decreases noticeably.



**Fig. 4.** The relationship between USBL positioning accuracy and heading accuracy

3.2 Analysis of Depth

To simulate the impact of depth on positioning accuracy and heading accuracy, the depth values are varied in the simulation. The relationship between different depths and positioning accuracy is shown in Fig. 5. Table 3 summarizes the positioning accuracy and heading accuracy at different positioning errors. The analysis based on Fig. 5, Fig. 6, and Table 3 is as follows:

The experiment simulates the motion of the underwater vehicle by varying the depth, with continuous movement from 0 to 3500 m at each depth to simulate the vehicle’s movement state. Four different depths were used for the simulation: 1000 m, 2000 m, 3000 m, and 4000 m. The positioning errors at different depths are shown in Fig. 5.

**Table 3.** The effect of depth on positioning accuracy and heading accuracy

| Depth | $\Delta x/m$ | $\Delta y/m$ | $\Delta r/m$ | Heading Accuracy/° |
|-------|--------------|--------------|--------------|--------------------|
| 0.3   | 1.21         | 0.19         | 1.22         | 8.81               |
| 0.6   | 1.35         | 0.22         | 1.37         | 10.17              |
| 1     | 1.52         | 0.34         | 1.56         | 12.95              |
| 1.5   | 1.71         | 0.59         | 1.81         | 16.69              |

As shown in Fig. 5, with the increase in depth, the horizontal positioning accuracy gradually decreases. The impact of depth on positioning accuracy is reflected in both the X and Y axes, with a more noticeable change in the Y axis. Under constant USBL

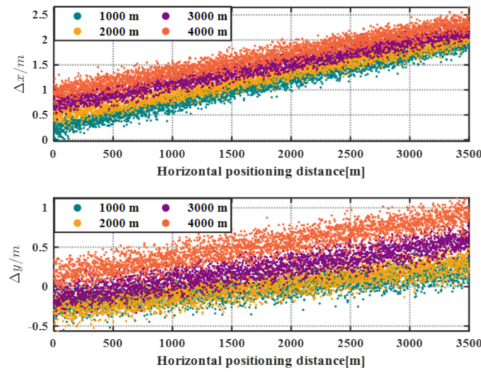


Fig. 5. Positioning errors at four different depths

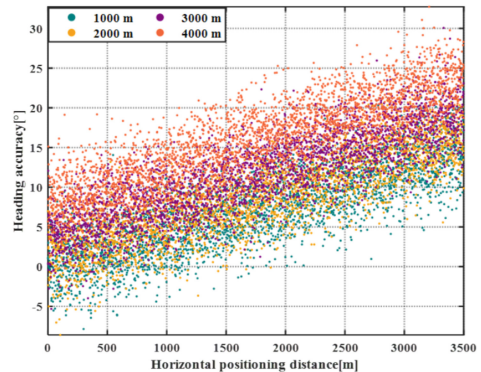


Fig. 6. The effect of different depths on heading accuracy

positioning accuracy, when the depth increases from 1000 m to 4000 m, the positioning error increases from 1.22 m to 1.81 m, an increase of 48.3%. Compared to the influence of USBL accuracy on positioning accuracy, the impact of depth on positioning accuracy is less significant.

As seen in Fig. 6, the scatter plots of heading errors at different depths show a clear stratification. With increasing depth, heading accuracy gradually deteriorates. When the depth increases from 1000 m to 4000 m, a fourfold increase in depth leads to a heading error increasing from  $8.81^\circ$  to  $16.69^\circ$ , nearly doubling the heading error. Therefore, while depth affects heading accuracy, its impact is not as pronounced.

3.3 Analysis of Baseline Length

Figure 7 simulates the trend of heading error with respect to horizontal distance for baseline lengths of 5 m, 10 m, and 20 m. It is evident that as the baseline length increases, the heading accuracy improves significantly. For a baseline length of 5 m, the heading accuracy is  $31.11^\circ$ , while for a baseline length of 20 m, the heading accuracy improves to  $7.92^\circ$ , representing a one-order magnitude improvement compared to the 5 m baseline.

Since the positioning errors on the X-axis and Y-axis remain the same for different baseline lengths, and other simulation parameters are kept constant, the increase in baseline length directly leads to an improvement in heading accuracy, as shown in Eq. (5). Additionally, a pattern can be observed: within a certain range of baseline lengths, and assuming other error factors are negligible, heading accuracy improves in a linear relationship with the increase in baseline length.

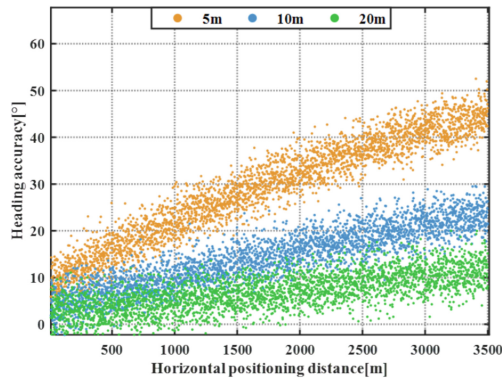


Fig. 7. Impact of different baseline lengths on heading accuracy.

## 4 Conclusion

Heading information is a fundamental prerequisite for the successful completion of tasks by underwater vehicles. To address the limitations of traditional underwater heading estimation methods, this paper proposes a real-time underwater heading estimation method based on a dual-beacon Ultra-Short Baseline (USBL) acoustic positioning system. Based on this method, for underwater vehicles already equipped with a USBL, it is only necessary to add a low-cost beacon to provide additional heading information, which is of significant importance for the navigation of underwater vehicles.

**Acknowledgments.** This study was funded in part by the National Key Research and Development Program (grant number 2021YFB3901300), and in part by the Heilongjiang Provincial Key R&D Plan Project (Innovation Base) (grant number JD2023SJ09).

**Disclosure of Interests.** The authors have no competing interests to declare that are relevant to the content of this article.

## References

1. Zeng, Z., Lyu, C., Bi, Y., Jin, Y., Lu, D., Lian, L.: Review of hybrid aerial underwater vehicle: cross-domain mobility and transitions control. *Ocean Eng.* **248**, 110840 (2022). <https://doi.org/10.1016/j.oceaneng.2022.110840>

2. Jin, K., Chai, H., Su, C., Xiang, M.: A performance-enhanced DVL/SINS integrated navigation system based on data-driven approach. *Meas. Sci. Technol.* **34**(9), 095120 (2023). <https://doi.org/10.1088/1361-6501/acd9e1>
3. Miao, L., Wu, Z., Zhou, Z.: Compass ellipse fitting error compensation method based on improved OPTICS algorithm. *J. Inertial Technol.* **30**(02), 174–180 (2022). <https://doi.org/10.13695/j.cnki.12-1222/o3.2022.02.006>
4. Xu, B., Wang, L., Li, S.: SINS/DVL tight integration navigation method under measured outliers and beam fault conditions. *J. Inertial Technol.* **29**(06), 746–755 (2021). <https://doi.org/10.13695/j.cnki.12-1222/o3.2021.06.007>
5. Wählin, A.K., et al.: Pathways and modification of warm water flowing beneath Thwaites Ice Shelf, West Antarctica. *Sci. Adv.* **7**(15), eabd7254 (2021). <https://doi.org/10.1126/sciadv.abd7254>
6. Yang, F., Zhang, X., Sui, H., Xin, M., Luo, Y., Shi, B.: Robust sequential adaptive Kalman filter algorithm for ultrashort baseline underwater acoustic positioning. *Meas. Sci. Technol.* **34**(3), 035005 (2022). <https://doi.org/10.1088/1361-6501/aca3c5>
7. Wang, Y., et al.: Design and experimental results of passive iUSBL for small AUV navigation. *Ocean Eng.* **248**, 110812 (2022). <https://doi.org/10.1016/j.oceaneng.2022.110812>
8. Zhao, L., Yang, F., Li, L., Ding, J., Zhao, Y.: GBAS ionospheric anomaly monitoring based on a two-step approach. *Sensors* **16**(6), 763 (2016). <https://doi.org/10.3390/s16060763>
9. Jin, B., Xu, X., Zhang, T.: Ultra-short baseline positioning technology and its application in marine engineering. *Navig. Positioning Timing* **5**(04), 8–20 (2018). <https://doi.org/10.19306/j.cnki.2095-8110.2018.04.002>





# Optimal Cooperative Control for Constrained Multi-agent Systems Using Adaptive Dynamic Programming Method

Zijie Guo, Wenshuai Lin, Xiaohong Zheng, and Zilong Zhang<sup>(✉)</sup>

School of Electronics and Information, Guangdong Polytechnic Normal University,  
Guangzhou 510665, China  
zilongzhang@gpnu.edu.cn

**Abstract.** This article proposes an adaptive learning-based optimal cooperative control method for nonlinear multi-agent systems (MASs) subject to output constraints. This constraint is more comprehensive, as the constrained boundaries are not only time-varying and asymmetric but also positive and negative switching. To address asymmetric time-varying output constraints, a system transformation method is designed to obtain equivalent unconstrained MASs utilizing a one-to-one state mapping. Subsequently, a neural network (NN)-based distributed observer is developed to simplify the cooperative control problem into multiple trajectory tracking problems. The proposed distributed observer overcomes the limitation of prior knowledge of leader dynamics, and ensures that the observer errors can converge in a finite time. For optimizing system performance, an adaptive dynamic programming algorithm is developed. Leveraging the powerful learning ability of the critic NNs, the approximate optimal controller for each agent can be obtained. Lyapunov analysis provides a rigorous proof of the boundedness of all signals in the closed-loop systems. The efficacy of the proposed approach can be substantiated through simulation results conducted on multiple single-link robot arm systems.

**Keywords:** Adaptive dynamic programming (ADP) · Asymmetric time-varying constraints · Distributed observer · Multi-agent systems (MASs)

## 1 Introduction

Multi-agent systems (MASs) have attracted considerable attention and become an important concern in the current industrial manufacture [1–4]. The cooperative control of MASs has transitioned from a centralized approach to a distributed one. To satisfy the higher practical requirements of engineering, achieving synchronization between followers and leader while minimizing control costs becomes imperative, giving rise to the optimal cooperative control problem

[5,6]. Adaptive dynamic programming (ADP), which uses neural network (NN) approximation method to efficiently solve the Hamilton-Jacobi-Bellman (HJB) equation, has demonstrated remarkable outcomes in the realm of optimal control for MASs. In [7], a distributed optimal cooperative control problem for nonlinear MASs is solved by an event-triggered ADP algorithm. Afterwards, an exploration in [8] was conducted on the utilization of actor-critic ADP algorithm to achieve the distributed optimal formation control for leader-follower MASs.

It has been demonstrated that distributed architectures are crucial for optimal cooperative control of MASs, as they offer improved stability and faster computational performance [9,10]. Distributed observer presents a novel approach to distributed control, which allows each follower to estimate the leader's state without necessitating actual information exchange among followers. Then, the intricate task of distributed cooperative control is streamlined by transforming it into multiple trajectory tracking control. In [11], a simplified distributed approach was employed to facilitate the formation control of a quadrotor team by utilizing distributed observers. This strategy effectively decomposes the complicated cooperative control problem into an independent position tracking problem for each quadrotor. Following this, [12] and [13] made improvements to distributed observers for achieving optimal synchronization within heterogeneous MASs against unknown leader dynamics. Nevertheless, the convergence rate of observer error is disregarded by these aforementioned methods. Recent studies have placed greater emphasis on addressing the finite-time stability problem to enhance transient performance.

Considering the imperative of ensuring system safety and adhering to physical limitations, many practical systems normally operate under state constraints. Any deviation from these constraints can result in unfavorable tracking performance or potential harm to the controlled system [14]. The barrier Lyapunov function (BLF) has gained significant recognition as an efficient strategy for handling state constraints. A logarithmic BLF was utilized to tackle symmetric state constraints in [15]. The BLF approaches are essentially indirect techniques for managing constraints, necessitating additional effort in selecting initial values for the constraints. To address this concern, an ADP algorithm has incorporated a state transformation approach to design the optimal controller of a nonlinear system subject to output constraints [16]. However, the constraint boundaries are required to be constant. To address a broader spectrum of practical requirements, it is essential to further explore time-varying and asymmetric constraints.

Drawing upon the aforementioned discussion, this paper delves into the adaptive learning-based optimal cooperative control problem for a class of MASs with asymmetric time-varying output constraints. The contributions of the developed adaptive learning-based optimal cooperative control are summarized as follows:

1. A system transformation technique utilizing a one-to-one state mapping is introduced to address asymmetric time-varying output constraints. The transformed equivalent MASs are optimized using the ADP algorithm, which not only optimizes the system performance, but also ensures that the output constraint is not violated.

2. An NN-based distributed observer is proposed for estimating the states of the unknown nonlinear leader. By leveraging the online learning capability of NNs, precise knowledge about the leader dynamics is no longer necessary. Furthermore, convergence of estimation errors towards a small vicinity around zero is achieved within a finite time.

Notation:  $T$  is the transposition symbol.  $\lambda_{\min}(\cdot)$  denotes the minimum eigenvalue and  $\lambda_{\max}(\cdot)$  denotes the maximum eigenvalues.  $\mathbb{R}$  is the Euclidean space.  $\tanh(\cdot)$  is the inverse hyperbolic tangent function.  $\text{sign}(\cdot)$  represents the sign function.  $\text{diag}\{\cdot\}$  is a diagonal matrix.

## 2 Preliminaries

A group of followers is represented by the set  $\Gamma = \{r_1, \dots, r_N\}$ , and all followers exchange information through a weighted undirected topology denoted as  $\Pi = (\Gamma, U, \Upsilon)$ .  $U \subseteq \Gamma \times \Gamma$  is the edge set, if  $(r_i, r_j) \in \Gamma$ , follower  $i$  and  $j$  can exchange information and are considered adjacent to one another. The adjacency matrix  $\Upsilon = [\nu_{ij}] \in \mathbb{R}^{N \times N}$  satisfies  $(r_i, r_j) \in \Gamma \Leftrightarrow \nu_{ij} = \nu_{ji} = 1$ , and  $(r_i, r_j) \notin \Gamma \Leftrightarrow \nu_{ij} = \nu_{ji} = 0$ . For agent  $i$ , we define its neighbor set as  $N_i$ . In addition, the topology graph is simple without any repeated edges or self-loops. The Laplacian matrix  $\Psi = [\varsigma_{ij}] \in \mathbb{R}^{N \times N}$  is symmetric and given by subtracting the adjacency matrix from a diagonal matrix  $\Phi = \text{diag}\{\vartheta_1, \dots, \vartheta_N\}$  where  $\vartheta_i = \sum_{j=1}^N \nu_{ij}$ . To illustrate the information interaction between followers and leader, a diagonal matrix denoted as  $\Xi = \text{diag}\{\lambda_1, \dots, \lambda_N\}$  is introduced, where  $\lambda_i = 1$  if follower  $i$  receives information from the leader; otherwise,  $\lambda_i = 0$ . According to previous research [17], it has been shown that  $\Lambda = \Psi + \Xi$  is positive definite.

## 3 Problem Description

The nonlinear dynamics of the  $i$ th follower is given by

$$\begin{aligned} \dot{p}_i &= v_i \\ \dot{v}_i &= \bar{h}_i(p_i, v_i) + \bar{g}_i(p_i, v_i)u_i \\ y_i &= p_i \end{aligned} \tag{1}$$

where  $i = 1, \dots, N$ .  $p_i = [p_{i1}, \dots, p_{in}]^T \in \mathbb{R}^n$  is the position vector and  $v_i = [v_{i1}, \dots, v_{in}]^T \in \mathbb{R}^n$  represents the velocity vector. The nonlinear continuous differentiable functions  $\bar{h}_i(\cdot) = [h_{i1}(\cdot), \dots, h_{in}(\cdot)]^T \in \mathbb{R}^n$  and nonlinear continuous control gain matrices  $\bar{g}_i(\cdot) = [g_{i1}(\cdot), g_{i2}(\cdot), \dots, g_{in}(\cdot)]^T \in \mathbb{R}^{n \times o}$  are known.  $u_i = [u_{i1}, \dots, u_{io}]^T \in \mathbb{R}^o$  is the input signal and  $y_i \in \mathbb{R}^n$  is the output signal. And the virtual leader is formulated as follows:

$$\dot{p}_0 = v_0, \quad \dot{v}_0 = \bar{h}_0(p_0, v_0) \tag{2}$$

where the reference position  $p_0 = [p_{01}, \dots, p_{0n}]^T \in \mathbb{R}^n$  and reference velocity  $v_0 = [v_{01}, \dots, v_{0n}]^T \in \mathbb{R}^n$ .  $y_0 = p_0 \in \mathbb{R}^n$  denotes the output signal.  $\bar{h}_0(\cdot) = [h_{01}(\cdot), \dots, h_{0n}(\cdot)]^T \in \mathbb{R}^n$  is smooth function vectors.

In this paper, the output signals are subject to  $a_j(t) < y_{ij} < A_j(t)$ ,  $j = 1, \dots, n$ , where  $a_j(t)$  and  $A_j(t)$  are boundary functions. It is noted that only condition  $a_j(t) < A_j(t)$  should be satisfied, thus the constrained boundaries can be asymmetric, time-varying and positive and negative switching. Additionally, it is reasonable to assume that  $A_j(t)$  and  $a_j(t)$  are smooth continuous.

This paper aims to develop an adaptive learning-based distributed optimal control policy. This optimal control policy ensures that all followers synchronize with the leader, while also guaranteeing adherence to the asymmetric time-varying output constraints.

## 4 Optimal Cooperative Controller Design

An equivalent MASs without any constraints on the output is firstly achieved through a system transformation approach. Then, desired trajectories for each follower is generated by distributed observer design, which transforms the cooperative control task to multiple general tracking control. Ultimately, an approximate optimal controller is derived using a single critic NN ADP algorithm.

### 4.1 System Transformation

The one-to-one nonlinear mapping is designed as

$$\xi_{ij} = Q(p_{ij}, A_j(t), a_j(t)), j = 1, \dots, n \quad (3)$$

where  $Q(\cdot) = \text{atanh}\left(\frac{2p_{ij}-A_j(t)-a_j(t)}{A_j(t)-a_j(t)}\right)$ . Let  $\xi_i = [\xi_{i1}, \dots, \xi_{in}]^T \in \mathbb{R}^n$  be the position vector of the transformed system. By differentiating (3), one has

$$\begin{aligned} \dot{\xi}_{ij} &= \frac{\partial Q(\cdot)}{\partial p_{ij}} v_{ij} + \frac{\partial Q(\cdot)}{\partial A_j} \dot{A}_j + \frac{\partial Q(\cdot)}{\partial a_j} \dot{a}_j \triangleq \mu_{ij} \\ \dot{\mu}_{ij} &= \tilde{h}_{ij}(\xi_i, \mu_i) + \tilde{g}_{ij}(\xi_i, \mu_i) u_i \end{aligned} \quad (4)$$

where

$$\begin{aligned} \tilde{h}_{ij}(\xi_i, \mu_i) &= \frac{\partial Q(\cdot)}{\partial p_{ij}} h_{ij}(p_i, q_i) + \left[ \frac{\partial Q(\cdot)}{\partial p_{ij}} \right]^{(1)} v_{ij} + \frac{\partial Q(\cdot)}{\partial A_j} \ddot{A}_j \\ &\quad + \left[ \frac{\partial Q(\cdot)}{\partial A_j} \right]^{(1)} \dot{A}_j + \frac{\partial Q(\cdot)}{\partial a_j} \ddot{a}_j + \left[ \frac{\partial Q(\cdot)}{\partial a_j} \right]^{(1)} \dot{a}_j, \\ \tilde{g}_{ij}(\xi_i, \mu_i) &= \frac{\partial Q(\cdot)}{\partial p_{ij}} g_{ij}(p_i, v_i), \\ \mu_i &= [\mu_{i1}, \dots, \mu_{in}]^T \in \mathbb{R}^n. \end{aligned}$$

Then, an equivalent MASSs using state mapping is formulated as

$$\dot{\chi}_i = F_i + G_i u_i \quad (5)$$

where  $\chi_i = [\xi_i^T, \mu_i^T]^T$ ,  $F_i = \left[ \mu_i^T, \widehat{h}_i^T(\xi_i, \mu_i) \right]^T$  and  $G_i = \left[ 0_{n \times o}^T, \widehat{g}_i^T(\xi_i, \mu_i) \right]^T$ .

And  $\widehat{h}_i(\xi_i, \mu_i) = [\tilde{h}_{i1}(\cdot), \dots, \tilde{h}_{in}(\cdot)]^T$ ,  $\widehat{g}_i(\xi_i, \mu_i) = [\tilde{g}_{i1}(\cdot), \dots, \tilde{g}_{in}(\cdot)]^T$ .

**Assumption 1** [18]:  $\left\| \widehat{g}_i(\xi_i, \mu_i) \right\| \leq G_M$ , where  $G_M$  is a positive constant.

Similarly, define  $\xi_{0j} = Q(p_{0j}, A_j(t), a_j(t))$ , one has

$$\begin{aligned} \dot{\xi}_{0j} &= \frac{\partial Q(\cdot)}{\partial p_{0j}} v_{0j} + \frac{\partial Q(\cdot)}{\partial A_j} \dot{A}_j + \frac{\partial Q(\cdot)}{\partial a_j} \dot{a}_j \triangleq \mu_{0j} \\ \dot{\mu}_{0j} &= \tilde{h}_{0j}(\xi_0, \mu_0) \end{aligned} \quad (6)$$

where  $\tilde{h}_{0j}(\xi_0, \mu_0) = \frac{\partial Q(\cdot)}{\partial p_{0j}} h_{0j}(p_0, v_0) + \left[ \frac{\partial Q(\cdot)}{\partial p_{0j}} \right]^{(1)} v_{0j} + \frac{\partial Q(\cdot)}{\partial a_j} \ddot{a}_j + \left[ \frac{\partial Q(\cdot)}{\partial a_j} \right]^{(1)} \dot{a}_j + \frac{\partial Q(\cdot)}{\partial A_j} \ddot{A}_j + \left[ \frac{\partial Q(\cdot)}{\partial A_j} \right]^{(1)} \dot{A}_j$ .

Letting  $\chi_0 = [\xi_0^T, \mu_0^T]^T$  with  $\xi_0 = [\xi_{01}, \dots, \xi_{0n}]^T$  and  $\mu_0 = [\mu_{01}, \dots, \mu_{0n}]^T$ , the leader is transformed as

$$\dot{\chi}_0 = M_0(\chi_0) \quad (7)$$

where  $M_0(\chi_0) = \left[ \mu_0^T, \widehat{h}_0^T(\xi_0, \mu_0) \right]^T$  with  $\widehat{h}_0(\xi_0, \mu_0) = [\tilde{h}_{01}(\xi_0, \mu_0), \dots, \tilde{h}_{0n}(\xi_0, \mu_0)]^T$ .

## 4.2 NN-Based Distributed Observer Design

In this subsection, we propose an NN-based distributed observer for reconstructing the leader's dynamics. Using NNs,  $M_0(\chi_0)$  is reconstructed as

$$M_0(\chi_0) = w_{m0}^T \varphi(\chi_0) + \iota_0 \quad (8)$$

where  $w_{m0} = [w_{m0}^1, \dots, w_{m0}^n] \in \mathbb{R}^{l_m \times n}$  indicates the ideal weighted matrix with  $w_{m0}^k \in \mathbb{R}^{l_m}$  being the  $k$ th column of  $w_{m0}$ ,  $k = 1, \dots, n$ .  $\varphi(\chi_0) \in \mathbb{R}^{l_m}$  represents the bounded nonlinear activation function, while  $l_m$  signifies the count of hidden neurons.  $\iota_0 \in \mathbb{R}^n$  is approximation error.

The NN-based distributed observer is designed by

$$\dot{s}_i = \hat{w}_{mi}^T \varphi(s_i) - \mathcal{B} \text{sign}(z_i) - \mathcal{D} z_i \quad (9)$$

where  $z_i = \sum_{l \in N_i} \nu_{il}(s_i - s_l) + \lambda_i(s_i - \chi_0)$ ,  $\hat{w}_{mi} = [\hat{w}_{mi}^1, \dots, \hat{w}_{mi}^n] \in \mathbb{R}^{l_m \times n}$  is the estimated weight matrix, and  $\hat{w}_{mi}^k \in \mathbb{R}^{l_m}$  is the  $k$ th column of  $\hat{w}_{mi}$ .  $\mathcal{B}$  and  $\mathcal{D}$  are positive design parameters.

Defining  $s_{ik}$  as the  $k$ th element of the vector  $s_i$ , we have

$$\dot{s}_{ik} = \hat{w}_{mi}^{kT} \varphi(s_i) - \mathcal{B} \text{sign}(z_{ik}) - \mathcal{D} z_{ik} \quad (10)$$

where  $z_{ik}$  is the  $k$ th elements of the  $z_i$ .

**Assumption 2** [12]. For the NN approximation error, it follows that  $\|\iota_0\| \leq \iota_{M0}$ ,  $w_{m0}^k$  is bounded as  $\|w_{m0}^k\| \leq w_{M0}$ . Moreover,  $\varphi(\cdot)$  is bounded as  $\|\varphi(\cdot)\| \leq b_\phi$ , where  $\iota_{M0}$ ,  $w_{M0}$  and  $b_\phi$  are positive parameters.

**Theorem 1.** *If the weights of NNs are adjusted according to  $\dot{\hat{w}}_{mi}^k = -\alpha_{mi} \frac{\varphi(s_i) z_{ik}}{1+z_{ik}^2} - \beta_{mi} \hat{w}_{mi}^k$ , where  $\alpha_{mi} > 0$ ,  $\beta_{mi} > 0$ , and the design parameter  $\mathcal{B} > 0$ ,  $\mathcal{D} > \frac{b_\phi^2}{\lambda_{\min}(\Lambda\Lambda)} + 1$ ,  $b_\phi$  denotes the Lipschitz constant associated with  $\phi_0(\cdot)$ . Then, the weight estimation error  $\tilde{w}_{mi} = w_{m0} - \hat{w}_{mi}$  and state estimation error  $e_i = s_i - \chi_0$  achieve semi-global practical finite-time stable.*

**Proof.** The proof process is omitted here, which is similar to the literature [19].

### 4.3 Distributed Optimal Controller Design

Establish the tracking error as  $\eta_i = \chi_i - s_i$ . The derivative of  $\eta_i$  is

$$\dot{\eta}_i = F_i - \hat{w}_{mi}^T \varphi(s_i) + G_i u_i + \varpi_i \quad (11)$$

where  $\varpi_i = \mathcal{B} \text{sign}(z_i) + \mathcal{D} z_i$ .

Define  $X_i = [\eta_i^T, s_i^T]^T$ , we can construct the following augmented system:

$$\dot{X}_i = P_i(X_i) + U_i(X_i) u_i + W_i \quad (12)$$

where  $P_i(X_i) = [(F_i - \hat{w}_{mi}^T \varphi(s_i))^T, \varphi^T(s_i) \hat{w}_{mi}]^T$ ,  $U_i(X_i) = [G_i^T, 0_{n \times o}]^T$  and  $W_i = [\varpi_i^T, -\varpi_i^T]^T$ .

Then, we construct the following cost function:

$$V_i(X_i) = \int_t^\infty e^{-\alpha_{Vi}(\tau-t)} (X_i^T \bar{Y}_i X_i + u_i^T R_i u_i) d\tau \quad (13)$$

where  $\alpha_{Vi} > 0$ ,  $\bar{Y}_i = \text{diag}\{Y_i, 0_{2n \times 2n}\}$ ,  $Y_i \in \mathbb{R}^{2n \times 2n}$ ,  $R_i \in \mathbb{R}^{o \times o}$ .

According to (13), Hamiltonian is formulated as

$$\begin{aligned} H(X_i, u_i, \nabla V_i) &= X_i^T \bar{Y}_i X_i + u_i^T R_i u_i + \nabla V_i^T(X_i) (P_i(X_i) \\ &\quad + U_i(X_i) u_i + W_i) - \alpha_{Vi} V_i(X_i) \end{aligned} \quad (14)$$

Using  $\frac{\partial H}{\partial u_i} = 0$ , the optimal control  $u_i^*(X_i)$  is derived as

$$u_i^*(X_i) = -\frac{1}{2} R_i^{-1} U_i^T(X_i) \nabla V_i^*(X_i) \quad (15)$$

Substituting (15) into (14), the HJB equation is given by

$$\begin{aligned} H(X_i, u_i^*, \nabla V_i^*) &= X_i^T \bar{Y}_i X_i + (\nabla V_i^*(X_i))^T (P_i(X_i) \\ &\quad + W_i) - u_i^{*T} R_i u_i^* - \alpha_{Vi} V_i^*(X_i) = 0 \end{aligned} \quad (16)$$

To solve this HJB equation, a single critic NN ADP algorithm is developed in next subsection.

#### 4.4 Single Critic NN ADP Algorithm

By leveraging the universal approximation property exhibited by NNs, we can express  $V_i^*(X_i)$  as  $V_i^*(X_i) = w_{ci}^T \delta(X_i) + \varsigma(X_i)$ . Here,  $w_{ci} \in \mathbb{R}^{n_{ci}}$  represents the optimal weight vector,  $\delta(X_i) \in \mathbb{R}^{n_{ci}}$  denotes the activation function,  $n_{ci}$  corresponds to the number of neurons.  $\varsigma(X_i) \in \mathbb{R}$  is approximation error.

Then,  $\nabla V_i^*(X_i)$  is represented as

$$\nabla V_i^*(X_i) = (\nabla \delta(X_i))^T w_{ci} + \nabla \varsigma(X_i) \quad (17)$$

where  $\nabla \delta(X_i) = \frac{\partial \delta(X_i)}{\partial X_i}$  and  $\nabla \varsigma(X_i) = \frac{\partial \varsigma(X_i)}{\partial X_i}$ . Since the ideal weight vector is not available, the estimated cost function is formulated as  $\hat{V}_i^*(X_i) = \hat{w}_{ci}^T \delta(X_i)$ , where  $\hat{w}_{ci}$  stands for the estimated value of the ideal weight vector. Then,  $\nabla \hat{V}_i^*(X_i) = \hat{w}_{ci}^T \nabla \delta(X_i)$ .

Substituting (17) into (15), the optimal controller is given by

$$u_i^* = -\frac{1}{2} R_i^{-1} U_i^T(X_i) ((\nabla \delta(X_i))^T w_{ci} + \nabla \varsigma(X_i)) \quad (18)$$

The, the approximate optimal control policy is

$$\hat{u}_i = -\frac{1}{2} R_i^{-1} U_i^T(X_i) (\nabla \delta(X_i))^T \hat{w}_{ci} \quad (19)$$

A Bellman error is defined as  $\varepsilon_{ci} = X_i^T \bar{Y}_i X_i + \hat{u}_i^T R_i \hat{u}_i - \alpha_{Vi} \hat{w}_{ci}^T \delta(X_i) + \hat{w}_{ci}^T \nabla \delta(X_i) (P_i(X_i) + U_i(X_i) \hat{u}_i + W_i)$ . Since  $H_i(X_i, u_i^*, \nabla V_i^*(X_i)) = 0$ , one has

$$\varepsilon_{ci} = -\tilde{w}_{ci}^T \kappa_i - \varepsilon_{Hi} \quad (20)$$

where  $\kappa_i = -\alpha_{Vi} \delta(X_i) + \nabla \delta(X_i) (P_i(X_i) + U_i(X_i) \hat{u}_i + W_i)$ . The weight estimation error is defined as  $\tilde{w}_{ci} = w_{ci} - \hat{w}_{ci}$ . Furthermore,  $\varepsilon_{Hi}$  denotes the reconstruction error, which is obtained using methods similar to those presented in [20].

To ensure precise estimation of optimal weights by training the critic NNs, a learning rule for  $\hat{w}_{ci}$  is designed as

$$\dot{\hat{w}}_{ci} = -\frac{\alpha_{ci} \kappa_i}{(1 + \kappa_i^T \kappa_i)^2} \varepsilon_{ci} \quad (21)$$

where  $\alpha_{ci}$  is design learning rate.

Based on (20)–(21),  $\tilde{w}_{ci}$  is updated by

$$\dot{\tilde{w}}_{ci} = -\frac{\alpha_{ci} \kappa_i}{(1 + \kappa_i^T \kappa_i)^2} (\kappa_i^T \tilde{w}_{ci} + \varepsilon_{Hi}) \quad (22)$$

#### 4.5 Stability Analysis

**Assumption 3** [21].  $w_{ci}$  is bounded by a positive constant  $\rho_w$ , such that  $\|w_{ci}\| \leq \rho_w$ .  $\varsigma(X_i)$  and  $\nabla \varsigma(X_i)$  are bounded, with positive constants  $\rho_\varsigma$  and  $\rho_{\nabla \varsigma}$  respectively satisfying  $\|\varsigma(X_i)\| \leq \rho_\varsigma$  and  $\|\nabla \varsigma(X_i)\| \leq \rho_{\nabla \varsigma}$ . There is a constant  $\rho_\delta > 0$  such that  $\|\nabla \delta(X_i)\| \leq \rho_\delta$ .  $\varepsilon_{Hi}$  is bounded, such that  $\|\varepsilon_{Hi}\| \leq E_M$ , where  $E_M$  is a positive constant.

**Theorem 2.** Consider the given augmented system (12) and the cost function (13), the approximate optimal control strategy is designed as (19). Additionally, the update rule for critic NN weights is given by (21), then the weight approximation error  $\tilde{w}_{ci}$  and the tracking error  $\eta_i$  are uniformly ultimately bounded (UUB).

**Proof.** Choose the Lyapunov function as  $L_i = V_i^*(X_i) + \frac{1}{2}\tilde{w}_{ci}^T\tilde{w}_{ci}$ . Take the time derivative of  $L_i$  as

$$\begin{aligned}\dot{L}_i &= \nabla V_i^*(X_i)(P_i(X_i) + U_i(X_i)\hat{u}_i + W_i) - \alpha_{V_i}V_i^*(X_i) \\ &\quad - \tilde{w}_{ci}^T \frac{\alpha_{ci}\kappa_i\kappa_i^T}{(1 + \kappa_i^T\kappa_i)^2}\tilde{w}_{ci} - \tilde{w}_{ci}^T \frac{\alpha_{ci}\kappa_i}{(1 + \kappa_i^T\kappa_i)^2}\varepsilon_{Hi}\end{aligned}\quad (23)$$

According to (16), one has

$$(\nabla V_i^*(X_i))^T(P_i(X_i) + W_i) - \alpha_{V_i}V_i^*(X_i) = -X_i^T\bar{Y}_iX_i + u_i^{*T}R_iu_i^* \quad (24)$$

Substituting (24) into (23) and Applying Young's inequality, we have

$$\begin{aligned}\dot{L}_i &\leq -X_i^T\bar{Y}_iX_i - u_i^{*T}R_iu_i^* - 2u_i^{*T}R_i(\hat{u}_i - u_i^*) \\ &\quad - \frac{\alpha_{ci}}{2}\lambda_{\min}\left(\frac{\kappa_i\kappa_i^T}{(1 + \kappa_i^T\kappa_i)^2}\right)\|\tilde{w}_{ci}\|^2 + \frac{\alpha_{ci}E_M^2}{2}\end{aligned}\quad (25)$$

Applying  $2a^Tb \leq \|a\|^2 + \|b\|^2$ ,  $a, b \in \mathbb{R}^n$ , it follows

$$-2u_i^{*T}R_i(\hat{u}_i - u_i^*) \leq u_i^{*T}R_iu_i^* + \frac{\lambda_{\max}(R_i)}{2}G_M^2\rho_\delta^2\|\tilde{w}_{ci}\|^2 + \frac{\lambda_{\max}(R_i)}{2}G_M^2\rho_{\nabla\varsigma}^2 \quad (26)$$

Substituting (26) into (25) has

$$\begin{aligned}\dot{L}_i &\leq -\lambda_{\min}(Y_i)\|\eta_i\|^2 - \left(\frac{\alpha_{ci}}{2}\lambda_{\min}\left(\frac{\kappa_i\kappa_i^T}{(1 + \kappa_i^T\kappa_i)^2}\right) - \frac{\lambda_{\max}(R_i)}{2}G_M^2\rho_\delta^2\right)\|\tilde{w}_{ci}\|^2 \\ &\quad + \frac{\lambda_{\max}(R_i)}{2}G_M^2\rho_{\nabla\varsigma}^2 + \frac{\alpha_{ci}E_M^2}{2}\end{aligned}\quad (27)$$

Choose appropriate  $\alpha_{ci}$  to ensure that  $\frac{\alpha_{ci}}{2}\lambda_{\min}\left(\frac{\kappa_i\kappa_i^T}{(1 + \kappa_i^T\kappa_i)^2}\right) - \frac{\lambda_{\max}(R_i)}{2}G_M^2\rho_\delta^2 > 0$ , then if the following condition holds:

$$\|\tilde{w}_{ci}\| > \sqrt{\frac{\alpha_{ci}E_M^2 + \lambda_{\max}(R_i)G_M^2\rho_{\nabla\varsigma}^2}{\alpha_{ci}\lambda_{\min}\left(\frac{\kappa_i\kappa_i^T}{(1 + \kappa_i^T\kappa_i)^2}\right) - \lambda_{\max}(R_i)G_M^2\rho_\delta^2}} \quad (28)$$

or

$$\|\eta_i\| > \sqrt{\frac{\alpha_{ci}E_M^2 + \lambda_{\max}(R_i)G_M^2\rho_{\nabla\varsigma}^2}{2\lambda_{\min}(Y_i)}} \quad (29)$$

we have  $\dot{L}_{1i} < 0$ , which means that all signals in closed-loop system are UUB.



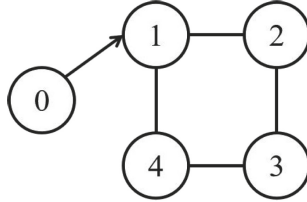
## 5 Simulation Results

The efficacy of the developed adaptive learning-based distributed optimal control policy is validated in this section using a MAS comprising four single-link robot arms. According to [22], each robot arm is described as

$$\begin{aligned}\dot{p}_i &= v_i \\ \dot{v}_i &= -4.905 \sin(p_i) - 0.2v_i + 0.1u_i \\ y_i &= p_i\end{aligned}\tag{30}$$

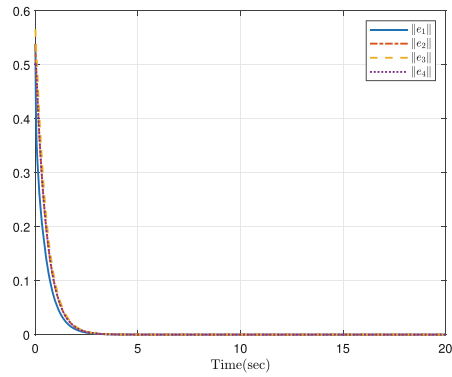
The initial states of each robotic arm are  $p_1(0) = 0.04$ ,  $v_1(0) = -0.06$ ,  $p_2(0) = 0.15$ ,  $v_2(0) = -0.05$ ,  $p_3(0) = 0.05$ ,  $v_3(0) = -0.05$  and  $p_4(0) = 0.08$ ,  $v_4(0) = -0.01$ . Figure 1 shows the communication connections between these single-link robot arms. The followers are denoted as 1, 2, 3, 4, while the leader is denoted as 0. Additionally, we describe the dynamics of the leader using  $\dot{p}_0 = v_0$ ,  $\dot{v}_0 = -p_0$ , and  $y_0 = p_0$ , where  $p_0(0) = 0$  and  $v_0(0) = 0.2$ . Furthermore, it is required that for each single-link robot arm, its output signal must satisfy  $0.2 \sin(t) - 0.1 - 0.1e^{-t} < y_i < 0.2 \sin(t) + 0.1 + 0.1e^{-t}$ .

Let the design parameters of distributed observer be  $\mathcal{B} = 0.1$ ,  $\mathcal{D} = 10$ ,  $\alpha_{mi} = 1$  and  $\beta_{mi} = 2$ . The parameters of cost function is selected as  $\alpha_{Vi} = 0.5$ ,  $Y_i = \text{diag}\{500, 500\}$  and  $R_i = 1$ . To obtain better learning effect of critic NNs, we set  $\alpha_{ci} = 2$ . And  $\delta(X_i) = [\eta_{i1}^2, \eta_{i2}^2, \eta_{i1}\eta_{i2}, \eta_{i1}^2\eta_{i2}^2, \eta_{i1}^2s_{i1}^2, \eta_{i2}^2s_{i2}^2]^T$ . Choose  $\hat{w}_{c1}(0) = \hat{w}_{c4}(0) = [200, 200, 50, 0, 0, 0]^T$  and  $\hat{w}_{c2}(0) = \hat{w}_{c3}(0) = [100, 250, 50, 0, 0, 0]^T$ .

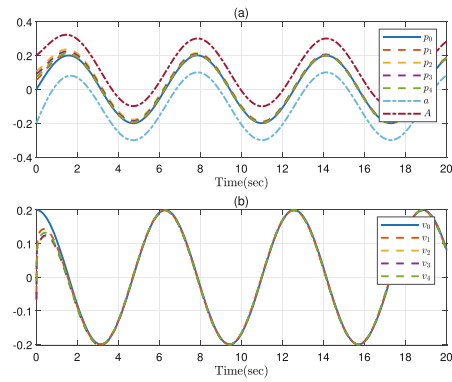


**Fig. 1.** Communication topology.

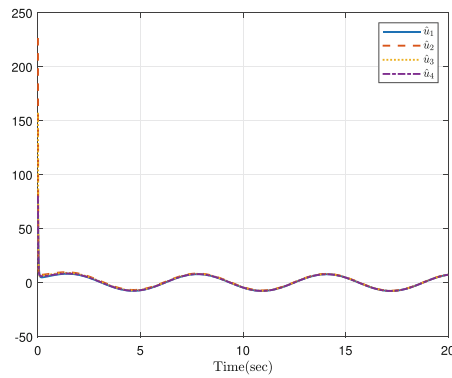
Firstly, the distributed observer (9) provides the respective reference trajectories for all the followers. The estimated errors of the distributed observers, as depicted in Fig. 2, indicate a rapid convergence of the estimation error  $e_i$  to zero in a finite time. The state trajectories of the followers are presented in Fig. 3, demonstrating that all followers successfully converge with the leader while maintaining the output signals within the asymmetric time-varying constraints. The approximate optimal control inputs for each single-link robot arm are illustrated in Fig. 4.



**Fig. 2.** Norm of estimated errors of distributed observers.



**Fig. 3.** Trajectories of the system states.



**Fig. 4.** Approximate optimal control inputs.

## 6 Conclusions

This study addresses the optimal cooperative control of MASs with asymmetric time-varying constraints using a single critic NN ADP algorithm. To tackle the asymmetric time-varying output constraints, a state mapping based on an inverse hyperbolic tangent function has been developed. Furthermore, an NN-based distributed observer has been proposed to reformulate the cooperative task as multiple general trajectory tracking problems. By effectively solving the HJB equation for each follower, the single critic NN ADP algorithm derives an optimal cooperative controller. It is demonstrated that this distributed observer offers a novel approach for distributed optimal control and simplifies the design process of optimal cooperative controllers. However, how to employ the distributed observer in solving optimal cooperative control of MASs with switching communication topology is still an open research direction.

**Acknowledgments.** This work was supported in part by the National Natural Science Foundation of China (6240314862273108), the Scientific Research Startup Project of Guangdong Polytechnic Normal University under Grant (991701004), the Young Innovative Talents Project of Guangdong Province (2021KQNCX045), the Provincial and Municipal Joint Project of Basic and Applied Basic Research Fund of Guangdong Province(2022A1515110644), the Basic and Applied Basic Research Project of Guangzhou Basic Plan (2023A04J0355) and the Enterprise Cooperation Project (1747352).

**Disclosure of Interests.** The authors state that they have no known financial conflicts of interest or personal relationships that might be perceived to have influenced the research presented in this paper.

## References

1. Zheng, X., Ma, H., Zhou, Q., Li, H.: Neural-based prescribed-time consensus control for multiagent systems via dynamic memory event-triggered mechanism. *SCIENCE CHINA-Technol. Sci.* (2024). <https://doi.org/10.1007/s11431-024-2694-x>
2. Peng, X.-J., He, Y., Li, H., Tian, S.: Consensus control and optimization of time-delayed multi-agent systems: analysis on different order-reduction methods. *IEEE Trans. Syst. Man Cybern.-Syst.* (2024). <https://doi.org/10.1109/TSMC.2024.3492335>
3. Zheng, S., Ma, H., Ren, H., Li, H.: Practically fixed-time adaptive consensus control for multiagent systems with prescribed performance. *SCIENCE CHINA-Technol. Sci.* (2024). <https://doi.org/10.1007/s11431-024-2780-3>
4. Ren, H., Zhang, C., Ma, H., Li, H.: Cloud-based distributed group asynchronous consensus for switched nonlinear cyber-physical systems. *IEEE Trans. Industr. Inf.* (2024). <https://doi.org/10.1109/TII.2024.3457811>

5. Luo, A., Ma, H., Ren, H., Li, H.: Estimator-based reinforcement learning consensus control for multiagent systems with discontinuous constraints. *IEEE Trans. Neural Netw. Learn. Syst.* (2024). <https://doi.org/10.1109/TNNLS.2024.3445880>
6. Wang, X., Guang, W., Huang, T., Kurths, J.: Optimized adaptive finite-time consensus control for stochastic nonlinear multiagent systems with non-affine nonlinear faults. *IEEE Trans. Autom. Sci. Eng.* (2023). <https://doi.org/10.1109/TASE.2023.3306101>
7. Fu, H., He, H., Chen, Y.: Event-triggered cooperative tracking control of multi-agent systems with a dynamic leader via approximate dynamic programming. *IEEE Trans. Artif. Intell.* **5**(6), 2752–2765 (2024)
8. Zhang, J., Fu, Y., Fu, J.: Optimal formation control of second-order heterogeneous multi-agent systems using adaptive predefined-time strategy. *IEEE Trans. Fuzzy Syst.* **32**(4), 2390–2402 (2024)
9. Sun, J., Ming, Z.: Cooperative differential game-based distributed optimal synchronization control of heterogeneous nonlinear multiagent systems. *IEEE Trans. Cybern.* **53**12, 7933–7942 (2023)
10. Luo, A., Zhou, Q., Ma, H., Li, H.: Observer-based consensus control for MASs with prescribed constraints via reinforcement learning algorithm. *IEEE Trans. Neural Netw. Learn. Syst.* (2023). <https://doi.org/10.1109/TNNLS.2023.3301538>
11. Zhao, W., Liu, H., Lewis, F.L.: Robust formation control for cooperative under-actuated quadrotors via reinforcement learning. *IEEE Trans. Neural Netw. Learn. Syst.* **32**(10), 4577–4587 (2021)
12. Wang, S., Wen, S., Yang, Y., Shi, K., Huang, T.: Suboptimal leader-to-coordination control for nonlinear systems with switching topologies: a learning-based method. *IEEE Trans. Neural Netw. Learn. Syst.* **34**(12), 10578–10588 (2023)
13. Xia, L., Li, Q., Song, R., Modares, H.: Optimal synchronization control of heterogeneous asymmetric input-constrained unknown nonlinear MASs via reinforcement learning. *IEEE/CAA J. Automatica Sinica* **9**(3), 520–532 (2022)
14. Cao, L., Ren, H., Meng, W., Li, H., Lu, R.: Distributed event triggering control for six-rotor UAV systems with asymmetric time-varying output constraints. *SCIENCE CHINA Inf. Sci.* **64**(7), 1–16 (2021). <https://doi.org/10.1007/s11432-020-3128-2>
15. Mishra, P.K., Verma, N.K.: On controller design for nonlinear systems with pure state constraints. *IEEE Trans. Circuits Syst. II Express Briefs* **69**(4), 2236C – 2240 (2022)
16. Fan, B., Yang, Q., Tang, X., Sun, Y.: Robust ADP design for continuous-time nonlinear systems with output constraints. *IEEE Trans. Neural Netw. Learn. Syst.* **29**(6), 2127–2138 (2018)
17. Li, K., Hua, C., Sun, C.: Finite-time observer-based leader-following consensus for nonlinear multiagent systems with input delays. *IEEE Trans. Cybern.* **51**(12), 5850–5858 (2021)
18. Yang, Y., Vamvoudakis, K.G., Modares, H., Yin, Y., Wunsch, D.C.: Safe intermittent reinforcement learning with static and dynamic event generators. *IEEE Trans. Neural Netw. Learn. Syst.* **31**(12), 5441–5455 (2020)
19. Guo, Z., Zhou, Q., Ren, H., Ma, H., Li, H.: ADP-based fault-tolerant consensus control for multiagent systems with irregular state constraints. *Neural Netw.* **180**, 106737 (2024)
20. Xue, S., Luo, B., Liu, D., Gao, Y.: Event-triggered integral reinforcement learning for nonzero-sum games with asymmetric input saturation. *Neural Netw.* **152**, 212–223 (2022)

21. Guo, Z., Ren, H., Li, H., Zhou, Q.: Adaptive-critic-based event-triggered intelligent cooperative control for a class of second-order constrained multiagent systems. *IEEE Trans. Artif. Intell.* **4**(3), 1654–1665 (2023)
22. Xue, S., Luo, B., Liu, D., Yang, Y.: Constrained event-triggered  $H_\infty$  control based on adaptive dynamic programming with concurrent learning. *IEEE Trans. Syst. Man Cybern. Syst.* **52**(1), 357–369 (2022)



# Pareto-Wise Ranking Generator for Multi-objective Coevolutionary Generative Adversarial Networks

Weifeng Guo<sup>1</sup>(✉) , Yijun Fu<sup>1</sup> , Ying Bi<sup>1</sup> , and Jing Liang<sup>2</sup>

<sup>1</sup> Zhengzhou University, Zhengzhou 450001, China  
guowf@zzu.edu.cn

<sup>2</sup> Henan Institute of Technology, Xinxiang 453003, China

**Abstract.** Existing optimization methods for Generative Adversarial Networks that use evolutionary algorithms frequently suffer from the problem of disordered ordering, which occurs when several sets of solutions produced by a generator during training have poor ordering. To address the aforementioned issue, we want to optimize the generation process by training a Pareto-wise end-to-end ordered generator and ranking its children via evolutionary algorithms. Finally, the MCE-GAN algorithm is proposed to decompose the problem of adversarial optimization of generative adversarial networks into two problems of generator and discriminator evolution, and to perform mutation operations on the generator and discriminator using evolutionary algorithms, where the generator selection method is performed by training a set of generated generators and then using non-dominated ordering. Extensive studies on a synthetic dataset and a benchmark picture dataset demonstrate that the proposed MCE-GAN produces competitive and superior results.

**Keywords:** Generative Adversarial Networks · Pareto-wise · Evolutionary Algorithms

## 1 Introduction

GAN [1] is a generative model that seeks to learn the data production process for generating new samples that are comparable to training data. GANs have a broad range of applications in generative modeling, including picture production [2], style migration [3], image complementation. They also have the advantage of producing high-quality and flexible samples.

GANs consist the generator and the discriminator. The training goal of the generator is to generate data to fool the discriminator, the discriminator is to distinguish between real and fake as much as possible. In practice, there are many problems of ineffective generative adversarial networks caused by pattern collapse or unstable training. In general, many recent optimizations on GANs optimize various adversarial training processes to overcome these model problems.

From the perspective of multi-objective optimization, generative adversarial network is a complex multi-objective optimization problem (e.g. the quality and diversity of the generated images). The main methods for multi-objective optimization are weighted summation [4], objective programming [5], pareto optimal solution [6] and evolutionary algorithms. In solving the problem of multi-objective GANs, Wang et al. [7] first proposed E-GAN. In addition Marco Baioletti et al. proposed MO-EGAN [8], which views the generator as a multi-objective optimization, and improves the performance of the generator by simultaneously optimizing the generated images.

Then we try to incorporate multi-objective optimization into the optimization of GANs. We propose a collaborative dual evolutionary GAN based on multi-objective optimization to enhance the diversity of result generation.

In conclusion, we suggest a new GAN technique based on multi-objective optimization to address the adversarial optimization challenges of GANs. To increase the diversity of the generated outcomes, we use a multi-objective algorithm for the generator, which selects created offspring based on many objectives. We undertake comprehensive experiments on many benchmark datasets, and the experiments reveal that our method works better than other methods, demonstrating MCE-GAN's superiority.

The rest is as follows. Section 2 is the related work. Section 3 details the proposed MCE-GAN. In Sect. 4, we give the experimental setup and comparison results of MCE-GAN with other methods. Finally, conclusions of this paper drawn in Sect. 5.

## 2 Related Work

### 2.1 Generating Adversarial Networks

GANs learn the distribution of data. The original form [1] of describing the max-min game between the two in the literature is:

$$\min_G \max_D \mathbb{E}_{\mathbf{x} \sim p_{\text{data}}(\mathbf{x})} [\log D(\mathbf{x})] + \mathbb{E}_{\mathbf{z} \sim p_{\mathbf{z}}(\mathbf{z})} [\log(1 - \log D(G(\mathbf{z})))] \quad (1)$$

There are several optimization methods to improve GANs: one is to modify the training objectives, such as Alec Radford et al.'s DCGAN [9] with convolutional method to improve the stability. The second approach considers the joint optimization of numerous discriminator or generator models as a multi-objective problem, such as the GMAN proposed by Durugkar et al. [10]. Third, stabilize the training process by introducing new optimization or regularization methods to reduce instability and pattern collapse, such as E-GAN uses evolutionary algorithms to optimize the generators and discriminators to adapt to the training environment and select the best generators.

However, the diversity of the generated solutions during the evolution of the aforementioned evolutionary GANs is yet to be improved.

## 2.2 Multi-objective Generating Adversarial Networks

Existing generative adversarial networks are optimized for a single objective, which is often insufficient to meet practical needs. For example, in an image generation task, a user may want the generated images to be both clear and lifelike, but also want them to vary in content, style, or diversity. This has led to the emergence of multi-objective generative adversarial networks.

Multi-objective optimization is a method for simultaneously considering multiple conflicting objectives in an optimization problem and finding the best balance between them. The optimization of one target results in the deterioration of other objectives in multi-objective optimization. The Pareto optimal solution formulation is as follows:

$$\min \bar{f}(\bar{x}) = [f_1(\bar{x}), f_2(\bar{x}), \dots, f_k(\bar{x})] \quad (2)$$

The core of multi-objective generative adversarial networks is the integration of multi-objective optimization theory into the GANs framework. Multi-objective loss function designs are: designing loss functions containing multiple objective terms; constructing multiple parallel discriminators to provide more feedback to the generator. There are also Pareto frontier approximation and hybrid optimization strategies, such as initializing parameters with evolutionary algorithms, dynamically tuning weights.

## 3 Proposed Method

We present the general algorithm. Subsequently, we present the algorithmic steps, describing in detail the evolutionary approach of the generator and the discriminator.

### 3.1 Main Framework

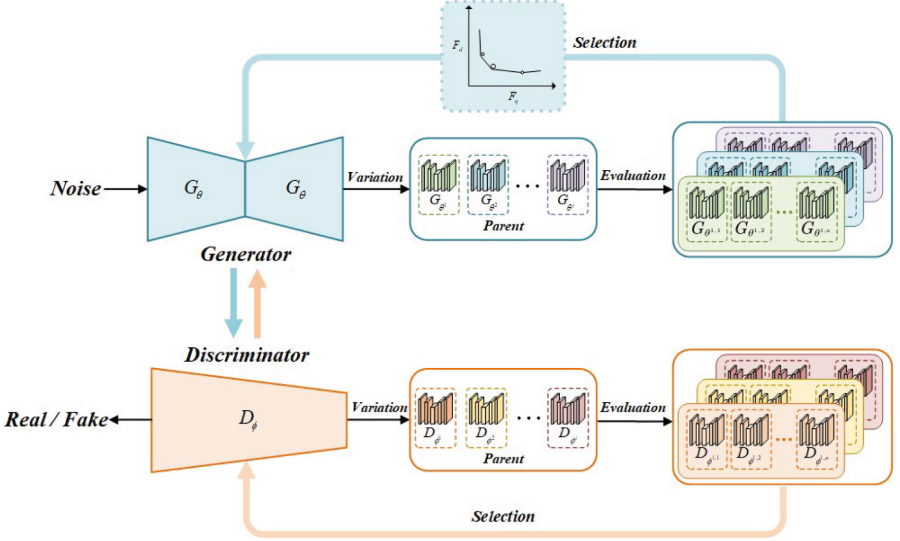
The main work of MCE-GAN is shown in Fig. 1. To improve the quality and diversity of the images generated by the generators and to build robust agent models with limited computational resources, we train a Pareto - wise classifier based on generator evolution, where the generator parameters are ranked according to the quality and diversity of the generated images. Thus, it enables generative adversarial network selection to choose competitive generator parameters without the need for a real evaluation of the real target or target alternatives. A complex generative adversarial network optimization is transformed into a simple Pareto dominance task. The pseudo-code of the MCE-GAN is shown in Algorithm 1, and the main steps are as follows:

- 1) Initialization. Specify the required parameters. Then, initialize the parameters of the generator  $\{\theta^1, \theta^2, \dots, \theta^J\}$  and the parameters of the discriminator  $\{\varphi^1, \varphi^2, \dots, \varphi^I\}$ .
- 2) Discriminator update loop setting. The purpose of this step is to distinguish the real data from the generated data better by continuously adjusting its parameters.



- 3) Selection of the best offspring of the discriminator. This step ensures that the discriminator can work with better parameters in subsequent iterations, thus improving the performance of the entire generative adversarial network.
- 4) Generator update loop setup. This step aims to fool the discriminator by continuously adjusting the parameters of the generator so that it can generate more realistic fake data.
- 5) Generator offspring non-dominated ordering and crowding alignment. This step helps to determine which generator offspring perform better under multiple objectives and maintains the diversity of generator offspring.
- 6) Generator best offspring selection (corresponding to the last part of the second big loop in the pseudo-code). This step ensures that the generator will work with better parameters in subsequent iterations, thus improving the performance of the GAN.

The key components of MCE-GAN are described in detail below.



**Fig. 1.** The main framework of MCEGAN.

### 3.2 Generator Multi-objective Evolution

We construct generators with the help of a multi-objective evolutionary algorithm and then mutate a batch of discriminators as an evolutionary environment. Generators are subpopulations of individuals that solve the subproblem of generating this one, which is used in collaboration with the discriminators.

$$\min \mathcal{L}_{G_\theta} = [\mathcal{L}_{G_{\theta^1}}, \mathcal{L}_{G_{\theta^2}}, \dots, \mathcal{L}_{G_{\theta^J}}]^T \quad (3)$$

**Algorithm 1:** MCE-GAN algorithm

---

**Require:**  $G_\theta$ : generator;  $D_\varphi$ : discriminator;  $T$ : iterations;  $K$ : discriminator's updating steps iteration;  $J$ : number for generator;  $I$ : number for discriminator;  $M$ : mutations for generator;  $N$ : mutations for discriminator,  $\gamma$ : hyper-parameter.

Initialize parameters  $\{\theta^1, \theta^2, \dots, \theta^J\}$   $\{\varphi^1, \varphi^2, \dots, \varphi^I\}$ ;

```

for  $t = 1, \dots, T$  do
  for  $k = 1, \dots, K$  do
    for  $i = 1, \dots, I$  do
      Sample a batch  $x_{real} \sim p_{data}$ ;
      Sample a batch  $z \sim p_z$ , use the generator to produce a batch of  $x_{fake}$ 
      with generators;
       $D_{\varphi^i} \leftarrow \text{Variation } \{N, D_{\varphi^{i,n}}\}$ ;
      Evaluate the progeny of the discriminator  $N$  ;
    end
    Choose the progeny with the most favorable results
     $\{D_{\varphi^{1,1}}, D_{\varphi^{2,1}}, \dots, D_{\varphi^{i,1}}\}$  as next parents;
  end
  for  $j = 1, \dots, J$  do
    Sample a batch of  $z \sim p_z$ ;
     $G_{\theta^j} \leftarrow \text{Variation } \{M, G_{\theta^j, M}\}$ ;
    Evaluate the progeny of the generator  $M$  ;
  end
   $PF \leftarrow \text{non dominated sorting } G_{\theta^j}$ ;
  crowding alignment assignment ( $PF$ );
  Choose the progeny with the most favorable results  $\{G_{\theta^{1,1}}, G_{\theta^{2,1}}, \dots, G_{\theta^{j,1}}\}$ 
  as next parents ;
end

```

---

On the basis of the algorithm, the generator subpopulation is evolved by mutation, evolution and selection. Figure 1 and Algorithm 1 illustrate the evolutionary process of the generator.

**Variation.** We specifically use the objective functions proposed in [7]. The mutation step is a form of asexual reproduction in which a single parent gives rise to  $Nm$  offspring by applying a step of Adam's algorithm with  $Nm$  distinct objective functions. In the evolution of the generator, we employ three complementary variants corresponding to the three different objectives of the generator, namely the Minimax variant [1], the Heuristic variant [1], and the Least-Square variant [12]. Different from E-GAN involving a generator that varies on a single specific discriminator, our generator variant trains multiple evolutionary discriminators.

$$\mathcal{M}_{\text{minimax}} = \frac{1}{2} \mathbb{E}_{z \sim q(z)} [\log(1 - D(G(Z)))] \quad (4)$$

$$\mathcal{M}_{\text{heuristic}} = -\frac{1}{2} \mathbb{E}_{z \sim q(z)} [\log(D(G(Z)))] \quad (5)$$

$$\mathcal{M}_{\text{least-square}} = \mathbb{E}_{z \sim q(z)} [\log(D(G(Z)) - 1)^2] \quad (6)$$

Thus, these variants offer diverse strategies for the generator, inject variant diversity into the training process, thereby capturing multiple modes and constructing complementary populations of generators for stable training.

**Evaluation.** After generating offspring through different mutations, we evaluated the quality (including quality and diversity) of the offspring individuals by  $F_G$ . Thus, our  $F_G$  is composed: the quality fitness score  $F_{G^q}$  and the diversity fitness score  $F_{G^d}$ .

$$\mathcal{F}_{G^q} = \sum_{j=1}^J w_i \mathbb{E}_{z \sim p_z} [D_{\varphi^j}(G_{\theta}(z))] \quad (7)$$

$$\mathcal{F}_{G^d} = -\sum_{j=1}^J w_i \left( \log \left\| \nabla D_{\varphi^j} - \mathbb{E}_{x \sim p_{\text{data}}} [\log D_{\varphi^j}(x)] - \mathbb{E}_{z \sim p_z} [\log(1 - D_{\varphi^j}(G_{\theta}(z)))] \right\| \right) \quad (8)$$

$$\text{minimize } \mathcal{F}_G(\theta) = \{\mathcal{F}_{G^q}(\theta), \mathcal{F}_{G^d}(\theta)\} \quad (9)$$

**Selection.** We describe how our algorithm selects generators. Our approach uses non-dominated ordering to filter the optimal generators. The generator children are categorized into non-dominated fronts according to the non-dominance relation, and then the crowding distance at the same non-dominated front is calculated.

### 3.3 Discriminator Evolution

Our algorithm evolves the discriminator through an evolutionary algorithm, which allows the discriminator to get a dynamic strategy that more effectively balances in training. On the basis of the evolutionary algorithm, the evolution of the discriminator subpopulation is performed by mutation, evolution and selection. Figure 1 and Algorithm 1 illustrate the evolutionary process of the discriminator. The objective function as follow:

$$\min \mathcal{L}_{D_{\varphi}} = [\mathcal{L}_{D_{\varphi^1}}, \mathcal{L}_{D_{\varphi^2}}, \dots, \mathcal{L}_{D_{\varphi^J}}]^{\varphi^T} \quad (10)$$

**Variation.** Unlike the generator variation, we choose two objective functions, Minimax variation and least squares variation, in the discriminator variation stage, where Minimax is common to both vanilla GAN and NS-GAN, and the size of the inputs also affects the generative performance, so we further use the least squares objective function for the variation, which is complementary to Minimax.

$$\mathcal{M}_{\text{minimax}} = \frac{1}{2} \mathbb{E}_{z \sim q(z)} [\log(1 - D(G(Z)))] \quad (11)$$

$$\mathcal{M}_{\text{least-square}} = \mathbb{E}_{z \sim q(z)} [\log(D(G(Z)) - 1)^2] \quad (12)$$

**Evaluation.** We optimize the negative logarithmic gradient as a fitness function (diversity fitness score), which is effective in determining whether the model has fallen into a pattern collapse because the gradient can reflect the training state of the GAN.

$$\mathcal{F}_D = -\log \left\| \nabla D_{\varphi^i} - \mathbb{E}_{x \sim p_{\text{data}}} [\log D_{\varphi^i}(x)] - \mathbb{E}_{z \sim p_z} [\log(1 - D_{\varphi^i}(G_{\theta}(z)))] \right\| \quad (13)$$

**Selection.** Similar to generator selection, following the evaluation, the new parent for the next evolution of the discriminator is selected.

## 4 Experiments and Evaluation

We present the evaluation metrics, parameter settings for the experiments, and analyze the generative performance of our model to validate our algorithm. At last, we compare the algorithmic performance of MCE-GAN with respect to other GANs to demonstrate the effectiveness of the improvements.

### 4.1 Evaluation Metric

In this study, we quantitatively evaluate the performance with the Inception Score (IS). It is worth mentioning that IS is a common metric in the field for GANs.

Specifically, IS obtains the conditional label distribution  $p(y | x)$  for each generated image. The unique feature of the IS metric is that it is able to accommodate both the key aspects of generativity and diversity at the same time. In general, higher values of IS imply better of the generated samples. In this study, we calculated the IS value to quantify the performance of the method. The expression is:

$$\text{IS} = \exp(\mathbb{E}_x \text{KL}(p(y | x) \parallel p(y))) \quad (14)$$

In addition, in order to evaluate our method in a more comprehensive and in-depth way, we also introduced the concept of human vision and evaluated it from a qualitative point of view. In this way, the combination of quantitative and qualitative evaluation provides a complete and accurate evaluation of the model.

### 4.2 Implementation Details

Specifically, two types of datasets are selected for the experiments: a well-constructed synthetic dataset, which is a mixture of 8 Gaussian, and a representative real dataset. By conducting experiments on these two types of datasets, the aim is to prove conclusively the validity of MCE-GAN.

Further, in order to make a more convenient and direct comparison with the existing research results, we adopt the same network structure as in the literature [7, 11], for the experiments on the real data. Moreover, we have selected MLP [11] as the model architecture of MCE-GAN for the experimental operation. The Table 1 lists the required model structure in a clear and exhaustive manner.

Meanwhile, during the experiments, the noise is sampled from the uniformly distributed  $p_z$  according to specific rules. The dimension of the sampled noise vector  $z$  is 100 dimensions for the real dataset and 256 dimensions for the synthetic dataset.

Finally, we chose the Adam with specific hyperparameter settings ( $\alpha = 0.0003$ ,  $\beta_1 = 0.4$ ,  $\beta_1 = 0.99$ ) to optimize our model. Moreover, all the operations in the experimental process are efficiently realized with the help of PyTorch, a powerful deep learning framework.

**Table 1.** The network structure used by MCE-GAN

| Generator                                | Discriminator                              |
|------------------------------------------|--------------------------------------------|
| <b>MLP</b>                               |                                            |
| <b>Input:</b> Noise                      | <b>Input:</b> Point                        |
| Layer 1: Fully Connected Layer (128)     | Layer 1: FC (128)                          |
| Layer 2: FC (128)                        | Layer 2: FC (128)                          |
| Layer 3: FC (128)                        | Layer 3: FC (128)                          |
| <b>Output:</b> Generated Point           | <b>Output:</b> True or False (probability) |
| <b>DCGAN</b>                             |                                            |
| <b>Input:</b> Noise                      | <b>Input:</b> Image (32*32*3)              |
| Layer 1: Fully Connected Layer (4*4*512) | Layer 1: ConvLayer (4*4*128)               |
| Layer 2: Transpose ConvLayer (4*4*512)   | Layer 2: ConvLayer (4*4*256)               |
| Layer 3: Transpose ConvLayer (4*4*256)   | Layer 3: ConvLayer (4*4*512)               |
| Layer 4: Transpose ConvLayer (4*4*128)   | Layer 4: Fully Connected Layer (1)         |
| Layer 5: Transpose ConvLayer (4*4*3)     | <b>Output:</b> True or False (probability) |
| <b>Output:</b> Generated Image (32*32*3) |                                            |

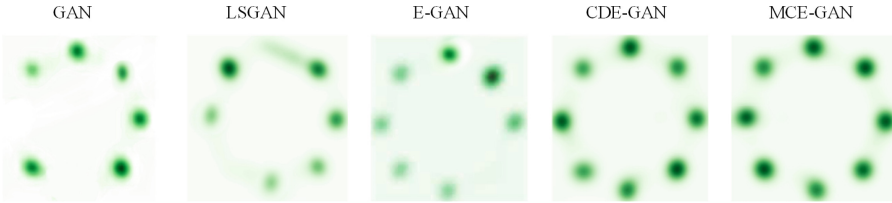
### 4.3 Model Evaluation

We conduct a qualitative and quantitative analysis of MCE-GAN’s generative performance. We compare and analyze Vanilla GAN (GAN), Least Squares GAN

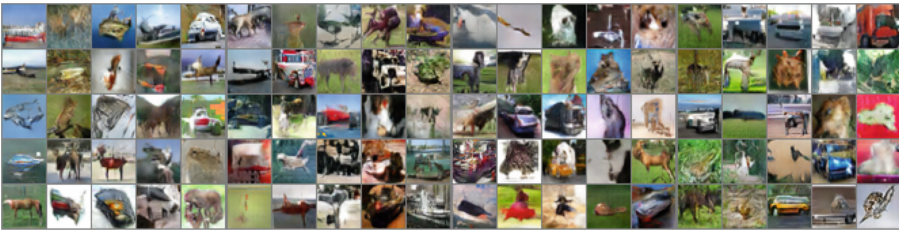
(LSGAN), Evolutionary GAN (E-GAN) [7], and CDE-GAN, all of which are similar to our technique.

(1) Qualitative evaluation. The diversity of GANs is assessed by learning a Gaussian mixed distribution, which intuitively displays whether GANs exhibit pattern collapse. Figure 3 displays the results of the KDE maps on various baselines.

To highlight the potential of MCE-GAN generating capability, Fig. 4 shows numerous samples generated by training our proposed model on the CIFAR-10 dataset. It should be mentioned that the samples used in this paper were selected rather randomly. On CIFAR-10, the MCE-GAN can create visually recognized images such as frogs, airplanes, and horses (Fig. 2).



**Fig. 2.** KDE plot of generated for 8 Gaussian distributions.



**Fig. 3.** Samples generated by our proposed MCE-GAN on the real dataset CIFAR-10.

(2) Quantitative evaluation. Since E-GAN, CDE-GAN is most similar to our method, we fuse it as a baseline to compare with our algorithm. In Table 2, E-GAN takes different number of generators  $J = 2, 4$ , CDE-GAN takes different number of discriminators  $I = 2, 4$ , and MCE-GAN takes different number of discriminators  $I = 2, 4$ . The results show that MCE-GAN exceeds others on the CIFAR-10 dataset. In addition, the other algorithms use the GP term during the training process, and their generative performance improves a little bit.

In addition, as the number of discriminators increases, the balance decreases. It is well tuned, and thus continues to achieve significant improvements.

**Table 2.** MCE-GAN comparison experiment results

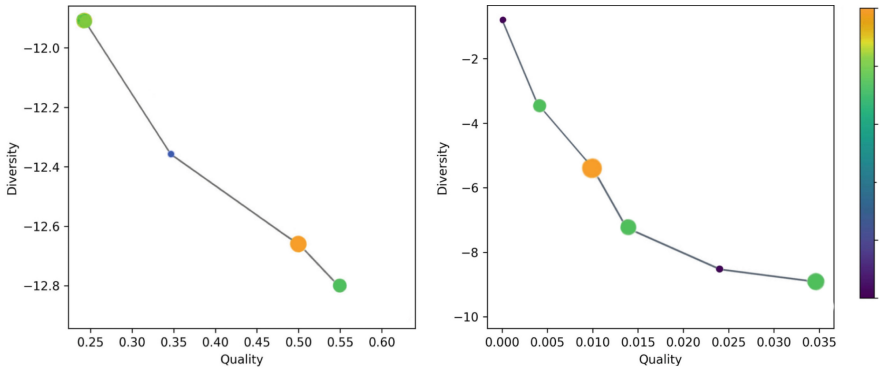
| Methods                     | IS              |
|-----------------------------|-----------------|
| E-GAN (J = 2, without GP)   | $6.70 \pm 0.06$ |
| E-GAN (J = 2, with GP)      | $7.23 \pm 0.08$ |
| E-GAN (J = 4, without GP)   | $6.93 \pm 0.06$ |
| E-GAN (J = 4, with GP)      | $7.31 \pm 0.08$ |
| CDE-GAN (I = 2, without GP) | $6.89 \pm 0.09$ |
| CDE-GAN (I = 2, with GP)    | $7.18 \pm 0.05$ |
| CDE-GAN (I = 4, without GP) | $7.06 \pm 0.08$ |
| CDE-GAN (I = 4, with GP)    | $7.48 \pm 0.10$ |
| MCEGAN (I = 2, without GP)  | $7.24 \pm 0.08$ |
| MCEGAN (I = 2, with GP)     | $7.41 \pm 0.07$ |
| MCEGAN (I = 4, without GP)  | $7.30 \pm 0.04$ |
| MCEGAN (I = 4, with GP)     | $7.51 \pm 0.05$ |

#### 4.4 Pareto-Front Analysis

We analyzed the Pareto frontier. The purpose of this experimental analysis is to determine whether the selection of a Pareto frontier for the population is necessary.

Indeed, if the best solution is in the middle of the front, this suggests that the application of the Pareto frontier is effective for selecting better generator parameters. However, if the best solution is located at the two extremes, it means that there is no need to optimize the generator with two functions.

So analyze the populations of the Pareto frontiers we found shown in Fig. 4: in the first row, we labeled the Pareto frontiers that found the best solution. We normalized the IS between bad and good results. Thus, the orange circle is



**Fig. 4.** Pareto frontiers of optimal solutions for different population sizes.

the population with the largest IS value and, conversely, the small blue circle corresponds to the population with the smallest IS value.

## 5 Conclusion

In this paper, we propose an MCE-GAN that incorporates Pareto frontier selection into the optimization of the generator as well as evolutionary computation of both parts of the GAN, in order to achieve the goal of co-evolution of the generator-discriminator and effectively overcome the problem of GAN pattern collapse and instability. The introduction of the Pareto frontier selection enhances the quality of generator selection, hence improving MCE-GAN's generation performance. Additionally, competitive results on a synthetic dataset and a real dataset are shown to demonstrate the usefulness of our technique.

**Acknowledgments.** This study was funded by the National Natural Science Foundation of China (62002329 (WFG)), and Henan Province Natural Science Foundation (242300421401), and China postdoctoral science foundation (2021M692915(WFG)), and Open Research Fund of State Key Laboratory of Digital Medical Engineering (2024-K07 (WFG)), and Frontier Exploration Projects of Longmen Laboratory (LMQYTSKT031).

**Disclosure of Interests.** The authors have no competing interests to declare that are relevant to the content of this article.

## References

1. Goodfellow, I., et al.: Generative adversarial nets. In: *Advances in Neural Information Processing Systems*, pp. 99–110 (2014)
2. Karras, T., Laine, S., Aila, T.: A style-based generator architecture for generative adversarial networks. In: *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 4401–4410 (2019)
3. Zhu, J., Park, T., Isola, P., Efros, A.A.: Unpaired image-to-image translation using cycle-consistent adversarial networks. In: *Proceedings IEEE International Conference on Computer Vision (ICCV)*, pp. 2242–2251 (2017)
4. Xia, W., Wu, Z.: An effective hybrid optimization approach for multi-objective flexible job-shop scheduling problems. *Comput. Ind. Eng.* **48**(2), 409–425 (2005)
5. Ghoseiri, K., Ghannadpour, S.: Multi-objective vehicle routing problem with time windows using goal programming and genetic algorithm. *Appl. Soft Comput.* **10**(4), 1096–1107 (2010)
6. Petchrompo, S., Coit, D., Brintrup, A., Wannakrairot, A., Parlikad, A.: A review of Pareto pruning methods for multi-objective optimization. *Comput. Ind. Eng.* **167**, 108022 (2022)
7. Wang, C., Xu, C., Yao, X., Tao, D.: Evolutionary generative adversarial networks. *IEEE Trans. Evol. Comput.* **23**(6), 921–934 (2019)
8. Ghoseiri, K., Ghannadpour, S.: Multi-objective evolutionary GAN. In: *Proceedings of the 2020 Genetic and Evolutionary Computation Conference Companion*, pp. 1824–1831 (2020)



9. Radford, A.: Unsupervised representation learning with deep convolutional generative adversarial networks. arXiv preprint [arXiv:1511.06434](https://arxiv.org/abs/1511.06434) (2015)
10. Durugkar, I., Gemp, I., Mahadevan, S.: Generative multi-adversarial networks. arXiv preprint [arXiv:1611.01673](https://arxiv.org/abs/1611.01673) (2016)
11. Gulrajani, I., Ahmed, F., Arjovsky, M., Dumoulin, V., Courville, A.: Improved training of Wasserstein GANs. In: Advances in Neural Information Processing Systems, vol. 30 (2017)
12. Xudong, M., Qing, L., Haoran, X., Raymond, Y., Zhen, W., Stephen, P.: Least squares generative adversarial networks. In: Proceedings of the IEEE International Conference on Computer Vision, pp. 2794–2802 (2017)
13. Arjovsky, M., Chintala, S., Bottou, L.: Wasserstein generative adversarial networks. In: Proceedings International Conference on Machine Learning, pp. 214–223 (2017)
14. Chen, S., et al.: CDE-GAN: cooperative dual evolution-based generative adversarial network. *IEEE Trans. Evol. Comput.* **25**(5), 986–1000 (2021)



# Research Progress on Control Technology of Tiltrotor Aircraft

Zhang Junhong, Li Hao, Zhao Letian<sup>(✉)</sup>, Liu Long, and Zhao Hai

The First Aircraft Institute of AVIC, Beijing 710089, China  
zletian2021@163.com

**Abstract.** Tiltrotor aircraft, with its vertical takeoff and landing capabilities and high-speed cruise efficiency, have demonstrated significant advantages in the aviation field, especially in complex terrains and special missions where they play an irreplaceable role. The excellent flight performance of tiltrotor aircraft makes it have great potential for development and application in the aviation field. However, the diversity of flight modes and the complexity of transition stages also bring many challenges to the control design of tiltrotor aircraft. This article focuses on the development of tiltrotor aircraft and the current research status of control technology. Based on the current situation, it looks forward to the next step of tiltrotor aircraft controller design and proposes some research direction suggestions.

**Keywords:** Tiltrotor Aircraft · Flight Control · Control Algorithm

## 1 Introduction

Helicopters and fixed wing aircraft are two typical types of aircraft. The helicopter has excellent maneuverability and can complete special flights such as vertical takeoff and landing, fixed-point hovering, rear flight, and side flight. It has unique advantages in complex environmental operations [1]. However, its rotor efficiency is relatively low, which limits its range and cruising speed, making it significantly inadequate for long-distance flight missions [2]. The advantages of fixed wing aircraft lie in their high flight speed, large payload capacity, ability to perform long-range missions, and outstanding performance in long-distance transportation and military strategic operations. However, it has strict requirements for takeoff and landing runways, which limits the expansion of its usage scenarios. The shortcomings of traditional helicopters and fixed wing aircraft greatly limit their application scope. To overcome this limitation, the aviation industry has been seeking a unique aircraft that combines the advantages of both, and tiltrotor aircraft has emerged. The tiltrotor aircraft successfully combines the advantages of helicopters and fixed wing aircraft. It has the ability to cruise at high speeds while also performing various maneuvers of helicopters during low-speed flight [3, 4].

As a special type of rotary wing aircraft, tiltrotor aircraft has the characteristics of vertical takeoff and landing, hovering in the air, excellent low-speed and low altitude flight capabilities of ordinary helicopters, and high-speed cruising flight of fixed wing propeller aircraft [5]. Tiltrotor aircraft are equipped with rotors on both sides of the wings

that can tilt with the nacelle [6]. Based on the difference in nacelle angle and aerodynamic characteristics, they can be divided into three modes: helicopter mode, fixed wing mode, and transition mode. When the short cabin is perpendicular to the body, the aircraft is in helicopter mode, achieving takeoff, landing, and maneuvering through the rotor. When the short cabin is parallel to the aircraft, it is in a fixed wing mode, with the wings providing lift and the rotors acting like propellers providing tension. The control force and torque are jointly provided by the total pitch and aerodynamic control surfaces [7, 8]. When the nacelle is between the helicopter and fixed wing modes, it is a transitional mode, in which the rotor and wing jointly provide lift, and the rotor control surface and aerodynamic control surface jointly participate in the guidance and control of the aircraft [9].

For the flight control system of a tiltrotor aircraft, all three working modes are indispensable. The research on helicopter and fixed wing modal flight control systems can draw on traditional methods, but the transition mode is no longer fully applicable due to its complex flight characteristics and control methods [10]. Based on this, this article summarizes the development and control technology of tiltrotor aircraft, analyzes and predicts its future development trends, and points out feasible research directions for the future, which can serve as inspiration for the development of tiltrotor aircraft control technology in China.

## 2 Development Status of Tiltrotor Aircraft

In the mid-20th century, Bell Airlines led the development of tiltrotor technology in the United States. In 1955, the XV-3 prototype was introduced (Fig. 1), and in 1958, it successfully conducted a tilting flight test, which not only verified the theoretical feasibility of this innovative technology, but also accumulated important experience for subsequent technological development [11]. Although the XV-3 had limitations in power and control, failing to fully meet military requirements, it paved the way for future technological advancements. After the XV-3, Bell Airlines launched the XV-15 validation aircraft (Fig. 2). Between 1977 and 1979, the XV-15 further explored the potential applications of tiltrotor technology in military and civilian fields through a series of flight tests, including hovering in helicopter mode and flight state transitions [12]. The outstanding performance of XV-15 at the Paris Air Show showed the broad prospects of tiltrotor aircraft, laying a solid foundation for the birth of V-22 “Osprey” tiltrotor aircraft.

The V-22 “Osprey” tiltrotor aircraft (Fig. 3) is the result of a collaboration between Boeing and Bell Helicopter. After years of rigorous testing and improvement, it completed key flight missions including vertical takeoff and landing, wing stall, single engine failure, and landing test flights from 1989 to 1990. The V-22, with its outstanding performance, has become the most widely used and numerous tiltrotor aircraft in the world [13]. The success of V-22 has also spawned multiple models to meet the diverse mission requirements of the army, navy, special operations, and transportation. However, so far V-22 has experienced multiple major accidents, resulting in over 40 fatalities. This not only causes losses to personnel, but also poses challenges to the future development of tiltrotor aircraft [14, 15].



**Fig. 1.** XV-3



**Fig. 2.** XV-15



**Fig. 3.** V-22

To compensate for the deficiencies and instability of the V-22, Bell Corporation launched the next-generation V-280 (Fig. 4) “Courage” tiltrotor helicopter program in the United States Army’s Joint Multi Purpose Rotorcraft (JMR) program [16]. The V-280 has undergone technological upgrades based on the V-22, enhancing maneuverability for low-speed flight and further improving transport efficiency and range. It is a tiltrotor aircraft platform with both transport and attack capabilities [17]. In early 1996, the Italian Agusta Helicopter Company and the American Bell Helicopter Company collaborated to develop the 6350 kg, 9-seater civilian tiltrotor aircraft AW609 (Fig. 5), which required 20 s for the rotor to tilt from vertical to horizontal position [18]. Eurocopter and Agusta of Italy have also jointly proposed the “Eurotilt” and “ERICA” plans with other countries, aiming to develop a 10 ton, 20 seat civilian tiltrotor aircraft. In 2020, Leonardo proposed the NextGenCTR civilian tiltrotor project (Fig. 6).



**Fig. 4.** V-280



**Fig. 5.** AW609



**Fig. 6.** NextGenCTR

Due to historical reasons, China started researching tiltrotor aircraft relatively late and has not yet developed a large tiltrotor aircraft capable of carrying passengers. However, at the 2013 China Helicopter Expo, the Aviation Industry Group showcased a model of the “Blue Whale” tiltrotor aircraft (Fig. 7), which is designed to carry a load of 20 tons, cruise at a speed of 538 km per hour, and have a combat radius of over 815 km. It plans to be manufactured using new composite materials and equipped with advanced optical flight control systems and fault prediction technology, with expected performance significantly surpassing traditional helicopters. At the same time, China Aerospace Science and Technology Corporation’s “Rainbow 10” tiltrotor unmanned aerial vehicle (Fig. 8) demonstrated its capabilities at the Zhuhai Air Show. This drone has a maximum takeoff weight of 350 kg during vertical takeoff and landing, and is capable of performing reconnaissance and strike missions, as well as providing information support. The flight control system is jointly developed by Nanjing University of Aeronautics and

Astronautics and Rainbow Company, and the drone body is manufactured by Rainbow UAV Company, achieving a high level of localization.



**Fig. 7.** Blue Whale



**Fig. 8.** Rainbow-10

### 3 Research on Control Technology of Tiltrotor Aircraft

At present, the controller design technology for traditional helicopters and fixed wing aircraft is relatively mature, such as traditional PID control, LQR control [19], neural network control [20], feature structure configuration [21], and robust control [22]. Tiltrotor aircraft have two sets of control modes: helicopter mode and airplane mode. In transition mode, the rotor/propeller nacelle tilts, exhibiting variable configuration and speed motion characteristics. As the model parameters change, the proportion of different control modes also constantly changes. In addition, due to the large number of maneuvers, there is a serious coupling relationship between them, which not only affects the balance, maneuverability, and stability in flight mode, but also directly affects the transition mode. Tiltrotor aircraft is a nonlinear, multivariable, and time-varying control object, and has characteristics such as mathematical model inaccuracy and environmental uncertainty [23]. Considering the above characteristics, it is very difficult to design controllers for tiltrotor aircraft during hovering, takeoff and landing, and transition. But this has also attracted numerous experts both domestically and internationally to conduct research in this field. According to the development of control theory, the flight control technology of tiltrotor aircraft can be mainly divided into two categories: linear control technology and nonlinear control technology [24]. Starting from these two aspects, analyze the existing control methods.

#### 3.1 PID Control

Due to the complexity of the model of a tiltrotor aircraft, the most commonly used method in the initial research stage of the tiltrotor aircraft flight controller is the PID control method. The PID control method has adjustable dynamic characteristics, and even without understanding the system model characteristics, good control performance can be achieved by adjusting parameters through experiments. Choi et al. [25] designed a traditional PID control method for the “Smart UAV” scaled down aircraft project at the Korea Aerospace Research Institute. Stability control was adopted in the pitch, roll, and

yaw channels, and the control variables were matched to the actuators through a control surface mixer. Different PID parameters were designed for different flight modes, and ground mooring tests and tilt flight tests were completed through gain scheduling. In terms of transition modes, based on the requirements of engineering applications, Xiao Siqu et al. [26] designed a tilting transition scheme after analyzing the aerodynamic and control characteristics of transition mode aircraft. On this basis, PID control method was used for simulation verification and successfully applied in flight testing. However, PID algorithm has limitations in anti-interference ability, difficulty in dealing with nonlinear systems, and complex handling of coupling problems.

### 3.2 Gain Scheduling

The basic idea of the gain scheduling algorithm is to linearize the system at multiple operating points and use classical or linear control methods to design controllers separately, and schedule the gain control at each operating point according to the scheduling rules. The gain scheduling algorithm is simple and reliable, and is widely used in the design of tiltrotor unmanned aerial vehicle controllers. In terms of transition modes, Kang et al. [27] designed a control system for the scaled “SmartUAV” using airspeed maintenance, altitude maintenance, and orientation control for the outer loop. The gain was configured under different flight modes, and the experiment achieved transitional mode flight, verifying the effectiveness of attitude control. Holsten et al. [28] used PID method for control and optimized the controller gain based on multi-objective parameter synthesis method (Fig. 9). Through experimental analysis of the time response of states at different operating points, the results showed that the performance met the optimization design criteria. The main disadvantage is that it is relatively cumbersome, and the generated control rate satisfies local characteristics, but in order to satisfy global characteristics, other techniques need to be combined. Moreover, the robustness of this method is limited, and it is difficult to suppress the interference caused by uncertainty under complex conditions [29].

### 3.3 Linear Quadratic Optimal Control

Linear quadratic optimal control is a design method for linear control systems based on state space models, which has good time-domain and frequency-domain characteristics. In early 2004, Vorwald and Chopra from the University of Maryland in the United States added a linear quadratic Gaussian regulator and a state observer to analyze and derive the dynamic equations of a tiltrotor. They separated stable and unstable modes in the modal space and designed controllers only for unstable modes. This processing method formed the basic prototype of LQG (Linear Quadratic Gaussian) control [30]. Papacristos et al. [31] designed a longitudinal control law for a small tiltrotor unmanned aerial vehicle (UAV) based on the theory of linear quadratic tracker (LQT), and achieved trajectory tracking control in the UAV helicopter mode (Fig. 10). Numerical simulations and hovering flight tests showed that the control device had good tracking performance. Onen et al. [32] used the LQR/LQT control method for attitude control of tiltrotor aircraft during vertical takeoff and landing phases. The effectiveness of the controller was verified through simulation and vertical takeoff and landing flight tests. In the transition mode,

Ma Jiahui [33] studied the control method of linear quadratic regulator (LQR) for the longitudinal and vertical flight attitude of XV-15 tiltrotor aircraft in the transition section. Simulation experiments showed that the controller had good control effect on the flight attitude of tiltrotor aircraft. The main drawbacks of this method are strong dependence on model accuracy, difficulty in selecting weight matrices, and limited robustness.

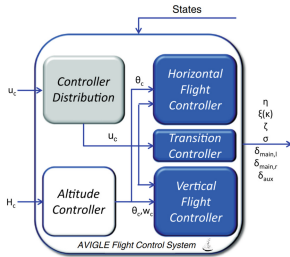


Fig. 9. Controller structure [28]

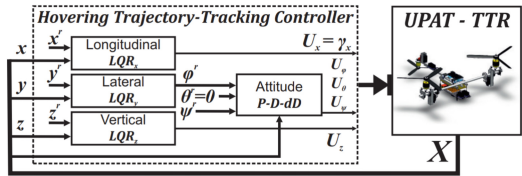


Fig. 10. UPAT-TTR's hovering trajectory-tracking control scheme [31]

### 3.4 Dynamic Inversion Control

Dynamic inversion eliminates all or part of the nonlinear characteristics in the original system through appropriate nonlinear transformations and nonlinear state feedback transformations, thereby transforming the original system into a linear system and applying mature linear system control methods to design controllers for the transformed system. On the basis of establishing a tilting three rotor model, Chen Chao et al. [34] used dynamic inverse control method in the attitude control inner loop to calculate the output of the controller. Simulation results showed that the designed control system can effectively track the desired input and suppress disturbances. Kim et al. [35] used the dynamic inversion method for the inner loop stabilization system and outer loop tracking system of a tiltrotor unmanned aerial vehicle, and compensated for the dynamic inversion error using a neural network. Simulation results showed that the system has good stability and tracking capability. In terms of transition modes, Calise et al. [36] conducted research on the high confidence simulation model GTRS (Generic TiltRotor Simulator) of the tiltrotor unmanned aerial vehicle XV-15. They used dynamic inverse design attitude and angular velocity basic model controllers, and compensated for errors between the model and the actual object using neural networks. The simulation experiment results showed that the attitude met the level 1 flight quality requirements of ADS-33D during the tilting process. However, it also suffers from model dependency and poor robustness, as well as complex control allocation.



### 3.5 Adaptive Control

The term “adaptive” was first proposed in “Engineering Control Theory” in 1954. The basic idea of adaptive control is to use adaptive algorithms to identify unknown parameters online and design control laws based on their estimated values, so that the actual closed-loop control system meets performance requirements [37]. In the mid-20th century, Professor Whitaker from the Massachusetts Institute of Technology in the United States first proposed applying the model reference adaptive control method based on parameter optimization design to the design of helicopter flight control systems [38]. Cardoso et al. [39] adopted an adaptive hybrid framework approach for a small tiltrotor unmanned aerial vehicle, which achieved smooth switching between various controllers by selecting a weight function. In terms of transition modes, Guo Jiandong [40] divided the tiltrotor unmanned aerial vehicle into several subsystems, and then modeled and designed controllers separately. Finally, a model adaptive control strategy was adopted to achieve a smooth transition of the controller (Fig. 11). Yildiz et al. [41] designed a cascaded control system consisting of a position loop and an attitude loop for a small tilting wing unmanned aerial vehicle with a quadcopter layout. Both loops used model reference methods to design controllers. Simulation experiments show that due to the introduction of adaptive methods, the interference caused by rotational inertia and mass uncertainty has been solved, with a transition time of 10 s and stable height and attitude. However, it has the problem of large computational complexity and the contradiction between convergence speed and stability.

### 3.6 Fuzzy Control

The basic idea of fuzzy control is to use computers to implement human control experience, summarize the strategies of human experts for specific controlled objects or processes into a series of control rules expressed in conditional statements, and obtain a set of control effects through fuzzy reasoning, which act on the controlled objects or processes, but precision and stability issues require special attention [42]. Kadmiry et al. [43] used fuzzy control techniques to design scheduling functions and obtained good simulation results. However, the gain scheduling method can only ensure the effectiveness near the parameter trajectory, and is used for controller design and ensuring stability close to all trajectories. It must sacrifice the local performance of the closed-loop system, which greatly reduces the control performance. In terms of transition modes, Chang et al. [44] studied the scaled “SmartUAV” and linearized the system at tilt angles of 0°, 30°, 60°, and 90°. A fuzzy inference method was used to establish a fuzzy model approximation of the original model using the linearized model. The outer loop of the controller maintained control of altitude, airspeed, and yaw angle, while the inner loop used a T-S fuzzy controller to control the roll and pitch channels, achieving transitional mode flight in 20 s. The tilt process had good tracking of altitude and velocity, and the yaw angle converged to the expected value when entering the fixed wing mode quickly. During the reverse tilt process, the velocity tracking was good, but the altitude deviated from the expected value by up to 40 m.

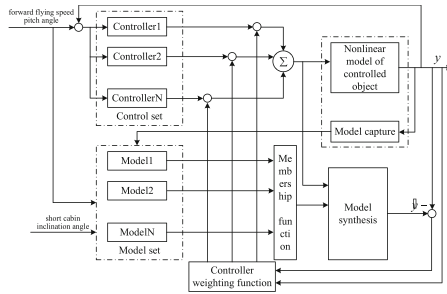


### 3.7 Sliding Mode

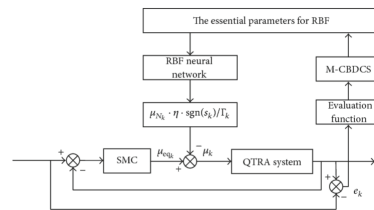
Sliding mode control is a concept developed from the variable structure control proposed by Soviet scholar Emelynaov in the 1950s. After decades of development, this method has become a very important branch of nonlinear control, the advantage of this method is that it has strong robustness to modeling errors, parameter uncertainties, and other disturbances [45]. Pan et al. [46] designed a novel control method based on active disturbance rejection sliding mode controller and observer. The attitude angle and angular velocity commands were obtained using a tracking differentiator, and various disturbances were estimated using a sliding mode extended observer. The stability of the controller was proved using Lyapunov method, and the robustness of the flight control system was demonstrated through simulation experiments in different modes. Xiong et al. [47] used sliding mode control method for attitude and position tracking control of tiltrotor unmanned aerial vehicles. Simulation results showed that the system has fast convergence and good robustness. Liu Haibo et al. [48] used a fractional order sliding mode controller based on a super twisting second-order sliding mode disturbance observer to perform attitude step control on the rotor mode of a quadcopter configuration tiltrotor unmanned aerial vehicle.

### 3.8 Neural Network Control

Neural network control algorithms have strong self-learning and nonlinear mapping capabilities, but their training time is long, they are sensitive to parameters, and there may be overfitting problems, Rysdyk et al. [20] proposed a control method for tiltrotor aircraft using neural network model inverse control, which has consistent response characteristics throughout the entire flight envelope. This method achieves attitude maintenance in the longitudinal channel and attitude angular rate maintenance in the lateral direction, improves the dynamic control effect of the roll channel and heading channel, and enables heading maintenance and coordinated turning flight at low airspeed. Yin et al. [49] combined sliding mode control with neural network methods to solve the convergence process oscillation problem that often occurs in sliding mode control design (Fig. 12). Kang et al. [50] designed a control system for the “TR-40”, consisting of a task scheduling layer, a navigation layer, a velocity and altitude maintenance layer, and an attitude control layer. The velocity and altitude maintenance control and attitude control used a dynamic inverse control method based on sigma pi through a network. The baseline model was a linearized model in the rotary wing mode. Simulation experiments showed that in the presence and absence of wind, the acceleration process during the tilting process took more than 60 s while raising, and the reverse tilting process took more than 70 s to complete at the same height. The maximum amplitude of the pitch angle did not exceed 20°, and the height tracked the expected value well without significant deviation.



**Fig. 11.** Multi model adaptive control law switching structure diagram [40]



**Fig. 12.** The control structure of the adaptive neural network sliding mode controller for QTRA [49]

## 4 Development Ideas for Control Technology of Tiltrotor Aircraft

Tiltrotor aircraft have the advantages of fast flight speed, low vibration and noise, strong payload capacity, and high flight efficiency. They break free from the limitations of takeoff and landing sites and can better adapt to complex and narrow terrains such as mountains and ships. The unique aerodynamic layout and excellent flight performance of tiltrotor aircraft make it have great potential for development and application. However, there are still many complex control problems that have not been fully resolved in the research and design of tiltrotor aircraft, especially in the field of aviation weapons and equipment in China, where tiltrotor aircraft still do not exist. The research and design of tiltrotor aircraft still have a long way to go. The following development directions are just some of our research and learning experiences, hoping to provide some ideas and inspirations for relevant researchers.

### 4.1 Meet the Adaptability Requirements of Complex Environments

Tiltrotor aircraft need to operate in diverse environments, including complex terrains such as ships, cities, canyons, and mountainous areas, as well as adverse weather conditions such as strong winds and turbulence. These environments pose challenges to flight stability and safety, therefore, control technology must be developed to adapt to these complex environments. By enhancing environmental awareness, improving aerodynamic models, and developing adaptive control algorithms, tiltrotor aircraft can maintain performance under various conditions, ensuring successful mission execution and flight safety.

### 4.2 High Reliability and High Security Requirements

Tilt rotor aircraft play an important role in critical missions such as military and rescue operations, and any control failure could lead to catastrophic consequences. Therefore, high reliability and high safety are the cornerstones of its control technology development. Through fault diagnosis, fault-tolerant control, and redundancy design, it can be ensured that the aircraft can operate safely even in the event of partial system failure, thereby protecting personnel and asset safety and ensuring the continuity of critical tasks.

### 4.3 Research on New Control Algorithms

With the advancement of computing technology and the development of control theory, new control algorithms can provide higher control accuracy and robustness. The flight envelope of a tiltrotor aircraft is wide, and traditional control methods are difficult to meet the control requirements within the envelope, especially during the transition process. Therefore, researching new control algorithms is aimed at addressing control challenges under model uncertainty, external disturbances, and complex flight conditions, and improving the performance and reliability of aircraft.

### 4.4 Multidisciplinary Integration Development

Tiltrotor aircraft is a highly integrated system that involves multiple fields such as aerodynamics, structure, and control. The integration of multiple disciplines can promote the exchange and integration of interdisciplinary knowledge, and drive the application of innovative technologies. In today's world, science and technology are developing rapidly, and a large number of emerging achievements have emerged in fields such as digital technology, electric drive technology, artificial intelligence technology, as well as new materials and processes for aviation. Therefore, interdisciplinary integration is crucial for improving the overall performance and competitiveness of tiltrotor aircraft.

## 5 Conclusion

This article comprehensively reviews the research progress of control technology for tiltrotor aircraft, from traditional control methods to modern control methods, each with its unique principles and application characteristics. Traditional control methods have certain advantages in simple operating conditions, but have limitations in complex flight environments. Modern control methods demonstrate better performance in dealing with uncertainty, nonlinearity, and multimodal switching problems. However, current control technologies still face many challenges in terms of adaptability to complex environments, reliability, and accuracy efficiency balance. The article also summarizes several development directions for the next step of research on tiltrotor aircraft control technology, which can provide reference for subsequent related research.

**Disclosure of Interests.** The authors have no competing interests to declare that are relevant to the content of this article.

## References

1. Ximing, W., Xiaowei, M.: A perspective of the future development of key helicopter technologies. *Acta Aerodynamica Sinica* **39**(3), 1–10 (2021)
2. Wei, Z., Mingchuan, Y., Feng, F., et al.: Requirement analyses and technical prospects of helicopter rotor airfoils. *Acta Aerodynamica Sinica* **39**(6), 61–69 (2021)
3. Weiwei, W., Cunwang, M., Lian, Z., et al.: Numerical simulation of continuous tilting transition of tiltrotor aircraft. *Adv. Aeronaut. Sci. Eng.* **12**(3), 55–64 (2021)

4. Xin, Y., Renliang, C.: Effect of tiltrotor nacelle tilting strategy on the pilot workload. *J. Aerosp. Power* **38**(8), 1889–1990 (2023)
5. Yunyun, C.: Research on mathematical modeling method for tilt rotor aircraft flight dynamics. Nanjing Univ. Aeronaut. Astronaut. (2012)
6. Wang, W., Pang, H., Wang, B.: Dynamics simulation analysis of tiltrotor aircraft nacelle tilt mechanism. *Adv. Aeronaut. Sci. Eng.* **7**(2), 235–240(2016)
7. Wineefeld, J., Kendal, F.: Unmanned systems integrated roadmap (2011)
8. Paw, Y.C.: Synthesis and validation of flight control for UAV. *Dissertations & Theses-Gradworks* (2009)
9. Kai, M.: Study on aerodynamic characteristics of tilting wing unmanned aerial vehicle in transition state. *Adv. Aeronaut. Sci. Eng.* 1–9 (2024). <http://kns.cnki.net/kcms/detail/61.1479.V.20241212.1109.002.html>
10. Xin, Y., Yanqin, Z., Renliang, C., et al.: Tiltrotor aircraft automatic conversion control adapted to lift and thrust regulation. *J. Aerosp. Power*, 1–13(2024). <https://doi.org/10.13224/j.cnki.jasp.20240479>
11. Thomason, T.: The bell helicopter XV-3 and XV-15 experimental aircraft-lessons learned (1990)
12. Bilger, J.M., Marr, R.L., Zahedi, A.: Results of structural dynamic testing of the XV-15 tilt rotor research aircraft. *J. Am. Helicopter Soc.* **27**(2), 58–65 (1982)
13. Rourke, R.O.: V-22 Osprey tilt-rotor aircraft: background and issues for congress. *Congressional Res. Serv. Rep.* (2011)
14. Chunming, L.I.U., Jiye, L.U.O.: Research of NASA large civil tilt rotor development. *Civ. Aircr. Des. Res.* **4**, 1–5+50 (2010)
15. Jaeger, D.B.: Accreditation of modeling and simulation test capabilities for CV-22 initial operational test and evaluation. In: International Conference on US Air Force T & E Days (2010)
16. Jiangtao, X.: Design of flight control law for the new configuration unmanned tilt-rotor aircraft in transition state. Nanjing Univ. Aeronaut. Astronaut. (2021)
17. Qiankun, L.: Design and simulation of electric drive unmanned tilting rotorcraft powertrain. Nanjing Univ. Aeronaut. Astronaut. (2020)
18. Huber, M.: AW609 flight testing achieves milestones. *Aviat. Int. News* **49**(11), 47 (2018)
19. Donadel, R.: Modeling and control of a tiltrotor unmanned aerial vehicle for path tracking (2015)
20. Rolf, T., Rysdyk, Anthony J, et al.: Adaptive model inversion flight control for tilt-rotor aircraft. *J. Guidance Control Dyn.* (1999)
21. Garrard, W.L., Low, E., Prouty, S.: Design of attitude and rate command systems for helicopters using eigenstructure assignment. *J. Guid. Control. Dyn.* **12**(6), 783–791 (1989)
22. Smerlas, A., Strange, M.E., Gubbels, A.W., et al.: Design and flight testing of an H Sub(infinity) controller for the NRC bell 205 experimental fly-by-wire helicopter (1998)
23. Jinghui, D.: Key technologies and development for high-speed helicopters. *Acta Aeronautica et Astronautica Sinica* **45**(9), 1–20(2024)
24. Zaibin, C.: Research on key technologies of flight control system of a tilt rotor UAV. *Univ. Chin. Acad. Sci.* (2020)
25. Choi, S.W., Kang, Y., Chang, S., et al.: Development and conversion flight test of a small tiltrotor unmanned aerial vehicle. *J. Aircr.* **47**(2), 730–732 (2012)
26. Siqi, X., et al.: Research on longitudinal control strategy design in transition section of tilt-rotor UAV. *Navig. Position Timing* **8**(6), 88–95(2021)
27. Holsten, J., Moormann, D.: Flight control law design criteria for the transition phase for a tiltwing aircraft using multi-objective parameter synthesis. *CEAS Aeronaut. J.* **6**(1), 17–30 (2014)

28. Kang, Y., Park, B., Yoo, C., et al.: Control law modification according to flight test of small scaled tilt rotor UAV. In: AIAA Guidance, Navigation Control Conference and Exhibit (2008)
29. Rugh, W.J., Shamma, J.S.: Research on gain scheduling. *Automatica* **36**(10), 1401–1425 (2000)
30. Vorwald, J., Chopra, I.: Stabilizing pylon whirl flutter on a tilt-rotor aircraft. In: AIAA (1991)
31. Papachristos, C., Alexis, K., Tzes, A.: Linear quadratic optimal trajectory-tracking control of a longitudinal thrust vectoring-enabled unmanned Tri-TiltRotor. In: IEEE (2013)
32. Onen, A.S., Cevher, L., Senipek, M., et al.: Modeling and controller design of a VTOL UAV. In: International Conference on Unmanned Aircraft Systems (2015)
33. Jiahui, M.: Design of transition control law of tilt-rotor aircraft based on LQR. *Autom. Appl.* **65**(1), 76–78+85(2024)
34. Chao, C., Lincheng, S., Daibing, Z., et al.: Mathematical modeling and control of a tiltrotor UAV. In: IEEE International Conference on Information & Automation (2017)
35. Kim, B.-M., Choi, K., Kim, B.S.: Trajectory tracking controller design using neural networks for tiltrotor UAV. In: AIAA Guidance, Navigation & Control Conference & Exhibit (2007)
36. Calise, A.J., Rysdyk, R.T.: Nonlinear adaptive flight control using neural networks. *Control Syst. IEEE* **18**(6), 14–25 (1998)
37. Ferrari, S., Stengel, R.F.: Online adaptive critic flight control. *J. Guid. Control. Dyn.* **27**(5), 777–786 (2004)
38. Bodson, M.: Reconfigurable nonlinear autopilot. *J. Guidance, Control, Dyn.* (2003)
39. Cardoso, D.N., Raffo, G.V., Esteban, S.: A robust adaptive mixing control for improved forward flight of a tilt-rotor UAV. In: 19th International Conference on Intelligent Transportation Systems (ITSC) (2016)
40. Jiandong, G.: Flight control of unmanned tiltrotor aircraft. Nanjing Univ. Aeronaut. Astronaut. (2013)
41. Yildiz, Y., Unel, M., Demirel, A.E.: Nonlinear hierarchical control of a quad tilt-wing UAV: an adaptive control approach. *Int. J. Adapt. Control Signal Process.* (2017)
42. Feng, G.: A survey on analysis and design of model-based fuzzy control systems. *IEEE Trans. Fuzzy Syst.* **14**(5), 676–697 (2006)
43. Kadmiry, B., Driankov, D.: A fuzzy gain-scheduler for the attitude control of an unmanned helicopter. *IEEE Trans. Fuzzy Syst.* **12**(4), 502–515 (2004)
44. Chang, S., Bang, H., Lee, W., et al.: Dynamic modeling and fuzzy control for a small tiltrotor unmanned aerial vehicle. *Proc. Inst. Mech. Eng. Part G J. Aeros. Eng.* **227**(G9), 1468–1487 (2013)
45. Barbot and J. Pierre: *Sliding Mode Control In Engineering* (2002)
46. Pan, Z., Wang, W., Song, S., et al.: Nonlinear attitude control of tiltrotor aircraft based on active disturbance rejection sliding mode method. In: Guidance, Navigation & Control Conference (2017)
47. Xiong, J.J., Zheng, E.H.: Position and attitude tracking control for a quadrotor UAV. *ISA Trans.* **53**(3), 725–731 (2014)
48. Liu, H., Wang, H., Shen, L.: Attitude control for vertical take off and landing mode of QTR based on fractional order sliding mode control. *Syst. Eng. Electron* **39**, 156–161 (2017)
49. Yin, Y., Niu, H., Liu, X.: Adaptive neural network sliding mode control for quad tilt rotor aircraft. *Complexity* (2017)
50. Kang, Y., Kim, N., Kim, B.S., et al.: Autonomous waypoint guidance for tilt-rotor unmanned aerial vehicle that has nacelle-fixed auxiliary wings. *Proc. Inst. Mech. Eng. Part G J. Aerosp. Eng.* **228**(g14), 2695–2717 (2014)



# Robust and Non-fragile Control for Non-uniform Sampled-Data Systems Under Try-Once-Discard Protocol

Jiyuan Zhang, Shipei Cai, and Jinling Liang

Southeast University, Nanjing 210096, People's Republic of China  
jinlliang@seu.edu.cn

**Abstract.** This paper investigates the robust/non-fragile control problem is studied for non-uniform sampled-data system under the scheduling of the try-once-discard (TOD) protocol. The uncertainties of the system and the fragility of the controller are considered. Both of them are assumed to be norm-bounded. To reduce communication load and prevent network congestion, the TOD protocol is applied to manage the data transmissions. We aim to develop a criterion such that the considered non-uniform sampling system with an exponential gain is exponentially stable for all possible system uncertainties. The desired criterion is derived via the Lyapunov theory and the Halanay-like inequality. Finally, a numerical example is presented.

**Keywords:** Sampled-data systems · Non-uniform sampling · Robust control · Non-fragile control · Try-once-discard protocol

## 1 Introduction

In the previous few decades, sampled-data control theory for networked control systems (NCSs) has garnered significant recognition owing to the widespread adoption of digital technology in controller implementation [1, 2]. The signals transmitted through communication networks are often constrained by limited network resources, and thus the sampling intervals in a NCS might be inevitably non-uniform. Correspondingly, the stability research of NCSs under non-uniform sampling has garnered growing academic focus, see e.g. [3, 4].

In practical applications, NCSs might suffer from uncertainties resulted from modeling errors, linear approximations, external disturbances and system failures. These uncertainties would lead to discrepancies between the actual system parameters and their model values, affect the system stability, and complicate the controller design [5, 6]. Till now, the robust sampled-data control for uncertain NCSs has been extensively studied. To cite some, in [7], the robust control has been explored for sampled-data systems where data loss occurs. In [8], the robust stability analysis have been considered for uncertain Markovian jump linear systems and the related controller design problem has been solved.

Another important issue in NCSs is the fragility (or possible parameter perturbations in controller implementation) of the controller, which is mainly caused from the limited word length of microprocessors used to implement digital controllers, and the errors arising from A/D and D/A conversions. To cope with the effects of these perturbations on the control performance, the non-fragile control issue has been widely studied, leading to many excellent results [9, 10]. In [11], a non-fragile observed-based robust controller has been obtained, which stabilizes systems with structural uncertainties. In [12], the  $H_\infty$  non-fragile control for linear systems has been analyzed, and the non-fragility against possible controller parameter perturbation has been achieved.

In NCSs, it is often the case that a communication network is used to transfer information from the sensors to the remote controller. With the aim of mitigating data congestion, various protocols are widely employed to schedule signal transmissions within the network [13, 14]. In NCSs, the try-once-discard (TOD) scheduling is commonly found, which determines the sequence in which sensor nodes are scheduled based on a set quadratic selection principle. By now, substantial research has been undertaken regarding the analysis and synthesis of NCSs under TOD protocol [15, 16]. For example, in [17], for nonlinear NCSs, TOD protocol and uniform quantization phenomena have been considered, the ultimate boundedness of the system has been studied. Nevertheless, effects of the TOD protocol on NCSs have not been fully considered under the sampling control framework, not to mention the simultaneous consideration of the system uncertainties, the controller fragility, and the non-uniform sampling. Shortening such a gap is what drives us to undertake this paper.

In view of the discussions mentioned priorly, we aim to explore the robust and non-fragile control problem for non-uniform sampling NCSs, where the TOD protocol is employed in the communication network. To do this, two substantial challenges are specified as follows: 1) how to analyze the exponential stability when the uncertain NCSs with non-uniform sampling and scheduling of the TOD protocol? and 2) how to obtain the maximum bound of the sampling intervals? Our fundamental objective is to give effective solutions to the two questions. Our major work is summarized as below: 1) the stability analysis issue is researched, initially, for non-uniform sampling NCSs with non-fragile robust controllers possessing exponential gains; and 2) a sufficient condition, which explicitly gives the maximum bound of the sampling intervals, is proposed, stabilizing the considered NCSs exponentially.

## 2 Problem Formulation

Consider a linear continuous-time system as follows:

$$\begin{cases} \dot{x}(t) = (A + \Delta A)x(t) + Bu(t), \\ x(0) = x_0, \end{cases} \quad (1)$$

where  $x(t) \in \mathbb{R}^n$  and  $u(t) \in \mathbb{R}^m$  are the state vector and the control input, matrices  $A$  and  $B$  are constant with their dimensions being consistent with the

requirements of the system.  $\Delta A$  describes the parameter uncertainty, which has the following form:

$$\Delta A = E_1 H_1(\sigma) F_1, \quad (2)$$

where  $E_1 \in \mathbb{R}^{n \times n}$  is a fixed matrix with full rank,  $F_1 \in \mathbb{R}^{n \times n}$  is also a fixed matrix, the matrix  $H_1(\sigma)$  is uncertain and satisfies  $H_1^T(\sigma) H_1(\sigma) \leq I$ , in which  $\sigma \in \Psi$ , and  $\Psi$  is a compact set in  $\mathbb{R}$ .

The system state in (1) is sampled under a non-uniform sampling scheme. We denote the sequence of sampling time instants as  $0 = t_0 < t_1 < t_2 < \dots < t_k < \dots$ . Then, the measurement equation is expressed as

$$y(t_k) = \bar{C}x(t_k), \quad (3)$$

where  $y(t_k) \in \mathbb{R}^n$  is the measurement output, and  $\bar{C}$  is the measurement matrix.

To prevent data congestion from the sensors to the remote controller, we introduce the TOD protocol which requires that only one of the sensor nodes can get the chance to transmit its corresponding data (i.e., a component of  $y(t_k)$ ) at each communication instant. Denote the transmitted component of  $y(t_k)$  as  $\xi(t_k)$  at time instant  $t_k$ . In line with the TOD scheduling protocol,  $\xi(t_k)$  is determined as follows:

$$\xi(t_k) = \arg \max_{1 \leq i \leq n} \|y_i(t_k) - \bar{y}_i(t_{k-1})\|, \quad (4)$$

where  $y_i(t_k)$  is the data of the  $i$ -th sensor node,  $\bar{y}_i(t_k)$  is the input signal of the controller at instant  $t_k$  satisfying the following updating rule for  $i = 1, 2, \dots, n$ :

$$\bar{y}_i(t_k) = \begin{cases} y_i(t_k), & \xi(t_k) = i \\ 0, & \text{otherwise.} \end{cases} \quad (5)$$

Define  $\bar{y}(t_k) = [\bar{y}_1(t_k) \ \bar{y}_2(t_k) \ \dots \ \bar{y}_n(t_k)]^T$ . From (5), one has

$$\bar{y}(t_k) = \Phi_{\xi(t_k)} y(t_k), \quad (6)$$

where  $\Phi_{\xi(t_k)} \triangleq \text{diag}_{1 \leq j \leq n} \{\delta(\xi(t_k) - j)I\}$  and  $\delta(\cdot)$  is the Kronecker delta function.

Based on signal  $\bar{y}(t_k)$  in (6), the output feedback controller is constructed as:

$$u(t) = \mathcal{K} \exp\{BK(t - t_k)\} \bar{y}(t_k), \quad t \in [t_k, t_{k+1}) \quad (7)$$

where  $\mathcal{K} \triangleq K + \Delta K$ ,  $K \in \mathbb{R}^{m \times n}$  is the gain matrix, and  $\Delta K \in \mathbb{R}^{m \times n}$  characterizes the fragility of the controller satisfying  $\Delta K = E_2 H_2(\sigma) F_2$ , in which  $E_2$ ,  $F_2$  represent fixed real constant matrices of suitable dimensions,  $H_2$  represents an uncertain matrix satisfying  $H_2^T(\sigma) H_2(\sigma) \leq I$ .

*Remark 1.* The controller with the exponential gain was proposed in [21]. Compared with the traditional controller, such a controller can improve the convergence performance. As such, we utilize this controller to tackle the robust control problem for NCSs under TOD protocol.



*Remark 2.* In practical control systems, parameter perturbations in controller implementation are often unavoidable due to hardware constraints and environmental factors. Non-fragile control aims to design controllers that maintain robustness/effectiveness despite variations in controller gains. In this paper, we introduce  $\Delta K$  to characterize the possible variations of the controller parameter. Obviously, when  $\Delta K = 0$ , controller (7) simplifies to the traditional one.

From (1), (6), and (7), the NCS (1) under the TOD scheduling can be rewritten in the form:

$$\begin{cases} \dot{x}(t) = (A + \Delta A)x(t) + BK \exp\{BK(t - t_k)\} \Phi_{\xi(t_k)} y(t_k), & t \in [t_k, t_{k+1}) \\ y(t_k) = \bar{C}x(t_k). \end{cases} \quad (8)$$

**Definition 1.** [18] Let  $c \geq 1$  and  $\omega > 0$  be constants such that the solution  $x(s)$  to (8) satisfies  $\|x(s)\| \leq c\|x(t_0)\| \exp\{-\omega(s - t_0)\}$  for  $s \geq t_0$ , then system (8) is called exponentially stable.

Our aim is to give a sufficient criterion which can ensure that the system (8) achieves exponential stability.

### 3 Main Results

To give the desired sufficient condition, there is a definition and some lemmas required.

**Definition 2.** [19] A function  $W(\cdot) : \mathbb{C} \rightarrow \mathbb{C}$  is called the Lambert  $W$  function if it satisfies the following condition:

$$W(z) \exp\{W(z)\} = z, \quad \forall z \in \mathbb{C}$$

where  $\mathbb{C}$  represents the set containing all complex numbers.

**Lemma 1.** [20] Given matrices  $G = G^T$ ,  $S$ ,  $R$ , and  $\Sigma$  with proper dimensions,  $G + S\Sigma R + R^T \Sigma^T S^T < 0$  is satisfied for all  $\Sigma^T \Sigma \leq I$  if and only if there exists a scalar  $\vartheta > 0$  such that  $G + \vartheta S S^T + \vartheta^{-1} R^T R < 0$ .

**Lemma 2.** [20] Given real matrices  $P_1$ ,  $P_2$  and  $P_3$  with  $P_1 = P_1^T$  and  $P_3 > 0$ . Then,  $P_1 + P_2 P_3^{-1} P_2^T < 0$  if and only if

$$\begin{bmatrix} P_1 & P_2 \\ P_2^T & -P_3 \end{bmatrix} < 0.$$

**Lemma 3.** [21] Consider a function  $v(\cdot) : [t_0, +\infty) \rightarrow [0, +\infty)$  that satisfies

$$\dot{v}(t) \leq -\lambda_1 v(t) + \lambda_2 \max_{s \in [b_1, t]} v(s), \quad t \in [b_1, b_2]$$

where  $\lambda_1, \lambda_2$  are two positive constants, and  $b_1, b_2$  belong to  $[t_0, +\infty)$ . Let  $\lambda_3 \triangleq \lambda_1 - W(\lambda_2(b_2 - b_1) \exp\{\lambda_1(b_2 - b_1)\}) / (b_2 - b_1)$  with  $W(\cdot)$  being the Lambert  $W$  function. If  $\lambda_1 > \lambda_2$ , then one has  $v(t) \leq \exp\{-\lambda_3(t - b_1)\} v(b_1)$  for  $t \in [b_1, b_2]$ .

**Lemma 4.** *Let a positive scalar  $\varepsilon$ , vectors  $\eta_1, \eta_2 \in \mathbb{R}^n$ , and an  $n \times n$  dimensional matrix  $H > 0$  be given. Then, one has*

$$2\eta_1^T \eta_2 \leq \varepsilon \eta_1^T H^{-1} \eta_1 + \frac{1}{\varepsilon} \eta_2^T H \eta_2. \quad (9)$$

*Proof.* Letting  $\tilde{H} \triangleq \frac{1}{\varepsilon} H$ , then (9) becomes  $2\eta_1^T \eta_2 \leq \eta_1^T \tilde{H}^{-1} \eta_1 + \eta_2^T \tilde{H} \eta_2$ . Defining  $W \triangleq \tilde{H}^{-\frac{1}{2}}(\eta_1 - \tilde{H} \eta_2)$  and noticing  $\tilde{H} = (\tilde{H}^{\frac{1}{2}})^T (\tilde{H}^{\frac{1}{2}})$ , one can easily conclude that

$$W^T W = (\eta_1^T - \eta_2^T \tilde{H}) \tilde{H}^{-1} (\eta_1 - \tilde{H} \eta_2) = \eta_1^T \tilde{H}^{-1} \eta_1 - 2\eta_1^T \eta_2 + \eta_2^T \tilde{H} \eta_2 \geq 0,$$

which finishes the proof of Lemma 4.

**Theorem 1.** *The exponential stability of system (8) can be guaranteed if 1) there exist matrices  $P > 0$ ,  $M > 0$ , scalars  $\varepsilon_1 > 0$ ,  $\varepsilon_2 > 0$ ,  $\varepsilon_3 > 0$  and  $\delta > 3$  such that*

$$\begin{bmatrix} \Omega + \varepsilon_1^{-1} F_1^T F_1 + \varepsilon_2^{-1} F_2^T F_2 & P E_1 & P B E_2 \\ * & -\varepsilon_1^{-1} I & 0 \\ * & * & -\varepsilon_2^{-1} I \end{bmatrix} < 0, \quad (10)$$

$$\begin{bmatrix} -(E_1 E_1^T)^{-1} + \varepsilon_3^{-1} F_2^T F_2 & K^T B^T & 0 \\ * & -P^{-1} & B E_2 \\ * & * & -\varepsilon_3^{-1} I \end{bmatrix} < 0, \quad (11)$$

where  $\Omega = (A + BK)^T P + P(A + BK) + M$ ; and 2) the allowable maximum and minimum transmission intervals (denoted by  $h_{\max}$  and  $h_{\min}$  respectively) satisfy

$$\begin{aligned} h_{\max} &\triangleq \sup_{k \in \mathbb{N}} (t_{k+1} - t_k) < \sqrt{\frac{1}{\kappa} \left( \frac{(\delta - 3)\gamma}{\delta} - \frac{\delta \rho^2}{\gamma \lambda_{\min}(P)} \right)}, \\ h_{\min} &\triangleq \inf_{k \in \mathbb{N}} (t_{k+1} - t_k) \\ &> \frac{\delta}{(\delta - 3)\gamma} W \left( \left( \kappa (h_{\max})^2 + \frac{\delta \rho^2}{\gamma \lambda_{\min}(P)} \right) h_{\max} \exp \left\{ \frac{(\delta - 3)\gamma}{\delta} h_{\max} \right\} \right), \end{aligned} \quad (12)$$

where

$$\begin{aligned} \gamma &\triangleq \frac{\lambda_{\min}(M)}{\lambda_{\max}(P)}, \quad \kappa \triangleq \frac{\delta}{\gamma \lambda_{\min}(P)} \|E_1^{-1} A\|^2 + \frac{\delta}{\gamma \lambda_{\min}(P)} \|F_1\|^2, \\ \rho &\triangleq \max_{\xi(t_k) \in \{1, 2, \dots, n\}} \|E_1^{-1} (I - \Phi_{\xi(t_k)} \bar{C})\|. \end{aligned}$$

*Proof.* Applying Lemma 2 to (10), we have

$$\begin{bmatrix} \Omega + \varepsilon_1^{-1} F_1^T F_1 & P E_1 \\ * & -\varepsilon_1^{-1} I \end{bmatrix} + \varepsilon_2 \begin{bmatrix} P B E_2 \\ 0 \end{bmatrix} \begin{bmatrix} P B E_2 \\ 0 \end{bmatrix}^T + \varepsilon_2^{-1} \begin{bmatrix} F_2^T \\ 0 \end{bmatrix} \begin{bmatrix} F_2^T \\ 0 \end{bmatrix}^T < 0.$$

It can be derived from Lemma 1 that

$$\begin{aligned} & \begin{bmatrix} \Omega + \varepsilon_1^{-1} F_1^T F_1 & P E_1 \\ * & -\varepsilon_1^{-1} I \end{bmatrix} + \begin{bmatrix} P B E_2 \\ 0 \end{bmatrix} H_2(\sigma) \begin{bmatrix} F_2 & 0 \end{bmatrix} \\ & + \begin{bmatrix} F_2^T \\ 0 \end{bmatrix} H_2^T(\sigma) \begin{bmatrix} E_2^T B^T P^T & 0 \end{bmatrix} < 0. \end{aligned}$$

Then, we further have

$$\begin{bmatrix} \Omega + \Delta K^T B^T P + P B \Delta K + \varepsilon_1^{-1} F_1^T F_1 & P E_1 \\ * & -\varepsilon_1^{-1} I \end{bmatrix} < 0.$$

On the basis of the above inequality, by using Lemma 2 for another time, it can be derived that  $\Omega + \Delta K^T B^T P + P B \Delta K + \varepsilon_1 P E_1 (P E_1)^T + \varepsilon_1^{-1} F_1^T F_1 < 0$ , which, together with Lemma 1, gives  $\Omega + \Delta K^T B^T P + P B \Delta K + P E_1 H_1(\sigma) F_1 + F_1^T H_1^T(\sigma) E_1^T P^T < 0$ . Then, from (2), it is evident that

$$\left( (A + \Delta A + B K)^T P + P (A + \Delta A + B K) \right) + M < 0. \quad (13)$$

Letting  $e(t) = x(t) - \exp\{B K(t - t_k)\} \Phi_{\xi(t_k)} \bar{C} x(t_k)$ , from (8), we have that  $\dot{x}(t) = ((A + \Delta A) + B K)x(t) - B K e(t)$ . Noticing that  $\dot{e}(t) = (A + \Delta A)x(t)$  and

$$e(t_k) = (I - \Phi_{\xi(t_k)} \bar{C}) x(t_k), \quad (14)$$

it can be obtained from the integral mean value theorem that

$$e(t) = (t - t_k)(A + \Delta A)x(\zeta) + e(t_k) \quad (15)$$

with  $\zeta \in [t_k, t]$  being an unknown scalar.

The Lyapunov function for system (8) can be formulated as  $V(t) = x^T(t) P x(t)$ . By taking derivative to  $V(t)$ , it is evident that

$$\begin{aligned} \dot{V}(t) &= x(t)^T [A + \Delta A + B K]^T P x(t) + x(t)^T P [A + \Delta A + B K] x(t) \\ &\quad - 2x(t)^T P B K e(t), \quad t \in [t_k, t_{k+1}). \end{aligned} \quad (16)$$

From (13) and (16), one has

$$\dot{V}(t) < -x^T(t) M x(t) - 2x^T(t) P B K e(t) \leq -\gamma V(t) - 2x^T(t) P B K e(t),$$

where  $\gamma = \lambda_{\min}(M)/\lambda_{\max}(P)$ .

Utilizing Lemma 2, it can be inferred from (11) that

$$\begin{bmatrix} -(E_1 E_1^T)^{-1} K^T B^T \\ * & -P^{-1} \end{bmatrix} + \varepsilon_3 \begin{bmatrix} 0 \\ B E_2 \end{bmatrix} \begin{bmatrix} 0 \\ B E_2 \end{bmatrix}^T + \varepsilon_3^{-1} \begin{bmatrix} F_2^T \\ 0 \end{bmatrix} \begin{bmatrix} F_2^T \\ 0 \end{bmatrix}^T < 0,$$

which, together with Lemma 1, leads to

$$\begin{aligned} & \begin{bmatrix} -(E_1 E_1^T)^{-1} K^T B^T \\ * & -P^{-1} \end{bmatrix} + \begin{bmatrix} 0 \\ B E_2 \end{bmatrix} H_2(\sigma) \begin{bmatrix} F_2 & 0 \end{bmatrix} \\ & + \begin{bmatrix} F_2^T \\ 0 \end{bmatrix} H_2^T(\sigma) \begin{bmatrix} 0 & E_2^T B^T \end{bmatrix} < 0. \end{aligned}$$

Recalling the known condition, we have

$$\begin{bmatrix} -(E_1 E_1^T)^{-1} (K + \Delta K)^T B^T \\ * \\ -P^{-1} \end{bmatrix} < 0,$$

by which we can further obtain  $E_1^T \mathcal{K}^T B^T P B \mathcal{K} E_1 < I$  from using Lemma 2.

According to (14) and (15), we have from Lemma 4 that

$$\begin{aligned} & -2x^T(t)PB\mathcal{K}e(t) \\ &= -2x^T(t)PB\mathcal{K}(t-t_k)Ax(\zeta) - 2x^T(t)PB\mathcal{K}(I - \Phi_{\xi(t_k)}\bar{C})x(t_k) \\ & \quad - 2x^T(t)PB\mathcal{K}(t-t_k)E_1H_1F_1x(\zeta) \\ &\leq \frac{3\gamma}{\delta}x^T(t)Px(t) + \frac{\delta}{\gamma}(t-t_k)^2x^T(\zeta)A^T(E_1E_1^T)^{-1}Ax(\zeta) + \frac{\delta}{\gamma}x^T(t_k) \\ & \quad \times (I - \Phi_{\xi(t_k)}\bar{C})^T(E_1E_1^T)^{-1}(I - \Phi_{\xi(t_k)}\bar{C})x(t_k) + \frac{\delta}{\gamma}(t-t_k)^2x^T(\zeta) \\ & \quad \times F_1^TH_1^TH_1F_1x(\zeta) \\ &\leq \frac{3\gamma}{\delta}x^T(t)Px(t) + \frac{\delta}{\gamma}(t-t_k)^2x^T(\zeta)A^T(E_1E_1^T)^{-1}Ax(\zeta) + \frac{\delta}{\gamma}x^T(t_k) \\ & \quad \times (I - \Phi_{\xi(t_k)}\bar{C})^T(E_1E_1^T)^{-1}(I - \Phi_{\xi(t_k)}\bar{C})x(t_k) + \frac{\delta}{\gamma}(t-t_k)^2x^T(\zeta) \\ & \quad \times F_1^TF_1x(\zeta) \\ &\leq \frac{3\gamma}{\delta}V(t) + \left(\frac{\delta}{\gamma\lambda_{\min}(P)}\|E_1^{-1}A\|^2h_{\max}^2 + \frac{\delta}{\gamma\lambda_{\min}(P)}\right. \\ & \quad \times \|E_1^{-1}(I - \Phi_{\xi(t_k)}\bar{C})\|^2 + \frac{\delta}{\gamma\lambda_{\min}(P)}\|F_1\|^2h_{\max}^2) \max_{s \in [t_k, t]} V(s), \end{aligned}$$

which leads to

$$\begin{aligned} \dot{V}(t) &\leq -\gamma V(t) + \frac{3\gamma}{\delta}V(t) + \left(\frac{\delta}{\gamma\lambda_{\min}(P)}\|E_1^{-1}A\|^2h_{\max}^2 + \frac{\delta}{\gamma\lambda_{\min}(P)}\right. \\ & \quad \times \|F_1\|^2h_{\max}^2 + \frac{\delta}{\gamma\lambda_{\min}(P)}\|E_1^{-1}(I - \Phi_{\xi(t_k)}\bar{C})\|^2) \max_{s \in [t_k, t]} V(s) \\ &\leq -\frac{(\delta-3)\gamma}{\delta}V(t) + \left(\kappa h_{\max}^2 + \frac{\delta\rho^2}{\gamma\lambda_{\min}(P)}\right) \max_{s \in [t_k, t]} V(s). \end{aligned} \quad (17)$$

By employing Lemma 3 on (17) and using (12), one can deduce that

$$V(t) \leq \exp\{-\omega(t-t_k)\}V(t_k), \quad t \in [t_k, t_{k+1})$$

where

$$\omega = \frac{(\delta-3)\gamma}{\delta} - \frac{1}{h_k}W\left(\left(\kappa h_{\max}^2 + \frac{\delta\rho^2}{\gamma\lambda_{\min}(P)}\right)h_k \exp\left\{\frac{(\delta-3)\gamma}{\delta}h_k\right\}\right).$$

Letting

$$\tilde{\omega} \triangleq \frac{(\delta - 3)\gamma}{\delta} - \frac{1}{h_{\min}} W\left(\left(\kappa h_{\max}^2 + \frac{\delta \rho^2}{\gamma \lambda_{\min}(P)}\right) h_{\max} \exp\left\{\frac{(\delta - 3)\gamma}{\delta} h_{\max}\right\}\right),$$

it is clear that  $\tilde{\omega} \leq \omega$ . Then, relying on the fact that the Lambert  $W$  function is monotonically increasing over  $[0, +\infty)$ , it can be observed that

$$V(t) \leq \exp\{-\tilde{\omega}(t - t_k)\} V(t_k), \quad t \in [t_k, t_{k+1}). \quad (18)$$

Considering the continuity of  $V(t)$  over  $[t_0, +\infty)$ , from (18), it is clear that

$$V(t) \leq \exp\{-\tilde{\omega}(t - t_0)\} V(t_0), \quad \forall t \geq t_0. \quad (19)$$

With (19), it is easy to see  $p_1 \|x(t)\|^2 \leq p_2 \exp\{-\tilde{\omega}(t - t_0)\} \|x(t_0)\|^2$  for all  $t \geq t_0$ , which means

$$\|x(t)\| \leq c \exp\left\{-\frac{\tilde{\omega}}{2}(t - t_0)\right\} \|x(t_0)\|, \quad t \geq t_0$$

where  $c = \sqrt{p_2/p_1}$ ,  $p_1 = \lambda_{\min}(P)$ ,  $p_2 = \lambda_{\max}(P)$ . Hence, the proof is finished.

## 4 Numerical Example

In this section, to verify the proposed sampled-data control scheme, a comprehensive numerical simulation is conducted, showing its effectiveness in achieving the desired control performance.

Consider a continuous-time system (1) characterized by the following parameters:

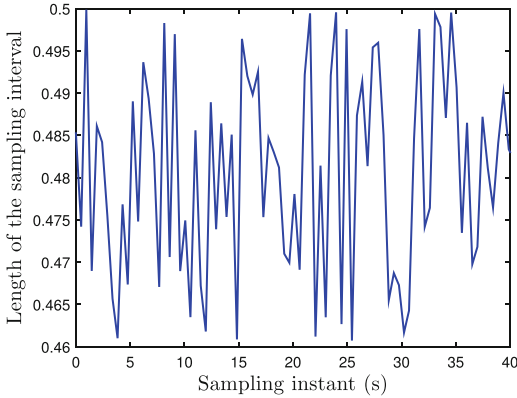
$$A = \begin{bmatrix} -0.8 & 0 \\ 1 & 0.1 \end{bmatrix}, \quad B = \begin{bmatrix} 0.2 & 0 \\ 0 & 0.2 \end{bmatrix}, \quad E_1 = \begin{bmatrix} 200 & 0 \\ 0 & 200 \end{bmatrix},$$

$$F_1 = \begin{bmatrix} 0.001 & 0 \\ 0 & 0.001 \end{bmatrix}, \quad E_2 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \quad F_2 = \begin{bmatrix} 0.1 & 0 \\ 0 & 0.1 \end{bmatrix}.$$

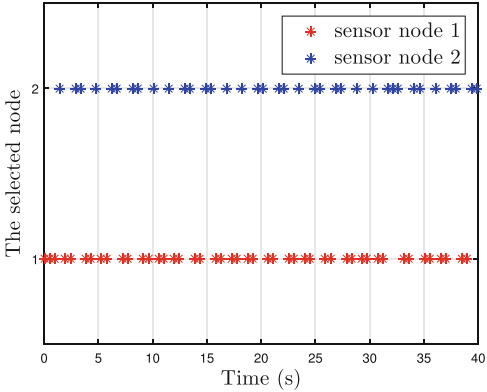
The measurement matrix is set to be  $\bar{C} = [C_1^T, C_2^T]^T$  with  $C_1 = [0.8 \ 0.4]$ ,  $C_2 = [0.4 \ 0.2]$ . The controller gain matrix is designed as

$$K = \begin{bmatrix} -16.6508 & -2.4974 \\ -2.5026 & -21.1508 \end{bmatrix}.$$

By solving LMIs (10) and (11), it can be obtained that  $P = \text{diag}\{0.0095, 0.0095\}$ ,  $M = \text{diag}\{0.0027, 0.0027\}$ ,  $\varepsilon_1 = 0.0040$ ,  $\varepsilon_2 = 281.5534$ , and  $\varepsilon_3 = 41.4957$ . Then, it can be given that  $\delta = 11.3258$ ,  $\gamma = 0.2798$ ,  $\kappa = 0.1797$ , and  $\rho = 0.0057$ . Based



**Fig. 1.** Sampling instants and the relating lengths of the sampling intervals.

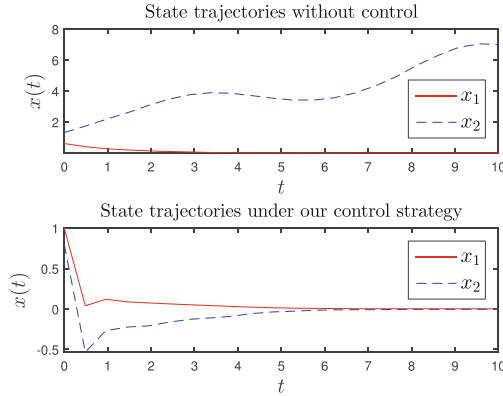


**Fig. 2.** Different sensor nodes chosen to access the network.

on these parameters, the allowable maximum bound of the sampling intervals calculated by Theorem 1 is  $h_{\max} = 0.6050$ .

The non-uniform sampling intervals are selected according to  $0.4600 \leq h_k \leq 0.5000$ ,  $k \in \mathbb{N}$ . Correspondingly, the sampling instants are shown in Fig. 1. Moreover, as illustrated in Fig. 2, different sensor nodes are chosen to access the network at different time instants based on the TOD protocol scheme.

In order to perform the simulation, we set  $x_0 = [1 \ 0.8]^T$ . Figure 3 depicts both the original state trajectories of system (8) without control and those under our control strategy. This illustrates the correctness of our stability criterion.



**Fig. 3.** State trajectories.

## 5 Conclusion

This article has dealt with the robust and non-fragile control issue for a class of uncertain NCSs with non-uniform sampling subject to the TOD protocol. A controller with exponential gain has been adopted for the considered NCSs. By utilizing the Lyapunov theory and the Halanay-like inequality, conditions that guarantee the closed-loop system exponentially stable has been established. Finally, an illustrated example has been presented. In future research, we will further investigate other network-induced issues in non-uniform sampling NCSs.

**Acknowledgments.** This study was funded by the National Natural Science Foundation of China (grant number 62373103) and the Jiangsu Provincial Scientific Research Center of Applied Mathematics (grant number BK20233002).

**Disclosure of Interests.** The authors have no competing interests to declare that are relevant to the content of this article.

## References

1. Hu, Z., Ren, H., Deng, F., Li, H.: Stabilization of sampled-data systems with noisy sampling intervals and packet dropouts via a discrete-time approach. *IEEE Trans. Autom. Control* **67**(6), 3204–3211 (2022)
2. Kao, C.-Y.: An IQC approach to robust stability of aperiodic sampled-data systems. *IEEE Trans. Autom. Control* **61**(8), 2219–2225 (2016)
3. Zhang, W.-A., Dong, H., Guo, G., Yu, L.: Distributed sampled-data  $H_\infty$  filtering for sensor networks with nonuniform sampling periods. *IEEE Trans. Industr. Inf.* **10**(2), 871–881 (2014)

4. Fujioka, H.: Stability analysis of systems with aperiodic sample-and-hold devices. *Automatica* **46**(3), 771–775 (2009)
5. Caravani, P.: Stabilization and disturbance attenuation of uncertain discrete-time linear systems. *Int. J. Syst. Sci.* **29**(11), 1313–1323 (1998)
6. Wang, M., Wang, M., Qiu, J., Wu, L.: Robust Exponential stabilization for sampled-data systems with variable sampling and packet dropouts. *Circuits Syst. Signal Process* **35**(10), 3482–3505 (2016)
7. Lee, T.H., Park, J.H., Lee, S.-M., Kwon, O.: Robust sampled-data control with random missing data scenario. *Int. J. Control* **1787**(9), 1957–1969 (2014)
8. Hu, L.-S., Shi, P., Frank, P.M.: Robust sampled-data control for Markovian jump linear systems. *Automatica* **42**(11), 2025–2030 (2006)
9. Zhang, Z., Zhang, H., Wang, Z., Shan, Q.: Non-fragile exponential  $H_\infty$  control for a class of nonlinear networked control systems with short time-varying delay via output feedback controller. *IEEE Trans. Cybern.* **47**(8), 2008–2019 (2017)
10. Yu, H., Hao, F.: Design of event conditions in event-triggered control systems: a non-fragile control system approach. *IET Control Theory Appl.* **10**(9), 1069–1077 (2016)
11. Yang, Y., He, Y.: Non-fragile observer-based robust control for uncertain systems via aperiodically intermittent control. *Inf. Sci.* **573**, 239–261 (2021)
12. Yang, G.-H., Wang, J.L.: Non-fragile  $H_\infty$  control for linear systems with multiplicative controller gain variations. *Automatica* **37**(5), 727–737 (2001)
13. Walsh, G.C., Ye, H., Bushnell, L.G.: Stability analysis of networked control systems. *IEEE Trans. Control Syst. Technol.* **10**(3), 438–446 (2002)
14. Li, X., Dong, H., Wang, Z., Han, F.: Set-membership filtering for state-saturated systems with mixed time-delays under weighted try-once-discard protocol. *IEEE Trans. Circ. Syst.-II: Express Briefs* **66**(2), 312–316 (2019)
15. Zou, L., Wang, Z., Gao, H.: Set-membership filtering for time-varying systems with mixed time-delays under round-robin and weighted try-once-discard protocols. *Automatica* **74**, 341–348 (2016)
16. Shen, Y., Wang, Z., Shen, B., Alsaadi, F.E., Alsaadi, F.E.: Fusion estimation for multi-rate linear repetitive processes under weighted try-once-discard protocol. *Inf. Fusion* **55**, 281–291 (2020)
17. Zou, L., Wang, Z., Han, Q.-L., Zhou, D.: Ultimate boundedness control for networked systems with try-once-discard protocol and uniform quantization effects. *IEEE Trans. Autom. Control* **62**(12), 6582–6588 (2017)
18. Chen, W.-H., Zheng, W.X.: An improved stabilization method for sampled-data control systems with control packet loss. *IEEE Trans. Autom. Control* **57**(9), 2378–2384 (2012)
19. Corless, R.M., Gonnet, G.H., Hare, D., Jeffrey, D.J., Knuth, D.E.: On the Lambert W function. *Adv. Comput. Math.* **5**(4), 329–359 (1996)



20. Boyd, S., El Ghaoui, L., Feron, E., Balakrishnan, V.: Linear Matrix Inequalities in System and Control Theory. SIAM Studies in Applied Mathematics, Pennsylvania (1994)
21. Zhang, W., Tang, Y., Zheng, W.X., Zou, Y.: Stability of sampled-data systems with packet losses: a nonuniform sampling interval approach. *IEEE Trans. Cybern.* **53**(12), 7648–7658 (2023)



# Tracking Control of 2-DOF Continuum Manipulator Using Fuzzy PID Method

Youchuan Wang<sup>1</sup>, Junfeng Sun<sup>1</sup>, Ancai Zhang<sup>1</sup>(✉), Wenbo Zheng<sup>1</sup>,  
and Haichuan Yang<sup>2</sup>(✉)

<sup>1</sup> School of Automation and Electrical Engineering, Linyi 276000, China  
zhangancai@lyu.edu.cn

<sup>2</sup> Graduate School of Technology, Industrial and Social Sciences, Tokushima University,  
Tokushima 770-8506, Japan  
yokaisen1994@gmail.com

**Abstract.** As a new type of bionic manipulator, the continuum manipulator has a good potential application prospect. This paper tackles the issue of trajectory tracking control for a continuum manipulator possessing two degrees of freedom (DOF). Given that the accuracy of tracking control system designed based on nominal model is commonly affected by modeling errors, external frictions and parameter perturbations. We develop a fuzzy PID control strategy to design the tracking controllers for the manipulator. This control strategy first design PID tracking controllers using the tracking error information. And then, the fuzzy logic technique is utilized to adjust the PID controller parameters adaptively. This enables the continuum manipulator to quickly track the desired motion trajectory. The numerical simulations show the validity of the presented method.

**Keywords:** Continuum manipulator · trajectory tracking · PID control · fuzzy control

## 1 Introduction

Manipulator is a sophisticated mechanical system designed to emulate the functionality of the human arm. It typically consists of the structural frames, driving systems, control systems, and end effectors. And it commonly fabricates from rigid materials. Manipulator can offer high-speed repetitive operations and is applicable for performing rapid and predictable work tasks. However, their rigid construction limits the ability of replicating the nuanced movements of the human arm. This reduces its adaptability for complex environments [1].

With the continuous advancement of technological innovation, the continuum manipulator have been intensively researched. Compared with traditional rigid mechanical manipulator, this kind of manipulator possesses greater flexibility and maneuverability [2]. By virtue of their flexible materials and structures, the continuum manipulator can realize multi-DOF and precise motion control [3]. It is widely used in many

different areas that include logistics, warehousing, medical treatments, and industrial manufacturing [4].

In the application process of continuum manipulator, the primary issue is to solve its tracking control problem. In the past few years, researchers have carried out intensively study on this control problem and developed many effective control methods. For example, a fixed-time sliding mode method in [5], a variable impedance control method in [6], a spatial extended state observer method in [7], a neural network-based adaptive control method in [8]. Although these methods can effectively solve the tracking control of continuum manipulator system, they are all designed based on the system's nominal mathematical mode. As a result, the tracking control performance is commonly affected by modeling errors, external frictions and parameter perturbations. This will seriously hinder its application in practical environments. In order to address this issue, this research has formulated a control strategy based on PID principles.

Owing to its superior performance and ease of design, the PID controller has been widely utilized in many control systems [9]. Nevertheless, the continuum manipulator is a typical nonlinear and strongly coupled mechanical system. The complexity of its dynamic model poses challenges to tune the PID controller parameters. To solve this, a fuzzy control strategy is used to adaptively adjust the parameters of PID controller by the information of tracking error and fuzzy control rules [10]. It ensures the tracking control objective of the continuum manipulator to quickly be achieved.

## 2 Model of 2-DOF Continuum Manipulator

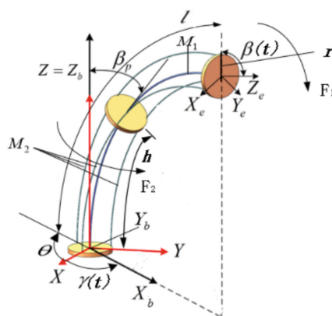


Fig. 1. Model of 2-DOF continuum manipulator

According to the architectural features of the continuum manipulator, it can be considered as a flexible beam that consists of a main skeleton and three sub-skeletons (See Fig. 1). The motion of the manipulator is achieved by applying two forces  $F_1$  and  $F_2$  to two secondary skeletons simultaneously and respectively.

Table 1 gives the physical meaning of parameters in Fig. 1. We choose  $\{X, Y, Z\}$  as the fixed coordinate system. Calculating the kinetic energy  $E_k$  and potential energy  $E_p$  yields

$$\begin{cases} E_k = \frac{1}{6}(4M_1l^2K_1 + 3M_1K_3 + 3M_2K_6)\dot{\beta}^2 + \frac{1}{2}(M_1l^2K_2 + M_1K_5 + M_2K_7)\dot{\gamma}^2 + \frac{1}{2}M_1K_4\dot{\beta}\dot{\gamma}, \\ E_p = 2E \cdot I\beta^2/l, \end{cases} \quad (1)$$

where

$$\begin{aligned} K_1 &= \beta^3 + 6\beta - 12\sin\beta + 6\beta\cos\beta/\beta^5, K_2 = 6\beta - 8\sin\beta + \sin 2\beta/\beta^3, \\ K_3 &= r^2[\cos^2\gamma + \cos^2(-\gamma + \theta) + \cos^2(\gamma + \theta)], K_4 = r^2\beta[-\sin 2\gamma + \sin 2(-\gamma + \theta) - \sin 2(\gamma + \theta)], \\ K_6 &= \frac{h^2\lambda(\lambda+1)(2\lambda+1)}{6\beta^2} + \frac{(2\lambda+1)l^2}{\beta^4} - \frac{l^2}{\beta^4} \frac{\sin\beta + \sin\frac{(\lambda+1)\beta}{\lambda}}{\sin\frac{\beta}{\lambda}} \\ &\quad - \frac{hl}{\beta^3} \frac{(\lambda+1)\sin\beta - \lambda\sin\frac{(\lambda+1)\beta}{\lambda}}{2\sin^2\frac{\beta}{2\lambda}}, \\ K_7 &= \frac{l^2}{\beta^2} \left[ \frac{3\lambda}{2} + \frac{3}{4} - \frac{\sin\beta + \sin\frac{(\lambda+1)\beta}{\lambda}}{\sin\frac{\beta}{\lambda}} + \frac{\sin 2\beta + \sin\frac{2(\lambda+1)\beta}{\lambda}}{4\sin\frac{2\beta}{\lambda}} \right], \end{aligned}$$

and  $E \cdot I$  is the stiffness of the *NiTi* alloy, and  $\lambda = l/h$  represents the aggregate quantity of spacers and end discs.

We choose the Lagrangian function for the continuum manipulator system to be  $L = E_k - E_p$ . Based on the Euler-Lagrange modelling method, it is not difficult to obtain the system's dynamic equations as [11]

$$\begin{bmatrix} D_{11}(x) & D_{12}(x) \\ D_{21}(x) & D_{22}(x) \end{bmatrix} \begin{bmatrix} \ddot{\beta} \\ \ddot{\gamma} \end{bmatrix} + \begin{bmatrix} C_1(x, \dot{x}) \\ C_2(x, \dot{x}) \end{bmatrix} + \begin{bmatrix} G_1(x) \\ 0 \end{bmatrix} = \begin{bmatrix} F_1 \\ F_2 \end{bmatrix}, \quad (2)$$

where  $x = [\beta, \gamma]^T$ , and

$$\begin{aligned} D_{11} &= \frac{4M_1l_2K_1+3M_1K_3+3M_2K_6}{3}, D_{12}(x) = \frac{M_1K_4}{2}, D_{21}(x) = \frac{M_1K_4}{2}, \\ D_{22}(x) &= M_1l^2K_2 + M_1K_5 + M_2K_7, C_1(x, \dot{x}) = C_{11}\dot{\beta}^2 + C_{12}\dot{\beta}\dot{\gamma} + C_{13}\dot{\gamma}^2, \\ C_2(x, \dot{x}) &= C_{21}\dot{\beta}^2 + C_{22}\dot{\beta}\dot{\gamma} + C_{23}\dot{\gamma}^2, C_{11}(x) = \frac{2}{3}M_1l^2\frac{\partial K_1}{\partial \beta} + \frac{1}{2}M_1\frac{\partial K_3}{\partial \beta} + \frac{1}{2}M_2\frac{\partial K_6}{\partial \beta}, \\ C_{12}(x) &= M_1\frac{\partial K_3}{\partial \beta}, C_{13}(x) = -\frac{1}{2}\left(M_1l^2\frac{\partial K_2}{\partial \beta} + M_1\frac{\partial K_5}{\partial \beta} + M_2\frac{\partial K_7}{\partial \beta} - M_1\frac{\partial K_4}{\partial \gamma}\right), \\ C_{21}(x) &= \frac{1}{2}\left(M_1\frac{\partial K_4}{\partial \beta} - M_1\frac{\partial K_3}{\partial \gamma}\right), C_{22}(x) = M_1l^2\frac{\partial K_2}{\partial \beta} + M_1\frac{\partial K_5}{\partial \beta} + M_2\frac{\partial K_7}{\partial \beta}, \\ C_{23}(x) &= \frac{M_1}{2}\frac{\partial K_5}{\partial \gamma}, G_1(x) = \frac{4E \cdot I}{l}\beta. \end{aligned}$$

For the nonlinear control system (2), the tracking control problem is commonly discussed. That is, a controller  $F = [F_1, F_2]^T$  is needed to design to drive the state variable  $x(t)$  track a desired trajectories  $x_d(t) = [\beta_d(t), \gamma_d(t)]^T$ .

**Table 1.** Variables of continuum manipulator

| Parameters                | Meanings                                                             |
|---------------------------|----------------------------------------------------------------------|
| $F_1, F_2$                | the driving forces                                                   |
| $M_1, M_2$                | the principal skeleton mass and sub-skeleton mass                    |
| $l$                       | the length of the flexible arm length                                |
| $r$                       | the distance from the main backbone to each sub-skeleton on the disk |
| $h$                       | the distance between adjacent disks                                  |
| $\theta = \frac{2}{3}\pi$ | the division angle                                                   |
| $\beta$                   | the bending angle at the end disk                                    |
| $\gamma$                  | the rotation angle of the bending plane                              |
| $\{X, Y, Z\}$             | the base coordinate system                                           |
| $\{X_b, Y_b, Z_b\}$       | the bending plane coordinate system                                  |
| $\{X_e, Y_e, Z_e\}$       | the end disk coordinate system                                       |

### 3 Design of Fuzzy PID Tracking Controller

The system's tracking error is labeled as  $e(t) = x(t) - x_d(t)$ . To make  $e(t) \rightarrow 0$ , the current most methods design the tracking control system based on nominal model (2). However, the tracking control accuracy is easily affected by modeling errors, external frictions and parameter perturbations. To address this issue, this section develops a fuzzy PID tracking control method. The method has the advantages of both traditional PID method and fuzzy logic method [12]. It can dynamically modify the PID controller's parameters based on the pre-established fuzzy logic rules and the system's current performance. This increases the flexibility and resilience of control system [13].

Firstly, we design the following PID controllers for (2) as

$$\begin{cases} F_1 = K_{P1}e_1(t) + K_{I1} \int e_1(t)dt + K_{D1} \dot{e}_1(t), \\ F_2 = K_{P2}e_2(t) + K_{I2} \int e_2(t)dt + K_{D2} \dot{e}_2(t). \end{cases} \quad (3)$$

where  $K_{Pj}$ ,  $K_{Ij}$  and  $K_{Dj}$  ( $j = 1, 2$ ) are positive constants. In order to guarantee the continuum manipulator quickly and smoothly track the desired trajectory, a fuzzy control system is designed to adjust the parameters  $K_{Pj}$ ,  $K_{Ij}$  and  $K_{Dj}$  in  $F_j$  ( $j = 1, 2$ ). We take  $F_1$  as an example to explain the design of fuzzy control system, whose structural diagram is shown in Fig. 2. It has two input ( $e_1(t)$ ,  $\dot{e}_1(t)$ ) and three outputs ( $K_{P1}$ ,  $K_{I1}$ ,  $K_{D1}$ ). When the entire control system runs, the input variables  $e_1(t)$  and  $\dot{e}_1(t)$  are fuzzified first. Based on fuzzy rule library, the values of  $K_{P1}$ ,  $K_{I1}$ ,  $K_{D1}$  are obtained by fuzzy inference and defuzzification processes. We elaborate on the design of each part in detail below.

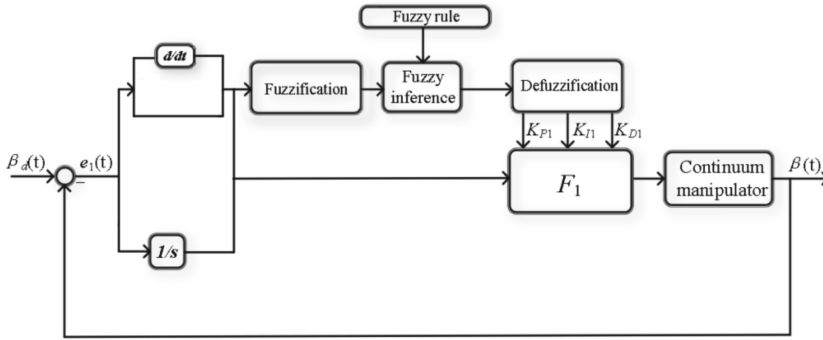


Fig. 2. Structural diagram of fuzzy PID control system.

### 3.1 Input Variables Fuzzification

Fuzzification is a critical step in converting the input variable values into fuzzy values in the fuzzy control system. this transformation is achieved through membership functions. it can quantitatively elucidate the relationship between input values and their corresponding fuzzy values.

In order to get the fuzzy values of  $e_1(t)$  and  $\dot{e}_1(t)$ , the fuzzification rules are designed by using the standard method as in [14], which consist of seven fuzzy subsets: {NB (negative large), NM (negative medium), NS (negative small), ZO (zero), PS (positive small), PM (positive medium), PB (positive large)}. In addition, the universes of  $e_1(t)$  and  $\dot{e}_1(t)$  are set to be  $[-300, 300]$ . And the membership function of these two variables are all chosen to be a bell-shaped Gaussian form function (see Fig. 3).

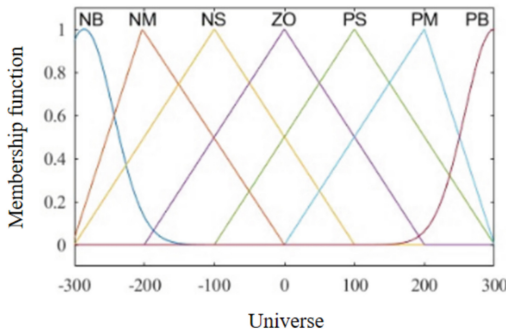


Fig. 3. Membership function of input variables.

### 3.2 Fuzzy Control Rules

The primary key step in designing fuzzy PID parameters is to construct fuzzy rules. Normally, the control rules use conditional statements to construct the relationship between

the input and the output fuzzy sets [15]. Here, the universes and membership functions of  $K_{P1}$ ,  $K_{I1}$  and  $K_{D1}$  are chosen as that of  $e_1(t)$  and  $\dot{e}_1(t)$ .

The structure of fuzzy control rules adopts “If-Then” type. The trigger conditions of  $e_1(t)$  and  $\dot{e}_1(t)$  are defined in “If” components. And the response actions of  $K_{P1}$ ,  $K_{I1}$  and  $K_{D1}$  are defined in “Then” components. Based on the motion characteristics of continuum manipulator, we design a comprehensive set of fuzzy control rules (see Table 2).

**Table 2.** Fuzzy control rules.

| $K_{P1}/K_{I1}/K_{D1}$ |    | $\dot{e}_1(t)$ |          |          |          |          |          |          |
|------------------------|----|----------------|----------|----------|----------|----------|----------|----------|
|                        |    | NB             | NM       | NS       | ZO       | PS       | PM       | PB       |
| $e_1(t)$               | NB | PB/NM/PM       | PM/NM/NS | PM/NB/NM | PB/NB/NM | PS/NS/NM | ZO/ZO/PS | ZO/NS/PS |
|                        | NM | PM/NM/PM       | PB/NB/NS | PM/NM/NB | PS/NS/NM | PS/NB/NB | ZO/ZO/NM | NM/ZO/ZO |
|                        | NS | PB/NB/ZO       | PM/NB/ZO | PB/NM/NB | PS/NS/NM | ZO/ZO/ZO | NS/PS/NS | NS/PS/ZO |
|                        | ZO | PM/NB/ZO       | PB/NB/NS | PM/NM/NM | ZO/ZO/ZO | NS/PS/NS | NB/PB/ZO | NM/PM/ZO |
|                        | PS | PS/NM/ZO       | ZO/ZO/ZO | ZO/NS/ZO | NS/ZO/ZO | NM/PM/ZO | NB/PB/ZO | NM/PB/ZO |
|                        | PM | PS/ZO/PB       | ZO/PS/NS | NB/PB/PM | NM/PS/PS | NM/PS/PS | NB/PM/PS | NB/PB/PS |
|                        | PB | ZO/ZO/PM       | ZO/PS/PB | NB/PM/PM | NM/PB/PM | NB/PB/PM | NM/PM/PM | NM/PM/PM |

### 3.3 Defuzzification

The control system requires precise control signal rather than fuzzy one [16]. So, it is essential to convert the fuzzy values into actual values.  $K_{P1}$ ,  $K_{I1}$  and  $K_{D1}$ . This section uses centroid method to achieve the defuzzification of output fuzzy variables. This method obtain the precise value of fuzzy variable by calculating the centroid of the fuzzy set [17]. The process of defuzzification can be described by

$$u = \frac{\sum \mu(x_i)x_i}{\sum \mu(x_i)}, \quad (4)$$

where  $u$  represents the precise value after defuzzification;  $x_i$  is the element of fuzzy variable, and  $\mu(x_i)$  is the membership degree of  $x_i$ .

## 4 Numerical Examples

To assess the efficacy of our newly developed control technique, this section conducts numerical simulation studies. The physical parameters of continuum manipulator used in simulations are

$$\begin{cases} M_1 = 1.086 \times 10^{-3} \text{ kg}, M_2 = 0.848 \times 10^{-3} \text{ kg}, h = 0.015 \text{ m}, r = 0.003 \text{ m}, \\ E = 65 \times 10^9 \text{ Pa}, I = 0.102 \times 10^{-12} \text{ m}^4, l = 0.15 \text{ m}. \end{cases} \quad (5)$$

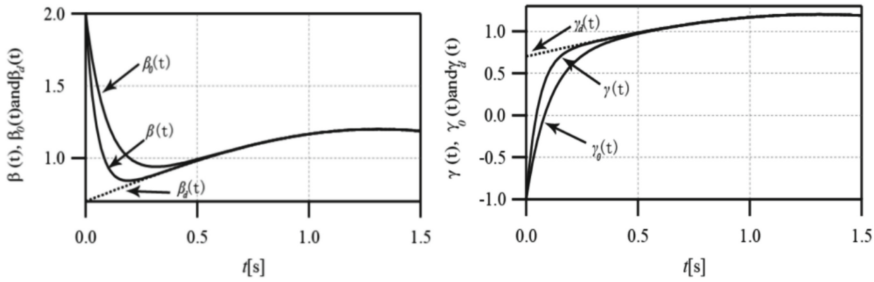
Note that the parameters  $K_1$ ,  $K_2$ ,  $K_6$  and  $K_7$  have singularities when  $\beta = 0$ . To address this issue, we use the least squares method to fit these parameters and obtain

$$\begin{cases} K_1 = -0.00426x_1^2 - 0.0027x_1 + 0.15085, \\ K_2 = -0.05567x_1^3 + 0.2328x_1^2 + 0.006216x_1 - 0.00406, \\ K_6 = -0.00043x_1^2 - 0.00031x_1 + 0.01435, \\ K_7 = -0.00394x_1^3 + 0.01575x_1^2 + 0.00131x_1 - 0.00047. \end{cases} \quad (6)$$

Moreover, the desired trajectories are assumed to be  $\beta_d(t) = \gamma_d(t) = \sin t$ . And the initial values of the manipulator system is chosen as

$$[\beta(0) \dot{\beta}(0) \gamma(0) \dot{\gamma}(0)]^T = [2 \ 2 \ 1 \ -1]. \quad (7)$$

Under these conditions, the simulation experimental results are given in Fig. 4, where  $\beta(t)$  and  $\gamma(t)$  are the fuzzy PID controller responses,  $\beta_0(t)$  and  $\gamma_0(t)$  are the traditional PID controller responses with  $K_{P1} = K_{P2} = 2.2 \times 10^6$ ,  $K_{I1} = K_{I2} = 20$ ,  $K_{D1} = K_{D2} = 2 \times 10^5$ . Note that our designed fuzzy PID control method enables the continuum manipulator quickly to track the desired trajectory. The convergence time of tracking error is less than 0.25 s. It is faster than that of using traditional PID method.



**Fig. 4.** Simulations of continuum manipulator by fuzzy PID and traditional PID methods.

## 5 Conclusion

For a 2-DOF continuum manipulator, the tracking control issue was examined in this research. A method for fuzzy PID control has been introduced. This technique is capable of dynamically modifying the PID controller parameters based on the system's instantaneous feedback and the fuzzy logic rules. It enables the continuum manipulator to quickly and accurately track the desired trajectory even though the modeling errors, external frictions and parameter perturbations exist. On the base of results in this study, we will further explore other model-free tracking control methods for the continuum manipulator in the future.

**Acknowledgments.** This work was partially supported by Taishan Scholar Program of Shandong Province (No. Tsqn202211240), by Natural Science Foundation Program of Shandong Province (No. ZR2024MF041), by Development Plan for Youth Innovation Teams in Universities in Shandong Province (No. 2023KJN049).



**Disclosure of Interests.** There are no potential interests to influence the results presented in this study.

## References

1. Xiang, Z.: Optimal design and dynamic simulation of mountain bike with rear suspension. *Chinese J. Mech. Eng.* **22**(01), 21–26 (2009)
2. Wang, Y., Zhou, Y., Wei, L., Ruya, L.: Design of a four-axis robot arm system based on machine vision. *Appl. Sci.* **13**(15), 8836 (2023)
3. Shoani, M., Ribuan, M., Faudzi, M.: Design and performance of a slender soft continuum manipulator for tall structure inspection. *Ind. Robot. Int. J. Robot. Res. Appl.* **50**(6), 944–958 (2023)
4. Russo, M., Sadati, S., Dong, X.: Continuum robots: an overview. *Adv. Intell. Syst.* **5**(5), 2200367 (2023)
5. Sun, J., Zhang, A., Wang, Z., Lei, Y., Pang, G.: Tracking control for 2-DOF continuum manipulator based on fixed-time sliding mode method. *The 14th Asian Control Conference*, Dalian, China, pp. 2254–228 (2024)
6. Spyrakos-Papastavridis, E., Dai, J.: Minimally model-based trajectory tracking and variable impedance control of flexible-joint robots. *IEEE Trans. Industr. Electron.* **68**(7), 6031–6041 (2020)
7. Fan, C., Liu, Y., Li, C., Yu, C., Liu, H.: Manipulator trajectory tracking of fuzzy control based on spatial extended state observer. *IEEE Access* **8**, 24296–24308 (2020)
8. Xu, K., Wang, Z.: The design of a neural network-based adaptive control method for robotic arm trajectory tracking. *Neural Comput. Appl.* **35**(12), 8785–8795 (2023)
9. Bruder, D., Fu, X., Gillespie, R.B., Remy, C.D., Vasudevan, R.: Koopman-based control of a soft continuum manipulator under variable loading conditions. *IEEE Robot. Autom. Lett.* **6**(4), 6852–6859 (2021)
10. Benbouhenni, H., Bizon, N., Mosaad, M.I.: Enhancement of the power quality of DFIG-based dual-rotor wind turbine systems using fractional order fuzzy controller. *Expert Syst. Appl.* **238**, 121695 (2024)
11. He, B., Wang, Z., Qiang, L., Hong, X., She, R.: An analytic method for the kinematics and dynamics of a multiple-backbone continuum robot. *Int. J. Adv. Rob. Syst.* **10**(84), 257–271 (2013)
12. Li, Y., Ang, K.H., Chong, G.C.Y.: PID control system analysis and design. *IEEE Control. Syst. Mag.* **26**(1), 32–41 (2006)
13. Parvaresh, A., Moosavian, S.: Modeling and model-free fuzzy control of a continuum robotic arm. In: *The 6th RSI International Conference on Robotics and Mechatronics*, Tehran, Iran, pp.501–506 (2018)
14. Precup, R.E., Nguyen, A.T., Blažič, S.: A survey on fuzzy control for mechatronics applications. *Int. J. Syst. Sci.* **55**(4), 771–813 (2024)
15. Ma, X.J., Sun, Z.Q., He, Y.Y.: Analysis and design of fuzzy controller and fuzzy observer. *IEEE Trans. Fuzzy Syst.* **6**(1), 41–51 (1998)
16. Nguyen, A.T., Taniguchi, T., Eciolaza, L.: Fuzzy control systems: past, present and future. *IEEE Comput. Intell. Mag.* **14**(1), 56–68 (2019)
17. Wang, L.: Stable adaptive fuzzy control of nonlinear systems. *IEEE Trans. Fuzzy Syst.* **1**(2), 146–155 (1993)

# Author Index

## A

Aiju, Li I-3

## B

Bi, Ying II-268

## C

Cai, Jiazhen I-319  
Cai, Shipai II-292  
Cao, Mingwei I-254  
Chen, Hang II-170  
Chen, Ken I-186  
Chen, Qili I-243, I-265  
Chen, Xiangyong II-55, II-161  
Chen, Zhuxiang I-75  
Cheng, Shaoqi I-272  
Chu, Zhongyi I-319  
Cong, Lanmei I-296  
Cui, Huixia I-152  
Cui, Jing I-319

## D

Deng, Fang II-222  
Du, Lipeng I-243  
Du, Pengcheng II-37  
Du, Yanwei I-214  
Du, Yu II-37

## F

Feng, Jiayue I-272  
Feng, Xiaoliang II-89  
Fu, Deqian II-182  
Fu, Yijun II-268  
Fu, Zunwei II-20, II-47

## G

Gao, Ji Xian I-144  
Gao, Jixian I-231  
Gong, Ping II-215  
Gong, Shuli II-235

Guo, Ming II-55  
Guo, Weifeng II-268  
Guo, Zijie II-255

## H

Hai, Zhao II-280  
Han, Weixin I-131  
Han, Ziyue I-296  
Hao, Li II-280  
He, Liping II-134, II-198  
He, Yuesheng I-63, II-149  
Hu, Ningning I-296  
Huo, Qingyun I-27

## J

Ji, Yang I-88  
Jia, Chun II-246  
Jia, Senping I-152  
Jiang, Du I-75, II-74  
Junhong, Zhang II-280

## K

Kang, Yuntong I-254

## L

Lei, Chen II-134, II-198  
Lei, Yidi II-170  
Letian, Zhao II-280  
Li, Boao II-74  
Li, Fanjun I-272  
Li, Jianmin I-283  
Li, Meng I-231  
Li, Rui II-120  
Li, Wenjie II-111  
Li, Yuxuan I-27  
Li, Zhangliang II-182  
Li, Zhaoxing I-178  
Liang, Jing II-268  
Liang, Jinling II-292  
Liang, Tian I-117

Lin, Wenshuai II-255  
 Liu, Jianhua II-161  
 Liu, Longjing II-198  
 Liu, Mingtao I-27  
 Liu, Shufen II-161  
 Liu, Ying II-74  
 Liu, Yunzhuo II-99  
 Liu, Ziqi II-182  
 Long, Liu II-280  
 Lyu, Yi I-186

## M

Ma, Bing I-15  
 Ma, Jianliang II-222  
 Ma, Xiaolei II-170  
 Ma, Zhaowei II-99  
 Mao, Yongqiang I-296  
 Ming, Lei I-308

## N

Niu, Yifeng II-99

## P

Pan, Guangyuan I-54, I-243  
 Pan, Qiang I-15  
 Peng, Jinzhu I-231

## Q

Qi, Mingdong I-214  
 Qiao, Xunyu I-96  
 Qiao, Zhongli II-182  
 Qin, Yaoyao II-198  
 Qingping, Xie I-308  
 Qiu, Jianlong II-55, II-182  
 Qiu, Sen I-254

## S

Shan, Liqiu I-39  
 Shan, Wu I-117  
 Shan, Xiaohan I-283  
 Shen, Huang I-88  
 Shen, Xinyu I-272  
 Shi, Shaoguang II-66  
 Sun, Junfeng II-304  
 Sun, Wenqi I-54

## T

Tang, Xinyang I-165  
 Tao, Bo I-75  
 Tian, Yunlong I-39

## W

Wang, Bing I-165  
 Wang, Chunxiang I-165  
 Wang, Dawei I-186  
 Wang, Fei I-63, II-149  
 Wang, Haozhu II-74  
 Wang, Jinjiang I-198  
 Wang, Keqing II-182  
 Wang, Mengyun II-99  
 Wang, Shuo II-89  
 Wang, Tong I-243  
 Wang, Xing I-39  
 Wang, Xixin II-66  
 Wang, Yaonan I-231  
 Wang, Yifang II-27  
 Wang, Yizhen II-47  
 Wang, Youchuan II-304  
 Wang, Yunyu I-27  
 Wang, Zhelong I-254  
 Wang, Zheng II-20  
 Wang, Zhujun II-235  
 Wei, Yuan I-283  
 Wen, Qinglin II-246  
 Wu, Chengcheng I-186  
 Wu, Mingkuo I-231

## X

Xue, Jinqiu II-3

## Y

Yajing, Li I-308  
 Yang, Chenxi I-165  
 Yang, Haichuan II-304  
 Yang, Haobin I-96  
 Yang, Jian II-222  
 Yang, Meng I-27  
 Yang, Ming I-63, I-165, II-149  
 Yang, Yicong I-265  
 Yang, Yuqi II-99  
 Yao, Haoran I-254  
 Yao, Shuanglong I-39  
 Yao, Ximing I-15  
 Yao, Yuan I-15  
 Yining, Du I-3

Yu, Hongliang [II-55](#)  
Yu, Junyao [II-235](#)  
Yuan, Wei [I-165](#)  
Yue, Zhengyuan [II-182](#)  
Yun, Juntong [I-75](#), [II-74](#)  
Yusen, Zhao [I-117](#)

**Z**

Zhang, Ancai [I-54](#), [II-235](#), [II-304](#)  
Zhang, Congcong [II-47](#), [II-66](#)  
Zhang, Fang Fang [I-144](#)  
Zhang, Guoqiang [I-198](#)  
Zhang, Haiying [I-96](#)  
Zhang, Hangwei [I-15](#)  
Zhang, Heng [I-214](#)  
Zhang, Jiachang [II-246](#)  
Zhang, Jiage [II-134](#)  
Zhang, Jiyuan [II-292](#)  
Zhang, Lele [II-222](#)  
Zhang, Rui [II-120](#)  
Zhang, Wenyin [II-3](#)

Zhang, Xin [I-27](#)  
Zhang, Xingzhao [I-178](#)  
Zhang, Yiman [I-186](#)  
Zhang, Yingxin [I-152](#)  
Zhang, Yuan [I-131](#)  
Zhang, Zhaohui [I-296](#)  
Zhang, Zhengqiang [II-111](#)  
Zhang, Zilong [II-255](#)  
Zhao, Feng [II-161](#)  
Zhao, Letian [II-170](#)  
Zhao, Lin [I-243](#)  
Zhao, Nan [I-231](#)  
Zheng, Wenbo [I-54](#), [II-304](#)  
Zheng, Xiaohong [II-255](#)  
Zhou, Kun [II-55](#)  
Zhou, Qi [I-198](#)  
Zhou, Xinhao [I-243](#)  
Zhou, Yangyang [II-170](#)  
Zhu, Yunlong [II-20](#)  
Zhu, Zhen Yu [I-144](#)  
Zong, Zihan [I-63](#), [II-149](#)