

Recent Advances on Mathematical Modeling and Control Methods for Complex Vehicle Systems

Guest Editors: Hamid Reza Karimi, Zheping Yan, Yuxin Zhao, Weichao Sun, and Mohammed Chadli





Recent Advances on Mathematical Modeling and Control Methods for Complex Vehicle Systems

Journal of Applied Mathematics

Recent Advances on Mathematical Modeling and Control Methods for Complex Vehicle Systems

Guest Editors: Hamid Reza Karimi, Zheping Yan, Yuxin Zhao,
Weichao Sun, and Mohammed Chadli



Copyright © 2014 Hindawi Publishing Corporation. All rights reserved.

This is a special issue published in "Journal of Applied Mathematics." All articles are open access articles distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Editorial Board

- Saeid Abbasbandy, Iran
Mina B. Abd-El-Malek, Egypt
Mohamed A. Abdou, Egypt
Subhas Abel, India
Mostafa Adimy, France
Carlos J. S. Alves, Portugal
Mohamad Alwash, USA
Igor Andrianov, Germany
Sabri Arik, Turkey
Francis T.K. Au, Hong Kong
Olivier Bahn, Canada
Roberto Barrio, Spain
Alfredo Bellen, Italy
Jafar Biazar, Iran
Hester Bijl, The Netherlands
Anjan Biswas, Saudi Arabia
Stephane P.A. Bordas, USA
James Robert Buchanan, USA
Alberto Cabada, Spain
Xiao Chuan Cai, USA
Jinde Cao, China
Alexandre Carvalho, Brazil
Song Cen, China
Qianshun S. Chang, China
Tai-Ping Chang, Taiwan
Shih-sen Chang, China
Rushan Chen, China
Xinfu Chen, USA
Ke Chen, UK
Eric Cheng, Hong Kong
Francisco Chiclana, UK
Jen-Tzung Chien, Taiwan
C. S. Chien, Taiwan
Han H. Choi, Republic of Korea
Tin-Tai Chow, China
M. S. H. Chowdhury, Malaysia
Carlos Conca, Chile
Vitor Costa, Portugal
Livija Cveticanin, Serbia
Eric de Sturler, USA
Orazio Descalzi, Chile
Kai Diethelm, Germany
Vit Dolejsi, Czech Republic
Bo-Qing Dong, China
Magdy A. Ezzat, Egypt
- Meng Fan, China
Ya Ping Fang, China
Antonio J. M. Ferreira, Portugal
Michel Fliess, France
M. A. Fontelos, Spain
Huijun Gao, China
Bernard J. Geurts, The Netherlands
Jamshid Ghaboussi, USA
Pablo González-Vera, Spain
Laurent Gosse, Italy
K. S. Govinder, South Africa
Jose L. Gracia, Spain
Yuantong Gu, Australia
Zhihong GUAN, China
Nicola Guglielmi, Italy
Frederico G. Guimarães, Brazil
Vijay Gupta, India
Bo Han, China
Maoan Han, China
Pierre Hansen, Canada
Ferenc Hartung, Hungary
Xiaoqiao He, Hong Kong
Luis Javier Herrera, Spain
J. Hoenderkamp, The Netherlands
Ying Hu, France
Ning Hu, Japan
Zhilong L. Huang, China
Kazufumi Ito, USA
Takeshi Iwamoto, Japan
George Jaiani, Georgia
Zhongxiao Jia, China
Tarun Kant, India
Ido Kanter, Israel
Abdul Hamid Kara, South Africa
Hamid Reza Karimi, Norway
Jae-Wook Kim, UK
Jong Hae Kim, Republic of Korea
Kazutake Komori, Japan
Fanrong Kong, USA
Vadim . Krysko, Russia
Jin L. Kuang, Singapore
Miroslaw Lachowicz, Poland
Hak-Keung Lam, UK
Tak-Wah Lam, Hong Kong
PGL Leach, Cyprus
- Yongkun Li, China
Wan-Tong Li, China
J. Liang, China
Ching-Jong Liao, Taiwan
Chong Lin, China
Mingzhu Liu, China
Chein-Shan Liu, Taiwan
Kang Liu, USA
Yansheng Liu, China
Fawang Liu, Australia
Shutian Liu, China
Zhijun Liu, China
Julián López-Gómez, Spain
Shiping Lu, China
Gert Lube, Germany
Nazim Idrisoglu Mahmudov, Turkey
Oluwole Daniel Makinde, South Africa
Francisco J. Marcellán, Spain
Guiomar Martín-Herrán, Spain
Nicola Mastronardi, Italy
Michael McAleer, The Netherlands
Stephane Metens, France
Michael Meylan, Australia
Alain Miranville, France
Ram N. Mohapatra, USA
Jaime E. Munoz Rivera, Brazil
Javier Murillo, Spain
Roberto Natalini, Italy
Srinivasan Natesan, India
Jiri Nedoma, Czech Republic
Jianlei Niu, Hong Kong
Roger Ohayon, France
Javier Oliver, Spain
Donal O'Regan, Ireland
Martin Ostoja-Starzewski, USA
Turgut Öziş, Turkey
Claudio Padra, Argentina
Reinaldo Martinez Palhares, Brazil
Francesco Pellicano, Italy
Juan Manuel Peña, Spain
Ricardo Perera, Spain
Malgorzata Peszynska, USA
James F. Peters, Canada
Mark A. Petersen, South Africa
Miodrag Petkovic, Serbia

Vu Ngoc Phat, Vietnam
Andrew Pickering, Spain
Hector Pomares, Spain
Maurizio Porfiri, USA
Mario Primicerio, Italy
Morteza Rafei, The Netherlands
Roberto Renò, Italy
Jacek Rokicki, Poland
Dirk Roose, Belgium
Carla Roque, Portugal
Debasish Roy, India
Samir H. Saker, Egypt
Marcelo A. Savi, Brazil
Wolfgang Schmidt, Germany
Eckart Schnack, Germany
Mehmet Sezer, Turkey
Naseer Shahzad, Saudi Arabia
Fatemeh Shakeri, Iran
Jian Hua Shen, China
Hui-Shen Shen, China
Fernando Simões, Portugal
Theodore E. Simos, Greece

Abdel-Maksoud A. Soliman, Egypt
Xinyu Song, China
Qiankun Song, China
Yuri N. Sotskov, Belarus
Peter J. C. Spreij, The Netherlands
Niclas Strömberg, Sweden
Ray KL Su, Hong Kong
Jitao Sun, China
Wenyu Sun, China
XianHua Tang, China
Alexander Timokha, Norway
Mariano Torrisi, Italy
Jung-Fa Tsai, Taiwan
Ch Tsitouras, Greece
Kuppalapalle Vajravelu, USA
Alvaro Valencia, Chile
Erik Van Vleck, USA
Ezio Venturino, Italy
Jesus Vigo-Aguiar, Spain
Michael N. Vrahatis, Greece
Baolin Wang, China
Mingxin Wang, China

Qing-Wen Wang, China
Guangchen Wang, China
Junjie Wei, China
Li Weili, China
Martin Weiser, Germany
Frank Werner, Germany
Shanhe Wu, China
Dongmei Xiao, China
Gongnan Xie, China
Yuesheng Xu, USA
Suh-Yuh Yang, Taiwan
Bo Yu, China
Jinyun Yuan, Brazil
Alejandro Zarzo, Spain
Guisheng Zhai, Japan
Jianming Zhan, China
Zhihua Zhang, China
Jingxin Zhang, Australia
Shan Zhao, USA
Chongbin Zhao, Australia
Renat Zhdanov, USA
Hongping Zhu, China

Contents

Recent Advances on Mathematical Modeling and Control Methods for Complex Vehicle Systems,

Hamid Reza Karimi, Zheping Yan, Yuxin Zhao, Weichao Sun, and Mohammed Chadli

Volume 2014, Article ID 872860, 2 pages

Robust Stability Clearance of Flight Control Law Based on Global Sensitivity Analysis, Liuli Ou, Lei Liu,

Shuai Dong, and Yongji Wang

Volume 2014, Article ID 153602, 10 pages

Mathematical Model of Movement of the Observation and Tracking Head of an Unmanned Aerial Vehicle Performing Ground Target Search and Tracking, Izabela Krzysztofik and Zbigniew Koruba

Volume 2014, Article ID 934250, 11 pages

Optimal Acceleration-Velocity-Bounded Trajectory Planning in Dynamic Crowd Simulation,

Fu Yue-wen, Li Meng, Liang Jia-hong, and Hu Xiao-qian

Volume 2014, Article ID 501689, 12 pages

Real-Time Control Strategy of Elman Neural Network for the Parallel Hybrid Electric Vehicle,

Ruijun Liu, Dapai Shi, and Chao Ma

Volume 2014, Article ID 596326, 11 pages

The Burgers Equation for a New Continuum Model with Consideration of Driver's Forecast Effect,

Lei Yu and Bingchang Zhou

Volume 2014, Article ID 539459, 7 pages

3D Road Scene Monitoring Based on Real-Time Panorama, Yuezhou Wu, Changjiang Liu, Shiyong Lan, and Menglong Yang

Volume 2014, Article ID 403126, 8 pages

An Aeromagnetic Compensation Algorithm for Aircraft Based on Fuzzy Adaptive Kalman Filter,

Yao Zhang, Yu-Xin Zhao, and Shuai Chang

Volume 2014, Article ID 405671, 9 pages

Research on Wavelet Singularity Detection Based Fault-Tolerant Federated Filtering Algorithm for INS/GPS/DVL Integrated Navigation System, Jianhua Cheng, Daidai Chen, Rene Landry Jr., Lin Zhao, and Dongxue Guan

Volume 2014, Article ID 825726, 8 pages

Probabilistic Harmonic Calculation in Distribution Networks with Electric Vehicle Charging Stations,

Jianxue Wang, Yanlin Cui, and Minghui Zhu

Volume 2014, Article ID 167565, 11 pages

Multifeature Fusion Vehicle Detection Algorithm Based on Choquet Integral, Wenhui Li, Peixun Liu,

Ying Wang, and Hongyin Ni

Volume 2014, Article ID 701058, 11 pages

Quad-Rotor Helicopter Autonomous Navigation Based on Vanishing Point Algorithm, Jialiang Wang,

Hai Zhao, Yuanguo Bi, Xingchi Chen, Ruofan Zeng, and Shiliang Shao

Volume 2014, Article ID 567057, 12 pages

Stepped Fault Line Selection Method Based on Spectral Kurtosis and Relative Energy Entropy of Small Current to Ground System, Xiaowei Wang, Xiangxiang Wei, Jie Gao, Yaxiao Hou, and Yanfang Wei

Volume 2014, Article ID 726205, 18 pages

Estimating Target Heights Based on the Earth Curvature Model and Micromultipath Effect in Skywave OTH Radar, Hou Chengyu, Wang Yuxin, and Chen Jiawei
Volume 2014, Article ID 424191, 14 pages

Research on Torque Ratio Based on the Steering Wheel Torque Characteristic for Steer-by-Wire System, Yandong Han, Lei He, Xiang Wang, and Changfu Zong
Volume 2014, Article ID 929164, 7 pages

Study on the Detectability of the Sky-Surface Wave Hybrid Radar, Hou Chengyu, Ke Guo, Shi Tiange, and Wang Yuxin
Volume 2014, Article ID 602075, 10 pages

Urban and Indoor Weak Signal Tracking Using an Array Tracker with MVA and Nonlinear Filtering, Jicheng Ding, Guoli Zhang, and Lin Zhao
Volume 2014, Article ID 107156, 10 pages

A Novel Method for Target Navigation and Mapping Based on Laser Ranging and MEMS/GPS Navigation, Jianhua Cheng, Rene Landry Jr., Daidai Chen, and Dongxue Guan
Volume 2014, Article ID 918423, 10 pages

Synchronous Routing for Personal Rapid Transit Pods, Mehdi Mrad, Olfa Chebbi, Mohamed Labidi, and Mohamed Ali Louly
Volume 2014, Article ID 623849, 8 pages

Nonlinear Dynamical Analysis of Hydraulic Turbine Governing Systems with Nonelastic Water Hammer Effect, Junyi Li and Qijuan Chen
Volume 2014, Article ID 412578, 11 pages

Performance of a Nonlinear Real-Time Optimal Control System for HEVs/PHEVs during Car Following, Kaijiang Yu and Junqi Yang
Volume 2014, Article ID 879232, 14 pages

Robust Control for Uncertain Linear System Subject to Input Saturation, Qingyun Yang and Mou Chen
Volume 2014, Article ID 803842, 12 pages

Support Vector Regression-Based Adaptive Divided Difference Filter for Nonlinear State Estimation Problems, Hongjian Wang, Jinlong Xu, Aihua Zhang, Cun Li, and Hongfei Yao
Volume 2014, Article ID 139503, 10 pages

A Multiagent Transfer Function Neuroapproach to Solve Fuzzy Riccati Differential Equations, Mohammad Shazri Shahrir, N. Kumaresan, Kuru Ratnavelu, and M. Z. M. Kamali
Volume 2014, Article ID 605625, 6 pages

Sign Stability for Switched Linear Systems and Its Application in Flight Control, Qing Wang, Tong Wang, Yanze Hou, and Chaoyang Dong
Volume 2014, Article ID 391617, 9 pages

Reliability Modeling and Evaluation of Electric Vehicle Motor by Using Fault Tree and Extended Stochastic Petri Nets, Bing Wang, Guangdong Tian, Yanping Liang, and Tiangang Qiang
Volume 2014, Article ID 638013, 9 pages

Research on Optimal Control for the Vehicle Suspension Based on the Simulated Annealing Algorithm, Jie Meng, Qingzhang Chen, and Ren He
Volume 2014, Article ID 420719, 5 pages

A Direct Attitude Determination Approach Based on GPS Double-Difference Carrier Phase Measurements, Jianhua Cheng, Jing Wang, and Lin Zhao
Volume 2014, Article ID 548083, 6 pages

An Adaptive Unscented Kalman Filtering Algorithm for MEMS/GPS Integrated Navigation Systems, Jianhua Cheng, Daidai Chen, Rene Jr. Landry, Lin Zhao, and Dongxue Guan
Volume 2014, Article ID 451939, 8 pages

Dynamic Surface and Active Disturbance Rejection Control for Path Following of an Underactuated UUV, Juan Li, Haitao Gao, Jiajia Zhou, and Zheping Yan
Volume 2014, Article ID 714079, 9 pages

Editorial

Recent Advances on Mathematical Modeling and Control Methods for Complex Vehicle Systems

Hamid Reza Karimi,¹ Zheping Yan,² Yuxin Zhao,² Weichao Sun,³ and Mohammed Chadli⁴

¹ Department of Engineering, Faculty of Engineering and Science, University of Agder, 4898 Grimstad, Norway

² College of Automation, Harbin Engineering University, Nantong Street #145, Nangang District, Harbin, China

³ School of Astronautics, Harbin Institute of Technology, P.O. Box 3015, Yikuang Street #2, Nangang District, Harbin, China

⁴ Laboratory of Modeling, Information & Systems, University of Picardie Jules Verne, Amiens, France

Correspondence should be addressed to Hamid Reza Karimi; hrkarimi@gmail.com

Received 31 October 2014; Accepted 31 October 2014; Published 27 November 2014

Copyright © 2014 Hamid Reza Karimi et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Complex vehicle systems such as highway vehicle, surface vehicle, underwater vehicle, space vehicle, and unmanned vehicle play extremely important roles in wide range of industrial sectors like transportation, mineral exploration, and space exploration. Due to the increasing demands on system performance, diversity, and complexity of the tasks, more and more attention has been paid to complex vehicle correlative technologies with full consideration of issues like mathematical modeling, mission planning, robust performance, coordination and formation, optimization, communication, navigation, prognosis, fault diagnosis and isolation, and so on. All aforementioned issues provide a basis for the design and operation of practical complex vehicle systems in order to achieve desired complex tasks. This special issue aims at providing an opportunity for scientists, engineers, and practitioners to propose their latest theoretical and technological achievements in systems control, especially in the aforesaid aspects.

This special issue contains twenty-nine papers, the contents of which are mainly summarized as follows.

“Urban and indoor weak signal tracking using an array tracker with MVA and nonlinear filtering” by J. Ding et al. focuses on the need for weak GPS signal tracking technique, where the tracking loop is unlocked and data bit edge position is unknown. A modified Viterbi algorithm (MVA) based on dynamic programming is developed, and it is applied to GPS bit synchronization to improve bit edge position detection probability. Meanwhile, two combination carrier

tracking schemes based on Central Difference Kalman Filter (CDKF) and MVA module are designed for tracking very weak GPS signal. The testing results indicate that the methods can successfully detect bit edge position with high detection probability whether or not the tracking loop is locked.

“Support vector regression-based adaptive divided difference filter for nonlinear state estimation problems” by H. Wang et al. presents a support vector regression-based adaptive divided difference filter (SVRADDF) algorithm for improving the low state estimation accuracy of nonlinear systems, which are typically affected by large initial estimation errors and imprecise prior knowledge of process and measurement noises. The derivative-free SVRADDF algorithm is significantly simpler to compute than other methods and is implemented using only functional evaluations.

“Robust stability clearance of flight control law based on global sensitivity analysis” by L. Ou et al. proposes a new clearance framework based on structural singular value μ theory and global uncertainty sensitivity analysis (SA) to validate the robust stability of the flight control system of hypersonic flight vehicle, which suffers from a large number of parametrical uncertainties. In this framework, SA serves as the preprocess of uncertain model to be analyzed to help engineers determine which uncertainties affect the stability of the closed loop system more slightly. Instead of analyzing the effect of uncertainties on μ which involves solving optimal problems repeatedly, a simpler stability analysis function is proposed. Based on this stability analysis function, Sobol

method, the most widely used global SA method, is extended and applied to the new clearance framework due to its suitability for system with strong nonlinearity and input factors varying in large interval, as well as input factors subjecting to random distributions.

“*Probabilistic harmonic calculation in distribution networks with electric vehicle charging stations*” by J. Wang et al. provides the integration of a mathematical simulation method for studying the working status of charging stations, which considers influencing factors including random leaving factor, electricity price, and waiting time. Based on the proposed simulation method, the probability distribution of the harmonic currents of EV charging stations is obtained and used in the calculation of the probability harmonic power flow. Then the impacts of EVs and EV charging stations on distribution networks can be analyzed. In case study, the proposed simulation and analysis method is implemented on the IEEE-34 distribution network. The influences of EV arrival rates, the penetration rate, and the accessing location of EV charging station are also investigated.

In the paper “*Sign stability for switched linear systems and its application in flight control*” by Q. Wang et al., the sign stability concept in ecological systems is introduced into the analysis and synthesis of switched linear system to explore new control design technique. The necessary and sufficient condition for sign stability of a switched linear system under arbitrary switching is achieved. A new approach for the stabilization of switched system is presented. Although the controllers are devised for each subsystem, respectively, the switched system is sign-stabilized by the constitution of CISSS. The provided method has natural robustness and more design freedoms than the familiar Lyapunov function method, which bears relative conservativeness as the requirement of solving LMIs. The presented technique is validated by an example of flight control within a large-scale flight envelop.

In the paper “*An aeromagnetic compensation algorithm for aircraft based on fuzzy adaptive Kalman filter*” by Y. Zhang et al., the classical aircraft aeromagnetic compensation algorithm based on Leliak model is analyzed and an aircraft aeromagnetic compensation algorithm based on fuzzy adaptive Kalman filter is proposed, which is a new approach for aircraft to achieve aeromagnetic compensation. Simulation results show that it has better compensation performance without relying on the aircraft attitude.

“*Nonlinear dynamical analysis of hydraulic turbine governing systems with nonelastic water hammer effect*” by J. Li and Q. Chen proposes a nonlinear mathematical model for hydroturbine governing system (HTGS), and all essential components of HTGS, that is, conduit system, turbine, generator, and hydraulic servosystem, are considered in the model. Using the proposed model, the existence and stability of Hopf bifurcation of an example HTGS are investigated. In addition, chaotic characteristics of the system with different system parameters are studied extensively and presented in the form of bifurcation diagrams, time waveforms, phase space trajectories, Lyapunov exponent, chaotic attractors, and Poincare maps. The simulation results provide a reasonable explanation for the sustained oscillation phenomenon commonly seen in operation of hydroelectric generating set.

“*Research on optimal control for the vehicle suspension based on the simulated annealing algorithm*” by J. Meng et al. presents a designed method to optimize the weight matrix of the LQR controller by using the simulated annealing algorithm. This method utilizes the random searching characteristics of the algorithm to optimize the weight matrices with the target function of suspension performance indexes. This method improves the design efficiency and control performance of the LQR control and solves the problem of the LQR controller when defining the weight matrices. And a simulation is provided for vehicle active chassis control. The result shows that the active suspension using LQR optimized by the genetic algorithm compared to the chassis controlled by the normal LQR and the passive one shows better performance. Meanwhile, the problem of defining the weight matrices is greatly solved.

“*Optimal acceleration-velocity-bounded trajectory planning in dynamic crowd simulation*” by F. Yue-wen et al. studies one special type of crowd which is composed of urgent individuals, normal individuals, and normal groups, where three steps are used to construct the crowd simulation in dynamic environment. The first one is that the urgent individuals move forward along a given path around dynamic obstacles and other crowd members. An optimal acceleration-velocity-bounded trajectory planning method is utilized to model their behaviors, which ensures that the durations of the generated trajectories are minimal and the urgent individuals are collision-free with dynamic obstacles (e.g., dynamic vehicles). In the second step, a pushing model is adopted to simulate the interactions between urgent members and normal ones, which ensures that the computational cost of the optimal trajectory planning is acceptable. The third step is obligated to imitate the interactions among normal members using collision avoidance behavior and flocking behavior.

Of course, the selected topics and papers are not a comprehensive representation of the area of this special issue. Nonetheless, they represent the rich and many-faceted knowledge that we have the pleasure of sharing with the readers. We hope that readers will find in this special issue not only the published papers on complex vehicle systems but also important questions to be resolved such as mathematical modeling, control strategies, and advanced computational algorithm.

Acknowledgments

We would like to express appreciation to the authors for their excellent contributions and patience in assisting us. The hard work of all reviewers on these papers is also very greatly acknowledged.

Hamid Reza Karimi
Zheping Yan
Yuxin Zhao
Weichao Sun
Mohammed Chadli

Research Article

Robust Stability Clearance of Flight Control Law Based on Global Sensitivity Analysis

Liuli Ou, Lei Liu, Shuai Dong, and Yongji Wang

Key Lab of Ministry of Education for Image Processing & Intelligent Control, School of Automation,
Huazhong University of Science and Technology, 1037 Luoyu Road, Wuhan 430074, China

Correspondence should be addressed to Lei Liu; liulei@hust.edu.cn

Received 16 June 2014; Revised 23 August 2014; Accepted 23 August 2014; Published 13 October 2014

Academic Editor: Weichao Sun

Copyright © 2014 Liuli Ou et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

To validate the robust stability of the flight control system of hypersonic flight vehicle, which suffers from a large number of parametrical uncertainties, a new clearance framework based on structural singular value (μ) theory and global uncertainty sensitivity analysis (SA) is proposed. In this framework, SA serves as the preprocess of uncertain model to be analysed to help engineers to determine which uncertainties affect the stability of the closed loop system more slightly. By ignoring these unimportant uncertainties, the calculation of μ can be simplified. Instead of analysing the effect of uncertainties on μ which involves solving optimal problems repeatedly, a simpler stability analysis function which represents the effect of uncertainties on closed loop poles is proposed. Based on this stability analysis function, Sobol's method, the most widely used global SA method, is extended and applied to the new clearance framework due to its suitability for system with strong nonlinearity and input factors varying in large interval, as well as input factors subjecting to random distributions. In this method, the sensitive indices can be estimated via Monte Carlo simulation conveniently. An example is given to illustrate the efficiency of the proposed method.

1. Introduction

In the past decades, clearance of flight control system has been paid great attention by the air force of many countries. Rational clearance improves not only the reliability and safety of flight control system, but also feedback valuable information to designers. Clearance methods split neatly into two types based on the different clearance principle. One is the so-called analytical-model-based method (AMBM) [1] and another is simulation-data-based method (SDBM) [2, 3]. AMBM analyses stable margin, robust stability and some other criteria based on advanced control theories which include (1) structural singular value (μ) analysis [4, 5]; (2) polynomial-based clearance [6]; (3) v-gap analysis; (4) bifurcation and continuation method [7]; and (5) optimization-based clearance [8]. AMBM has been improved and extended to more criteria and more types of air vehicles [9–14]. SDBM demands high fidelity simulation platform and combines experiment design techniques with decision sciences. Generally speaking, AMBM is suitable for early stages of the whole controller designing process, whereas SDBM is suitable for late stages. These two types of methods are complementary.

For hypersonic flight vehicle (HFV) whose maximum flight speed may reach to about 20 Mach, the principal criterion to validate is robust stability of its attitude control system because of the existence of uncertain coefficients. The uncertainties are mainly caused by the following three factors. First, the aerodynamic coefficients of HFV, obtained via wind tunnel test and computational fluid mechanics tools, often suffer from much more serious deviations than those of supersonic and subsonic vehicles. Second, it is still unknown how the external environment affects the HFV due to lack of flight experience. Finally, erosion and corrosion may damage the aerodynamic configuration of the vehicle and hence change the aerodynamic coefficients. Many advanced methods [15–17] have the potential to improve the ability and performance of the attitude control system, which may also lead to more complicated controllers and require more rational clearance.

Among these model-based clearance methods mentioned above, μ analysis is the favourable method for robust stability criterion due to its nonconservatism. Calculation of μ with pure real uncertainties [18] had always been a challenge until the optimization-based approach [19] was proposed. Hybrid

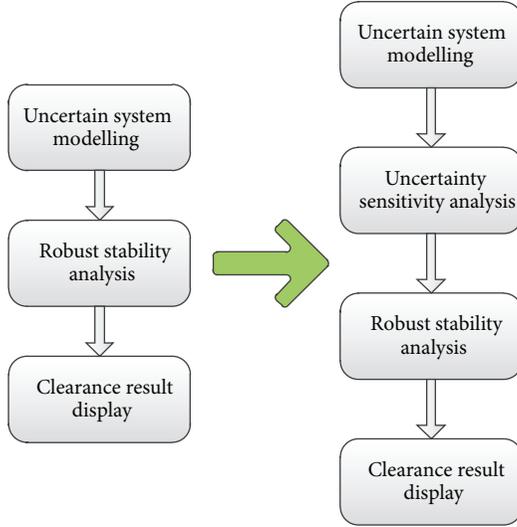


FIGURE 1: The new clearance process.

optimization [8], which combines sequential quadratic programming (SQP) with intelligent optimization, can produce an exact estimation of μ . However, the computational burden is still heavy due to iteratively sweeping frequencies. In particular, when the number of uncertainties is big, it will take a long time to get μ values at a series of frequency points.

In order to reduce the computational burden of μ , engineers usually ignore some unimportant uncertain factors according to their experience. How to find those minor uncertain factors of the HFV when we have little experience? The answer is sensitivity analysis (SA). SA methods can be broken down into local SA [20, 21] and global SA [22–25]. Global SA has been applied widely in industrial production, ecological engineering, hydrology systems, and so on [26] due to its two main advantages. First, it is suitable for input factors varying in broad range and system model with strong nonlinearity. Second, it can consider input factors obeying certain random distributions.

In this work, we proposed a new clearance framework for robust stability clearance of the control system of HFV based on μ theory and global uncertainty SA (GUSA), which is shown in Figure 1. The GUSA serves as the preprocess of the μ analysis to help engineers to determine those minor uncertainties. Because measuring the direct effect of uncertainties on μ involves computing μ repeatedly, a novel indirect GA index based on our newly defined stability analysis function is adopted. This function represents the effect of uncertainties on dominant closed loop poles. With this new GA index, we can extend the widely used Sobol's method [23] and apply it to analyse the effect of uncertainties on stability of the flight control system. It is worth noting that the framework is suitable for not only μ analysis but also other robust analysis approaches. Furthermore, GUSA can provide valuable information about the effect of aerodynamic coefficients on stability for subsequent design and validation.

The paper is organized as follows. In Section 2, the uncertain model of the attitude control system of HFV

is established and μ framework is presented. The Sobol's method is extended and applied to GUSA in Section 3. Section 4 gives an example to illustrate the efficiency of the proposed method. In Section 5, some conclusions are stated.

2. Problem Description

The motion equations of the hypersonic flight vehicle are given as follows [27]:

$$\begin{aligned}\dot{\alpha} &= \frac{C_A q S_m \sin \alpha - C_N q S_m \cos \alpha}{mv \cos \beta} - \frac{\omega_x \cos \alpha \sin \beta}{\cos \beta} \\ &\quad + \frac{\omega_y \sin \alpha \tan \beta + \omega_z - g_y}{v \cos \beta} \\ \dot{\beta} &= \frac{C_A q S_m \cos \alpha \sin \beta + C_N q S_m \sin \alpha \sin \beta + C_Z q S_m \cos \beta}{mv} \\ &\quad + \frac{g_z + \omega_x \sin \alpha + \omega_y \cos \alpha}{v} \\ \dot{\gamma} &= \omega_x \\ \dot{\omega}_x &= \frac{C_{mx} q S_m l_k}{J_x} + \frac{C_{lp} q S_m l_k^2 \omega_x}{2v J_x} + \frac{(J_y - J_z) \omega_y \omega_z}{J_x} \\ \dot{\omega}_y &= \frac{C_{my} q S_m l_k}{J_y} + \frac{C_{nr} q S_m l_k^2 \omega_y}{2v J_y} + \frac{(J_z - J_x) \omega_z \omega_x}{J_y} \\ \dot{\omega}_z &= \frac{C_{mz} q S_m l_k}{J_z} + \frac{C_{mq} q S_m l_k^2 \omega_z}{2v J_z} + \frac{(J_x - J_y) \omega_x \omega_y}{J_z},\end{aligned}\quad (1)$$

where

- α, β, γ = attack, sideslip, and roll angle;
- $\omega_z, \omega_y, \omega_x$ = pitch, yaw, and roll rate;
- C_A, C_N, C_Z = aerodynamic coefficients;
- C_{mx}, C_{my}, C_{mz} = aerodynamic moment coefficients;
- C_{lp}, C_{nr}, C_{mq} = damping moment coefficients;
- J_x, J_y, J_z = inertia moment in x, y, z axes;
- S_m, l_k = wing area and span;
- q = dynamic pressure;
- m = mass;
- v = velocity; and
- $\delta_\phi, \delta_\psi, \delta_\gamma$ = elevator, rudder, and aileron deflection.

These aerodynamic coefficients are nonlinear functions of the deflection angles δ_ϕ, δ_ψ , and δ_γ . Consider

$$C_i = f_i(\delta_\phi, \delta_\psi, \delta_\gamma), \quad i = A, N, Z, mx, my, mz, lp, nr, mq \quad (2)$$

whose nominal values are given in the form of look-up tables. C_i, q, J_x, J_y , and J_z are uncertain coefficients varying in some certain intervals around their nominal values.

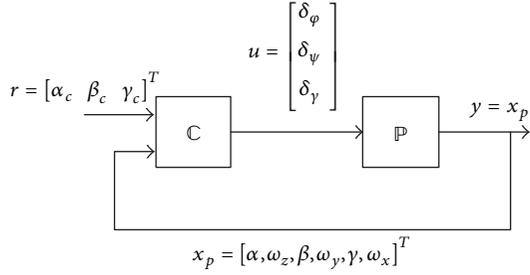


FIGURE 2: Closed loop system of the attitude control of HFV.

Linearizing (1) and (2) along the nominal flight trajectory according to the methods introduced in [4, 18], the uncertain model \mathbb{P} can be written as

$$\dot{x}_p = A_p x_p + B_p u, \quad (3)$$

where $x_p = [\alpha \ \omega_z \ \beta \ \omega_y \ \gamma \ \omega_x]^T$ is the state vector, $u = [\delta_\varphi \ \delta_\psi \ \delta_\gamma]^T$ is the control input, and $A_p \in \mathfrak{R}^{6 \times 6}$ and $B_p \in \mathfrak{R}^{6 \times 3}$ are uncertain matrices in the affine form as follows:

$$A_p = A_{p0} + \sum_{i=1}^k \sigma_i A_{pi} \quad (4)$$

$$B_p = B_{p0} + \sum_{i=1}^k \sigma_i B_{pi},$$

where k is the number of normalized uncertainties after combination and linearization and $\sigma_i \in [-1, 1]$ are uncertain parameters that represent the uncertainties of these coefficients. The uncertain model (4) is not unique via different modelling methods. Because of different designing and manufacturing environment for different vehicles, uncertainties may obey certain random distribution. Consider

$$\sigma_i \sim R, \quad (5)$$

for instance, the normal distribution $\sigma_i \sim N(\mu, v^2)$ or the uniform distribution $\sigma_i \sim U(a, b)$.

Suppose that a state feedback controller \mathbb{C} , which can stabilize the nominal system of (3), is

$$\dot{x}_c = A_c x_c + B_{c1} x_p + B_{c2} r \quad (6)$$

$$u = C_c x_c + D_{c1} x_p + D_{c2} r,$$

where $x_c \in \mathfrak{R}^{d \times 1}$ is the state vector of controller, $r = [\alpha_c \ \beta_c \ \gamma_c]^T$ is the command attitude angles to be tracked, $A_c \in \mathfrak{R}^{d \times d}$, $B_{c1} \in \mathfrak{R}^{d \times 6}$, $B_{c2} \in \mathfrak{R}^{d \times 3}$, $C_c \in \mathfrak{R}^{3 \times d}$, $D_{c1} \in \mathfrak{R}^{3 \times 6}$, and $D_{c2} \in \mathfrak{R}^{3 \times 3}$.

Then, the closed loop system, as shown in Figure 2, can be written as

$$\begin{bmatrix} \dot{x}_p \\ \dot{x}_c \end{bmatrix} = (A_0 + \Delta A) \begin{bmatrix} x_p \\ x_c \end{bmatrix} + (B_0 + \Delta B) r, \quad (7)$$

where

$$A_0 = \begin{bmatrix} A_{p0} + B_{p0} D_{c1} & B_{p0} C_c \\ B_{c1} & A_c \end{bmatrix}$$

$$B_0 = \begin{bmatrix} B_{p0} D_{c2} \\ B_{c2} \end{bmatrix}$$

$$\Delta A = \sum_{i=1}^k \sigma_i A_i = \sum_{i=1}^k \left(\sigma_i \begin{bmatrix} A_{pi} + B_{pi} D_{c1} & B_{pi} C_c \\ \mathbf{0} & \mathbf{0} \end{bmatrix} \right)$$

$$\Delta B = \sum_{i=1}^k \sigma_i B_i = \sum_{i=1}^k \left(\sigma_i \begin{bmatrix} B_{pi} D_{c2} \\ \mathbf{0} \end{bmatrix} \right).$$

(8)

It is clear that the stability of the system (7), suffering from affine interval uncertainties σ_i , is determined by the system matrix $A_0 + \Delta A$. In μ framework, the linear fraction representation (LFR) of (7), denoted as $\mathbb{F}(M, \Delta)$, should be modelled first according to the approach presented in [4, 18]. Consider

$$\left. \begin{aligned} \begin{bmatrix} \dot{x}_p \\ \dot{x}_c \end{bmatrix} &= \begin{bmatrix} A_{p0} + B_{p0} D_{c1} & B_{p0} C_c \\ B_{c1} & A_c \end{bmatrix} \begin{bmatrix} x_p \\ x_c \end{bmatrix} \\ &+ [U_1 \ U_2 \ \dots \ U_k] w \\ z &= [V_1^T \ V_2^T \ \dots \ V_k^T]^T \begin{bmatrix} x_p \\ x_c \end{bmatrix} \\ w &= \Delta z = \text{diag}(\sigma_i I_{r_i}) z \end{aligned} \right\} M, \quad (9)$$

where M represents known part of the system with input w and output z , $\Delta = \text{diag}\{\sigma_i I_{r_i}\}$ represents uncertainties of diagonal structure, $\text{rank}(\cdot)$ is the rank of a matrix, and $U_i \in \mathfrak{R}^{(n+d) \times r_i}$ and $V_i \in \mathfrak{R}^{r_i \times (n+d)}$ are obtained by full rank decomposition. Consider

$$U_i V_i = \begin{bmatrix} A_{pi} + B_{pi} D_{c1} & B_{pi} C_c \\ \mathbf{0} & \mathbf{0} \end{bmatrix} \quad (10)$$

$$r_i = \text{rank} \left(\begin{bmatrix} A_{pi} + B_{pi} D_{c1} & B_{pi} C_c \\ \mathbf{0} & \mathbf{0} \end{bmatrix} \right).$$

At a specific frequency point ω , structural singular value of $\mathbb{F}(M, \Delta)$ (9) is defined as

$$\mu_\Delta(M(j\omega)) = \frac{1}{\min \{\|\Delta\|_\infty \mid \det(I - M(j\omega)\Delta) = 0\}} > 0, \quad (11)$$

where $\|\cdot\|_\infty$ is the infinite norm of a matrix. Define $\mu^* = \max_\omega \mu_\Delta(M(j\omega))$. Since $\|\Delta\|_\infty \leq 1$, the sufficient condition of robust stability is $\max\{\|\Delta\|_\infty\} < 1/\mu^*$; that is, $\mu^* < 1$. Furthermore, the system stays stable even if the uncertain intervals extend to $\sigma_i \in [-1/\mu^*, 1/\mu^*]$. For simplicity, we denote $\mu_\Delta(M(j\omega))$ as μ hereinafter.

According to definition (11), if the problem

$$\begin{aligned} \min \quad & \bar{\gamma} \\ \text{s.t.} \quad & (1) \|\Delta\|_\infty \leq \bar{\gamma} \\ & (2) \det(I - M(j\omega)\Delta) = 0 \end{aligned} \quad (12)$$

has an optimal solution $\bar{\gamma}^*$, then the estimation of μ is $1/\bar{\gamma}^*$. By combining global intelligent algorithm (such as particle swarm optimization (PSO) and differential evolution (DE) [28]) with local SQP, this problem can be iteratively solved [8]. In order to get the estimation of $\mu^* = \max_{\omega} \mu$, we need to solve the optimal problem (12) at a series of frequency points.

When the number of uncertainties is large, the dimension of the searching space is also large. In order to cover the whole space uniformly, the number of particles of PSO (or analogy of other intelligent algorithms) should exponentially increase with respect to the number of uncertainties. Then, the computational burden will increase exponentially. Therefore, our aim is to find those σ_i s which affect the stability dramatically and ignore remaining unimportant σ_i s when calculating μ^* . If we measure the effect of uncertainties on μ^* directly, we need to calculate μ^* repeatedly. So we introduce an indirect stability analysis function.

As we know, a system is stable if and only if all of its closed loop poles lie in the left half of the complex plan. Therefore, we define the nonlinear stability analysis function as follows:

$$f(\sigma_i) = \max_j \{ \text{Re} [\lambda_j (A_0 + \Delta A)] \}, \quad (13)$$

where $\lambda_j(\cdot)$ is the j th eigenvalue of a matrix and $\text{Re}[\cdot]$ is the real component of a complex. The effect of σ_i on the stability of the closed loop system can be represented by (13) and SA can be invoked. It is clear that the stability analysis function (13) is also suitable for other robust stability clearance method.

All local SA methods need to calculate the differential or derivative; that is, they are based on the analytical model. However, the nonlinear function (13) does not satisfy this requirement. So we will use the global Sobol's method to deal with this problem in the upcoming sections.

3. Global Uncertainty Sensitivity Analysis

3.1. Sobol's Method. Consider the nonlinear model in the form

$$\begin{aligned} f: \mathfrak{R}^k &\longrightarrow \mathfrak{R} \\ \mathbf{x} &\longrightarrow y = f(\mathbf{x}), \end{aligned} \quad (14)$$

where $\mathbf{x} = [x_1, x_2, \dots, x_k]$ are k independent input factors and y is the scalar output. If \mathbf{x} is defined over $\Omega = \{\mathbf{x} \mid 0 \leq x_i \leq 1, i = 1, 2, \dots, k\}$, then the function can be decomposed as the following form [23]:

$$\begin{aligned} f &= f_0 + \sum_i f_i + \sum_i \sum_{j>i} f_{ij} \\ &+ \sum_{i_1} \sum_{i_2>i_1} \sum_{i_3>i_2} f_{i_1 i_2 i_3} + \dots + f_{12\dots k}, \end{aligned} \quad (15)$$

where $f_i = f_i(x_i)$, $f_{ij} = f_{ij}(x_i, x_j)$, and so on. The uniqueness of (15) is guaranteed by

$$\int_0^1 f_{i_1, i_2, \dots, i_s}(x_{i_1}, x_{i_2}, \dots, x_{i_s}) dx_{i_w} = 0, \quad (16)$$

where $1 \leq i_1 < i_2 < \dots < i_s \leq k$ and $i_w \in \{i_1, i_2, \dots, i_s\}$.

Denote $\mathbf{x}_{i_1 i_2 \dots i_s} = [x_{i_1}, x_{i_2}, \dots, x_{i_s}]$, $1 \leq i_1 < i_2 < \dots < i_s \leq k$, and $\mathbf{x}_{\sim i_1 i_2 \dots i_s}$ as the complement of $\mathbf{x}_{i_1 i_2 \dots i_s}$; then we have

$$\begin{aligned} \mathbf{x}_{\sim i} &= [x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_k] \\ \mathbf{x}_{\sim ij} &= [x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_{j-1}, x_{j+1}, \dots, x_k] \end{aligned} \quad (17)$$

and so on. Define the corresponding domain as $\Omega_{i_1 i_2 \dots i_s}$ and $\Omega_{\sim i_1 i_2 \dots i_s}$; then we have

$$\begin{aligned} \Omega_{i_1 i_2 \dots i_s} &= \{ \mathbf{x}_{i_1 i_2 \dots i_s} \mid 0 \leq x_l \leq 1, l = i_1, i_2, \dots, i_s \} \\ \Omega_{\sim i_1 i_2 \dots i_s} &= \{ \mathbf{x}_{\sim i_1 i_2 \dots i_s} \mid 0 \leq x_l \leq 1, l \neq i_1, i_2, \dots, i_s \}. \end{aligned} \quad (18)$$

These functions in (15) are obtained by

$$\begin{aligned} f_0 &= \int_{\Omega} f(\mathbf{x}) d\mathbf{x} \\ f_i &= \int_{\Omega_{\sim i}} f(\mathbf{x}) d\mathbf{x}_{\sim i} - f_0 \\ f_{ij} &= \int_{\Omega_{\sim ij}} f(\mathbf{x}) d\mathbf{x}_{\sim ij} - f_i - f_j - f_0 \end{aligned} \quad (19)$$

and similarly for higher orders, where $\int_{\Omega_{i_1 \dots i_r}} f(\mathbf{x}) d\mathbf{x}_{i_1 \dots i_r} = \int_0^1 \int_0^1 \dots \int_0^1 f(\mathbf{x}) dx_{i_1} dx_{i_2} \dots dx_{i_r}$.

The variance of y is

$$V = V(y) = \int_{\Omega} f^2(\mathbf{x}) d\mathbf{x} - f_0^2. \quad (20)$$

And the partial variance is defined as

$$V_{i_1 i_2 \dots i_s} = \int_{\Omega_{i_1 i_2 \dots i_s}} f_{i_1 i_2 \dots i_s}^2 d\mathbf{x}_{i_1 i_2 \dots i_s}. \quad (21)$$

Then the decomposition (15) leads to the following decomposition of the variance of y

$$V = \sum_{i=1}^k V_i + \sum_{1 \leq i < j \leq k} V_{ij} + \dots + V_{12\dots k}. \quad (22)$$

The normalized sensitivity index is defined as

$$\begin{aligned} S_i &= \frac{V_i}{V} \\ S_{ij} &= \frac{V_{ij}}{V} \\ S_{i_1 i_2 \dots i_s} &= \frac{V_{i_1 i_2 \dots i_s}}{V}, \end{aligned} \quad (23)$$

where S_i are the so-called first-order sensitivity indices and S_{ij} and $S_{i_1 i_2 \dots i_s}$ are the second-order and the s th-order indices, respectively. $S_{i_1 i_2 \dots i_s}$ indicates sensitivity of y to the interaction among $x_{i_1}, x_{i_2}, \dots, x_{i_s}$. So the total sensitivity indices S_{Ti} for the input x_i is defined as the sum of all indices relating to

it. For instance, if $k = 3$, we have $S_{T_1} = S_1 + S_{1_2} + S_{1_3} + S_{1_{23}}$. It is easy to find that

$$S_{T_i} = 1 - \frac{V_{\sim i}}{V} \quad (24)$$

and $0 \leq S_{T_i} \leq 1$, where $V_{\sim i}$ is the $(k - 1)$ th indices of the remaining $k - 1$ factors except the i th one.

Because these global sensitivity indices are based on variance, enough samples are needed. The best and most widely used method is Monte Carlo (MC) simulations.

Supposing $\mathbf{X} \in \mathfrak{R}^{N \times k}$ is the matrix of N groups of k input factors

$$\mathbf{X} = (\mathbf{x}_{ij})_{N \times k} = [\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_k] = [\mathbf{x}_1^T, \mathbf{x}_2^T, \dots, \mathbf{x}_N^T]^T, \quad (25)$$

where $\mathbf{X}_i \in \mathfrak{R}^{N \times 1}$ and $\mathbf{x}_i \in \mathfrak{R}^{1 \times k}$. $\mathbf{Y} \in \mathfrak{R}^{N \times 1}$ is the corresponding output

$$\begin{aligned} \mathbf{Y} &= f(\mathbf{X}) = [f(\mathbf{x}_1), f(\mathbf{x}_2), \dots, f(\mathbf{x}_N)]^T \\ &= [y_1, y_2, \dots, y_N]^T, \end{aligned} \quad (26)$$

and \mathbf{P} and \mathbf{Q} are two samples of \mathbf{X} generated by MC. Then we define that

$$\mathbf{P}_Q^{(i)} = [\mathbf{P}_1 \ \dots \ \mathbf{P}_{i-1} \ \mathbf{Q}_i \ \mathbf{P}_{i+1} \ \dots \ \mathbf{P}_k] \quad (27)$$

is the combined matrix with column i from \mathbf{Q} and all other $k - 1$ columns from \mathbf{P} , where \mathbf{P}_i and \mathbf{Q}_i are samples of \mathbf{X}_i . Similarly, we define

$$\mathbf{Q}_P^{(i)} = [\mathbf{Q}_1 \ \dots \ \mathbf{Q}_{i-1} \ \mathbf{P}_i \ \mathbf{Q}_{i+1} \ \dots \ \mathbf{Q}_k]. \quad (28)$$

The estimation below can be obtained [24, 25] as follows:

$$\begin{aligned} \hat{f}_0 &= \frac{1}{2N} \sum_{j=1}^N [f(\mathbf{p}_j) + f(\mathbf{q}_j)] \\ \hat{V} &= \frac{1}{2N} \sum_{j=1}^N [f^2(\mathbf{p}_j) + f^2(\mathbf{q}_j)] - \hat{f}_0^2 \\ \hat{V}_i &= \frac{1}{N} \sum_{j=1}^N f(\mathbf{q}_j) f((\mathbf{P}_Q^{(i)})_j) - \hat{f}_0^2 \\ \hat{V}_{\sim i} &= \frac{1}{N} \sum_{j=1}^N f(\mathbf{q}_j) f((\mathbf{Q}_P^{(i)})_j) - \hat{f}_0^2 \\ \hat{S}_i &= \frac{\hat{V}_i}{\hat{V}}, \quad \hat{S}_{T_i} = 1 - \frac{\hat{V}_{\sim i}}{\hat{V}}, \end{aligned} \quad (29)$$

where \mathbf{p}_j , \mathbf{q}_j , $(\mathbf{P}_Q^{(i)})_j$, and $(\mathbf{Q}_P^{(i)})_j$ are, respectively, the j th row of \mathbf{P} , \mathbf{Q} , $\mathbf{P}_Q^{(i)}$, and $\mathbf{Q}_P^{(i)}$.

3.2. Uncertainties Sensitivity Analysis Based on Sobol's Method. It is worth noting that the input factors of the nonlinear model

considered in Sobol's method are limited to $0 \leq x_i \leq 1$. However, the normalized uncertainties of the flight control system is $\sigma_i \in [-1, 1]$. One manner is to renormalize the uncertainties to the interval $[0, 1]$. Here we will introduce an easier method.

Considering the nonlinear model (14), if its domain of definition is

$$\Pi = \{\mathbf{x} \mid -0.5 \leq x_i \leq 0.5, i = 1, 2, \dots, k\}, \quad (30)$$

we can get the below theorem.

Theorem 1. *The nonlinear model (14) defined over Π can be decomposed uniquely to (15) with its items (34) if*

$$\int_{-1/2}^{1/2} f_{i_1, i_2, \dots, i_s}(x_{i_1}, x_{i_2}, \dots, x_{i_s}) dx_{i_w} = 0 \quad (31)$$

hold for all $1 \leq i_1 < i_2 < \dots < i_s \leq k$ and $i_w \in \{i_1, i_2, \dots, i_s\}$.

Proof. From (31), we can obtain that the members in (15) are orthogonal. That is, for any

$$1 \leq i_1 < i_2 < \dots < i_s \leq k \quad (32)$$

$$1 \leq j_1 < j_2 < \dots < j_r \leq k,$$

if $(i_1, i_2, \dots, i_s) \neq (j_1, j_2, \dots, j_r)$, there exists

$$\int_{\Pi} f_{i_1, i_2, \dots, i_s} f_{j_1, j_2, \dots, j_r} d\mathbf{x} = 0. \quad (33)$$

Then, we have

$$\begin{aligned} \int_{\Pi} f(\mathbf{x}) d\mathbf{x} &= f_0 \\ \int_{\Pi_{\sim i}} f(\mathbf{x}) d\mathbf{x}_{\sim i} &= f_i(x_i) + f_0 \\ \int_{\Pi_{\sim ij}} f(\mathbf{x}) d\mathbf{x}_{\sim ij} &= f_{ij}(x_i, x_j) + f_i(x_i) + f_j(x_j) + f_0 \end{aligned} \quad (34)$$

and so on for high order terms. Obviously, (34) is similar to (19). \square

Remark. In fact, Sobol's method can be extend to any nonlinear function defined over

$$\Gamma = \{\mathbf{x} \mid \underline{x} \leq x_i \leq \bar{x}, i = 1, 2, \dots, k\}, \quad (35)$$

where $\bar{x} - \underline{x} = 1$.

The uncertain matrix ΔA of the system matrix in (7) can be rewritten as

$$\begin{aligned} \Delta A &= \begin{bmatrix} \sum_i \frac{\sigma_i}{2} (2A_{pi} + 2B_{pi}D_{c1}) & \sum_i \frac{\sigma_i}{2} 2B_{pi}C_c \\ \mathbf{0} & \mathbf{0} \end{bmatrix} \\ &= \sum_i 2\bar{\sigma}_i \begin{bmatrix} A_{pi} + B_{pi}D_{c1} & B_{pi}C_c \\ \mathbf{0} & \mathbf{0} \end{bmatrix}, \end{aligned} \quad (36)$$

where $\bar{\sigma}_i = \sigma_i/2 \in [-0.5, 0.5]$ are the normalized uncertainties. It is clear that the uncertain space of $\bar{\sigma}_i$ is Π defined in (30).

Substituting (7) and (36) into the nonlinear function (13), we can get

$$f(\bar{\sigma}) = \max_j \left\{ \operatorname{Re} \left[\lambda_j \left(A_0 + \sum_i \bar{\sigma}_i \begin{bmatrix} 2A_{pi} + 2B_{pi}D_{c1} & 2B_{pi}C_c \\ \mathbf{0} & \mathbf{0} \end{bmatrix} \right) \right] \right\}, \quad (37)$$

where $\bar{\sigma} = (\bar{\sigma}_1, \bar{\sigma}_2, \dots, \bar{\sigma}_k)$, $\bar{\sigma} \in \Pi$. The Sobol's method can be applied to analyse the effect of uncertainty terms $\bar{\sigma}_i$ on the real part of the closed loop poles. The sample matrices of $\bar{\sigma}$ can be generated randomly according to the distribution of σ .

Generally, the whole clearance process can be summarized into the following.

Step 1: establish the uncertain model of the attitude control system and normalize the uncertainties to $[-0.5, 0.5]$.

Step 2: construct stability analysis function as (37) according to the specific uncertain model.

Step 3: generate the sample matrices P and Q of N groups of uncertainties $\bar{\sigma}$.

Step 4: calculate output samples of the stability analysis functions (37) according to MC simulations with input samples P and Q .

Step 5: calculate the SA index according to (29) and ignore the minor uncertainties with relatively small SA index.

Step 6: calculate the singular structure value according to (12).

4. Examples

In this section, a numerical example is given to reveal the effectiveness of the proposed method. Consider the nominal flight trajectory of a HFV depicted as Figure 3, from which 30 operation points for robust stability clearance have been selected. At first, we will illustrate the application of the new clearance process with one single point. Then, we will analyse all these operation points together.

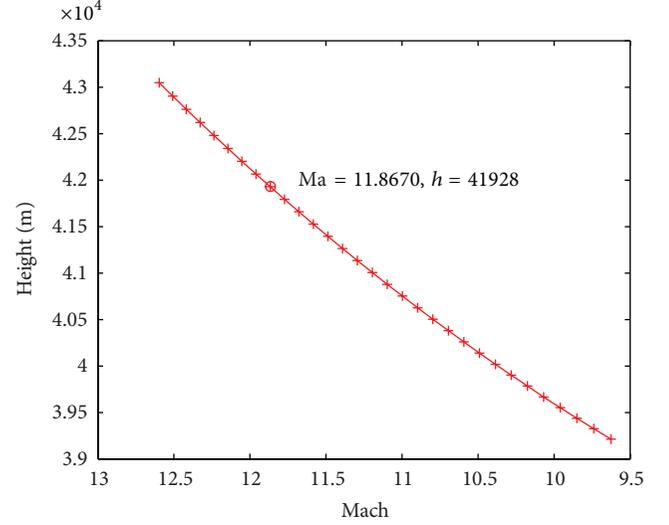


FIGURE 3: Nominal trajectory of the HFV.

At the 9th point ($Ma = 11.8670$, $h = 41928$), the nominal model of the attitude control equation in the form of (3) is

$$A_{p0} = \begin{bmatrix} -0.0138 & 1 & 0 & 0 & 0 & 0 \\ -7.6433 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & -0.0082 & 0.9725 & 0 & 0.23 \\ 0 & 0 & -5.3385 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & -33.8537 & 0 & 0 & 0 \end{bmatrix} \quad (38)$$

$$B_{p0} = \begin{bmatrix} -0.0005 & 0 & 0 \\ -2.5109 & 0 & 0 \\ 0 & -0.0001 & -0.0001 \\ 0 & -0.2711 & -1.2441 \\ 0 & 0 & 0 \\ 0 & -1.1006 & -52.7017 \end{bmatrix}.$$

Because the pitch motion has no coupling with yaw and roll motion, the controller can be designed separately. For convenience, a low order controller that can stabilize the nominal system has been designed as follows:

$$u = \begin{bmatrix} L_1 & \mathbf{0} \\ \mathbf{0} & L_2 \end{bmatrix} r - \begin{bmatrix} K_1 & \mathbf{0} \\ \mathbf{0} & K_2 \end{bmatrix} x_p, \quad (39)$$

where

$$L_1 = -99.9845$$

$$L_2 = \begin{bmatrix} -3.3664 & 0.6189 \\ -0.2088 & -9.9808 \end{bmatrix} \quad (40)$$

$$K_1 = [-98.1456 \quad -7.9508]$$

$$K_2 = \begin{bmatrix} -1.6586 & -3.1971 & 0.6189 & 0.0781 \\ -1.7428 & 2.4028 & -9.9808 & -0.6311 \end{bmatrix}.$$

Then the system matrix of the closed loop system is

$$A = \begin{bmatrix} A_{clo1} & \mathbf{0} \\ \mathbf{0} & A_{clo2} \end{bmatrix} + \Delta A, \quad (41)$$

where

$$A_{clo1} = \begin{bmatrix} -0.0587 & 0.9964 \\ -254.0734 & -19.9634 \end{bmatrix}$$

$$A_{clo2} = \begin{bmatrix} -0.0085 & 0.9726 & -0.0012 & 0.2329 \\ -7.9562 & 2.1227 & -12.2491 & -0.7640 \\ 0 & 0 & 0 & 1 \\ -127.5277 & 123.1129 & -525.3240 & -33.1741 \end{bmatrix}$$

$$\Delta A = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ \Delta a_{21} & \Delta a_{22} & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & \Delta a_{43} & \Delta a_{44} & \Delta a_{45} & \Delta a_{46} \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & \Delta a_{63} & \Delta a_{64} & \Delta a_{65} & \Delta a_{66} \end{bmatrix}$$

$$\Delta a_{21} = -1.5287\sigma_1 + 49.286\sigma_4$$

$$\Delta a_{22} = 3.9927\sigma_4$$

$$\Delta a_{43} = -2.6692\sigma_2 + 0.0899\sigma_5 + 0.4336\sigma_6$$

$$\Delta a_{44} = 0.1733\sigma_5 - 0.5979\sigma_6$$

$$\Delta a_{45} = -0.0336\sigma_5 + 2.4834\sigma_6$$

$$\Delta a_{46} = -0.0042\sigma_5 + 0.1570\sigma_6$$

$$\Delta a_{63} = -6.7707\sigma_3 + 1.4604\sigma_7 + 64.2940\sigma_8$$

$$\Delta a_{64} = 2.8150\sigma_7 - 88.6422\sigma_8$$

$$\Delta a_{65} = -0.5449\sigma_7 + 368.2036\sigma_8$$

$$\Delta a_{66} = -0.0688\sigma_7 + 23.2820\sigma_8$$

$$\sigma_i \sim U(-1, 1), \quad i = 1, 2, \dots, 8.$$

(42)

Calculate the global sensitivity indices of the function (37) with respect to $\bar{\sigma}_i = \sigma_i/2 \sim U(-0.5, 0.5)$ according to the steps presented in Section 3. In order to illustrate the consistency and repeatability of the proposed method, 10 MC simulations have been conducted. The obtained global sensitivity indices S_{Ti} are depicted in Figures 4 and 5. It is evident that (1) the results of 10 simulations are nearly consistent; (2) S_{T1} , S_{T3} , and S_{T4} are extremely close to 0 and obviously smaller than other indices; and (3) when the sample number N increases from 20000 to 50000, the results converge accordingly. Varying σ_i in the interval $[-1, 1]$ singly and fixing $\sigma_j = 0$, $j \neq i$, the distribution of the dominant closed loop poles is shown as Figure 6. This can directly validate the SA results.

Since $S_{Ti} \approx 0$, $i = 1, 3, 4$, we can ignore σ_i , $i = 1, 3, 4$ in the analysis of robust criterion. μ is calculated via hybrid optimization method at 136 frequency points $\omega = [0.01, 0.02, \dots, 0.09, 0.1, 0.2, \dots, 3, 4, 5, \dots, 100]$ in the MATLAB environment. Contrastive results are shown in Figures 7 and 8. When only considering the remaining 5 major uncertainties (σ_i , $i = 2, 5, 6, 7, 8$), the μ curve is very close to

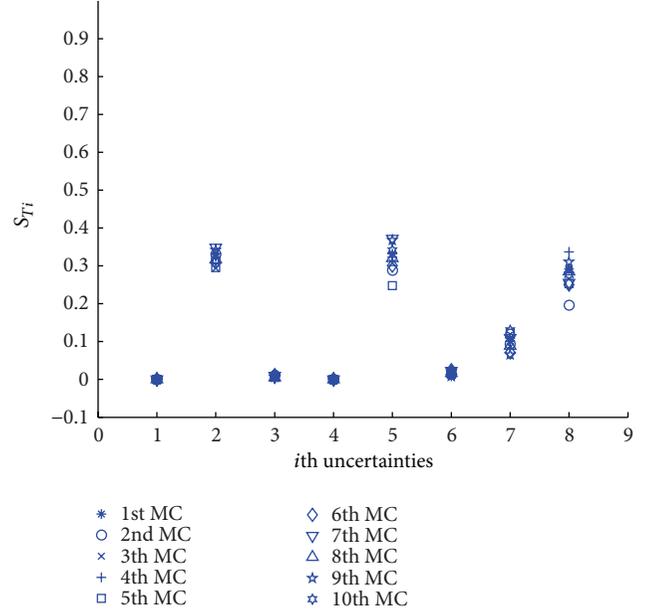


FIGURE 4: The global uncertainty sensitivity indices ($N = 20000$).

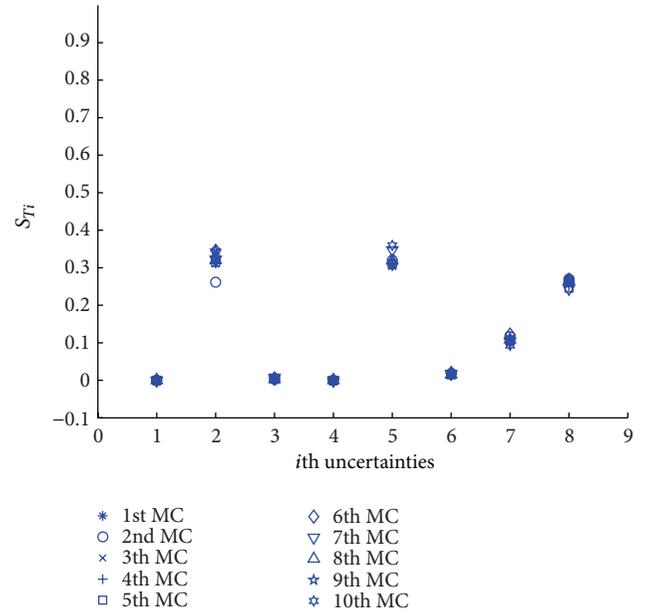


FIGURE 5: The global uncertainty sensitivity indices ($N = 50000$).

the situation of considering all 8 uncertainties. The computational time, however, decreases from 3.3387 s to 2.2903 s when the number of uncertainties decreases from 8 to 5. When taking into account all the operation points together, the proposed clearance frame can save a large amount of time. Because $\mu^* \approx 1.1 > 1$, the controller cannot stabilize the system in the whole uncertain space. However, we can get the valuable information that the approximate allowable uncertain space is $|\sigma_i| < 1/1.1 \approx 0.9$ and feed it back to designers.

The global sensitivity indices S_{Ti} of all operation points are plotted in Figure 9. Note that $S_{Ti} \approx 0$, $i = 1, 3, 4$, holds for

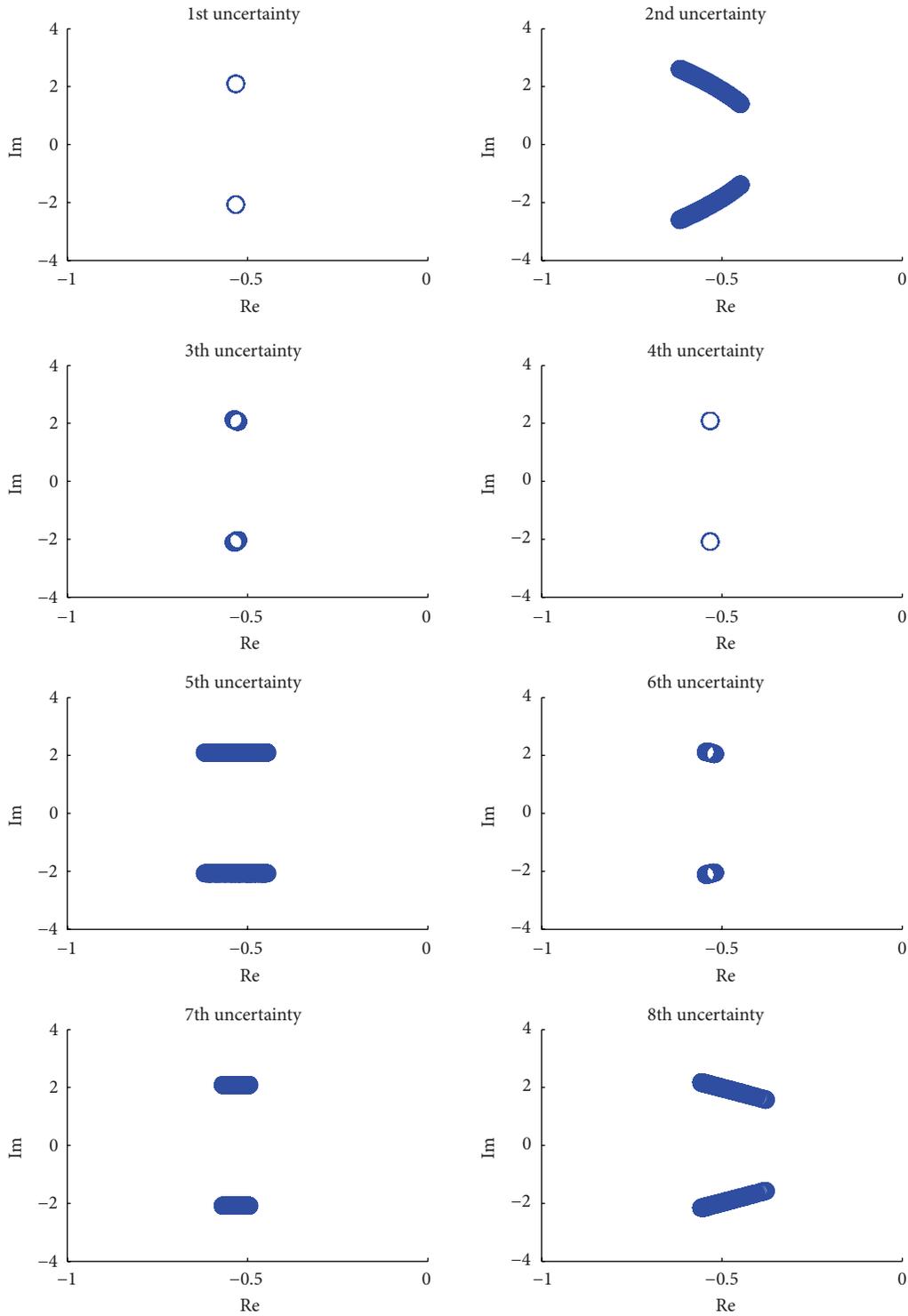


FIGURE 6: Dominant poles distributions for systems with one uncertainty.

all operation points. That is, because the flight trajectory in Figure 3 is relatively smooth and the uncertain model varies slowly. Therefore, we just need to analyse a few operation points when the trajectory is smooth. A comparison of computational time of μ is also depicted in Figure 10. With

the proposed clearance framework, we can save more than 30 s totally while the GUSA just takes 3.6645 s.

In summary, GUSA provides a quantitative index for engineers to measure the effect of uncertainties on dominant closed loop poles. It is worth noting that the GUSA index

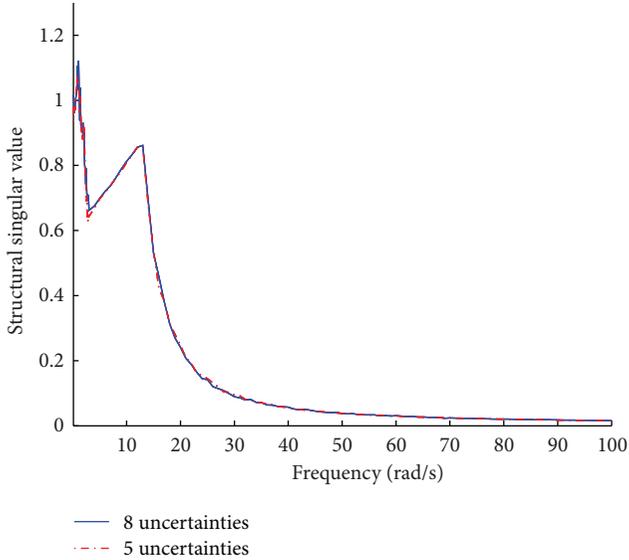


FIGURE 7: Structural singular value of the uncertain systems.

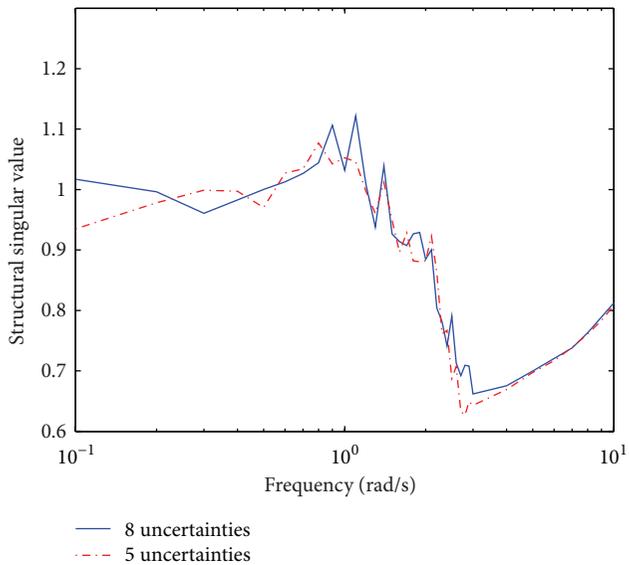


FIGURE 8: Structural singular value in low frequency interval of the uncertain systems.

is relative index. It means that we can order the effect of all uncertainties based on the GUSA index, but we cannot find a threshold value to ignore unimportant uncertainties. There is a tradeoff between precision and computational time when ignoring unimportant uncertainties. For instance, if we need to reduce more computational time, we can ignore one more uncertainty σ_6 .

5. Conclusion

In this paper, a new μ framework based on GUSA is proposed. GUSA serves as the preprocessing step of μ analysis and obviously release the computational burden. Based on our

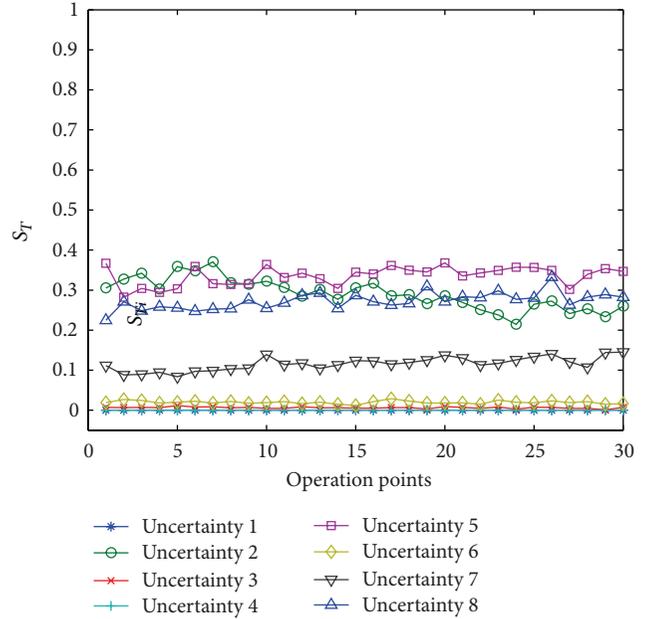


FIGURE 9: GSA of all operation points.

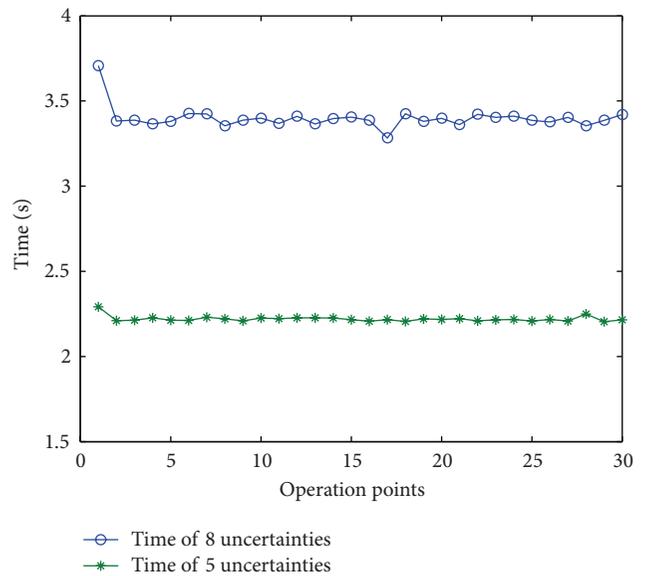


FIGURE 10: Computational time of μ for all operation points.

proposed stability analysis function, Sobol's method is extended and applied to determine the major uncertainties because of its maturity and flexibility. An example is given to illustrate the efficiency of the proposed method. In the potential applications, this work provides an easy-to-follow procedure for engineers to digest and expand.

Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

Acknowledgments

This work is supported in part by the National Nature Science Foundation of China (nos. 61203081 and 61174079), Doctoral Fund of Ministry of Education of China (no. 20120142120091), and Precision manufacturing technology and equipment for metal part (no. 2012DFG70640).

References

- [1] M. Selier, U. Korte, C. Fielding, and R. Luckner, "New analysis techniques for clearance of flight control laws," in *Proceedings of the AIAA Guidance, Navigation, and Control Conference and Exhibit*, 14, p. 11, Austin, Tex, USA, August 2003.
- [2] J. Hanson, "Advanced guidance and control project for reusable launch vehicles," in *Proceedings of the AIAA Guidance, Navigation, and Control Conference and Exhibit*, 2000.
- [3] J. Hanson and R. Jones, "Advanced guidance and control methods for reusable launch vehicles: test results," in *Proceedings of the AIAA Guidance, Navigation, and Control Conference and Exhibit*, Monterey, Calif, USA, 2002.
- [4] D. G. Bates, R. Kureemun, and T. Mannchen, "Improved clearance of a flight control law using μ -analysis techniques," *Journal of Guidance, Control, and Dynamics*, vol. 26, no. 6, pp. 869–884, 2003.
- [5] K. Zhou, J. C. Doyle, and K. Glover, *Robust and Optimal Control*, Prentice Hall, Upper Saddle River, NJ, USA, 1996.
- [6] F. Corraro and M. Virgilio, "A polynomial based clearance method," in *Proceedings of the AIAA Guidance, Navigation, and Control Conference and Exhibit*, Austin, Tex, USA, 2003.
- [7] M. H. Lowenberg and P. P. Menon, in *Proceedings of the AIAA Guidance, Navigation and Control Conference and Exhibit*, Hilton Head, SC, USA, 2007.
- [8] P. P. Menon, J. Kim, D. G. Bates, and I. Postlethwaite, "Clearance of nonlinear flight control laws using hybrid evolutionary optimization," *IEEE Transactions on Evolutionary Computation*, vol. 10, no. 6, pp. 689–698, 2006.
- [9] A. Yazıcı, A. Karamancıoğlu, and R. Kasimbeyli, "A nonlinear programming technique to compute a tight lower bound for the real structured singular value," *Optimization and Engineering*, vol. 12, no. 3, pp. 445–458, 2011.
- [10] P. Seiler, A. Packard, and G. J. Balas, "A gain-based lower bound algorithm for real and mixed μ problems," *Automatica*, vol. 46, no. 3, pp. 493–500, 2010.
- [11] L. Liu, *Advanced Verification and Clearance Techniques for Modern Flight Control Systems*, National Defense Industry Press, Beijing, China, 2010.
- [12] S. Juliana, Q. P. Chu, and J. A. Mulder, "Reentry flight clearance using interval analysis," *Journal of Guidance, Control, and Dynamics*, vol. 31, no. 5, pp. 1295–1307, 2008.
- [13] U. Tancredi, M. Grassi, F. Corraro, and E. Filippone, "Robustness analysis for terminal phases of reentry flight," *Journal of Guidance, Control, and Dynamics*, vol. 32, no. 5, pp. 1679–1683, 2009.
- [14] P. P. Menon, D. G. Bates, and I. Postlethwaite, "Nonlinear robustness analysis of flight control laws for highly augmented aircraft," *Control Engineering Practice*, vol. 15, no. 6, pp. 655–662, 2007.
- [15] W. Sun, Y. Zhao, J. Li, L. Zhang, and H. Gao, "Active suspension control with frequency band constraints and actuator input delay," *IEEE Transactions on Industrial Electronics*, vol. 59, no. 1, pp. 530–537, 2012.
- [16] W. Sun, H. Gao, and O. Kaynak, "Adaptive backstepping control for active suspension systems with hard constraints," *IEEE/ASME Transactions on Mechatronics*, vol. 18, no. 3, pp. 1072–1079, 2013.
- [17] W. Sun, Z. Zhao, and H. Gao, "Saturated adaptive robust control for active suspension systems," *IEEE Transactions on Industrial Electronics*, vol. 60, no. 9, pp. 3889–3896, 2013.
- [18] T. Mannchen, D. G. Bates, and I. Postlethwaite, "Modeling and computing worst-case uncertainty combinations for flight control systems analysis," *Journal of Guidance, Control, and Dynamics*, vol. 25, no. 6, pp. 1029–1039, 2002.
- [19] M. J. Hayes, D. G. Bates, and I. Postlethwaite, "New tools for computing tight bounds on the real structured singular value," *Journal of Guidance, Control, and Dynamics*, vol. 24, no. 6, pp. 1204–1213, 2001.
- [20] C. Zhao and F. He, *Multivariable Robust Control Systems*, Harbin Institute of Technology Press, Harbin, China, 2010.
- [21] Z. H. Zhang and Y. H. Suo, "Stability and sensitivity analysis of a plant disease model with continuous cultural control strategy," *Journal of Applied Mathematics*, vol. 2014, Article ID 207959, 15 pages, 2014.
- [22] A. Saltelli, M. Ratto, T. Andres et al., *Global Sensitivity Analysis: The Primer*, John Wiley & Sons, Chichester, UK, 2008.
- [23] I. Sobol, "Sensitivity analysis for non-linear mathematical models," *Mathematical Modeling & Computational Experiment*, vol. 1, pp. 407–414, 1993.
- [24] J. Jacques, C. Lavergne, and N. Devictor, "Sensitivity analysis in presence of model uncertainty and correlated inputs," *Reliability Engineering and System Safety*, vol. 91, no. 10-11, pp. 1126–1134, 2006.
- [25] A. Saltelli, P. Annoni, I. Azzini, F. Campolongo, M. Ratto, and S. Tarantola, "Variance based sensitivity analysis of model output. Design and estimator for the total sensitivity index," *Computer Physics Communications*, vol. 181, no. 2, pp. 259–270, 2010.
- [26] J. X. Shen and W. Q. Li, "Sensitivity analysis of wavelet neural network model for short-term traffic volume prediction," *Journal of Applied Mathematics*, vol. 2013, 10 pages, 2013.
- [27] H. Li, *Guidance and Control Technology of Hypersonic Flight Vehicle*, China Astronautic Publishing House, Beijing, China, 2012.
- [28] Y. R. Huang, *Intelligent Optimization and Its Application*, National Defence Industry Press, Beijing, China, 2008.

Research Article

Mathematical Model of Movement of the Observation and Tracking Head of an Unmanned Aerial Vehicle Performing Ground Target Search and Tracking

Izabela Krzysztofik and Zbigniew Koruba

Department of Applied Computer Science and Armament Engineering, Faculty of Mechatronics and Machine Design, Kielce University of Technology, 7 Tysiąclecia PP Street, 25-314 Kielce, Poland

Correspondence should be addressed to Izabela Krzysztofik; pssik@tu.kielce.pl

Received 24 April 2014; Revised 31 July 2014; Accepted 2 August 2014; Published 8 September 2014

Academic Editor: Zheping Yan

Copyright © 2014 I. Krzysztofik and Z. Koruba. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

The paper presents the kinematics of mutual movement of an unmanned aerial vehicle (UAV) and a ground target. The controlled observation and tracking head (OTH) is a device responsible for observing the ground, searching for a ground target, and tracking it. The preprogrammed movement of the UAV on the circle with the simultaneous movement of the head axis on Archimedes' spiral during searching for a ground target, both fixed (bunkers, rocket missiles launching positions, etc.) and movable (tanks, infantry fighting vehicles, etc.), is considered. Dynamics of OTH during the performance of the above mentioned activities is examined. Some research results are presented in a graphical form.

1. Introduction

Due to the advantages of unmanned aerial vehicles (UAV) in relation to manned aircrafts, the multitask UAVs have become the basic equipment of a modern army [1]. They can carry out various tasks, such as aerial and radiolocation reconnaissance, observation of the battlefield, radioelectronic fight, adjustment of the artillery fire, target identification, laser indication of the target, assessment of the effects of striking other types of weapons, and imitation of air targets. UAVs can also have a wide civilian application, for example, observation and control of pipelines, electric tractions and road traffic. Because of that, more than 250 UAV models are developed and manufactured all over the world [2]. Many scientific institutions are engaged in developing and identifying models of dynamics and flight control of UAVs. For instance, papers [3, 4] present the comprehensive research on modelling the dynamics of flight of K100 UAV and Thunder Tiger Raptor 50 V2 Helicopter, respectively. Mathematical models of UAV with 6 degrees of freedom (6-DoF)

introduced on the basis of the Newton's rules of dynamics were presented, whereas [5] presents the methodology of modelling the dynamics of UAV flight with the use of neural networks. Paper [6] presents the mathematical model and the experimental research on SUAVI. That vehicle can take off and land vertically. The model was carried out with the use of the Newton-Euler formalism. Moreover, the controllers for controlling the height and stabilizing the vehicle's location have been developed. Paper [7] describes the problem of flight dynamics of the UAV formation with the use of models with 3 and 6 degrees of freedom.

An operation of UAV during the completion of a task is a complex process requiring comprehensive technical means and systems. The control system of UAV is one of the most important systems. Its task is to measure, evaluate, and control flight parameters, as well as properly control the flight and observation systems. Thanks to the use of complex microprocessor systems, the comprehensive automation of the mentioned processes is possible. Formerly, during the completion of a mission it was necessary to maintain bilateral

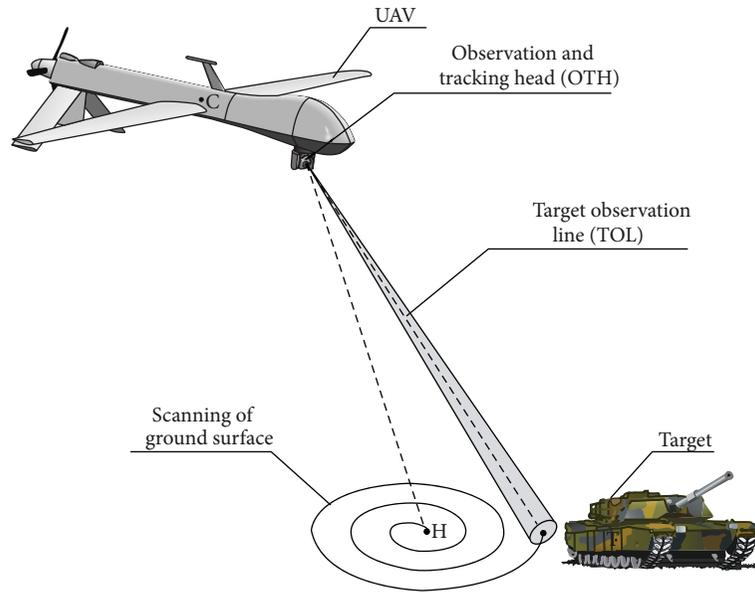


FIGURE 1: General view of the process of scanning the ground surface from UAV deck.

communications with the ground control point. In modern UAVs, autonomy plays an important role during detecting and tracking a ground target.

Paper [8] presents the model of the UAV autopilot and its sensors in the Matlab-Simulink environment for simulation research. Papers [9–11] pertain to planning and optimizing the trajectory of movement of the vehicle. The aspects of designing PID controllers for the systems of UAV flight control have been discussed in papers [12, 13].

From the quoted review of the literature it appears that the research mainly concentrates on UAV dynamics and control, without the consideration of the model of movement of the observation and tracking head (OTH) which is one of the most significant elements of the unmanned aerial vehicle. OTH is used for automatic searching for and tracking of targets intended for destruction [14]. This paper presents the mathematical model of kinematics of the movement of UAV, the ground target, and the dynamics of the controlled head located on the UAV deck. It needs to be emphasized that the problem of modelling, examining the dynamics, and control of such heads in the conditions of interferences from its base (deck of the manoeuvring UAV) is still topical. The paper examines this type of OTH with the built-in television and thermographic cameras and the laser illuminator.

2. Model of Movement Kinematics of UAV, OTH and Ground Target

General view of the process of searching for a target on the surface of the ground by OTH from UAV deck is shown in Figure 1, whereas the operation algorithm of OTH during scanning the ground surface and tracking a target detected on it is shown in Figure 2.

During the search for a ground target from UAV deck, the axis of OTH should perform the desired movements and circle strictly defined lines on the ground with the use of its extension. The optical system of OTH, having a certain viewing angle, may in this way encounter a light or infrared signal emitted by the moving object. Therefore, one should choose the kinematic parameters of mutual movement of UAV deck and OTH in such a way that the likelihood of detecting a target was the highest [15]. After locating a target, OTH goes to the tracking mode; that is, from that moment its axis has a specific location in space being pointed at the target.

Figure 3 shows the kinematics of mutual movement of the head axis and UAV during searching for a target on the ground surface. Individual coordinate systems have the following meaning:

$Ox_g y_g z_g$ —Earth-fixed reference system,

$Cx_a y_a z_a$ —coordinate system connected with target observation line (TOL),

$Cx_c y_c z_c$ —movable coordinate system connected with UAV velocity vector.

The velocity of changing vector \vec{R} in time, during searching for and tracking a target, is as follows:

$$\begin{aligned} \frac{d\vec{R}}{dt} = & \Pi(t_o, t_d) \cdot (\vec{V}_c - \vec{V}_h) \\ & + [\Pi(t_d, t_t) + \Pi(t_t, t_e)] \cdot (\vec{V}_c - \vec{V}_t), \end{aligned} \quad (1)$$

where \vec{R} is vector of the mutual distance of points C and H (during scanning the space) or points C and T (during

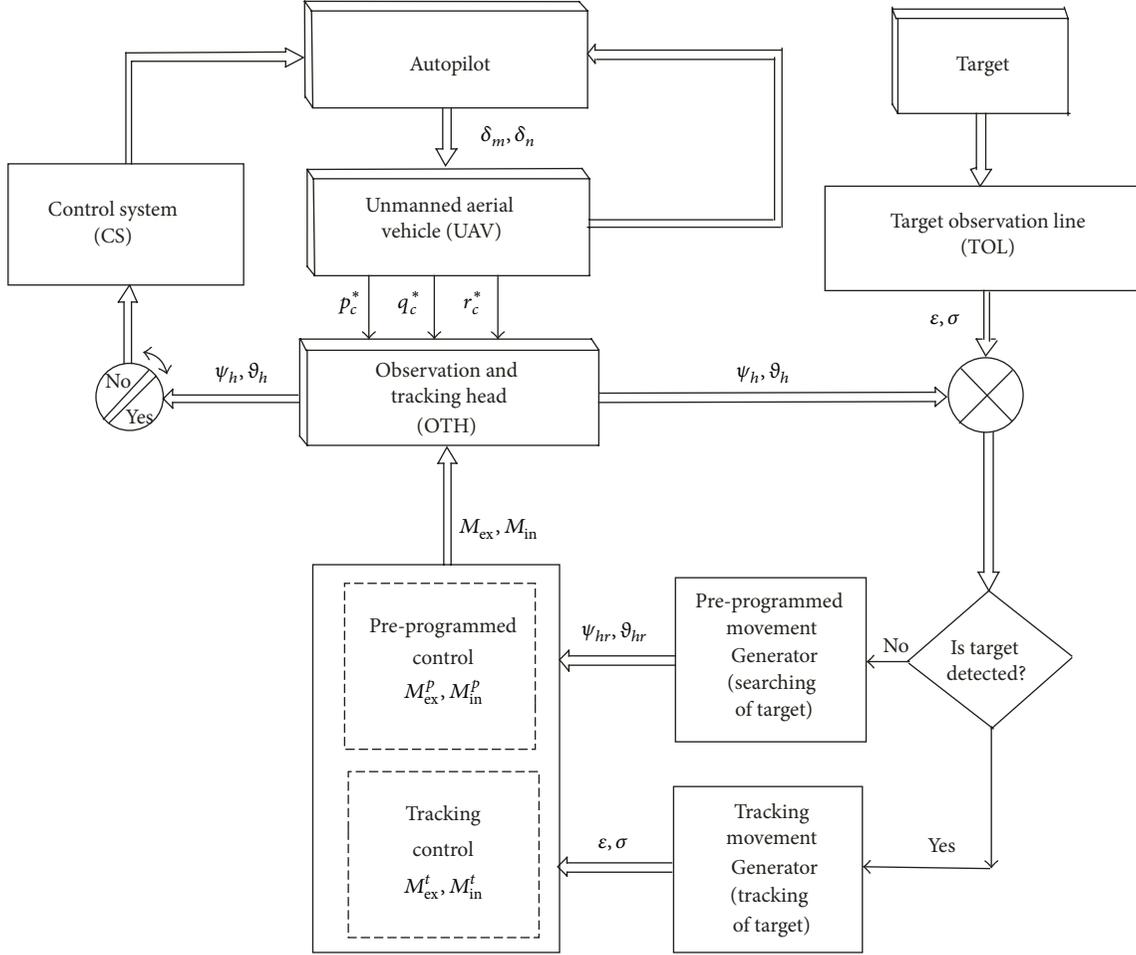


FIGURE 2: Diagram of operation of OTH mounted on UAV deck.

tracking); $\vec{V}_c, \vec{V}_h, \vec{V}_t$ are vectors of the velocity of movement of the centre of mass of UAV, of points H and T , respectively; $\Pi(t_o, t_d), \Pi(t_d, t_t), \Pi(t_t, t_e)$ are functions of rectangular impulse; t_o is the moment the scanning the ground is started; t_d is the moment a target is detected; t_t is the moment the tracking the target is started; t_e is the moment the scanning process and tracking the target is finished.

We project the left side of (1) on the axes of the coordinate system $Cx_a y_a z_a$ (Figure 3) connected with vector \vec{R} . In the result we get

$$\begin{aligned} \frac{d\vec{R}}{dt} &= \frac{dR}{dt} \vec{i}_a + \begin{vmatrix} \vec{j}_a & \vec{k}_a \\ \omega_{Rxa} & \omega_{Rya} & \omega_{Rza} \\ R & 0 & 0 \end{vmatrix} \\ &= \vec{i}_a \frac{dR}{dt} + \vec{j}_a R \frac{d\sigma}{dt} \cos \varepsilon - \vec{k}_a R \frac{d\varepsilon}{dt}, \end{aligned} \quad (2)$$

where $\omega_{Rxa}, \omega_{Rya}, \omega_{Rza}$ are components of angular velocity of TOL and σ, ε are angles of deflection and inclination of vector \vec{R} .

Next, we project the right side of (1) (i.e., velocities \vec{V}_c, \vec{V}_h and \vec{V}_t) on the axes of the system $Cx_a y_a z_a$ and get

$$\begin{aligned} \frac{dR}{dt} &= \Pi(t_o, t_d) \cdot (V_{cxa} - V_{hxa}) \\ &\quad + [\Pi(t_d, t_t) + \Pi(t_t, t_e)] \cdot (V_{cxa} - V_{txa}), \\ R \frac{d\sigma}{dt} \cos \varepsilon &= \Pi(t_o, t_d) \cdot (V_{cya} - V_{hya}) \\ &\quad + [\Pi(t_d, t_t) + \Pi(t_t, t_e)] \cdot (V_{cya} - V_{tya}), \\ R \frac{d\varepsilon}{dt} &= \Pi(t_o, t_d) \cdot (V_{cza} - V_{hza}) \\ &\quad + [\Pi(t_d, t_t) + \Pi(t_t, t_e)] \cdot (V_{cza} - V_{tza}). \end{aligned} \quad (3)$$

Individual components of velocity vectors \vec{V}_c, \vec{V}_h and \vec{V}_t in the system $Cx_a y_a z_a$ are as follows:

$$V_{cxa} = V_c [\cos \varepsilon \cos \gamma_c \cos (\sigma - \chi_c) + \sin \varepsilon \sin \gamma_c], \quad (4a)$$

$$V_{cya} = -V_c \cos \gamma_c \sin (\sigma - \chi_c), \quad (4b)$$

$$V_{cza} = V_c [\sin \varepsilon \cos \gamma_c \cos (\sigma - \chi_c) - \cos \varepsilon \sin \gamma_c], \quad (4c)$$

The path of UAV flight during scanning and tracking is as follows :

$$\begin{aligned} \frac{dr_c}{dt} = & \Pi(t_o, t_d) V_c \cos(\theta_c - \chi_c^s) + \Pi(t_d, t_t) V_c \cos(\theta_c - \chi_c^d) \\ & + \Pi(t_t, t_e) V_c \cos(\theta_c - \chi_c^t), \end{aligned} \quad (14a)$$

$$\begin{aligned} \frac{d\theta_c}{dt} = & \Pi(t_o, t_d) \frac{V_c}{r_c} \sin(\theta_c - \chi_c^s) + \Pi(t_d, t_t) \frac{V_c}{r_c} \sin(\theta_c - \chi_c^d) \\ & + \Pi(t_t, t_e) \frac{V_c}{r_c} \sin(\theta_c - \chi_c^t), \end{aligned} \quad (14b)$$

$$\begin{aligned} r_{cx} &= r_c \cos \theta_c, \\ r_{cy} &= r_c \sin \theta_c, \end{aligned} \quad (15)$$

where r_c is vector of location of UAV mass centre (point C) and θ_c is angle of deflection of vector r_c .

Path of movement of point H is as follows:

$$\frac{dR_h}{dt} = \Pi(t_o, t_d) V_h \cos(\theta_h - \chi_h), \quad (16a)$$

$$\frac{d\theta_h}{dt} = \Pi(t_o, t_d) V_h \sin(\theta_h - \chi_h), \quad (16b)$$

$$\begin{aligned} R_{hx} &= R_h \cos \theta_h, \\ R_{hy} &= R_h \sin \theta_h, \end{aligned} \quad (17)$$

where R_h is vector of location of point H and θ_h is angle of deflection of vector R_h .

Path of movement of the target is as follows:

$$\frac{dR_t}{dt} = \Pi(t_d, t_e) V_t \cos(\theta_t - \chi_t), \quad (18a)$$

$$\frac{d\theta_t}{dt} = \Pi(t_d, t_e) V_t \sin(\theta_t - \chi_t), \quad (18b)$$

$$\begin{aligned} R_{tx} &= R_t \cos \theta_t, \\ R_{ty} &= R_t \sin \theta_t, \end{aligned} \quad (19)$$

where R_t is vector of location of point T and θ_t is angle of deflection of vector R_t .

Desired angles ϑ_{hr} , ψ_{hr} and angular velocities $\dot{\vartheta}_{hr}$, $\dot{\psi}_{hr}$ of deflection of OTH axis can be determined from the following relationships:

$$\begin{aligned} \vartheta_{hr} &= \arctg \frac{r_x}{H_c}, & \frac{d\vartheta_{hr}}{dt} &= \frac{H_c (dr_x/dt)}{H_c^2 + (r_x)^2}, \\ \psi_{hr} &= \arctg \frac{r_y}{H_c}, & \frac{d\psi_{hr}}{dt} &= \frac{H_c (dr_y/dt)}{H_c^2 + (r_y)^2}, \end{aligned} \quad (20)$$

where

$$\begin{aligned} r_x &= r \cos \sigma, & r_y &= r \sin \sigma; \\ \frac{dr_x}{dt} &= \frac{dr}{dt} \cos \sigma - r \frac{d\sigma}{dt} \sin \sigma; \\ \frac{dr_y}{dt} &= \frac{dr}{dt} \sin \sigma + r \frac{d\sigma}{dt} \cos \sigma. \end{aligned} \quad (21)$$

Values (20) are used for determining the preprogrammed controls influencing the head.

2.1. Scanning the Ground Surface during UAV Flight on a Circle. We assume that during scanning the set area of the ground, UAV flight is at a constant altitude H_c in a horizontal plane on the circle of the set radius r_c with constant velocity V_c (Figure 3).

Then the preprogrammed UAV flight can be determined from the following relationships:

$$V_c = \omega_c \cdot r_c, \quad \chi_c = \omega_c t. \quad (22)$$

At the same time, we control the head axis in such a way that it drawn on the surface of the ground a curve in the shape of Archimedean spiral with angular velocity

$$\omega_h = \frac{2\pi V_h}{\rho_h}. \quad (23)$$

Velocity V_h is to be chosen in such a way that OTH mounted on UAV deck could in the set time t_s scan densely enough the set surface of the area in the shape of a circle of the radius R_s .

If the angle of vision of the head's optical system amounts to ϕ_h then lens coverage embraces the surface similar in shape to a circle of the radius:

$$\rho_h = \frac{1}{2} H_c \phi_h. \quad (24)$$

After detecting a target at the moment t_d , UAV passes into the tracking flight according to the relationship (12), while the target is lit with a laser beam for the period of time t_l . Hence, the total time of the process of detecting, tracking, and lighting the target amounts to $t_e = t_d + t_l$.

3. The Model of Dynamics of the Controlled Observation and Tracking Head

A spatial model of dynamics of the head presented in Figure 4 was adopted for the paper. OTH comprises two basic parts: external frame and internal frame with camera. Movement of the head is determined with the use of two angles: angle of head deflection ψ_h and angle of head inclination ϑ_h [16].

The following coordinate systems, shown in Figure 5, have been introduced:

$Cx_d y_d z_d$ —the movable system connected with the UAV deck,

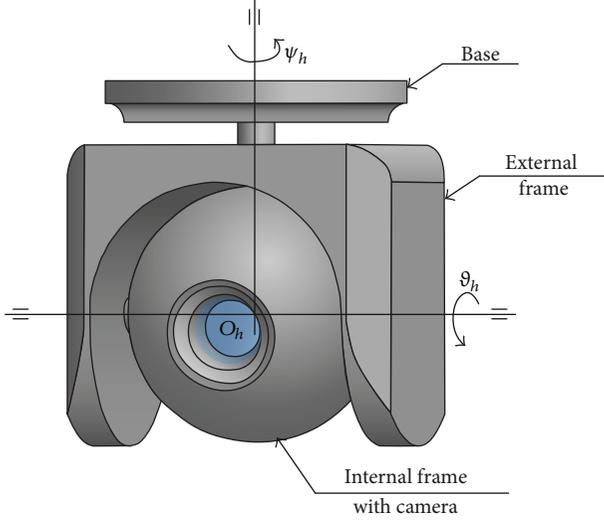


FIGURE 4: General view of the observation and tracking head.

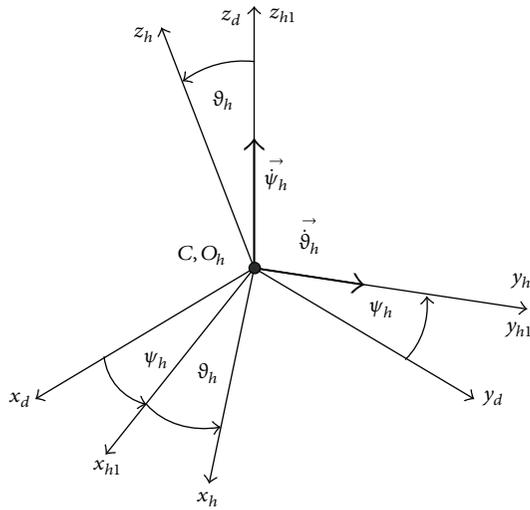


FIGURE 5: Transformations of the coordinate systems.

$O_h x_{h1} y_{h1} z_{h1}$ —the movable system connected with the external frame of the head,

$O_h x_h y_h z_h$ —the movable system connected with internal the frame of the head (including camera).

Components of angular velocity of the external frame of the head are

$$\omega_{x_{h1}} = p_c^* \cos \psi_h + q_c^* \sin \psi_h, \quad (25a)$$

$$\omega_{y_{h1}} = -p_c^* \sin \psi_h + q_c^* \cos \psi_h, \quad (25b)$$

$$\omega_{z_{h1}} = \dot{\psi}_h + r_c^*, \quad (25c)$$

where p_c^*, q_c^*, r_c^* are components of angular velocity of UAV.

Components of angular velocity of the internal frame of the head are

$$\omega_{x_h} = \omega_{x_{h1}} \cos \vartheta_h - \omega_{z_{h1}} \sin \vartheta_h, \quad (26a)$$

$$\omega_{y_h} = \omega_{y_{h1}} + \dot{\vartheta}_h, \quad (26b)$$

$$\omega_{z_h} = \omega_{x_{h1}} \sin \vartheta_h + \omega_{z_{h1}} \cos \vartheta_h. \quad (26c)$$

Components of linear velocity of displacement of mass centre of the external frame of the head are

$$V_{x_{h1}} = V_c \cos \psi_h, \quad (27a)$$

$$V_{y_{h1}} = -V_c \sin \psi_h, \quad (27b)$$

$$V_{z_{h1}} = 0. \quad (27c)$$

Components of linear velocity of displacement of mass centre of the internal frame of the head are

$$V_{x_h} = V_{x_{h1}} \cos \vartheta_h + l_h \omega_{y_{h1}} \sin \vartheta_h, \quad (28a)$$

$$V_{y_h} = V_{y_{h1}} + l_h (\omega_{z_{h1}} + \omega_{z_h}), \quad (28b)$$

$$V_{z_h} = V_{x_{h1}} \sin \vartheta_h - l_h (\omega_{y_{h1}} \cos \vartheta_h + \omega_{y_h}). \quad (28c)$$

Using the second order Lagrange equations, the following equations of the head movement have been derived:

$$\begin{aligned} & J_{z_{h1}} \frac{d}{dt} \omega_{z_{h1}} - J_{x_{h1}} \frac{d}{dt} (\omega_{x_{h1}} \sin \vartheta_h) + J_{z_h} \frac{d}{dt} (\omega_{z_h} \cos \vartheta_h) \\ & + m_{rw} l_h \frac{d}{dt} [V_{y_h} (1 + \cos \vartheta_h)] - (J_{x_{h1}} - J_{y_{h1}}) \omega_{x_{h1}} \omega_{y_{h1}} \\ & - J_{x_h} \omega_{x_h} \omega_{y_{h1}} \cos \vartheta_h + J_{y_h} \omega_{y_h} \omega_{x_{h1}} - J_{z_h} \omega_{z_h} \omega_{y_{h1}} \sin \vartheta_h \\ & + m_{in} l_h [V_{x_h} \omega_{x_{h1}} \sin \vartheta_h - V_{y_h} \omega_{y_{h1}} \sin \vartheta_h \\ & - V_{z_h} \omega_{x_{h1}} (1 + \cos \vartheta_h)] \\ & + m_{in} l_h [V_{x_{h1}} (\omega_{z_{h1}} + \omega_{z_h}) + V_{y_{h1}} \omega_{y_h} \sin \vartheta_h] \\ & = M_{ex} + M_{ex}^g - M_{dex}, \\ & J_{y_h} \frac{d}{dt} \omega_{y_h} - m_{in} l_h \frac{d}{dt} V_{z_h} + (J_{x_h} - J_{z_h}) \omega_{x_h} \omega_{z_h} \\ & - m_{in} l_h [V_{y_h} \omega_{x_h} - V_{x_h} \omega_{y_h}] = M_{in} + M_{in}^g - M_{din}, \end{aligned} \quad (29)$$

where ψ_h, ϑ_h are angles of location of OTH axis in space; m_{in} is the mass of internal frame with camera; l_h is the distance of mass centre of internal frame (camera) from the centre of movement; $J_{x_{h1}}, J_{y_{h1}}, J_{z_{h1}}$ are moments of inertia of external frame; $J_{x_h}, J_{y_h}, J_{z_h}$ are moments of inertia of internal frame (with camera); M_{ex}, M_{in} are moments of controlling forces influencing, respectively, external and internal frame; M_{ex}^g, M_{in}^g are moments of the force of gravity influencing

respectively: external and internal frame; $M_{\text{ex}}^g = 0$; $M_{\text{in}}^g = m_{\text{in}} g l_h \cos \vartheta_h$; and $\vec{M}_{d_{\text{ex}}}$, $\vec{M}_{d_{\text{in}}}$ are moments of friction forces in the bearings of, respectively, external and internal frame.

We assume viscous friction:

$$M_{d_{\text{ex}}} = \eta_{\text{ex}} \dot{\psi}_h; \quad M_{d_{\text{in}}} = \eta_{\text{in}} \dot{\vartheta}_h, \quad (30)$$

where η_{ex} is friction coefficient in suspension bearing of external frame and η_{in} is friction coefficient in suspension bearing of internal frame.

Control moments M_{ex} , M_{in} of the head will be presented as follows [17–19]:

$$\begin{aligned} M_{\text{ex}} &= \Pi(t_o, t_d) \cdot M_{\text{ex}}^p(t) + \Pi(t_t, t_e) \cdot M_{\text{ex}}^t(t), \\ M_{\text{in}} &= \Pi(t_o, t_d) \cdot M_{\text{in}}^p(t) + \Pi(t_t, t_e) \cdot M_{\text{in}}^t(t), \end{aligned} \quad (31)$$

where M_{ex}^p , M_{in}^p are preprogrammed control moments and M_{ex}^t , M_{in}^t are tracking control moments.

Preprogrammed control moments M_{ex}^p , M_{in}^p , which set the head axis into the required motion, are determined from the following relationships:

$$\begin{aligned} M_{\text{ex}}^p(t) &= \Pi(t_o, t_d) \cdot [k_{\text{ex}}(\psi_h - \psi_{hr}) + h_{\text{ex}}(\dot{\psi}_h - \dot{\psi}_{hr})], \\ M_{\text{in}}^p(t) &= \Pi(t_o, t_d) \cdot [k_{\text{in}}(\vartheta_h - \vartheta_{hr}) + h_{\text{in}}(\dot{\vartheta}_h - \dot{\vartheta}_{hr})], \end{aligned} \quad (32)$$

where k_{ex} , k_{in} are gain coefficients of the OTH control system and h_{ex} , h_{in} are attenuation coefficients of the OTH control system.

Angles ϑ_{hr} , ψ_{hr} and their derivatives will be determined from the relationships (20).

At the moment when the target will appear in the head coverage, that is,

$$|\vec{R}_t - \vec{R}_h| \leq \rho_h, \quad (33)$$

head control passes to the tracking mode.

Then, tracking control moments have the following form:

$$\begin{aligned} M_{\text{ex}}^t(t) &= \Pi(t_t, t_e) \cdot [k_{\text{ex}}(\psi_h - \sigma) + h_{\text{ex}}(\dot{\psi}_h - \dot{\sigma})], \\ M_{\text{in}}^t(t) &= \Pi(t_t, t_e) \cdot [k_{\text{in}}(\vartheta_h - \varepsilon) + h_{\text{in}}(\dot{\vartheta}_h - \dot{\varepsilon})]. \end{aligned} \quad (34)$$

Angles σ , ε will be determined from the relationships (3).

Coefficients k_{ex} , k_{in} , h_{ex} , h_{in} are chosen in an optimum way due to the minimum deviation between the real and set path [15].

It should be emphasized that the mathematical model of movement of the controlled observation and tracking head described with (29) allows for conducting a number of simulation tests of searching for and tracking a ground target from the UAV deck. Thanks to that, one can know about the areas of stability and permissible controls with the influence of kinematic excitations from the UAV deck.

4. Results of Computer Simulations

The presented algorithms of control of the movement of OTH performing the search and tracking of a ground target

from the UAV deck have been tested on the example of a hypothetical UAV system. The following parameters were adopted:

movement parameters of UAV

$$H_c = 1500 \text{ m}, \quad r_c = 500 \text{ m}, \quad V_c = 75 \text{ m/s}; \quad (35)$$

movement parameters of the head

$$\begin{aligned} t_s &= 10 \text{ s}, \quad t_l = 20 \text{ s}, \\ R_s &= 500 \text{ m}, \quad \phi_h = 1 \text{ deg}; \end{aligned} \quad (36)$$

movement parameters of the target

$$\begin{aligned} V_t &= 25 \text{ m/s}, \quad \chi_t = \omega_t \cdot t, \\ \omega_t &= 0.025 \text{ rad/s}; \end{aligned} \quad (37)$$

mass parameters of the head

$$\begin{aligned} m_{\text{in}} &= 3.375 \text{ kg}, \quad J_{x_{h1}} = 0.22 \text{ kgm}^2, \\ J_{y_{h1}} &= 0.114 \text{ kgm}^2, \quad J_{z_{h1}} = 0.117 \text{ kgm}^2, \\ J_{x_h} &= 0.061 \text{ kgm}^2, \quad J_{y_h} = 0.035 \text{ kgm}^2, \\ J_{z_h} &= 0.029 \text{ kgm}^2; \end{aligned} \quad (38)$$

the distance of mass centre of internal frame from the centre of movement

$$l_h = 0.002 \text{ m}; \quad (39)$$

friction coefficients in suspension bearings of the head

$$\eta_{\text{ex}} = \eta_{\text{in}} = 0.01 \text{ Nms/rad}. \quad (40)$$

Kinematic excitations of the base (UAV) were adopted in the following form:

$$\begin{aligned} p_c^* &= p_{c0}^* \sin(\nu \cdot t), \quad q_c^* = q_{c0}^* \cos(\nu \cdot t), \\ r_c^* &= r_{c0}^* \sin(\nu \cdot t), \quad p_{c0}^* = q_{c0}^* = r_{c0}^* = 0.5 \text{ rad/s}, \\ \nu &= 5 \text{ rad/s}. \end{aligned} \quad (41)$$

For nonoptimum control, regulator coefficients amounted to

$$\begin{aligned} k_{\text{ex}} &= -5.0, \quad h_{\text{ex}} = -1.5, \\ k_{\text{in}} &= -5.0, \quad h_{\text{in}} = -1.5. \end{aligned} \quad (42)$$

For optimum control, regulator coefficients amounted to

$$\begin{aligned} k_{\text{ex}} &= -20.0, \quad h_{\text{ex}} = -5.0, \\ k_{\text{in}} &= -20.0, \quad h_{\text{in}} = -5.0. \end{aligned} \quad (43)$$

Figure 6 presents the results of computer simulation of movement kinematics of UAV as well as the head axis during

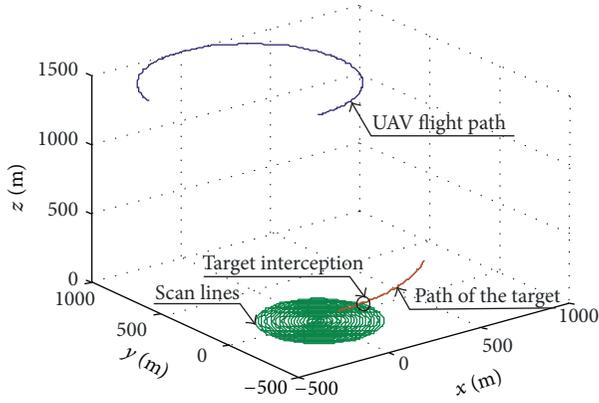


FIGURE 6: Path of movement of UAV, head axis, and the target during searching for and tracking the target.

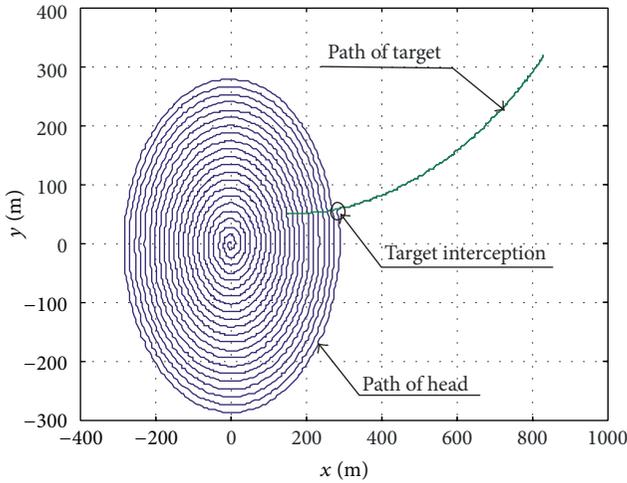


FIGURE 7: Path of movement of point H and the target on the surface of the ground.

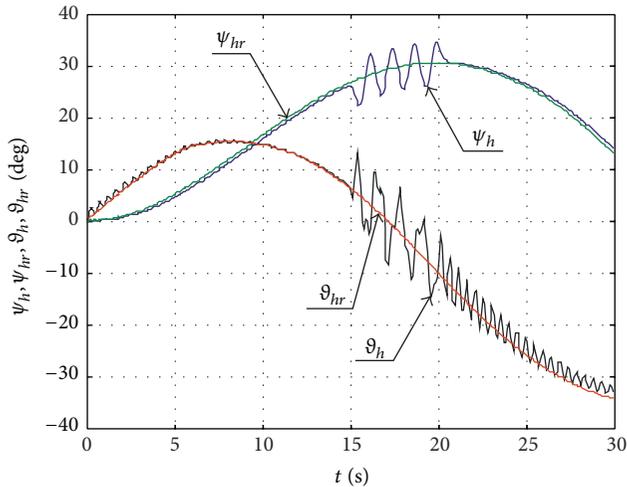


FIGURE 8: Time-dependent real ϑ_h, ψ_h and desired $\vartheta_{hr}, \psi_{hr}$ angles specifying the location of the head axis for nonoptimum controls with the influence of disturbances.

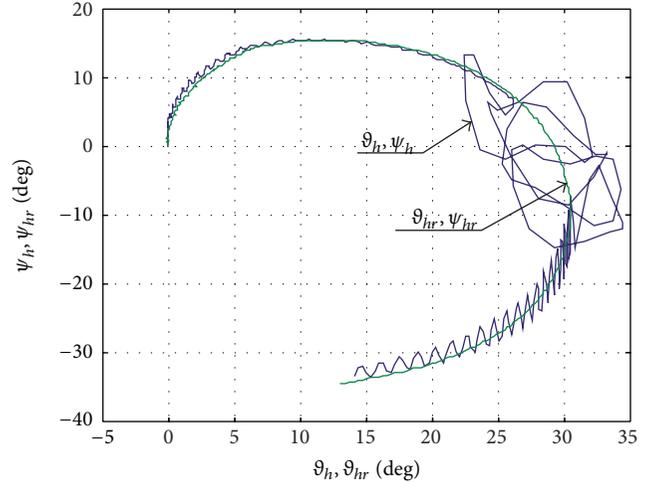


FIGURE 9: Real and desired trajectory of movement of the head axis for nonoptimum controls with the influence of disturbances.

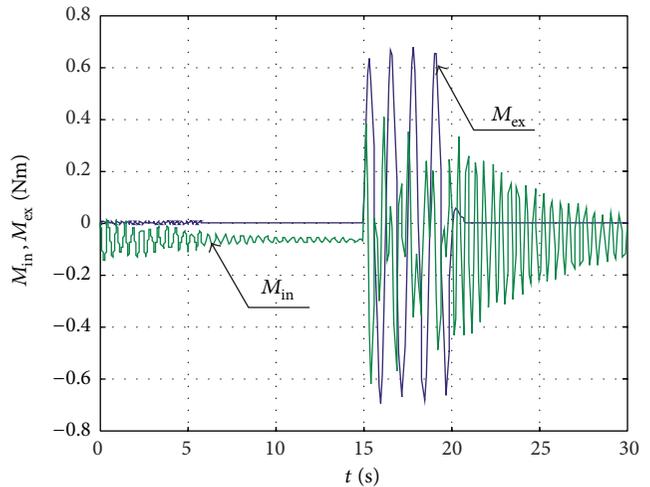


FIGURE 10: Time-dependent control moments M_{in} and M_{ex} for nonoptimum parameters of the regulator with the influence of disturbances.

searching the surface of the ground and tracking (laser lighting) a ground target.

Figure 7 presents the trajectory of movement of point H during scanning for and the movement of the target on the surface of the ground.

Figures 8–10 present the desired and the real angles of deflection and inclination of the head axis, as well as control moments for the case when the head is influenced by kinematic excitations from UAV deck and the parameters of the regulator are not optimum.

Figures 11, 12, and 13 present the desired and the real angles of deflection and inclination of the head axis, as well as control moments for the case when the head is not influenced by kinematic excitations from UAV deck and the parameters of the regulator are optimum.

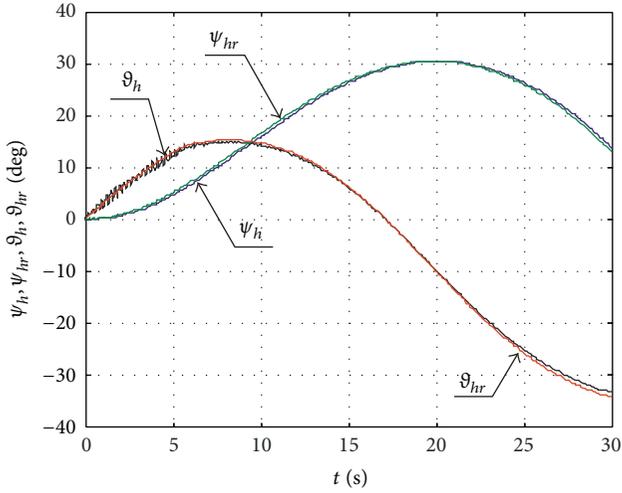


FIGURE 11: Time-dependent real ϑ_h, ψ_h and desired $\vartheta_{hr}, \psi_{hr}$ angles specifying the location of the head axis for optimum controls without the influence of disturbances.

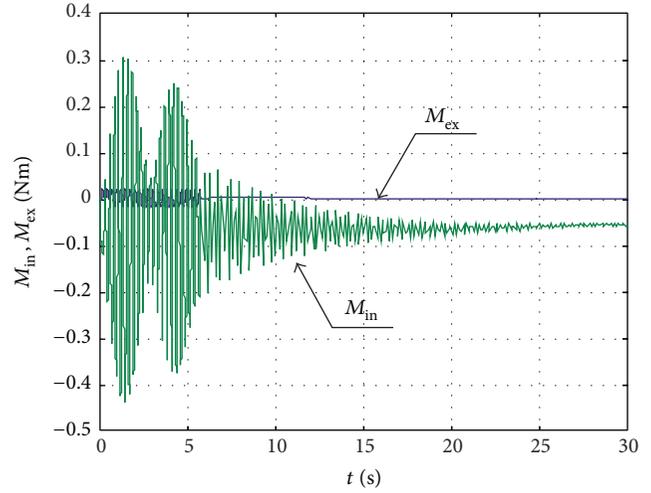


FIGURE 13: Time-dependent control moments M_{in} and M_{ex} for optimum parameters of the regulator without the influence of disturbances.

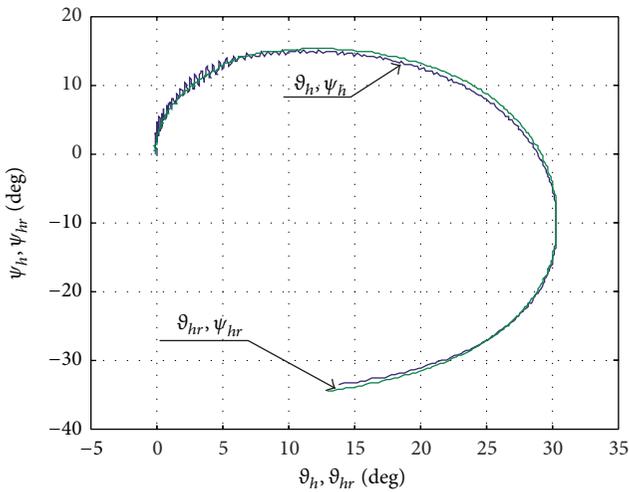


FIGURE 12: Real and desired trajectory of movement of the head axis for optimum controls without the influence of disturbances.

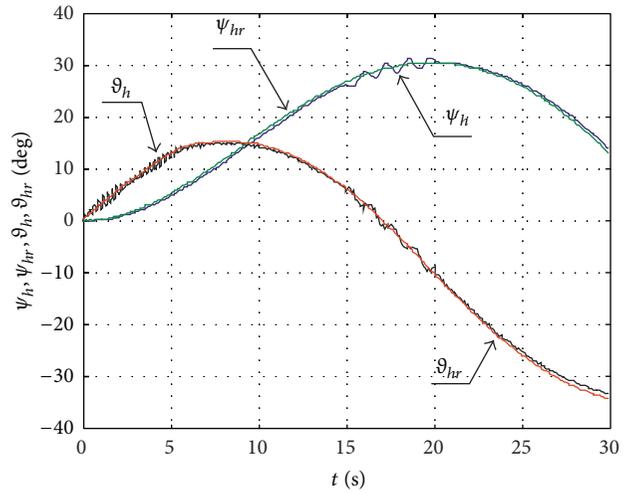


FIGURE 14: Time-dependent real ϑ_h, ψ_h and desired $\vartheta_{hr}, \psi_{hr}$ angles specifying the location of the head axis for optimum controls with the influence of disturbances.

Figures 14–16 present the desired and the real angles of deflection and inclination of the head axis, as well as control moments for the case when the head is influenced by kinematic excitations from UAV deck and the parameters of the regulator are optimum.

Kinematic excitations from UAV deck adversely affect the operation of the head. It can particularly be seen in Figures 11 and 14. In case of the nonoptimum choice of the parameters of the regulator, the deviations of the head axis from the set location are particularly visible (Figures 8 and 9). Smaller values of deviations can of course be achieved for optimum head controls (Figures 14 and 15). Control moments take small values.

From the presented theoretical analysis and the simulation research of the process of scanning by OTH from UAV

deck of the surface of the ground and then tracking the detected ground target, it can be inferred that

- (i) it is possible to search for a target on the area of any size, which is only limited to the durability and range of UAV flight;
- (ii) scanning is sufficiently precise;
- (iii) the programme of scanning the surface of the ground is simple;
- (iv) there occur relatively small values of OTH axis angle deviations from the nominal position;
- (v) full autonomy of UAV during the mission of searching and laser lighting of the detected ground target secures the point of control against detection and destruction by an enemy;

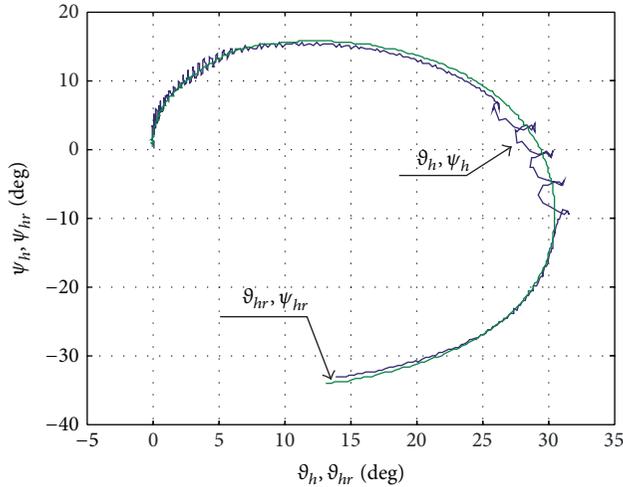


FIGURE 15: Real and desired trajectory of movement of the head axis for optimum controls with the influence of disturbances.

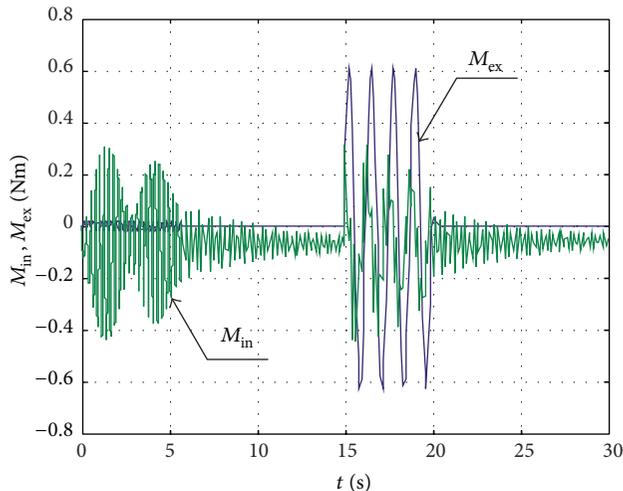


FIGURE 16: Time-dependent control moments M_{in} and M_{ex} for optimum parameters of the regulator without the influence of disturbances.

- (vi) the interference of an operator in controlling UAV may only be limited to cases of the vehicle's total coming off from the set path or the loss of the target from OTH lens coverage (due to gusts of wind, explosions of missiles, etc.). Hence, the possibility of automatic sending of information about such occurrences to the control point and the possible taking of control over UAV flight by the operator should be introduced.

5. Conclusion

The considerations presented in this paper will allow us to conduct the research on the dynamics of OTH when manoeuvring UAV on which this device is mounted. It may enable the construction engineers to choose such parameters of OTH so that the transient processes occurring in the conditions of kinematic influence of UAV deck and at the

moment of passing from scanning to tracking mode were minimized and disappeared in the shortest possible time.

Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

Acknowledgment

The work reported herein was undertaken as part of a research project supported by the National Centre for Research and Development over the period 2011–2014.

References

- [1] R. Austin, *Unmanned Aircraft Systems: UAVS Design, Development and Deployment*, John Wiley & Sons, Chichester, UK, 2010.
- [2] F. Kamrani and R. Ayani, *UAV Path Planning in Search Operations, Aerial Vehicles*, InTech, 2009, edited by T. M. Lam.
- [3] X. Q. Chen, Q. Ou, D. R. Wong et al., "Flight dynamics modelling and experimental validation for unmanned aerial vehicles," in *Mobile Robots-State of the Art in Land, Sea, Air, and Collaborative Missions*, X. Q. Chen, Ed., InTech, 2009.
- [4] S. Bhandari, R. Colgren, P. Lederbogen, and S. Kowalchuk, "Six-DoF dynamic modeling and flight testing of a UAV helicopter," in *Proceedings of the AIAA Modeling and Simulation Technologies Conference*, pp. 992–1008, San Francisco, Calif, USA, August 2005.
- [5] J. Manerowski and D. Rykaczewski, "Modelling of UAV flight dynamics using perceptron artificial neural networks," *Journal of Theoretical and Applied Mechanics*, vol. 43, no. 2, pp. 297–307, 2005.
- [6] K. T. Öner, E. Çetinsoy, E. Sirimoğlu et al., "Mathematical modeling and vertical flight control of a tilt-wing UAV," *Turkish Journal of Electrical Engineering and Computer Sciences*, vol. 20, no. 1, pp. 149–157, 2012.
- [7] T.-V. Chelaru, V. Pana, and A. Chelaru, "Flight dynamics for UAV formations," in *Proceedings of the 10th WSEAS International Conference on Automation & Information*, pp. 287–292, 2009.
- [8] P. Kanovsky, L. Smrcek, and C. Goodchild, "Simulation of UAV Systems," *Acta Polytechnica*, vol. 45, no. 4, pp. 109–113, 2005.
- [9] H. Ergezer and K. Leblebicioglu, "Planning unmanned aerial vehicles path for maximum information collection using evolutionary algorithms," in *Proceedings of the 18th IFAC World Congress*, pp. 5591–5596, Milano, Italy, 2011.
- [10] C. Sabo and K. Cohen, "Fuzzy logic unmanned air vehicle motion planning," *Advances in Fuzzy Systems*, vol. 2012, Article ID 989051, 14 pages, 2012.
- [11] H. Chen, K. Chang, and C. S. Agate, "Tracking with UAV using tangent-plus-lyapunov vector field guidance," in *Proceedings of the 12th International Conference on Information Fusion (FUSION '09)*, pp. 363–372, Seattle, Wash, USA, July 2009.
- [12] K. Turkoglu, U. Ozdemir, M. Nikbay et al., "PID parameter optimization of an UAV longitudinal flight control system," *Proceedings of World Academy of Science: Engineering & Technology*, vol. 2, no. 9, pp. 24–29, 2008.

- [13] B. Kada and Y. Ghazzawi, "Robust PID controller design for an UAV flight control system," in *Proceedings of the World Congress on Engineering and Computer Science (WCECS '11)*, vol. 2, San Francisco, Calif, USA, October 2011.
- [14] I. Moir and A. Seabridge, *Military Avionics Systems*, John Wiley & Sons, Chichester, UK, 2007.
- [15] J. Awrejcewicz and Z. Koruba, *Classical Mechanics: Applied Mechanics and Mechatronics*, vol. 30, Springer, New York, NY, USA, 2012.
- [16] I. Krzysztofik, "The dynamics of the controlled observation and tracking head located on a moving vehicle," *Solid State Phenomena*, vol. 180, pp. 313–322, 2012.
- [17] Z. Koruba, "Optimal control of the searching and tracking head (Sth) for self propelled anti aircraft vehicle," *Solid State Phenomena*, vol. 180, pp. 27–38, 2012.
- [18] R. Langton, *Stability and Control of Aircraft Systems*, John Wiley & Sons, Chichester, UK, 2007.
- [19] J. E. Takosoglu, P. A. Laski, and S. Blasiak, "A fuzzy logic controller for the positioning control of an electro-pneumatic servo-drive," *Proceedings of the Institution of Mechanical Engineers I: Journal of Systems and Control Engineering*, vol. 226, no. 10, pp. 1335–1343, 2012.

Research Article

Optimal Acceleration-Velocity-Bounded Trajectory Planning in Dynamic Crowd Simulation

Fu Yue-wen,¹ Li Meng,² Liang Jia-hong,¹ and Hu Xiao-qian¹

¹ College of Information System and Management, National University of Defense Technology, Changsha, Hunan 410073, China

² Army Officer Academy of PLA, Hefei 230031, China

Correspondence should be addressed to Fu Yue-wen; thdinf@126.com

Received 24 April 2014; Accepted 28 July 2014; Published 31 August 2014

Academic Editor: Yuxin Zhao

Copyright © 2014 Fu Yue-wen et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Creating complex and realistic crowd behaviors, such as pedestrian navigation behavior with dynamic obstacles, is a difficult and time consuming task. In this paper, we study one special type of crowd which is composed of urgent individuals, normal individuals, and normal groups. We use three steps to construct the crowd simulation in dynamic environment. The first one is that the urgent individuals move forward along a given path around dynamic obstacles and other crowd members. An optimal acceleration-velocity-bounded trajectory planning method is utilized to model their behaviors, which ensures that the durations of the generated trajectories are minimal and the urgent individuals are collision-free with dynamic obstacles (e.g., dynamic vehicles). In the second step, a pushing model is adopted to simulate the interactions between urgent members and normal ones, which ensures that the computational cost of the optimal trajectory planning is acceptable. The third step is obligated to imitate the interactions among normal members using collision avoidance behavior and flocking behavior. Various simulation results demonstrate that these three steps give realistic crowd phenomenon just like the real world.

1. Introduction

Crowds, ubiquitous in the real world from groups of humans to flocks of insects, are vital features to model in a virtual environment. The simulation of pedestrians is a difficult challenge that is beginning to capture the attention of researchers and practitioners in evacuation simulation and urban planning. The field of computer graphics, in which virtual human animation has been an important research interest for decades, has contributed technologies fundamental to the computer-assisted visualization, including the automatic animation of pedestrian [1–3]. Until now, various simulation models and architectures have been developed. Whether the behaviors of the crowd are realistic or not has many relations with the methods used in constructing the crowd models. This paper focuses on one special crowd phenomenon which is significant in real world. The crowd which is composed of urgent members and normal ones has the following characteristics: the urgent members are the people who move forward along a fixed, given path for something urgent to do. The normal ones are the people who

move in normal ways, just like most of the people in their daily lives. In real world, urgent people do not allow others to get in their way. Consequently, if the normal members are in a predefined range of the urgent ones, the latter will push them away. However, we must ensure that the urgent members are collision-free with the dynamic obstacles in the environment. Moreover, the interactions among the normal members are important issues to improve the authenticity of the crowd behaviors. For this special crowd, all of these propose demands for the studies of its behaviors.

Given the observations in real world, how is it possible that the urgent people can safely navigate through crowds and dynamic obstacles along a predefined path? The key insight is that the urgent people typically engage in moving forward just considering the dynamic obstacle: they push the normal members away from their trajectories to make room for their navigations.

In our work, we have been developing an autonomous, self-animating model of pedestrian crowd capable of performing a broad variety of natural activities in urban

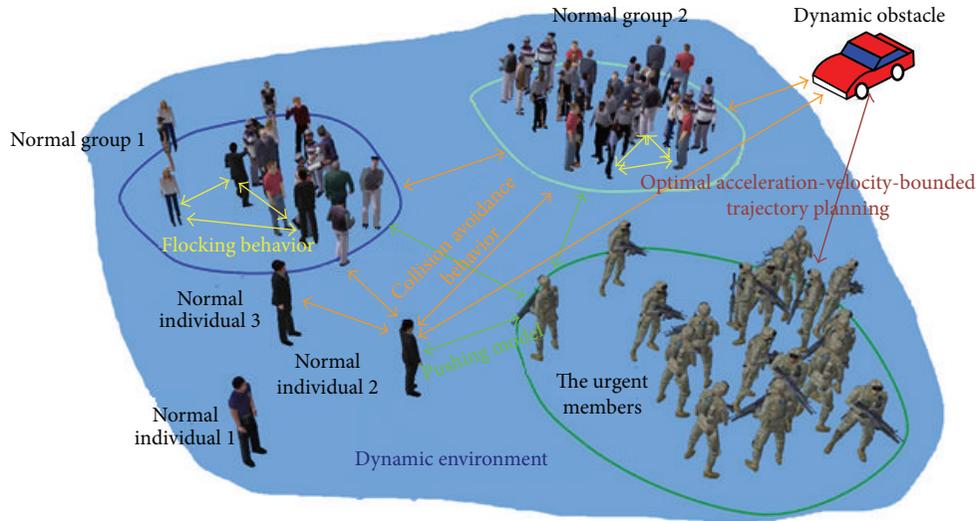


FIGURE 1: Components of the crowd behaviors.

dynamic environments. Because the acceleration and velocity of real person are bounded, we have adopted an optimal acceleration-velocity-bounded trajectory planning method in planning area and a comprehensive artificial life approach (e.g., collision avoidance method and flocking method) to addressing the problem of pedestrian behaviors. Our approach is inspired most heavily by the work of Johnson and Hauser [4] on optimal trajectory planning and by Foudil and Nouredine [5] on collision avoidance modeling in crowd simulation with priority rules. Also, Olfati-Saber's work on flocking model is used for reference [6].

In summary, our main contributions exist in three aspects.

- (1) Construct the behaviors of the urgent members with an optimal acceleration-velocity-bounded trajectory planning method. This novel method ensures that the urgent members reach their destinations as soon as possible and are collision-free with the dynamic obstacles.
- (2) Simulate the interactions between the urgent members and the normal ones. A novel pushing model is adopted to ensure that the normal members will not get in the way of the urgent ones. In this way, the normal members will not be treated as dynamic obstacles which promotes that the computational cost of the trajectory planning is acceptable for real time application.
- (3) Imitate the interactions among the normal members. These interactions play a major role in synthesizing the animations of the crowd. The key insight is that the normal members typically engage in joint collision avoidance: they adapt their moving to each other to make room for navigation. So the collision avoidance behavior and the flocking behavior are exploited.

The above three aspects are combined together to generate realistic behaviors for our special crowd. The components of the crowd are represented in Figure 1.

As illustrated in Figure 1, the crowd behaviors consist of four parts: optimal trajectory planning, pushing model, collision avoidance behavior, and flocking behavior, each of which is responsible for dealing with different interactions. By designing and combining the respective behaviors of these four parts, realistic and complex behavior phenomena are achieved.

The organization of the rest of this paper is as follows. At first, Section 2 describes background and related works in trajectory planning, collision avoidance model, and flocking behavior. In Section 3, we give the detailed process of the optimal acceleration-velocity-bounded trajectory planning method and the pushing model. In Section 4, the collision avoidance model and flocking behavior for crowd simulation are presented briefly. After performing experiments with 3D virtual environments, simulation results and conclusions are stated at the end of this paper.

2. Related Work

In this paper, the behavior of the urgent members is a trajectory planning problem which has a vast difference with path planning. In respect of optimal trajectory planning problem with dynamic constraints in dynamic environments, Fraichard first addressed optimal trajectory planning in dynamic environments, that is, motion planning for robot subject to dynamic constraints and moving in a workspace with moving obstacles [7]. A novel concept of state-time space was introduced and a near-time-optimal trajectory was obtained by searching a set of canonical trajectories. The state-time space for planning in a dynamic environment is the Cartesian product of configuration space and time, that is, $C \times T$. Fraichard defined the canonical trajectories as the ones

which had piecewise constant acceleration and changed its value at discrete times [7]. Fiorini and Shiller also presented a two-stage approach to plan a local time-optimal trajectory for a manipulator arm in the dynamic environments [8]. An optimization-based planning was used by Park et al. to handle the planning problem in arbitrary dynamic environments [9]. However, the dynamic constraints were ignored in their work [9]. Fiorini and Shiller presented a global optimal motion planning for Mars Rover which accounted for traversability and dynamic stability [10]. Recently, Maček et al. adopted rapidly exploring random trees and B-splines to develop a collision-free trajectory planning method in dynamic urban-like scenarios. This method can run automatically with the natural acceleration/deceleration phases according to the dynamic obstacles along the predefined path. The advantage of this method is that the vehicle velocities can be searched with a minimum time. Maček et al. have used kinodynamic planning techniques to explore the vehicle's action space [11]. While these above algorithms are able to find paths as well as velocity profiles, they have the defect of ignoring the acceleration constraint or obtaining near-optimal trajectories.

More recently, Johnson and Hauser have presented an optimal, exact, polynomial-time planner for optimal bounded-acceleration trajectories along a fixed, given path with dynamic obstacles [4]. In their method, the planner uses a velocity interval propagation technique to compute reachable velocity sets at obstacle vertices in the path-time space. In this process, the acceleration and velocity constraints are considered in path-velocity plane. This approach gives us great promotions in fulfillment of this paper to generate the optimal trajectories of urgent members. However, in this method the goal velocity is an interval rather than a fixed value. This is not the case in real life; for example, the goal velocity of the virtual human at the goal position is equal to zero. So the optimal trajectory planning in this paper is more complex and intractable.

In respect of collision avoidance behaviors for individuals and groups, techniques for simulating a crowd as a single entity have been proposed, as well as those which consider each person in the crowd separately. Collision avoidance is one of the crucial problems in crowd simulation. Without collision avoidance, crowd simulation does not look realistic. In crowd simulation, static and dynamic objects are the two types of obstacle that must concern us. The motionless objects can be seen as static obstacles, while dynamic objects refer to human, animals, and vehicles in the urban environments [12]. However, devising collision avoidance behavior among a number of individuals is more complex than designing the one between just two individuals. Morini et al. presented a short-term avoidance algorithm [13]. Aiming at the applications of densely populated urban environment, Tecchia and Chrysanthou proposed a simple and fast collision detection method. In their method, two different approaches for detecting and avoiding collision of the moving objects were used. Feurtey proposed an algorithm for collision avoidance which was planned from their current position to the goal based on the predicting and modifying trajectories. So in their method, the agents could predict their way by using the position and speed information of the obstacles [13]. Musse and Thalmann

developed a multiresolution collision avoidance behavior for group interrelationship in multiresolution [14]. Foudil and Noureddine proposed a novel collision avoidance behavior for crowd simulation with priority rules. This method also provides us with great inspirations for our paper to generate various collision avoidance strategies for normal members [5].

In respect of flocking behavior for groups, many researchers used flocking algorithm to model flock, herd, and large crowds of people. Flocking is one of the pioneer techniques used to model a crowd of animals or humans. Flocking is basically composed of three rules: separation-steer away from neighbors and obstacle, cohesion-steer towards the centre of nearby entities, and alignment-steer in the average velocity and direction with the neighbors. In recent years, various distributed control algorithms have been proposed for the motion of multiple dynamic agents and the multiagent systems. For example, the flocking problem with one virtual leader was considered by Shi et al. [15]. Su et al. investigated the problem of controlling a group of autonomous agents to track multiple virtual leaders. They proved that the agents with the same virtual leader attained the same velocity, and the center and average velocity of the whole group converged to the weighted center and average velocity of all the leaders [16]. Olfati-Saber proposed a groundbreaking work on flocking for multiagent dynamic systems. He proposed three types of flocking algorithm. The first one is the basic flocking algorithm with only separation, cohesion, and alignment. Then one single leader agent was added to obtain the second algorithm. His third algorithm was obligated to deal with flocking problem in the presence of multiple obstacles. Also, a systematic method is provided for analyzing the split/rejoin maneuver of flocking [6].

All these above techniques can be seen as the lamp to light up our researches in generating realistic crowd behaviors of this paper.

3. Optimal Acceleration-Velocity-Bounded Trajectory Planning

3.1. Preliminaries. Suppose the virtual humans are simplified into cylinder with a unified radius. It is assumed that they move in the workspace $\mathcal{W} = R^2$. The configuration of a virtual human is uniquely defined by the triple $(x, y, \theta) \in C$, where $C = R^2 \times [0, 2\pi]$ denotes the configuration space. Suppose one urgent individual \mathcal{A} is given a fixed parameterized geometric path S which is a continuous sequence of configurations, that is, a straight line or a curve in the configuration space. In this paper, we assume that S must be piecewise of class C^1 (a curve is of class C^n if it is differentiable n times and if its n th derivative is continuous). Because the urgent individual moves along S , we can reduce its configuration to a single variable s which satisfies $q(s) : [0, 1] \rightarrow C$. $s(t)$ is a time scaling function $s : [0, t_f] \rightarrow [0, 1]$ which assigns value s to each time $t \in [0, t_f]$. $s \cdot L_S$ represents the distance traveled along S in which L_S represents the total length of S . Moreover, the time scaling $s(t)$ should be twice-differentiable and monotonic ($\dot{s}(t) \geq 0$ for all $t \in [0, t_f]$).

The twice-differentiability of $s(t)$ ensures that the urgent individual's acceleration $\ddot{q}(t)$ is well defined and bounded [17].

The dynamic planning problem occurs in the path-velocity-time state space, that is, $s \times \dot{s} \times t$ space which is denoted as PVT [4]. So, the state of one urgent individual $x = (s, \dot{s}, t)$ consists of variable $s \in [0, 1]$, velocity \dot{s} , and time $t \in [0, t_{\max}]$. For a given $q(s)$, the following holds:

$$\begin{aligned} \dot{q} &= \frac{dq}{ds} \dot{s}, \\ \ddot{q} &= \frac{d^2q}{ds^2} \dot{s}^2 + \frac{dq}{ds} \ddot{s}. \end{aligned} \quad (1)$$

The dynamic constraints of the urgent individuals must satisfy velocity and acceleration bounds:

$$\begin{aligned} 0 &\leq \sqrt{\dot{q}_x^2 + \dot{q}_y^2} \leq v_{\max}, \quad \dot{q}_x \geq 0, \quad \dot{q}_y \geq 0, \\ 0 &\leq \sqrt{\ddot{q}_x^2 + \ddot{q}_y^2} \leq a_{\max}. \end{aligned} \quad (2)$$

Substituting (2) into (1), we get the following dynamic constraints on variable s :

$$\dot{s} \in [\dot{s}_{\min}(s, a_{\max}, v_{\max}), \dot{s}_{\max}(s, a_{\max}, v_{\max})], \quad (3)$$

$$\ddot{s} \in [\ddot{s}_{\min}(s, \dot{s}, a_{\max}, v_{\max}), \ddot{s}_{\max}(s, \dot{s}, a_{\max}, v_{\max})] \quad (4)$$

with $\dot{s}_{\min}(s, a_{\max}, v_{\max}) \geq 0$ and $\ddot{s}_{\min}(s, \dot{s}, a_{\max}, v_{\max}) < 0 < \ddot{s}_{\max}(s, \dot{s}, a_{\max}, v_{\max})$. A trajectory $x(t) = (s(t), \dot{s}(t), t)$ defined over $t \in [t_0, t_f]$ in state space is dynamically feasible if and only if (3) and (4) are satisfied at the same time. Then, we say that a goal state x_f is dynamically reachable from initial state x_0 if a dynamically feasible trajectory $x(t)$ over interval $[t_0, t_f]$ such that $x(t_0) = x_0$ and $x(t_f) = x_f$ is existed. Note that there is a one-to-one correspondence between the trajectory $x(t)$ and its time scaling function $s(t)$.

Let O_i $i \in \{1, \dots, m\}$ be the set of moving obstacles. Let $O_i(t)$ denote the region of \mathcal{W} occupied by O_i at time t and $\mathcal{A}(q(s(t)))$ the region of \mathcal{W} occupied by \mathcal{A} at position $q(s(t))$ along S . Then the obstacle region in state space is defined as

$$X_{\text{obs}} = \{(s, \dot{s}, t) \mid \exists i \in \{1, \dots, m\}, \mathcal{A}(q(s(t))) \cap O_i(t) \neq \emptyset\}. \quad (5)$$

Obviously, the X_{obs} corresponds to $(s(t), t)$ points in $s \times t$ space. The trajectory $x(t)$ is collision-free if $(s(t), t) \notin X_{\text{obs}}$ for $t \in [t_0, t_f]$.

Above all, we can formally state the problem which is to be solved. Let $(s_0, \dot{s}_0, 0)$ be the start state of \mathcal{A} and (s_f, \dot{s}_f, t_f) its goal state.

A trajectory $\Gamma : [0, 1] \rightarrow \text{PVT}$ is a solution to the problem if and only if

- (1) $\Gamma(0) = (s_0, \dot{s}_0, 0)$, $\Gamma(1) = (s_f, \dot{s}_f, t_f)$, and $\dot{s}_0, \dot{s}_f \geq 0$,
- (2) Γ is dynamically feasible and collision-free for $t \in [0, t_f]$ and $t_f \leq t_{\max}$. t_{\max} is the maximal time given for the planner,

- (3) \dot{s}_f is a specific value (e.g., $\dot{s}_f = 0$) rather than a velocity interval.

We are interested in finding an exact time-optimal trajectory, that is, a trajectory such that t_f should be minimal. Moreover, \dot{s}_f is an interval in Johnson's method rather than a specific value. So the optimal trajectory planning in this paper is more complex and intractable.

3.2. Optimal Acceleration-Velocity-Bounded Trajectory Planning. The method that we have developed in order to solve the above problem is initially motivated by the work described in [4]. We improve Johnson's method in order to find out the exact time-optimal trajectory between an initial and a specific goal state. The flowchart of our method is illustrated in Figure 2.

Next we will describe every part of Figure 2 in detail.

3.2.1. Compute Reachable Velocity Sets

(I) Compute Reachable Velocity Sets from One Specified Initial State. The goal of this section is to facilitate the computing reachable velocity intervals from one specified initial state. Without loss of generality, we suppose the specified initial state is $x_I = (s_I, \dot{s}_I, t_I)$ which can be any state in $\{(s_0(t_0), \dot{s}_0(t_0), t_0), \dots, (s_{2m+1}(t_{2m+1}), \dot{s}_{2m+1}(t_{2m+1}), t_{2m+1})\}$. The dynamically reachable set of velocities $R(t; x_I)$ from x_I at any point in time $t \geq t_I$ is the set of velocities attainable at t from x_I via dynamically feasible trajectories. It can be computed in $s \times \dot{s}$ space. In order to simplify the computation, we assume that S is a straight line. Then (3) and (4) can be simplified into

$$\dot{s} \in \left[0, \frac{v_{\max}}{L_S}\right],$$

$$\text{that is, } \dot{s}_{\min}(s, a_{\max}, v_{\max}) = 0; \quad (6)$$

$$\dot{s}_{\max}(s, a_{\max}, v_{\max}) = \frac{v_{\max}}{L_S},$$

$$\ddot{s} \in \left[-\frac{a_{\max}}{L_S}, \frac{a_{\max}}{L_S}\right],$$

$$\text{that is, } \ddot{s}_{\min}(s, \dot{s}, a_{\max}, v_{\max}) = -\frac{a_{\max}}{L_S}; \quad (7)$$

$$\ddot{s}_{\max}(s, \dot{s}, a_{\max}, v_{\max}) = \frac{a_{\max}}{L_S}.$$

Johnson showed that $R(t; x_I)$ is bounded by at most six curves [4] which are denoted in Figure 3 in our paper.

In Figure 3, parabolic segment (a) corresponds to the process in which \mathcal{A} firstly takes the minimal acceleration for time t_s and then switches to maximal acceleration for time $t - t_s$. It must be noted that the terminal velocity of the process of deceleration and the one of the process of acceleration must

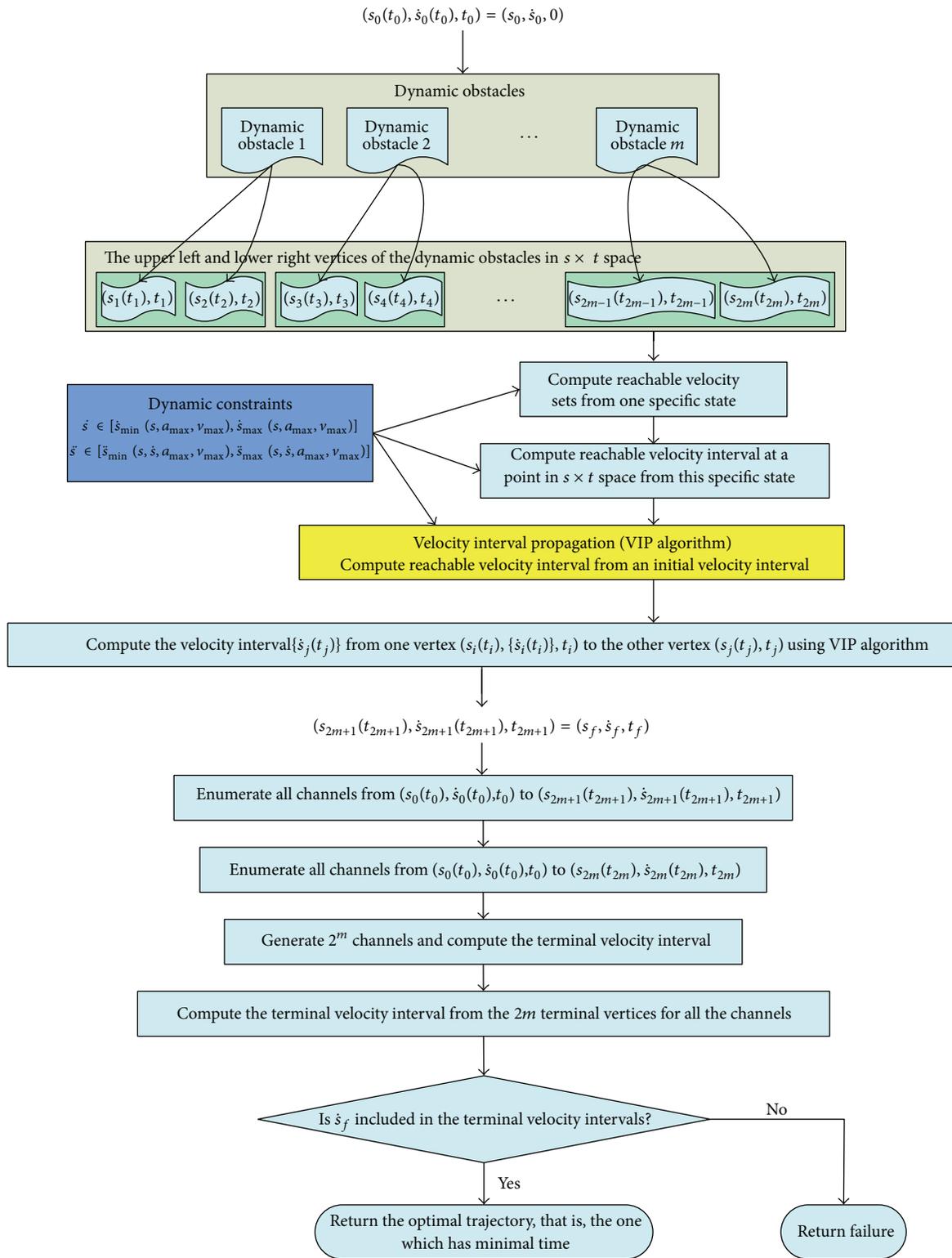


FIGURE 2: Flowchart of optimal trajectory planning.

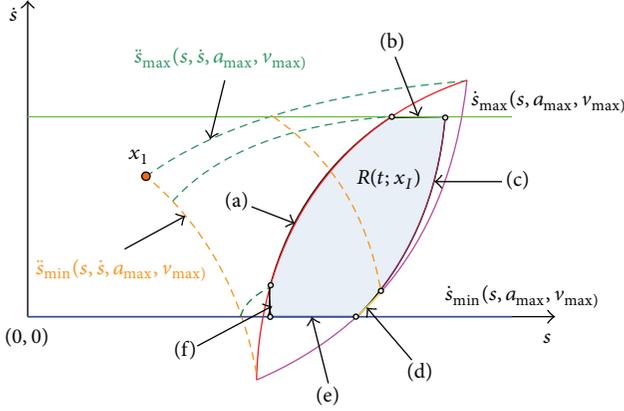


FIGURE 3: Reachable velocity sets from one initial state.

satisfy formulation (6). The states on (a) can be formulated as in the following equation:

$$\begin{aligned} x_{t_I+t} = & \left(s_I + \dot{s}_I t_s + \frac{1}{2} \ddot{s}_{\min} t_s^2 + (\dot{s}_I + \ddot{s}_{\min} t_s)(t - t_s) \right. \\ & + \frac{1}{2} \ddot{s}_{\max} (t - t_s)^2, \dot{s}_I + \ddot{s}_{\min} t_s + \ddot{s}_{\max} (t - t_s), \\ & \left. t_I + t \right), \\ \dot{s}_I + \ddot{s}_{\min} t_s \geq & \dot{s}_{\min}, \quad \dot{s}_I + \ddot{s}_{\min} t_s + \ddot{s}_{\max} (t - t_s) \leq \dot{s}_{\max}. \end{aligned} \quad (8)$$

Similarly, parabolic segment (d) corresponds to the process, takes the maximal acceleration for time t_s , and then switches to minimal acceleration for time $t - t_s$. The states on (d) can be formulated as

$$\begin{aligned} x_{t_I+t} = & \left(s_I + \dot{s}_I t_s + \frac{1}{2} \ddot{s}_{\max} t_s^2 + (\dot{s}_I + \ddot{s}_{\max} t_s)(t - t_s) \right. \\ & + \frac{1}{2} \ddot{s}_{\min} (t - t_s)^2, \dot{s}_I + \ddot{s}_{\max} t_s + \ddot{s}_{\min} (t - t_s), \\ & \left. t_I + t \right), \\ \dot{s}_I + \ddot{s}_{\max} t_s \leq & \dot{s}_{\max}, \quad \dot{s}_I + \ddot{s}_{\max} t_s + \ddot{s}_{\min} (t - t_s) \geq \dot{s}_{\min}. \end{aligned} \quad (9)$$

Parabolic segment (c) corresponds to the process, takes the maximal acceleration until reaching $\dot{s}_{\max}(s, a_{\max}, v_{\max})$ at the time t_v , progresses with $\dot{s}_{\max}(s, a_{\max}, v_{\max})$ until time t_s elapse, and then decelerates for time $t - t_s$. The states on (c) can be formulated as

$$\begin{aligned} x_{t_I+t} = & \left(s_I + \dot{s}_I t_v + \frac{1}{2} \ddot{s}_{\max} t_v^2 + \dot{s}_{\max} (t_s - t_v) + \dot{s}_{\max} (t - t_s) \right. \\ & + \frac{1}{2} \ddot{s}_{\min} (t - t_s)^2, \dot{s}_{\max} + \ddot{s}_{\min} (t - t_s), t_I + t \left. \right), \\ t_v = & \left(\frac{\dot{s}_{\max} - \dot{s}_I}{\ddot{s}_{\max}} \right); \quad \dot{s}_I + \ddot{s}_{\max} t_s \geq \dot{s}_{\max}. \end{aligned} \quad (10)$$

Similarly, parabolic segment (f) corresponds to the process, takes the minimal acceleration until reaching $\dot{s}_{\min}(s, a_{\max}, v_{\max})$, progresses with $\dot{s}_{\min}(s, a_{\max}, v_{\max})$ until time t_s elapse, and then accelerates for time $t - t_s$. The states on (f) can be formulated as

$$\begin{aligned} x_{t_I+t} = & \left(s_I + \dot{s}_I t_v + \frac{1}{2} \ddot{s}_{\min} t_v^2 + \dot{s}_{\min} (t_s - t_v) + \dot{s}_{\min} (t - t_s) \right. \\ & + \frac{1}{2} \ddot{s}_{\max} (t - t_s)^2, \dot{s}_{\min} + \ddot{s}_{\max} (t - t_s), t_I + t \left. \right), \\ t_v = & \left(\frac{\dot{s}_{\min} - \dot{s}_I}{\ddot{s}_{\min}} \right); \quad \dot{s}_I + \ddot{s}_{\min} t_s \leq \dot{s}_{\min}. \end{aligned} \quad (11)$$

Line segment (b) corresponds to the process, takes the minimal acceleration for time t_s , and then switches to maximal acceleration until reaching $\dot{s}_{\max}(s, a_{\max}, v_{\max})$, finally progressing with $\dot{s}_{\max}(s, a_{\max}, v_{\max})$ until time t elapse. The states on (b) can be formulated as

$$\begin{aligned} x_{t_I+t} = & \left(s_I + \dot{s}_I t_s + \frac{1}{2} \ddot{s}_{\min} t_s^2 + (\dot{s}_I + \ddot{s}_{\min} t_s) t_v \right. \\ & + \frac{1}{2} \ddot{s}_{\max} t_v^2 + \dot{s}_{\max} (t - t_s - t_v), \dot{s}_{\max}, t_I + t \left. \right), \\ t_v = & \left(\frac{\dot{s}_{\max} - \dot{s}_I - \ddot{s}_{\min} t_s}{\ddot{s}_{\max}} \right); \quad \dot{s}_I + \ddot{s}_{\min} t_s \geq \dot{s}_{\min}; \\ & \dot{s}_I + \ddot{s}_{\min} t_s + \ddot{s}_{\max} t_v \geq \dot{s}_{\max}. \end{aligned} \quad (12)$$

Similarly, line segment (e) corresponds to the process, takes the maximal acceleration for time t_s , and then switches to minimal acceleration until reaching $\dot{s}_{\min}(s, a_{\max}, v_{\max})$, finally progressing with $\dot{s}_{\min}(s, a_{\max}, v_{\max})$ until time t elapse. The states on (e) can be formulated as

$$\begin{aligned} x_{t_I+t} = & \left(s_I + \dot{s}_I t_s + \frac{1}{2} \ddot{s}_{\max} t_s^2 + (\dot{s}_I + \ddot{s}_{\max} t_s) t_v \right. \\ & + \frac{1}{2} \ddot{s}_{\min} t_v^2 + \dot{s}_{\min} (t - t_s - t_v), \dot{s}_{\min}, t_I + t \left. \right), \\ t_v = & \left(\frac{\dot{s}_{\min} - \dot{s}_I - \ddot{s}_{\max} t_s}{\ddot{s}_{\min}} \right); \quad \dot{s}_I + \ddot{s}_{\max} t_s \leq \dot{s}_{\max}; \\ & \dot{s}_I + \ddot{s}_{\max} t_s + \ddot{s}_{\min} t_v \leq \dot{s}_{\min}. \end{aligned} \quad (13)$$

All these six curves enclose the reachable velocity set $R(t; x_I)$ from one single initial state $x_I = (s_I, \dot{s}_I, t_I)$ which is graphically showed as the shadow region in Figure 3.

(II) Compute Reachable Velocity Interval at a Point in $s \times t$ Space from $x_I = (s_I, \dot{s}_I, t_I)$. After obtaining the reachable velocity set in $s \times \dot{s}$ space, we can then compute the reachable velocity interval at a point $(s_G(t_G), t_G)$ in $s \times t$ space from the initial state $x_I = (s_I(t_I), \dot{s}_I, t_I)$. We use $\{\dot{s}_G(t_G)\}$ to represent the reachable velocity interval at $(s_G(t_G), t_G)$. We use $R(t_G - t_I; x_I)$ to compute $\{\dot{s}_G(t_G)\}$ in the following formulation:

$$\{\dot{s}_G(t_G)\} = [\underline{\dot{s}}_G, \bar{\dot{s}}_G] = R(t_G - t_I; x_I) \cap \{(s, \dot{s}) \mid s = s_G\} \quad (14)$$

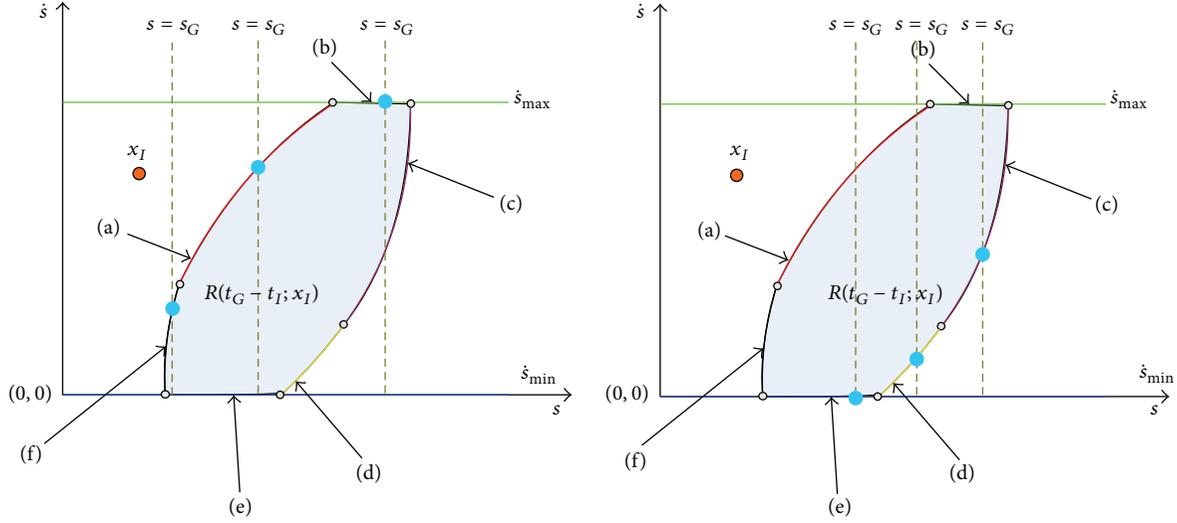


FIGURE 4: Reachable velocity sets from one initial state.

in which \bar{s}_G and \underline{s}_G represent the upper and lower bounds of velocity interval $\{\dot{s}_G(t_G)\}$, respectively. \bar{s}_G is defined as the intersection of a boundary curve of $R(t_G - t_I; x_I)$ with $s = s_G$.

We only need to consider three boundary curves in computing \bar{s}_G (curves (a), (b), and (f)) and \underline{s}_G (curves (c), (d), and (e)), as illustrated in Figure 4.

In this process, we determine the switching time t_s by substituting $s = s_G$ into (8), (11), and (12) for \bar{s}_G and (9), (10), and (13) for \underline{s}_G , respectively, and solving the resulting quadratic equation. At the same time, the constraints in (8)~(13) are validated to determine if some trajectories exist from $x_I = (s_I, \dot{s}_I, t_I)$ to $(s_G(t_G), \{\dot{s}_G(t_G)\}, t_G)$.

(III) *VIP Algorithm Computing Reachable Velocity Interval from an Initial Velocity Interval.* After computing the reachable velocity interval from an initial specified velocity, we can extend it to compute the reachable velocity interval from an initial velocity interval which forms the backbone of optimal trajectory planner. It takes $(s_G(t_G), \{\dot{s}_G(t_G)\}, t_G)$ and a target point $(s_H(t_H), t_H)$ in $s \times t$ space as input. It outputs a range of target velocities $\{\dot{s}_H(t_H)\} = [\underline{\dot{s}}_H, \bar{\dot{s}}_H]$ reachable from $(s_G(t_G), \{\dot{s}_G(t_G)\}, t_G)$. $\{\dot{s}_H(t_H)\} = [\underline{\dot{s}}_H, \bar{\dot{s}}_H]$ is computed in the following steps.

Step 1. Compute the reachable velocity interval $V_1 = R(t_H - t_G; x_G) \cap \{(s, \dot{s}) \mid s = s_H\}$ for initial velocity $\dot{s}_G = \bar{\dot{s}}_G$ from $(s_G(t_G), \bar{\dot{s}}_G, t_G)$.

Step 2. Compute the reachable velocity interval $V_2 = R(t_H - t_G; x_G) \cap \{(s, \dot{s}) \mid s = s_H\}$ for initial velocity $\dot{s}_G = \underline{\dot{s}}_G$ from $(s_G(t_G), \underline{\dot{s}}_G, t_G)$.

Step 3. The goal of this step is maximizing the terminal velocity at $(s_H(t_H), t_H)$. It does so by constructing a parabolic trajectory with acceleration $\ddot{s}_{\max}(s, \dot{s}, a_{\max}, v_{\max})$ that interpolates $(s_G(t_G), t_G)$ and $(s_H(t_H), t_H)$. If the initial velocity

$\dot{s}_{G,\max}^{\text{interpolate}}$ of this interpolating parabolic trajectory belongs to $\{\dot{s}_G(t_G)\}$, then $V_3 = \{\dot{s}_{G,\max}^{\text{interpolate}}\}$.

Step 4. The goal of this step is minimizing the terminal velocity at $(s_H(t_H), t_H)$. It also does so by constructing a parabolic trajectory with acceleration $\ddot{s}_{\min}(s, \dot{s}, a_{\max}, v_{\max})$ that interpolates $(s_G(t_G), t_G)$ and $(s_H(t_H), t_H)$. If the initial velocity $\dot{s}_{G,\min}^{\text{interpolate}}$ of this interpolating parabolic trajectory belongs to $\{\dot{s}_G(t_G)\}$, then $V_4 = \{\dot{s}_{G,\min}^{\text{interpolate}}\}$.

Step 5. Finally, the output velocity interval $\{\dot{s}_H(t_H)\} = [\underline{\dot{s}}_H, \bar{\dot{s}}_H] = V_1 \cup V_2 \cup V_3 \cup V_4$. Note that either V_1, V_2, V_3 , or V_4 may be empty.

Because the goal of the method in this section is propagating the velocity intervals, from one velocity interval to the other as long as we know the initial and terminal point in $s \times t$ space, we call this method velocity interval propagation (VIP algorithm). Steps 1~4 can be represented as $\{\dot{s}_H(t_H)\} = \text{VIP}((s_G(t_G), \{\dot{s}_G(t_G)\}, t_G), (s_H(t_H), t_H))$ in short.

3.2.2. *Computing the Optimal Trajectory.* Suppose there are m dynamic obstacles in the environment; then only $2m$ vertices are important to our optimal planning which correspond to the upper left and lower right vertex of the obstacles [4]. Let $(s_1(t_1), t_1), \dots, (s_{2m}(t_{2m}), t_{2m})$ be the sequence of upper left and lower right obstacle vertices such that $t_0 \leq t_i \leq t_f$ for all $i = 1, \dots, 2m$. Then we sort these $2m$ vertices by increasing t coordinate and denote $(s_0(t_0), \dot{s}_0(t_0), t_0) = (s_0, \dot{s}_0, 0)$ and $(s_{2m+1}(t_{2m+1}), \dot{s}_{2m+1}(t_{2m+1}), t_{2m+1}) = (s_f, \dot{s}_f, t_f)$. Our planner consists of four stages.

In the first stage, using the VIP algorithm, we can get the velocity interval from any vertex of one dynamic obstacle to any vertex of another one for $(s_0(t_0), t_0), \dots, (s_{2m}(t_{2m}), t_{2m})$; that is, we carry out $\{\dot{s}_j(t_j)\} = \text{VIP}((s_i(t_i), \dot{s}_i(t_i), t_i), (s_j(t_j), t_j))$ for all $i, j \in [1, 2m]; i \neq j$.

In the second stage, we get 2^m channels and 2^m terminal velocity intervals of these channels. These terminal velocity intervals can be merged into $2m$ velocity intervals $\{\{\dot{s}_{1,\text{union}}(t_1)\}, \dots, \{\dot{s}_{2m,\text{union}}(t_{2m})\}\}$ corresponding to the $2m$ vertices with union operation.

In the third stage, we use VIP algorithm to connect $(s_{2m}(t_{2m}), t_{2m})$ with $(s_{2m+1}(t_{2m+1}), t_{2m+1})$ by taking these $2m$ velocity intervals as input to determine the terminal velocity interval at given terminal time. We use discrete $t_{i,f}^k$ to represent the given terminal time:

$$t_{i,f}^k = t_i + k \cdot \Delta t \quad k = 1, \dots, \left\lfloor \frac{t_{\max} - t_i}{\Delta t} \right\rfloor. \quad (15)$$

Then we can get the terminal velocity interval at every given terminal time $t_{i,f}^k$ by calling $\{\dot{s}_k(t_{i,f}^k)\} = \text{VIP}((s_i(t_i), \{\dot{s}_{i,\text{union}}(t_i)\}, t_i), (s_k(t_{i,f}^k), t_{i,f}^k)) \quad i = 1, \dots, 2m; k = 1, \dots, \lfloor (t_{\max} - t_i)/\Delta t \rfloor$.

In the fourth stage, we judge if \dot{s}_f is in these $\sum_{i=1}^{2m} \lfloor (t_{\max} - t_i)/\Delta t \rfloor$ terminal velocity intervals or not. If true, we return the trajectory with minimal terminal time $t_{i,f}^k$ which is the optimal trajectory and its final velocity equals \dot{s}_f . If not, the planner returns failure. It means that no trajectory exists between the initial state and the goal state of \mathcal{A} which satisfies the dynamic constraints and is collision-free with dynamic obstacles.

Complexity Analysis. The computational cost of the planner mainly comes from the first and the third stage. In the first stage, we must call the VIP algorithm for 2^m times. In the third stage, VIP algorithm is called $\sum_{i=1}^{2m} \lfloor (t_{\max} - t_i)/\Delta t \rfloor$ times. Then the complexity of our planner is $O(2^m + \sum_{i=1}^{2m} \lfloor (t_{\max} - t_i)/\Delta t \rfloor) < O(2^m + 2m \lfloor (t_{\max} - t_i)/\Delta t \rfloor)$ which is equivalent to $O(2^m)$. This computational cost is vast for big m values. However, there are only several dynamic obstacles in our environment and this is why we apply pushing behavior to deal with the interaction between the urgent members and the normal members.

3.3. Pushing Model. Pushing model is used to model the interactions between the urgent members and the normal ones; that is, the normal members will be pushed away when they are in a predefined range of the urgent ones. We use the four following rules to model pushing behavior between any normal individual i and urgent one j :

$$\begin{aligned} \text{If } d_{ij} &= 0 & \text{Then } V_{ij}^{\rightarrow} &= -V_{\max} \\ \text{If } 0 < d_{ij} &< \mu & \text{Then } V_{ij}^{\rightarrow} &= -M \cdot V_{\max} \cdot e^{-d_{ij}} + V_0 \\ \text{If } |V_{ij}^{\rightarrow}| &> \sqrt{V_{\max}^2 - \left(V_{ij}^{\perp}\right)^2} & \text{Then } V_{ij}^{\rightarrow} &= -\sqrt{V_{\max}^2 - \left(V_{ij}^{\perp}\right)^2} \\ \text{If } d_{ij} &\geq \mu & \text{Then } V_{ij}^{\rightarrow} &= V_0 \end{aligned} \quad (16)$$

in which d_{ij} denotes the distance between i and j . μ represents the threshold of the interaction distance. V_{ij}^{\rightarrow} represents

the velocity of i at the direction of vector \vec{ij} . V_{ij}^{\perp} represents the velocity of i at the perpendicular direction of vector \vec{ij} . V_{\max} indicates the maximal velocity at which a normal individual moves. And V_0 represents the velocity of i before the interaction between i and j happens.

This novel pushing model ensures that the urgent members will never collide with the normal ones. So we do not need to consider the normal members in the process of optimal trajectory planning whose cost time will greatly reduce owing to only the dynamic obstacles being considered.

4. Collision Avoidance Behaviors and Flocking Behaviors

4.1. Collision Avoidance Behaviors. In this section, we present the collision avoidance behaviors on the basis of Foudil and Nouredine's work. As a general approach, this method can be combined with different crowd and multiagent simulation algorithms. In each time step of the simulation, we need to predict if every agent will collide with other agents in the crowd. Then we must determine the type of collision which may happen. In real life, there are three possible types of collision [5]. The first type is *toward collision behavior*. It happens when two agents are moving head-on toward each other. The second one is *away collision*. It happens when one agent who is behind the other one is moving with a bigger velocity and whose moving direction is consistent with the line from the back agent to the front one. The third collision behavior is *glancing collision* which happens when two agents are walking in roughly the same direction.

These three types of collision behaviors are shown in Figure 5.

We use a series of rules to realize these three collision avoidance behaviors separately. In *toward collision behavior*, we use the following strategy to avoid the forthcoming collision: the agent who has a low priority can select waiting or changing its moving direction and the agent who has an upper priority selects keeping its moving on with unchanged direction and velocity. In respect to *away collision*, the back agent can select slowing its velocity or changing its moving direction to any side. We deal with *glancing collision* with the same manner as the *toward collision behavior*.

Particularly, when these three collision avoidance behaviors conflict with each other, we resolve this problem by predefining the priority order of them. In this paper, the priority order of these three types of collision avoidance behaviors is toward collision, away collision, and glancing collision from high to low.

4.2. Flocking Behaviors. Flocking behaviors are used to model the collective gathering behaviors of the normal groups. Aiming at the agents group consisting of one leader agent and N follower ones, we use a distributed control method to simulate the flocking behavior [15, 16]. The motion of each virtual agent is described by two integrators as

$$\begin{aligned} \dot{\mathbf{p}}_i(t) &= \mathbf{v}_i(t) \\ \dot{\mathbf{v}}_i(t) &= \mathbf{a}_i(t) \end{aligned} \quad i = 1, 2, \dots, N, \quad (17)$$

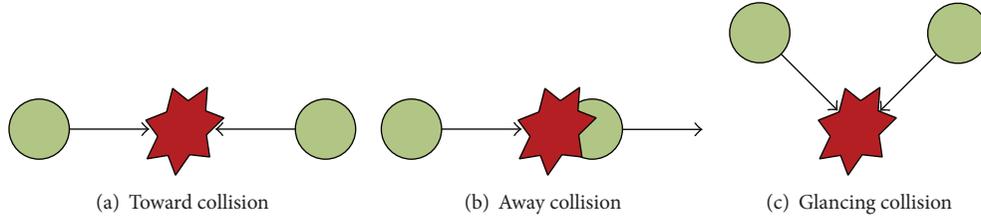


FIGURE 5: Three types of the collision behaviors.

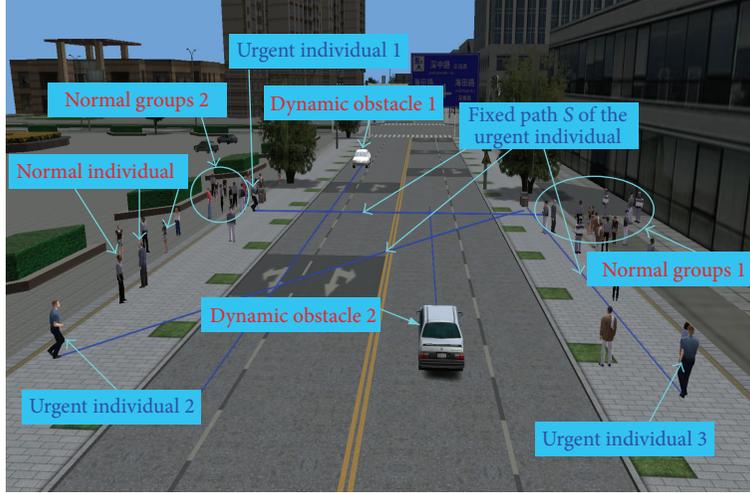


FIGURE 6: The dynamic environment of the experiments.

where \mathbf{p}_i , \mathbf{v}_i , and \mathbf{a}_i are denoted as the position, velocity, and acceleration of agent i , respectively. Similarly, the virtual leader has the following dynamics of motion:

$$\begin{aligned}\dot{\mathbf{p}}_\gamma(t) &= \mathbf{v}_\gamma(t), \\ \dot{\mathbf{v}}_\gamma(t) &= \mathbf{a}_\gamma(t),\end{aligned}\quad (18)$$

where \mathbf{p}_γ , \mathbf{v}_γ , and \mathbf{a}_γ represent the position, velocity, and acceleration of leader agent, respectively. The control algorithm of the flocking method is given as

$$\begin{aligned}\mathbf{a}_i &= - \sum_{j \in \mathcal{N}_i(t)} \nabla_{\mathbf{p}_i} \psi_\alpha(\|\mathbf{p}_i - \mathbf{p}_j\|_\sigma) + \sum_{j \in \mathcal{N}_i(t)} a_{ij}(t) (\mathbf{v}_j - \mathbf{v}_i) \\ &+ \mathbf{a}_\gamma^i - c_1 (\mathbf{p}_i - \mathbf{p}_\gamma^i) - c_2 (\mathbf{v}_i - \mathbf{v}_\gamma^i)\end{aligned}\quad (19)$$

$$i = 1, \dots, N \quad c_1, c_2 > 0,$$

where $\mathcal{N}_i(t) = \{j : \|\mathbf{p}_i - \mathbf{p}_j\| \leq R, j = 1, 2, \dots, N, j \neq i\}$ is the set of spatial neighbors of agent i . R represents the interaction range. $\|\cdot\|_\sigma$ is the Euclidean norm and $\|\mathbf{z}\|_\sigma = (1/\varepsilon)[\sqrt{1 + \varepsilon\|\mathbf{z}\|^2} - 1]$ $\varepsilon > 0$ for a vector \mathbf{z} . ψ_α is the artificial nonnegative smooth pairwise potential function whose characteristics depend on the relative distances between agent i and its neighbors: ψ_α reaches its maximal value as $\|\mathbf{p}_i - \mathbf{p}_j\|_\sigma \rightarrow 0$, and ψ_α acquires its unique minimal value at a predefined distance $\|d\|_\sigma$. When $\|\mathbf{p}_i - \mathbf{p}_j\|_\sigma < \|d\|_\sigma$, agent

i attains repulsion force from j by $\nabla_{\mathbf{p}_i} \psi_\alpha(\|\mathbf{p}_i - \mathbf{p}_j\|_\sigma)$. When $\|\mathbf{p}_i - \mathbf{p}_j\|_\sigma = \|d\|_\sigma$, the repulsion force and attraction force between i and j become balance. When $\|d\|_\sigma < \|\mathbf{p}_i - \mathbf{p}_j\|_\sigma < \|R\|_\sigma$, agent i attains attraction force from j . And when $\|\mathbf{p}_i - \mathbf{p}_j\|_\sigma \geq \|R\|_\sigma$, ψ_α becomes constant; then no force exists between i and j . c_1 and c_2 are positive constants which reflect the influence degree from leader's position and velocity. $a_{ij}(t)$ is the adjacent weight coefficient.

5. Experimental Results and Discussion

In our simulation, we employ a human animation software package called DI-Guy, which is commercially available from Boston Dynamics Inc. We control the moving of the urgent members and normal members using SDK interface by C++ programs. The scenario can be depicted as follows: in a city populated with a majority of civilians and 2 moving cars. The civilians consist of three urgent individuals and 50 normal members comprising 30 normal individuals and 2 normal groups with 10 members, respectively. The cars' moving characteristics are known previously. The priority values of the 50 normal agents are selected randomly from the interval $[0, 1]$. The interaction range $R = 10$ m and the predefined distance $d = 8$ m. The values of a_{ij} and ψ_α are set the same as those in [6], and $c_1 = 0.81$, $c_2 = 2\sqrt{c_1}$, and $N = 10$. The dynamic environment of the experiments is depicted in Figure 6.

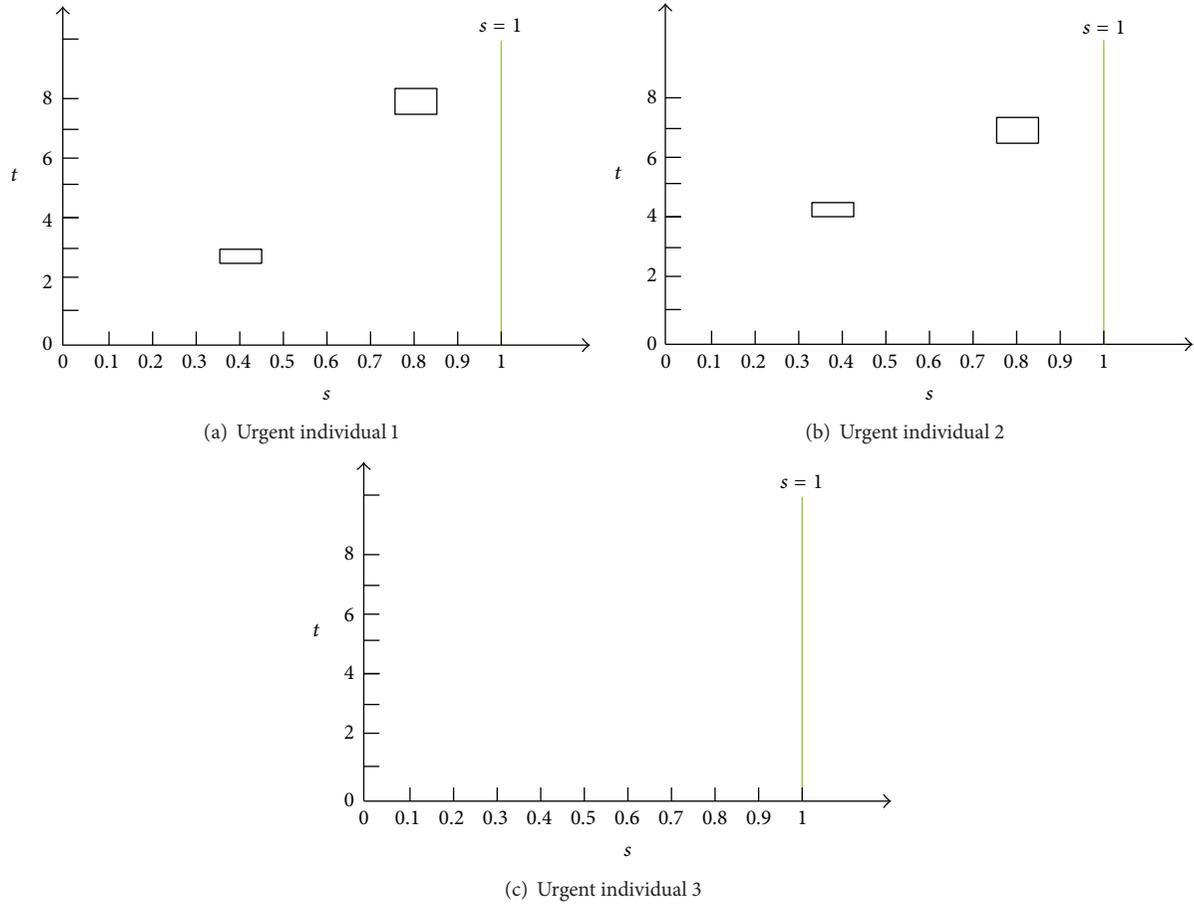


FIGURE 7: The $s \times t$ spaces of urgent individuals.

In Figure 6, we set $a_{\max} = 1.2 \text{ m/s}^2$ and $v_{\max} = 8 \text{ m/s}$. Length of the specified paths of urgent individuals is $L_{S1} = 20 \text{ m}$ and $L_{S2} = 30 \text{ m}$ and $L_{S3} = 25 \text{ m}$. The first dynamic obstacle (moving cars) moves at a constant speed of 10 m/s (Car1) and the second moves at a constant speed of 5 m/s (Car2). The $s \times t$ spaces of the three urgent individuals in the dynamic environment are depicted in Figure 7.

Figure 8 illustrates the reachable sets and the optimal trajectory for each urgent individual.

The visualization simulation results of the crowd behaviors in 3D space are illustrated in Figure 9.

From Figure 8, we can see that the gradient in $s \times t$ space is $+\infty$ which means that the terminal velocity of the urgent individual is zero, and this satisfies the requirement of the problem definition in Section 3.1; for example, in Figure 8(b), urgent individual 2 takes the maximal acceleration until reaching A point before Car1 arrives at his path and then switches to deceleration for the path A-B in case of colliding with Car2. After that, he continues to take the maximal acceleration for path B-C and then decelerates to zero from C point to the terminal point of the whole path. Figure 9 depicts realistic and believable crowd behaviors in city environment: the urgent individual can reach the destination along a given path in minimal time and be collision-free with dynamic cars. Simultaneously, the agents of the two normal groups show

real gathering behaviors according to the flocking method and the normal individuals try to keep away from the urgent ones and avoid collision with each other.

The computation cost of our approach is illustrated in Figure 10 when we change the numbers of the dynamic obstacles. Similarly, Figure 11 depicts the cost time of our approach according to the number of the normal members in the crowd.

From Figure 10, we can see that the cost time of our approach is influenced vastly by the number of dynamic obstacles in the environment. However the experiment results in Figure 11 show that the cost time of our approach changes slightly according to the number of the normal members. This means that the increase in the number of the normal members does not generate much computational cost. And all the above results are in accordance with previous analyses.

6. Conclusions

Although there have been some research studies on optimal trajectory planning for various purposes, few efforts have been conducted to simulate the realistic crowd behaviors with it, such as pedestrian navigation behavior in dynamic environment. In this paper, we first present an optimal

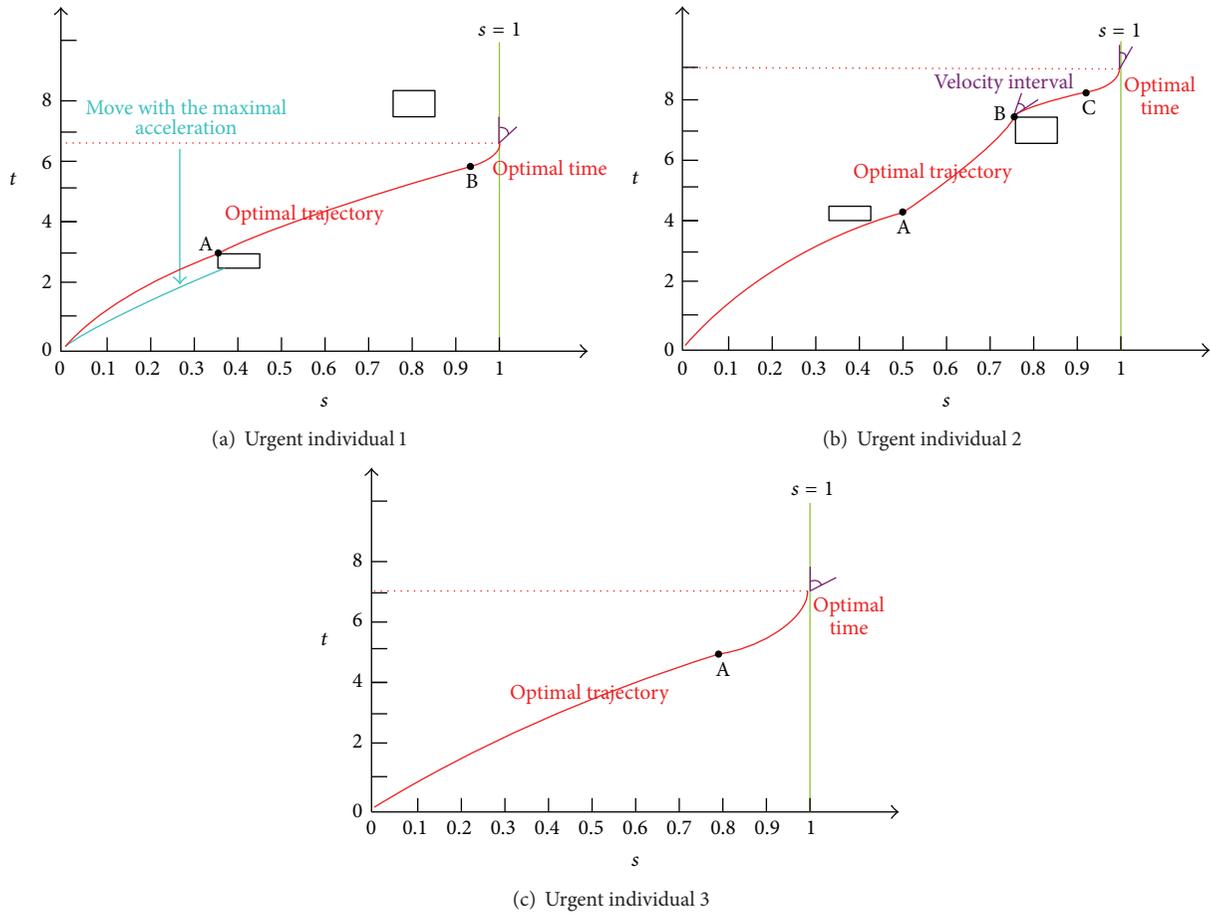


FIGURE 8: The optimal trajectory, time, and the velocity interval of urgent individuals.

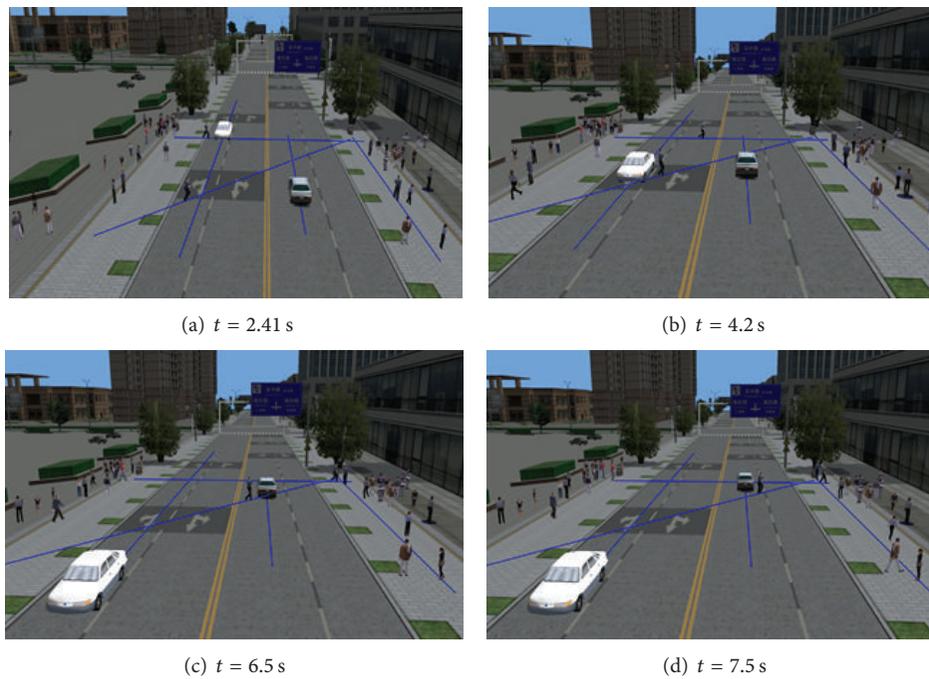


FIGURE 9: Visualization simulation results of the crowd behaviors.

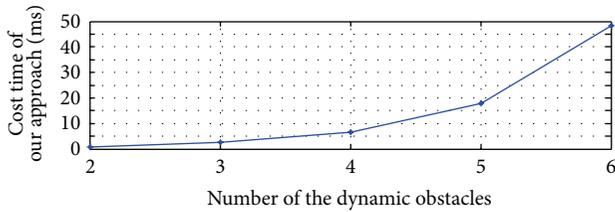


FIGURE 10: Cost time of our approach according to the number of the dynamic obstacles.

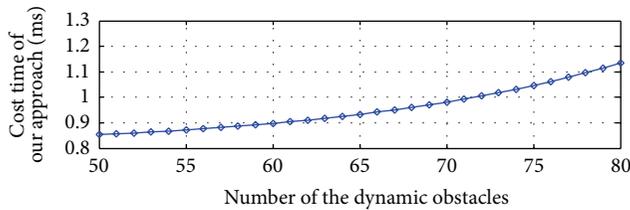


FIGURE 11: Cost time of our approach according to the number of the normal members in the crowd.

acceleration-velocity-bounded trajectory planning method along a fixed, given path with dynamic obstacles. We use it to generate the optimal trajectory of the urgent members under the dynamic constraints. This planner ensures that the moving time of the urgent members is minimal by using a velocity interval propagation algorithm to compute reachable velocity sets at obstacle vertices in $s \times t$ space. Moreover, the cost time of this optimal trajectory planning method is acceptable by applying a novel pushing model. Finally, combining with the collision avoidance behavior and the flocking behavior, the crowd simulation with dynamic vehicles is implemented. So, the potential of our approach for planning the optimal trajectories and modeling the social behaviors of the crowd is promising.

Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

Acknowledgment

This paper is supported by the National Natural Science Foundation of China (Grant nos. 61170160 and 61374185).

References

- [1] D. Thalmann and S. Raupp Musse, *Crowd Simulation*, Springer, London, UK, 2013.
- [2] N. Pelechano, J. Allbeck, and N. I. Badler, *Virtual Crowds: Methods, Simulation, and Control*, Morgan & Claypool Publishers, 2008.
- [3] S. Lemercier, A. Jelic, R. Kulpa et al., "Realistic following behaviors for crowd simulation," *Computer Graphics Forum*, vol. 31, no. 2, pp. 489–498, 2012.

- [4] J. Johnson and K. Hauser, "Optimal acceleration-bounded trajectory planning in dynamic environments along a specified path," in *Proceedings of IEEE International Conference on Robotics and Automation*, pp. 2035–2041, 2012.
- [5] C. Foudil and D. Noureddine, "Collision avoidance in crowd simulation with priority rules," *European Journal of Scientific Research*, vol. 15, no. 1, pp. 6–17, 2006.
- [6] R. Olfati-Saber, "Flocking for multi-agent dynamic systems: algorithms and theory," *IEEE Transactions on Automatic Control*, vol. 51, no. 3, pp. 401–420, 2006.
- [7] T. Fraichard, "Trajectory planning in a dynamic workspace: a "state-time space" approach," *Advanced Robotics*, vol. 13, no. 1, pp. 75–94, 1999.
- [8] P. Fiorini and Z. Shiller, "Time optimal trajectory planning in dynamic environments," in *Proceedings of IEEE International Conference on Robotics and Automation*, pp. 1553–1558, 1996.
- [9] C. Park, J. Pan, and D. Manocha, "Real-time optimization-based planning in dynamic environments," in *Proceedings of the International Symposium on Combinatorial Search (SOCS '12)*, 2012.
- [10] P. Fiorini and Z. Shiller, "Motion planning in dynamic environments using velocity obstacles," *International Journal of Robotics Research*, vol. 17, no. 7, pp. 760–772, 1998.
- [11] K. Maček, D. Vasquez, T. Fraichard, and R. Siegwart, "Safe vehicle navigation in dynamic urban scenarios," in *Proceedings of the 11th International IEEE Conference on Intelligent Transportation Systems (ITSC '08)*, pp. 482–489, December 2008.
- [12] N. A. B. M. Abdullah, B. B. Abdullah, and S. Kari, "A review of collision avoidance technique for crowd simulation," in *Proceedings of the International Conference on Information and Multimedia Technology*, pp. 388–392, 2009.
- [13] F. Morini, B. Yersin, J. Maim, and D. Thalmann, "Real-time scalable motion planning for crowds," in *Proceedings of the International Conference on Cyber Worlds (CW '07)*, pp. 144–151, Hannover, Germany, October 2007.
- [14] S. R. Musse and D. Thalmann, "A model of human crowd behavior: group inter-relationship and collision detection analysis," in *Proceedings of the Euro Graphics Workshop*, 1997.
- [15] H. Shi, L. Wang, and T. Chu, "Flocking of multi-agent systems with a dynamic virtual leader," *International Journal of Control*, vol. 82, no. 1, pp. 43–58, 2009.
- [16] H. Su, X. Wang, and W. Yang, "Flocking in multi-agent systems with multiple virtual leaders," *Asian Journal of Control*, vol. 10, no. 2, pp. 238–245, 2008.
- [17] H. Choset, *Principles of Robot Motion-Theory, Algorithm and Implementation*, MIT Press, 2005.

Research Article

Real-Time Control Strategy of Elman Neural Network for the Parallel Hybrid Electric Vehicle

Ruijun Liu, Dapai Shi, and Chao Ma

School of Transportation and Vehicle Engineering, Shandong University of Technology, No. 12 Zhangzhou Road, Zibo, Shandong 255049, China

Correspondence should be addressed to Dapai Shi; sdapai@163.com

Received 15 April 2014; Revised 30 June 2014; Accepted 1 July 2014; Published 26 August 2014

Academic Editor: H. R. Karimi

Copyright © 2014 Ruijun Liu et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Through researching the instantaneous control strategy and Elman neural network, the paper established equivalent fuel consumption functions under the charging and discharging conditions of power batteries, deduced the optimal control objective function of instantaneous equivalent consumption, established the instantaneous optimal control model, and designs the Elman neural network controller. Based on the ADVISOR 2002 platform, the instantaneous optimal control strategy and the Elman neural network control strategy were simulated on a parallel HEV. The simulation results were analyzed in the end. The contribution of the paper is that the trained Elman neural network control strategy can reduce the simulation time by 96% and improve the real-time performance of energy control, which also ensures the good performance of power and fuel economy.

1. Introduction

Under the dual pressure of environmental pollution and energy crisis, hybrid vehicles have advantages of both conventional vehicles and electric vehicles, which have characteristics of energy conservation, environmental protection, diverse shapes, and strong implementation. Hybrid vehicles have become an effective way to solve the problem of energy crisis and environmental protection and also have been one of the most perspective vehicle models.

According to different connective ways of power system, hybrid electric vehicle (HEV) can be mainly divided into four styles: series, parallel, series-parallel, and complex. The dynamic structure diagram is shown in Figure 1.

Engine output energy of series HEV is transformed two times and the efficiency of the motor and generator is relatively low, so series HEV loses more energy and leads to lower efficiency than vehicles of internal combustion engine. Parallel HEV (PHEV) is equipped with series and parallel power systems, and their structures and control systems are more complex and have higher cost. Complex HEV structures and control systems are most complex and have highest cost. However, parallel hybrid power system

can adapt to various road conditions and is widely used by enterprises [1, 2].

As the core of multiple energy control system, the energy control strategy determines performances of PHEV. Based on vehicle's torque, energy control strategies of PHEV are mainly divided into four types [3, 4]: static logic threshold energy control strategy, instantaneous optimal energy control strategy, global optimal energy control strategy, and neural network energy control strategy. The static logic threshold control strategy cannot guarantee the optimal fuel economy of PHEV, does not adapt to dynamic conditions, and cannot make the whole system to achieve maximum efficiency. Besides, its threshold parameters are set by engineering experience [5, 6]. The global optimal control strategy can achieve global optimal fuel consumption of HEV. Its defects are complex algorithm, large amount of calculation, and knowing the whole condition in advance [7, 8]. Neural network energy control strategy can adapt to diverse conditions with good robustness and obtain global fuel consumption optimum by engineering experience [9, 10]. Instantaneous optimal energy control strategy can realize minimum equivalent fuel consumption of PHEV in each control cycle which is widely used to distribute PHEV energy [11].

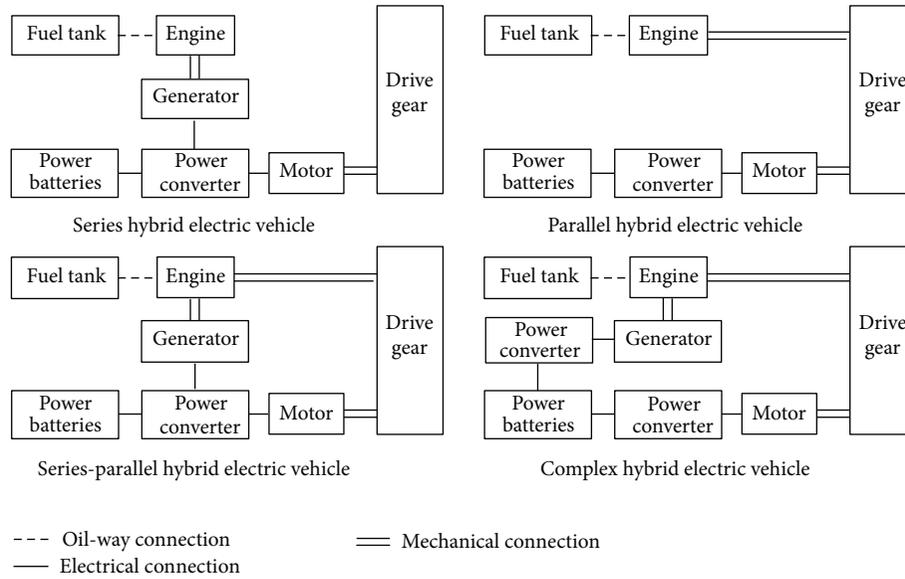


FIGURE 1: Classification of hybrid electric vehicles.

The basic principle of instantaneous optimal control strategy is based on the model of the optimal curve of engines; the object function of the whole power system was optimized on the specific operating points of parallel HEV. On the basis of the instantaneous optimal operating points, it makes power of variable states redistributed and make the loss of energy minimized in the energy flow process at any time (see Figure 2). Instantaneous optimal control strategy has good fuel economy at any time and bad real-time performance. Its real-time performance is influenced by these factors which are the accuracy of various components battery ages and engine and motor characteristics [12–14]. So it is difficult to improve the real-time performance of instantaneous optimal control strategy by changing these factors.

The hybrid vehicles possess good power performance and fuel economy and obtain rapid allocation energy by finding a new energy control strategy. Elman neural network is a feedback neural network and has a very strong computing ability and stability [15]. The instantaneous rules of the instantaneous optimal control strategy are used to train Elman neural network, establish Elman neural network controller, and improve the real-time performance of energy control [14–16].

Based on the research of the instantaneous optimal control strategy, the strategy possesses good fuel economy and makes energy distributed reasonably. However, its real-time performance is poor. In order to solve bad real-time defects of instantaneous control strategy, instantaneous optimal control rules are used to train the Elman neural network control strategy and improve the real-time performance of the trained Elman energy control strategy on the premise that it can guarantee advantages of the instantaneous optimal control strategy. The results show that the trained Elman neural network control strategy can replace the instantaneous

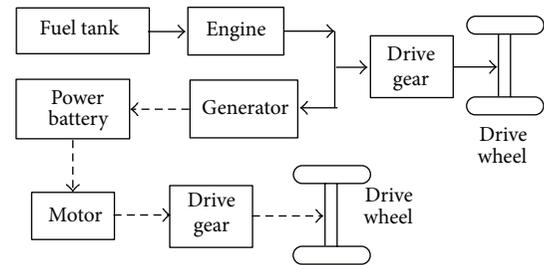


FIGURE 2: Energy flow diagram of parallel hybrid electric vehicles.

optimal control strategy, optimize power distribution, and make the simulation time reduced by 60%.

2. Research Energy Control Strategy

2.1. Instantaneous Optimal Control Strategy. Instantaneous optimal control strategy is defined as follows. In order to achieve the minimum fuel consumption of HEV, the optimal output power of the engine and electric motor is calculated in each control cycle of hybrid power system. Working conditions of HEV and calculation expressions of the equivalent fuel consumption are different in every time. So an optimal objective function should be established [17]. Working condition of hybrid system is divided into two cases of power battery charging and discharging. Then, objective functions of the instantaneous equivalent minimum fuel consumption were established on two working conditions.

Here, the full line represents the circulation and transformation of fuel chemical energy in the hybrid power system. The dotted line represents electric current circulation and transformation in the hybrid power system.

2.1.1. Calculate the Equivalent Fuel Consumption of Battery Discharging. When the power battery takes part in driving hybrid cars, its SOC value will reduce and deviate from the target of SOC value. In order to compensate for the used electricity and restore SOC value of power batteries, the engine drives the motor to charge power batteries in the future time [18]. The charging time is divided into n control cycles, and the motor power of each cycle is $P_{mc.chg.i}$ ($i = 1, 2, 3 \dots, n$) [5].

The relationship between the motor power (P_{mc}) with driving vehicle and the motor power ($P_{mc.chg.i}$) with power batteries charging is

$$P_{mc} = \bar{\eta}_{mc}^2 \bar{\eta}_{dischg} \bar{\eta}_{chg} \sum_{i=1}^n P_{mc.chg.i}, \quad (1)$$

where $\bar{\eta}_{mc}$ is the average efficiency of motor; $\bar{\eta}_{chg}$ is the average efficiency of power battery charging; $\bar{\eta}_{dischg}$ is the average efficiency of power battery discharging.

When the motor drives the vehicle, the energy consumption of power batteries can be converted into the engine fuel consumption. The equivalent fuel consumption rate of the motor is

$$b_{mc.eq} = \frac{\sum_{i=1}^n (P_{fcY.i} b_{Y.i} - P_{fcN.i} b_{N.i})}{\bar{\eta}_{mc}^2 \bar{\eta}_{dischg} \bar{\eta}_{chg} \sum_{i=1}^n P_{mc.chg.i}}, \quad (2)$$

where $P_{fcN.i}$, when the engine does not charge power batteries, is the engine power of the i control period; $b_{N.i}$, when the engine does not charge power batteries, is the fuel consumption rate of the engine; $P_{fcY.i}$, when the engine charges power batteries, is the engine power of the i control period; $b_{Y.i}$, when the engine charges power batteries, is the fuel consumption rate of the engine.

Let \bar{b}_{chg} be the average fuel consumption rate when the engine charges power batteries:

$$\bar{b}_{chg} = \sum_{i=1}^n \frac{P_{fcY.i} b_{Y.i} - P_{fcN.i} b_{N.i}}{P_{mc.chg.i}}. \quad (3)$$

Merge (3) and (2):

$$b_{mc.eq} = \frac{\bar{b}_{chg}}{\bar{\eta}_{mc}^2 \bar{\eta}_{dischg} \bar{\eta}_{chg}}. \quad (4)$$

When the motor drives the vehicle after a period of Δt , equivalent instantaneous fuel consumption of the motor is

$$m_{mc.eq} = \frac{\bar{b}_{chg} P_{mc}}{\bar{\eta}_{mc}^2 \bar{\eta}_{dischg} \bar{\eta}_{chg}}. \quad (5)$$

2.1.2. Calculate the Fuel Consumption of Batteries Charging. When the power battery is charged by the engine, its SOC value will rise and even exceed the target of the SOC value. In order to maintain SOC values, power battery energy will be consumed in future [18]. Discharging time is divided into n control cycles, and the motor power of the each control cycle is $P_{mc2.i}$ ($i = 1, 2, 3 \dots, n$) [17].

In a certain period of discharging time, the relationship between motor power ($P_{mc2.i}$) with driving vehicle and motor power ($P_{mc.chg2}$) with charging power batteries is

$$\sum_{i=1}^n P_{mc2.i} = P_{mc.chg2} \bar{\eta}_{mc2}^2 \bar{\eta}_{chg2} \bar{\eta}_{dischg2}, \quad (6)$$

where $P_{mc.chg2}$ is the motor power when the power battery is charged; $\bar{\eta}_{mc2}$ is the average efficiency of motor; $\bar{\eta}_{chg2}$ is the average efficiency of the power battery charging; $\bar{\eta}_{dischg2}$ is the average efficiency of the power battery discharging.

When the motor drives the vehicle, the relationship between the motor power battery energy consumption and the fuel consumption rate is

$$b_{mc.eq2} = \frac{P_{fcY2} b_{Y2} - P_{fcN2} b_{N2}}{P_{mc.chg2} \bar{\eta}_{mc2}^2 \bar{\eta}_{chg2} \bar{\eta}_{dischg2}}, \quad (7)$$

where P_{fcY2} , when the engine charges the power battery, is the power of the engine; b_{Y2} , when the engine charges the power battery, is the fuel consumption rate of the engine; P_{fcN2} , when the engine does not charge the power battery, is the power of the engine; b_{N2} , when the engine does not charge the power battery, is the fuel consumption rate of the engine.

Let b_{chg2} be the fuel consumption rate when the engine charges the power battery:

$$b_{chg2} = \frac{P_{fcY2} b_{Y2} - P_{fcN2} b_{N2}}{P_{mc.chg2}}. \quad (8)$$

Simplify the (8) formula:

$$b_{mc.eq2} = \frac{b_{chg2}}{\bar{\eta}_{mc2}^2 \bar{\eta}_{dischg2} \bar{\eta}_{chg2}}. \quad (9)$$

When the motor charges power batteries after a period of Δt , the objective function of equivalent instantaneous fuel consumption of the motor is

$$m_{mc.eq2} = \frac{\bar{b}_{chg2} P_{mc.chg2}}{\bar{\eta}_{mc2}^2 \bar{\eta}_{dischg2} \bar{\eta}_{chg2}}. \quad (10)$$

2.1.3. Deduce the Objective Function of the Instantaneous Optimal Control Strategy. Set two new variables:

$$f_{eq.dischg} = \frac{\bar{b}_{chg}}{\bar{\eta}_{mc}^2 \bar{\eta}_{dischg} \bar{\eta}_{chg}}, \quad f_{eq.chg} = \frac{b_{chg2}}{\bar{\eta}_{mc2}^2 \bar{\eta}_{dischg2} \bar{\eta}_{chg2}}, \quad (11)$$

where b_{chg2} is the fuel consumption rate when the engine charges power batteries at the present moment; \bar{b}_{chg} is the average fuel consumption rate when the engine charges power batteries in the future time.

The instantaneous control objective function of the lowest fuel consumption is

$$M = \sum \text{Min} \left\{ m_{fc} [T_{fc}(t), \omega(t)] \Delta t + m_{mc.eq} [T_{mc}(t), \omega_{mc}(t)] \Delta t \right\}, \quad (12)$$

where $\omega_{fc}(t)$ is engine speed; $\omega_{mc}(t)$ is motor speed; T_{fc} is output torque of the engine; T_{mc} is output torque of the motor. Consider

$$m_{mc.eq} [P_{mc}(t)] = \begin{cases} f_{eq.dischg} P_{mc} & T_{mc} > 0 \text{ (power batteries discharge)} \\ f_{eq.chg} P_{mc.chg2} & T_{mc} < 0 \text{ (power batteries charge)}. \end{cases} \quad (13)$$

P_{mc} and $P_{mc.chg2}$ are both motor power, so they can be unified as P_{mc} . By calculating, the improved instantaneous control objective function of the minimum fuel consumption is

$$M = \sum \text{Min} \{ m_{fc} [P_{fc}(t)] \Delta t + [\lambda f_{eq.dischg} + (1 - \lambda) f_{eq.chg}] P_{mc} \Delta t \}, \quad (14)$$

where $\lambda = (1 + \text{sign}(T_{mc}))/2$.

2.1.4. Improve the Objective Function of the Instantaneous Optimal Control Strategy. SOC value change of batteries and braking energy recovery both have a certain effect on energy control. The optimal function of the working point needs to be improved.

(1) *Revise the SOC Value Function of Power Batteries.* When power batteries work, their SOC value is maintained at the high efficient range by the reset function, in order to reduce the loss energy in the process of charging and discharging power batteries and make hybrid system keep better performances. The working principle of SOC reset function is as follows. When the SOC value is more than the target region, the hybrid power system will give priority to consuming power battery energy. It does not stop until SOC value decreases to the target region under the effect of the reset function. When the battery SOC value is lower than the target region, the hybrid system will give priority to consuming fuel energy to drive the vehicle and recover the value of SOC. It does not stop until SOC value returns to the target region under the effect of the reset function.

Set K_{SOC} be the variable in the reset function, and the value table between K_{SOC} and SOC value is shown in Table 1.

Based on the Matlab platform and Table 1 data, the fitting curve between K_{SOC} and SOC value is constructed by using fitting curve toolbox, as shown in Figure 3.

Polynomial function of the fitting curve is

$$K_{soc} = -8866x^7 + 33460x^6 - 52310x^5 + 43730x^4 - 21010x^3 + 5766x^2 - 833.5x + 49.77, \quad (15)$$

where x : is SOC value.

TABLE 1: The value table between K_{soc} and SOC.

SOC	0.11	0.14	0.2	0.25	0.3	0.35
K_{soc}	7	2	1.05	1.05	1.03	1.04
SOC	0.4	0.45	0.5	0.55	0.59	0.64
K_{soc}	1.02	1.02	1	0.99	0.98	0.96
SOC	0.7	0.74	0.81	0.85	0.9	
K_{soc}	0.96	0.94	0.94	0.8	0.2	

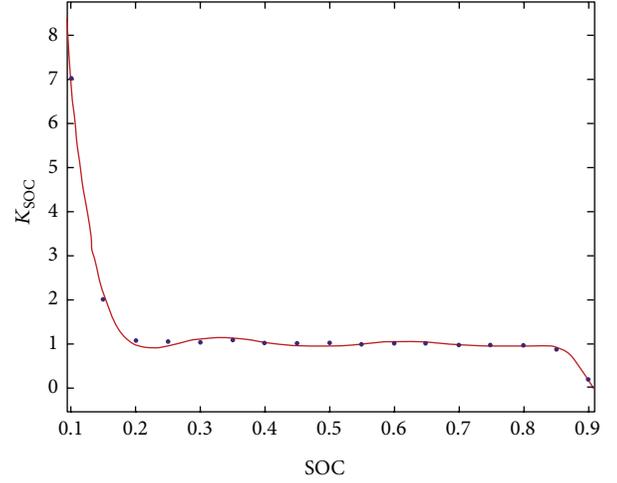


FIGURE 3: SOC and K_{SOC} relationship curve.

Considering the influence of power battery SOC, the formula of instantaneous equivalent fuel consumption of HEV can be expressed as $K_{soc} m_{mc.eq}$.

(2) *Revise the Objective Function of Braking Recovery Energy.* The adopted method which revises the equivalent fuel consumption function of power battery energy is as follows: the average braking power is calculated at a period of time before the current moment. The power is used as the standard of fuel consumption correction of the braking recovery power in the next moment [19–22].

The statistical time range is divided into n ($n > 0$) cycles. In each cycle, let braking power be a fixed value. Therefore, the average braking power of the whole time can be expressed as

$$\bar{P}_{braking} = \frac{\sum_{i=1}^n P_{braking.i}}{n}. \quad (16)$$

When power batteries discharge, the objective function of instantaneous equivalent fuel consumption is

$$m_{mc.eq} = \frac{\bar{b}_{chg} (P_{mc} + \bar{P}_{braking})}{\bar{\eta}_{mc} \bar{\eta}_{dischg} \bar{\eta}_{chg}}. \quad (17)$$

In summary, taking the influence of power battery SOC and brake energy recovery into consideration, the final

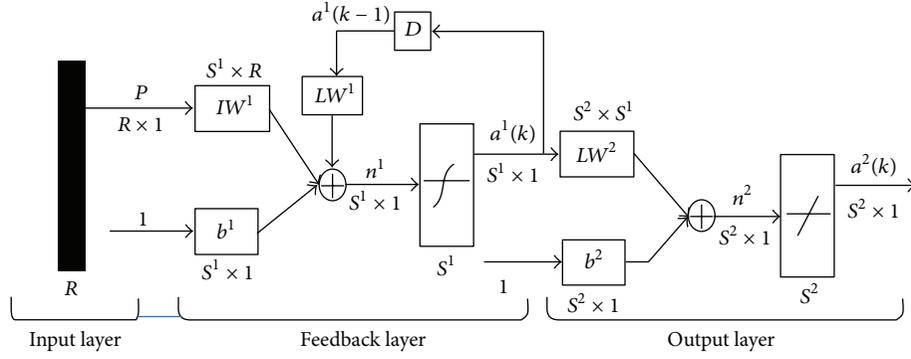


FIGURE 4: Elman neural network model.

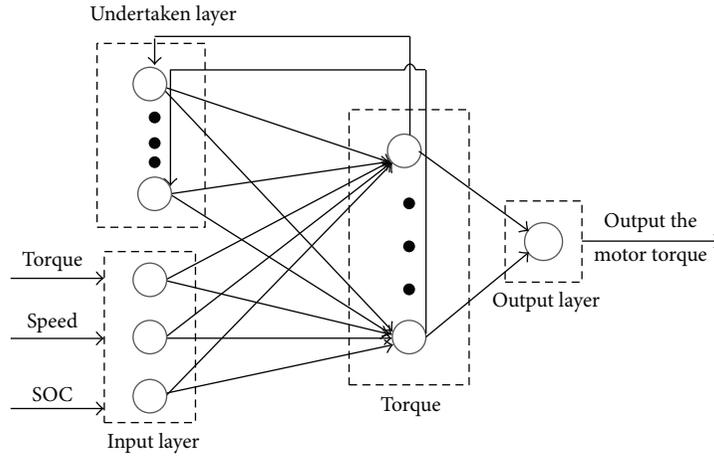


FIGURE 5: Structure diagram of Elman neural network.

objective function of instantaneous optimal control strategy is

$$M = \sum \text{Min} \{ m_{fc} [P_{fc}(t)] \Delta t + K_{soc} m_{mc,eq} [P_{mc}(t) + \bar{P}_{braking}(t)] \Delta t \}. \quad (18)$$

2.2. Elman Neural Network Control Strategy

2.2.1. The Structure of Elman Neural Network. Elman neural network is put forward by Jeffrey L. Elman in 1990 and is a typical local recessionary grid, as shown in Figure 4 [23].

P is the input of the neural network and its size is $R \times 1$; b^1 is the neuronal threshold vector of the feedback layer and its size is $S^1 \times 1$; IW^1 is the connective weight vector of the neurons and the input vector in the input layer and its size is $S^1 \times R$; n^1 is the middle operational result of the neurons in the feedback layer, namely, weighted sum of the connective weight vector and the threshold vector, and its size is $S^1 \times 1$; a^1 is the output vector of the feedback layer's neuron in the K iteration and its size is $S^1 \times 1$; D is feedback node. It is the same way that b^2 , LW^2 , n^2 , and a^2 are related to parameters of the output layer.

The inputs of structure diagram of Elman neural network is the required torque, speed, and SOC value of power battery

and its output is motor torque. The structure diagram of Elman neural network is shown in Figure 5 [24].

The structure diagram of Elman neural network contains input layer, hidden layer, undertaken layer, and output layer. Let the input vector of the input layer be three-dimensional vector u ; the output vector of the output layer is one-dimensional vector y ; the output vector of the hidden layer is n -dimensional vector x ; the output vector of the undertaken layer is n -dimensional vector r ; w^1 , w^2 , and w^3 are respective connective weights of the hidden layer to the output layer, the input layer to the hidden layer, and the undertaken layer to the hidden layer; $g(\cdot)$ is the driving function of the output neurons; $f(\cdot)$ is the driving function of hidden layer; $h(\cdot)$ is the driving function of undertaken layer; $\text{net}(\cdot)$ is the net input driving function of a certain layer; A shows the input layer; B shows the undertaken layer; K shows the iterative sequence.

Define two functions:

$$v_i(k) = \begin{cases} u_n(k), & \text{if } i \in A, \\ r_n(k), & \text{if } i \in B, \end{cases} \quad (19)$$

$$w^i(k) = \begin{cases} w^2, & \text{if } i \in A, \\ w^3, & \text{if } i \in B. \end{cases}$$

The input and output functions of the N neuron of the hidden layer are

$$\begin{aligned} \text{net}_n(k+1) &= \sum_{i \in AUB} w^i(k) v_i(k), \\ x_n(n+1) &= f(\text{net}_n(k+1)). \end{aligned} \quad (20)$$

The input and output functions of the N neuron of the hidden layer are

$$\begin{aligned} \text{net}_n(k) &= \sum_{i \in AUB} w^i(k-1) v_i(k-1), \\ r(k) &= h(\text{net}_n(k)). \end{aligned} \quad (21)$$

The input and output functions of the output layer's neuron are

$$\begin{aligned} \text{net}_n(k+1) &= \sum_{i \in AUB} w^i(k+1) x_n(k+1), \\ y(k+1) &= g(\text{net}(k+1)). \end{aligned} \quad (22)$$

2.2.2. *Select the Parameters of Elman Neural Network.* The neuron number is determined by following formula [25]:

$$k = \sqrt{m+n} + \beta, \quad (23)$$

where m is the number of the input vector; n is the neuron number of the output vector; β is a constant, (1~10).

The excitation function of Elman neural network of the feedback layer selects the Tansig function [26]:

$$\text{tansig}(x) = \frac{2}{1 + e^{-2x}} - 1. \quad (24)$$

2.2.3. *Learning and Training Mechanism of Elman Neural Network.* Elman neural network is trained by Levenberg-Marquardt algorithm. The error index function of Levenberg-Marquardt arithmetic is

$$E(w) = \frac{1}{2} \sum_{i=1}^p \|y_i - y_i'\|^2 = \frac{1}{2} \sum_{i=1}^p e_i^2(w), \quad (25)$$

where p is the sample number; e^i is the systemic error; y_i' is the actual output of the network.

The formula of the adjusting weight is

$$w^{k+1} = w^k + \Delta w. \quad (26)$$

The computing formula of the increment weight is

$$\Delta w = [J^T(w) J(w) + uI]^{-1} J^T(w) e(w), \quad (27)$$

where u is learning rate; I is the unit matrix; $J(w)$ is the Jacobian matrix. Consider

$$J(w) = \begin{bmatrix} \frac{\partial e_1(w)}{\partial w_1} & \frac{\partial e_1(w)}{\partial w_2} & \dots & \frac{\partial e_1(w)}{\partial w_n} \\ \frac{\partial e_2(w)}{\partial w_1} & \frac{\partial e_2(w)}{\partial w_1} & \dots & \frac{\partial e_2(w)}{\partial w_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial e_n(w)}{\partial w_1} & \frac{\partial e_n(w)}{\partial w_2} & \dots & \frac{\partial e_n(w)}{\partial w_n} \end{bmatrix}. \quad (28)$$

Neural network weight is adjusted by the Levenberg-Marquardt algorithm, and the adjusting process is shown in Figure 6.

3. Modeling of Energy Control Strategy

3.1. *Modeling of Instantaneous Optimal Control Strategy.* On the platform of Matlab/simulink, the instantaneous control model is established and mainly includes two parts: the calculation model of average braking energy of recovered power and the calculative model of optimal working point [25, 26].

First, the calculation model of average braking energy of the recovery power is established as shown in Figure 7. Inputs of the model are the required torque and rotational speed and its output is the calculated average recovery power of braking energy in the transmission system.

Second, the calculation model of the revised SOC value of power battery is established, as shown in Figure 8. The model can maintain the SOC value of power batteries at a high efficient range.

Third, the calculation model of the optimal operating point is established, as shown in Figure 9. Inputs of the model are the required torque, rotational speed, modified $S_{\text{target-motor}}$ value of power batteries, and the average braking recovery power and the output torque of the motor is the optimal allocation power between the engine and motor in the transmission system.

Finally, the whole simulation diagram of the instantaneous control strategy is established, as shown in Figure 10.

3.2. *The Elman Neural Network Model.* Elman neural network is gradually learning to do something by imitating human brain. Its essence is to improve the learned knowledge and the neurons weight [27–30]. The flow diagram of the trained Elman neural network is shown in Figure 11.

4. Simulation Results and Analysis

The basic vehicle parameters are shown in Table 2.

Traffic parameters of simulation experiments are described in Table 3.

The original control model is replaced by the instantaneous optimal control model in ADVISOR 2002. Then the trained Elman neural network controller is imported to the software [31, 32].

Simulation results are shown in Figure 13. HWFET_mdf and NEDC_mdf show simulation results of the instantaneous optimal control strategy of the PHEV model (Figure 12); HWFET_net and NEDC_net show simulation results of the Elman neural network controller of the PHEV model.

Compared with the instantaneous optimal control strategy, the Elman neural network strategy can slow down the decline of SOC value and maintain SOC value at high efficient range on the NEDC working condition in Figure 13(a). Meanwhile, the Elman neural network strategy has approximate SOC value change compared with the instantaneous optimal control strategy on the HWFET working in Figure 13(b).

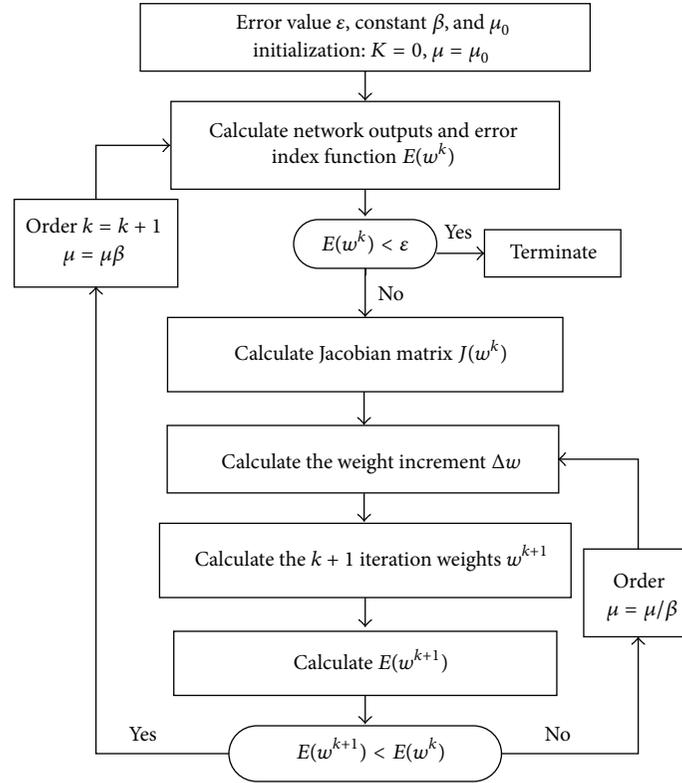


FIGURE 6: Process map of the adjusting network weight.

TABLE 2: Parallel hybrid electric vehicle parameters.

Vehicle		
Curb weight	1605 kg	
Face area	2.65 m ²	
Wheel base	2.775 m	
Height of the center of mass	0.5 m	
Front axle load distribution ratio	0.51	
Coefficient drag	0.32	
Engine		
Peak power	118 kW	
Displacement	2.5 L	
Power battery pack		
Voltage	244.8/650 V	
Style	NI-MH	
Volume	6.5 Ah	
Mold number	34	
Motor		
Peak power	105 kW	
Style	PMSM	

As shown in Figure 13(a), Elman neural network strategy can make the engine produce more torque than the instantaneous optimal control strategy at the beginning of 600 s on the NEDC working condition. It can make the vehicle start, accelerate, and climb better. After the vehicle starts, the two control strategies play the same role on the

TABLE 3: Traffic parameters of simulation experiments.

Parameter	NEDC	HWFET
Idle time (s)	298	6
Top speed (km/h)	10.93	16.51
Cycle time (s)	1184	765
Average speed (km/h)	33.21	77.58
Maximum acceleration (m/s ²)	1.06	1.43
Maximum deceleration (m/s ²)	-1.39	-1.48
Park time (time)	13	1
Traveling distance (km)	120	96.4

vehicle energy control. As seen in Figure 13(b), Elman neural network strategy can make the engine produce slightly more torque than instantaneous optimal control strategy on the HWFET working condition, while these two strategies have similar effect on the engine torque control.

As seen in Figure 15, Elman neural network strategy can make motor produce slightly more torque than the instantaneous optimal control strategy at some moments on the NEDC and HWFET working condition, while the two strategies have the similar effect on the motor torque control.

As shown in Table 4, compared with the instantaneous optimal neural strategy, fuel consumption of Elman neural network strategy only increases about 0.5 (L/100 km) on the NEDC and HWFET working condition, which implies that Elman neural network controller can also have the

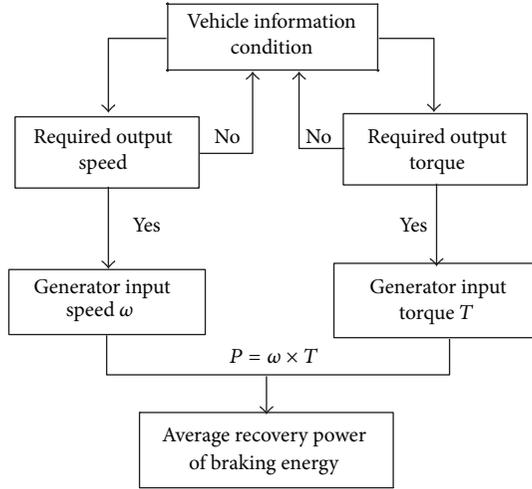


FIGURE 7: Flow diagram of average recovery power of braking energy.

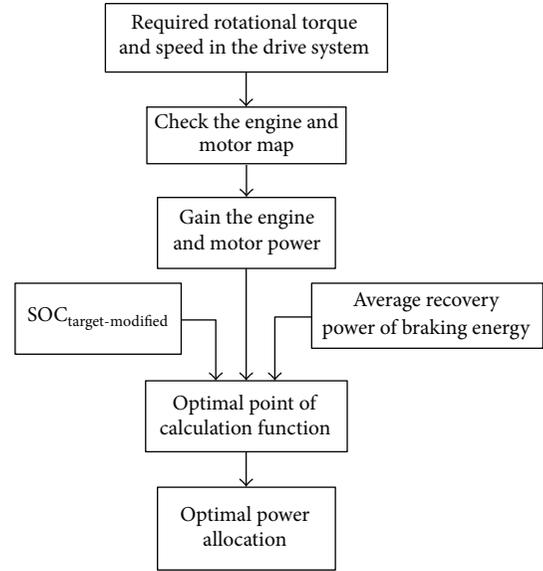


FIGURE 9: Flow diagram of the optimal operating point.

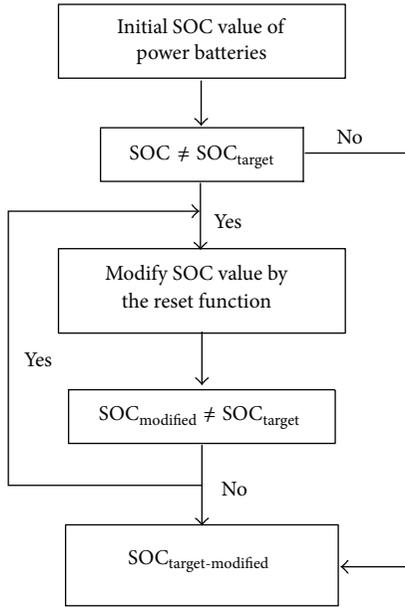


FIGURE 8: Flow diagram of the revised SOC value of power batteries.

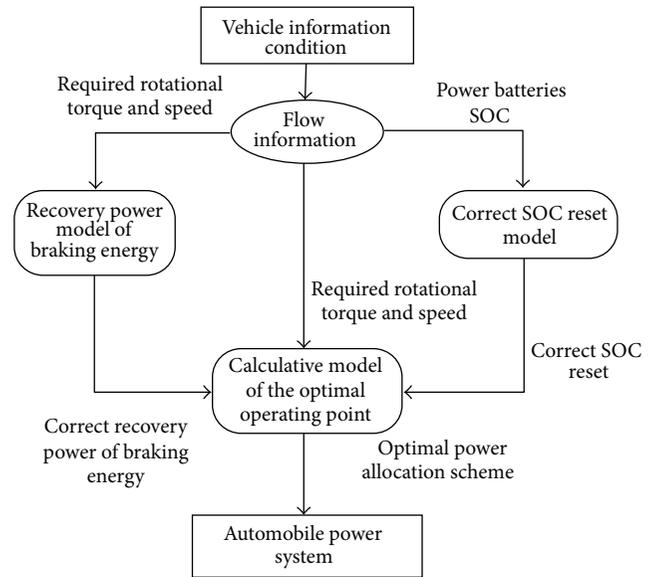


FIGURE 10: Total simulation diagram of the instantaneous control strategy.

TABLE 4: Fuel consumption of 100 km (L/100 km).

Strategy	Road	
	NEDC	HWFEF
Instantaneous optimal control	9.4	6.5
Elman neural network	9.8	7

advantage of low fuel consumption. The slight increase in fuel consumption can be accepted since it has a little effect on the fuel economy of the whole vehicle. As seen in Table 5, Elman neural network strategy makes the simulation time decreased greatly compared with the instantaneous optimal neural strategy and improves the response time of the vehicle greatly.

TABLE 5: Simulation time (s).

Strategy	Road	
	NEDC	HWFEF
Instantaneous optimal control	471.3	315.8
Elman neural network	15.6	10.2

In conclusion, as seen in Figures 13, 14, and 15, Elman neural network strategy can replace the instantaneous optimal control strategy to maintain SOC value at the high efficient range and achieve a reasonable distribution of the torque between the engine and the motor. The significance of the

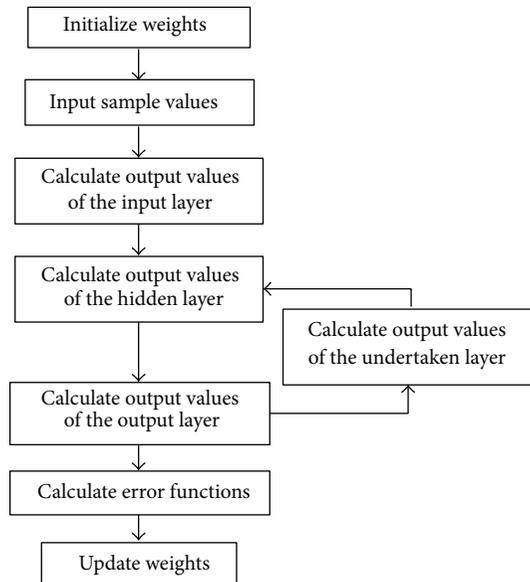


FIGURE 11: The training flow diagram of the trained Elman neural network.

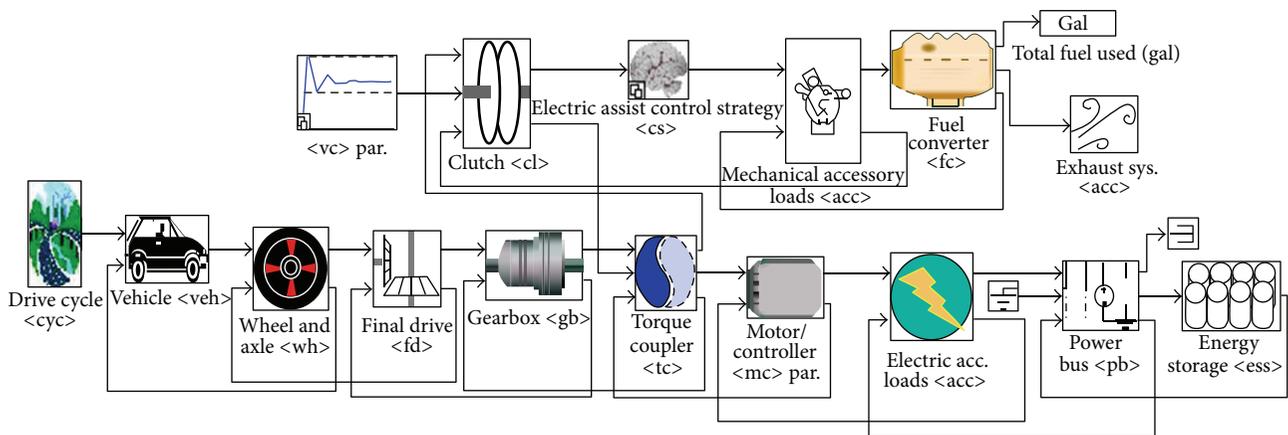


FIGURE 12: Simulation model of the parallel hybrid electric vehicle.

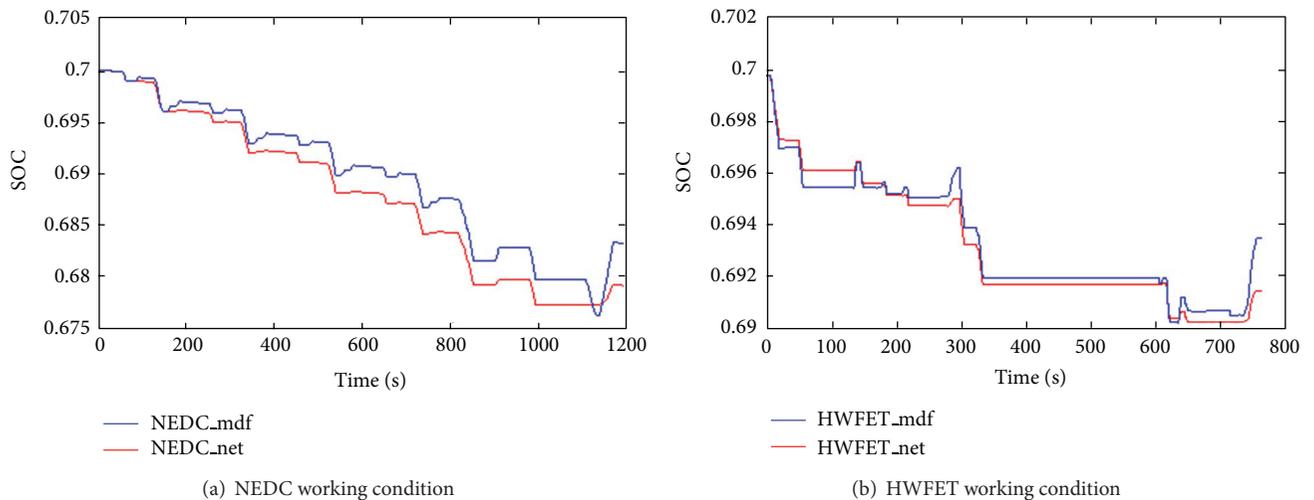


FIGURE 13: Contrast power battery SOC.

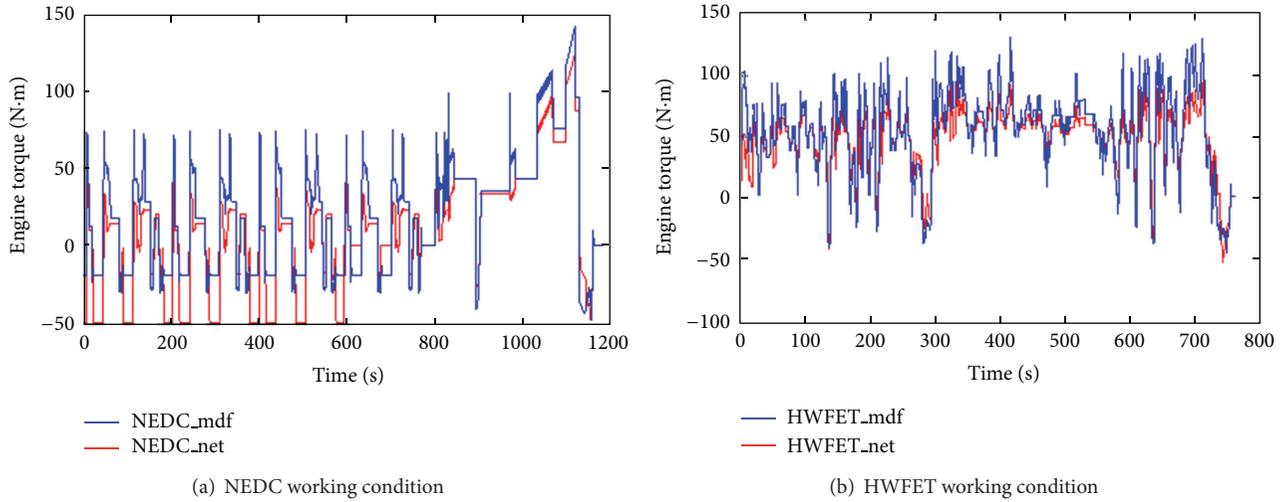


FIGURE 14: Contrast engine torque.

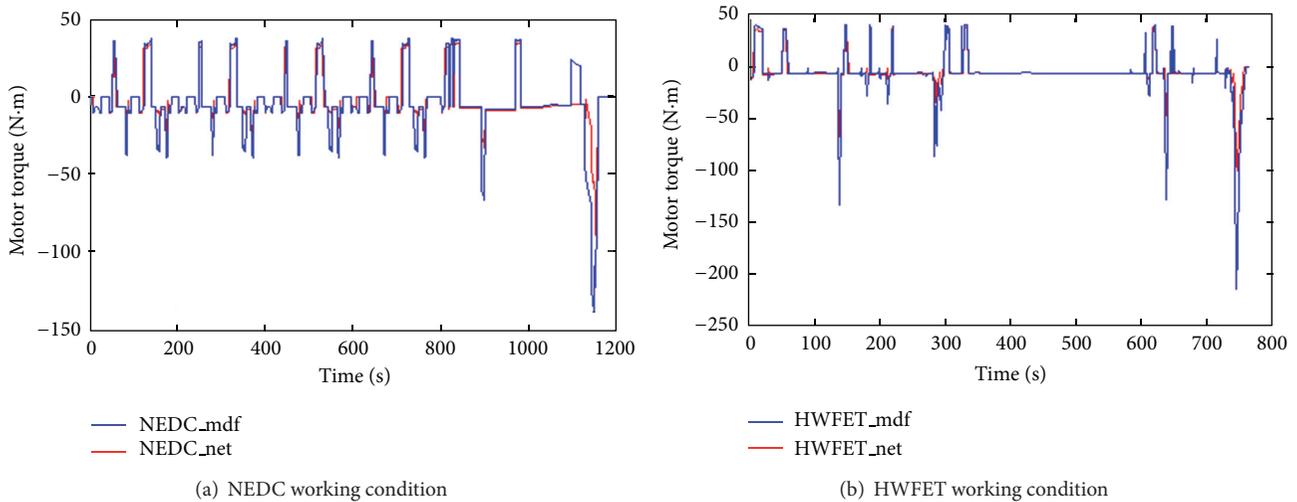


FIGURE 15: Contrast the motor torque.

paper is that the simulation time of energy control is reduced by 96%.

5. Conclusion

Through the research on the instantaneous optimal strategy and Elman neural network control strategy, we deduce the objective functions of instantaneous optimal control and establish the instantaneous control model and design the Elman controller. Based on the ADVISOR 2002 platform, two control strategies were simulated on a hybrid electric vehicle.

It is seen from the simulation results that the trained Elman neural network strategy shows similar control ability on the vehicle energy distribution compared with the instantaneous optimal control strategy, which ensures good performances of power and fuel economy of HEV, reduces the control reaction time greatly, and overcomes the disadvantage of poor real-time performance of the instantaneous optimal

control strategy. The research significance of the paper is that the simulation time of energy control is reduced by 96%.

Future works are listed as below.

- (1) Simulation and experiment should be improved by adding more design parameters, such as vehicle emission.
- (2) It is necessary to do lots of experiments to enrich simulation results.
- (3) Actual road condition is more complex than the simulation road condition, so control strategies need to be tested in the actual road conditions.

Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

References

- [1] S. G. Cui, *Automotive Technology Of New Energy*, Beijing University of Technology press, Beijing, China, 2009.
- [2] R. Hodkinson and J. Fentos, *Lightweight Electric/Hybrid Vehicle Design*, SAE International, Warrendale, Pa, USA, 2001.
- [3] J. Wu, *Optimization of energy management strategy for parallel hybrid electric vehicle [Ph.D. thesis]*, Shang Dong University, Ji Nan, China, 2008.
- [4] C. Lin, H. Peng, J. W. Grizzle, and J. Kang, "Power management strategy for a parallel hybrid electric truck," *IEEE Transactions on Control Systems Technology*, vol. 11, no. 6, pp. 839–849, 2003.
- [5] T. Hofman, M. Steinbuch, R. van Druten, and A. Serrarens, "A Rule-based energy management strategies for hybrid vehicles," *Electric and Hybrid Vehicles*, vol. 1, no. 1, pp. 71–94, 2007.
- [6] N. Jalil, N. A. Kheir, and M. Salman, "A rule-based energy management strategy for a series hybrid vehicle," in *Proceedings of the American Control Conference*, pp. 689–693, Albuquerque, NM, USA, 1997.
- [7] S. Delprat, J. Lauber, T. M. Guerra, and J. Rimaux, "Control of a parallel hybrid powertrain: optimal control," *IEEE Transactions on Vehicular Technology*, vol. 53, no. 3, pp. 872–881, 2004.
- [8] S. Delprat, T. M. Guerra, and J. Rimaux, "Optimal control of a parallel hybrid powertrain from global optimization to real time control strategy," in *Proceedings of the 18th International Electric Vehicle Symposium*, Berlin, Germany, 2001.
- [9] M. H. Asemani and V. J. Majd, "Stability of output-feedback DPDC-based fuzzy synchronization of chaotic systems via LMI," *Chaos, Solitons and Fractals*, vol. 42, no. 2, pp. 1126–1135, 2009.
- [10] G. D. Tian, J. W. Chu, Y. Lu, H. Ke, X. Zha, and G. Xu, "Expected energy analysis for industrial process planning problem with fuzzy time parameters," *Computers & Chemical Engineering*, vol. 35, no. 12, pp. 2905–2912, 2011.
- [11] G. Paganelli and S. Delprat, "Equivalent consumption minimization strategy for parallel hybrid powertrains," in *Proceedings of the IEEE Vehicular Technology Conference*, vol. 4, pp. 2076–2081, 2002.
- [12] J. Park, Y. Park, and J. H. Park, "Real-time powertrain control strategy for series-parallel hybrid electric vehicles," *SAE Paper* 3172, 2007.
- [13] A. Sciarretta, M. Back, and L. Guzzella, "Optimal control of parallel hybrid electric vehicle," *IEEE*, vol. 12, no. 3, pp. 352–363, 2004.
- [14] J. Park, Y. Park, and J. H. Park, "Real-time powertrain control strategy for series-parallel hybrid electric vehicles," *SAE Paper*, vol. 1, no. 3172, 2007.
- [15] H. R. Karimi and H. J. Gao, "New delay-dependent exponential H_∞ synchronization for uncertain neural networks with mixed time delays," *IEEE Transactions on Systems, Man, and Cybernetics B: Cybernetics*, vol. 40, no. 1, pp. 173–185, 2010.
- [16] R. Razavian, N. L. Azad, and J. Mcphee, "On real-time optimal control of a series Hybrid Electric Vehicle with an ultracapacitor," in *Proceedings of the American Control Conference*, pp. 547–552, 2012.
- [17] D. Sinoquet, G. Rousseau, and Y. Milhau, "Design optimization and optimal control for hybrid vehicles," *Optimization and Engineering*, vol. 12, no. 1-2, pp. 199–213, 2011.
- [18] G. Paganelli, G. Ercole, A. Brahma, Y. Guezennec, and G. Rizzoni, "General supervisory control policy for the energy optimization of charge-sustaining hybrid electric vehicles," *JSAE Review*, vol. 22, no. 4, pp. 511–518, 2001.
- [19] W. H. Sun, H. J. Gao, and O. Kaynak, "Adaptive backstepping control for active suspension systems with hard constraints," *IEEE/ASME Transactions on Mechatronics*, vol. 18, no. 3, pp. 1072–1079, 2013.
- [20] W. Sun, Z. Zhao, and H. Gao, "Saturated adaptive robust control for active suspension systems," *IEEE Transactions on Industrial Electronics*, vol. 60, no. 9, pp. 3889–3896, 2013.
- [21] W. Sun, H. Gao Sr., and O. Kaynak, "Finite frequency H_∞ control for vehicle active suspension systems," *IEEE Transactions on Control Systems Technology*, vol. 19, no. 2, pp. 416–422, 2011.
- [22] W. Sun, Y. Zhao, J. Li, L. Zhang, and H. Gao, "Active suspension control with frequency band constraints and actuator input delay," *IEEE Transactions on Industrial Electronics*, vol. 59, no. 1, pp. 530–537, 2012.
- [23] F. J. Lin, Y. C. Hung, and S. Y. Chen, "FPGA-based computed force control system using Elman neural network for linear ultrasonic motor," *IEEE Transactions on Industrial Electronics*, vol. 56, no. 4, pp. 1238–1253, 2009.
- [24] G. Pajares, M. Guijarro, and A. Ribeiro, "A Hopfield Neural Network for combining classifiers applied to textured images," *Neural Networks*, vol. 23, no. 1, pp. 144–153, 2010.
- [25] G. D. Tian, Y. M. Liu, Q. Z. Tian, and J. W. Chu, "Evaluation model and algorithm of product disassembly process with stochastic feature," *Clean Technologies and Environmental Policy*, vol. 14, no. 2, pp. 345–356, 2012.
- [26] G. D. Tian, M. C. Zhou, J. W. Chu, and Y. M. Liu, "Probability evaluation models of product disassembly cost subject to random removal time and different removal labor cost," *IEEE Transactions on Automation Science and Engineering*, vol. 9, no. 2, pp. 288–295, 2012.
- [27] M. Zapateiro, N. Luo, H. R. Karimi, and J. Vehí, "Vibration control of a class of semiactive suspension system using neural network and backstepping techniques," *Mechanical Systems and Signal Processing*, vol. 23, no. 6, pp. 1946–1953, 2009.
- [28] W. Pawlus, H. R. Karimi, and K. G. Robbersmyr, "Data-based modeling of vehicle collisions by nonlinear autoregressive model and feedforward neural network," *Information Sciences*, vol. 235, pp. 65–79, 2013.
- [29] J. Guo and W. J. Sun, *Neural Network Theory and the Realization of MATLAB7*, Publishing House of Electronics Industry, Beijing, China, 2005.
- [30] K. L. Zhou and Y. H. Kang, *Neural Network Model and MATLAB Simulation Programming*, Tsinghua University press, Beijing, China, 2005.
- [31] T. Markel, A. Brooker, T. Hendricks et al., "ADVISOR: a systems analysis tool for advanced vehicle modeling," *Journal of Power Sources*, vol. 110, no. 2, pp. 255–266, 2002.
- [32] X. H. Zeng and Q. N. Wang, "Develop control strategy model of hybrid electric vehicle based on ADVISOR2002," *Automotive Engineering*, vol. 26, no. 5, pp. 423–427, 2004.

Research Article

The Burgers Equation for a New Continuum Model with Consideration of Driver's Forecast Effect

Lei Yu¹ and Bingchang Zhou²

¹ College of Automation, Northwestern Polytechnical University, Xi'an, Shaanxi 710129, China

² School of Science, Northwestern Polytechnical University, Xi'an, Shaanxi 710129, China

Correspondence should be addressed to Lei Yu; yuleijk@126.com

Received 18 June 2014; Revised 12 August 2014; Accepted 12 August 2014; Published 24 August 2014

Academic Editor: Weichao Sun

Copyright © 2014 L. Yu and B. Zhou. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

A new continuum model with consideration of driver's forecast effect is obtained to study the density wave problem and the stop-and-go phenomena. The stability condition of the new model is derived by using linear analysis. The triangular shock wave, one type of density wave, which is determined by Burgers equation in the stable region, is discussed in great detail with reductive perturbation method. The local cluster appears when we perform the numerical simulations for the new model. It also proves that the driver's forecast effect has the positive effect of reducing the local cluster.

1. Introduction

Traffic jams, the typical signature of the complex behavior of vehicular traffic, have been studied by various traffic models [1–8]. From different theoretical basis, there are microscopic and macroscopic models to describe traffic flow. The dynamical aspects of microscopic models are based on the description of the individual vehicles' situation. This process is largely determined by the drivers' behavior and the physical performance of vehicles. The macroscopic models describe traffic streams as a compressible fluid obeying global rules. This coarse-grained means is needed for understanding the collective behavior of traffic, designing efficient control strategies, developing macroscopic traffic simulation, and so forth.

Bando et al. [9] propose the optimal velocity (OV) model to characterize the car-following behavior. Although the OV model is shown to have the universal structure in describing many properties of traffic flow, many approaches to extending the model toward more realistic traffic model have been pursued. Helbing and Tilch [10] develop a generalized force model with a velocity difference term added into the OV model. Xue et al. [11] extend the OV model to take into account the effect of the relative velocity.

The pioneer work of continuum traffic flow models is the LWR model [12, 13]. Although the LWR model can reproduce most basic traffic flow phenomena such as traffic congestion formation and dissipation in heavy traffic, this model can not describe nonequilibrium traffic flow dynamics and does not have the ability to explain the amplification of small disturbances in heavy traffic. To overcome the deficiencies in the LWR model, various macroscopic traffic models have been proposed. Payne [14] introduces a high-order continuum model which can describe the amplification of small disturbances in heavy traffic and allow fluctuations of speed around the equilibrium value. Thus, the Payne model is suitable to describe nonequilibrium situations such as stop-and-go traffic. However, a fundamental principle of traffic flow, that is, a car being influenced only by the motion of cars ahead of it, not by the motion of cars behind it, is violated in the Payne model because one characteristic speed of this model is greater than the macroscopic flow speed. Notice that cars are anisotropic particles and respond only to frontal stimuli. Jiang et al. [15] develop a macroscopic continuum model based on a car-following theory. This model overcomes the characteristic speed problem that exists in many high-order continuum models such as the Payne model.

However, the consistency between microscopic and macroscopic models has been proved [16]. There are also some papers to derive macroscopic continuum models from micromodels (mainly the car-following models), such as [14, 15, 17].

There are many studies to reveal nonlinear phenomena of vehicular traffic, such as stop-and-go, phase transition, self-organized, and the nonlinear waves. Kurtze and Hong [18] derive the Korteweg-de Vries (KdV) equation from one continuum traffic flow model. Zhou et al. [19] obtain the KdV equation and the modified Korteweg-de Vries (mKdV) equation from the continuum traffic flow model derived from a car-following model. Ou [17] obtains the Burgers equation and the KdV equation from a continuum version of the full velocity difference car-following model. Yu et al. [20] investigate density waves in an optimal velocity model with reaction-time delay of drivers and derive the Burgers, KdV, and mKdV equations. However, the nonlinear waves results of the continuum traffic flow models are far less than those of the car-following models. The reason is probably the complex partial differential forms of the continuum traffic flow models.

In fact, the future traffic situation can be forecasted by the intelligent transportation system (ITS) based on the current traffic status, so the driver may be guided by the forecast information to adjust his/her current acceleration. However, few models consider the drivers forecast effect. Recently, a new car-following model with the driver's forecast effect is proposed by Tang et al. [21]. Similar to the other models which consider the information of ITS, the model presented by Tang et al. [21] can improve the stability of traffic flow and reduce traffic jams.

In this paper, a new macroversion is obtained based on the anisotropic continuum model proposed by Tang et al. [21]. The density wave problem and the stop-and-go phenomena are studied. In Section 2, the stability condition of the model is derived. In Section 3, the triangular shock wave, which is determined by the Burgers equation in the stable region, is discussed in great detail by using nonlinear analysis. In Section 4, the simulation results are given. These results prove that the driver's forecast effect has the positive effect of reducing the local cluster. Finally, a summary is given.

2. The Model and Its Stability Analysis

The new car-following model with the consideration of the driver's forecast effect can be written as follows [21]:

$$\begin{aligned} \frac{dv_n(t)}{dt} = & \kappa (V(\Delta x_n(t)) - v_n(t)) \\ & + \beta \kappa (V(\Delta x_n(t + \tau)) - v_n(t + \tau)), \end{aligned} \quad (1)$$

where κ is the reactive coefficient, β is the coefficient of the driver's forecast effect, and τ is the time-step of the driver forecast. Using the transformation between microvariables

and macrovariables, the corresponding anisotropic macro-continuum model of (1) is obtained [21], that is, the following equations:

$$\begin{aligned} \frac{\partial \rho}{\partial t} + \frac{\partial(\rho v)}{\partial x} &= 0, \\ \frac{\partial v}{\partial t} + v \frac{\partial v}{\partial x} &= \frac{(1 + \beta)(V_e(\rho) - v)}{T + \beta\tau} + \beta\tau u'_e(h) c_0 \frac{\partial v}{\partial x}, \end{aligned} \quad (2)$$

where T is the reactive time given by the inverse of the coefficient κ , $c_0 = \epsilon/(\beta\tau + T) > 0$ is the propagating velocity of the small perturbation, ϵ is the distance between the following and leading vehicles in micromodel, and

$$u_e(h) = V_e(\rho), \quad h = 1/\rho. \quad (3)$$

To derive the stability condition, the Burgers equation, and its shock solution of the macroversion of (1), the following new macromodel is derived from (2) according to (3); that is,

$$\frac{\partial \rho}{\partial t} + \frac{\partial(\rho v)}{\partial x} = 0, \quad (4)$$

$$\frac{\partial v}{\partial t} + v \frac{\partial v}{\partial x} = \gamma (V_e(\rho) - v) - \beta\tau c_0 \rho^2 V'_e(\rho) \frac{\partial v}{\partial x},$$

where $\gamma = (1 + \beta)/(T + \beta\tau)$. Obviously, model (4) is anisotropy.

The linear stability theory will be applied to derive the linear stability condition of model (4). Assume traffic to be initially in a state differing infinitesimally from the uniform steady flow. Similar to the decomposition of the flow of (4) into a linear combination of Fourier modes in [18], we have

$$\begin{aligned} \rho(x, t) &= \rho_0 + \sum_k \hat{\rho}_k \exp(ikx + \sigma_k t), \\ v(x, t) &= v_0 + \sum_k \hat{v}_k \exp(ikx + \sigma_k t), \end{aligned} \quad (5)$$

where ρ_0 and v_0 are the uniform steady states of (4). Substituting (5) into (4), linearizing and neglecting the higher-order terms of the small perturbations $\hat{\rho}_k$ and \hat{v}_k , we have

$$(\sigma_k + ikv_0) \hat{\rho}_k + ik\rho_0 \hat{v}_k = 0, \quad (6)$$

$$\begin{aligned} \sigma_k \hat{v}_k + ikv_0 \hat{v}_k &= \gamma (V'_e(\rho_0) \hat{\rho}_k - \hat{v}_k) \\ &- \beta\tau c_0 \rho_0^2 V'_e(\rho_0) ik \hat{v}_k. \end{aligned} \quad (7)$$

From (6), we have

$$\hat{\rho}_k = -\frac{ik\rho_0 \hat{v}_k}{\sigma_k + ikv_0}. \quad (8)$$

Substituting (8) into (7), the following quadratic equation is obtained:

$$\begin{aligned} (\sigma_k + ikv_0)^2 + (\gamma + \beta\tau c_0 \rho_0^2 V'_e(\rho_0) ik) (\sigma_k + ikv_0) \\ + \gamma \rho_0 V'_e(\rho_0) ik = 0. \end{aligned} \quad (9)$$

Consider the long wave expansion of σ_k in (9), which is determined order by order around $ik \approx 0$ [8]. By expanding $\sigma_k = \sigma_1(ik) + \sigma_2(ik)^2 + \dots$ and separating the real part and the imaginary part, the coefficients of ik and $(ik)^2$ are derived as follows:

$$\begin{aligned}\sigma_1 &= -v_0 - \rho_0 V_e'(\rho_0), \\ \sigma_2 &= \frac{1}{\gamma} \rho_0^2 V_e'^2(\rho_0) (\omega \rho_0 - 1),\end{aligned}\quad (10)$$

where $\omega = \beta \tau c_0$.

The linear stability condition of (4) is decided by the real part of σ_k . In fact, the uniform traffic flow described by (4) is stable against all infinitesimal disturbances when $\sigma_2 > 0$; that is,

$$\omega \rho_0 > 1. \quad (11)$$

It means that the traffic flow described by (4) is stable when (11) holds.

3. Nonlinear Analysis

To investigate the system behavior in the stable region, we consider the long wavelength modes on coarse-grained scales. The reductive perturbation method as in [8] is applied to (4). Introduce the slow scales for space variable x and time variable t and define slow variables X and T as follows:

$$X = \varepsilon(x - bt), \quad T = \varepsilon^2 t, \quad (12)$$

where b is a constant to be determined and $0 < \varepsilon \ll 1$. We set the density and velocity as

$$\begin{aligned}\rho(x, t) &= \rho_0 + \varepsilon \hat{\rho}(X, T), \\ v(x, t) &= v_0 + \varepsilon \hat{v}(X, T).\end{aligned}\quad (13)$$

Substituting (12) and (13) into (4) and making the Taylor expansions to ε^3 , the following nonlinear partial differential equations are derived:

$$\begin{aligned}\varepsilon^2 (\rho_0 \partial_X \hat{v} + v_0 \partial_X \hat{\rho} - b \partial_X \hat{\rho}) \\ + \varepsilon^3 (\partial_T \hat{\rho} + \hat{\rho} \partial_X \hat{v} + \hat{v} \partial_X \hat{\rho}) = 0, \\ \varepsilon \gamma (\hat{v} - V_e'(\rho_0) \hat{\rho}) \\ + \varepsilon^2 \left[\hat{\rho} \hat{v} + (v_0 - b + \omega \rho_0^2 V_e'(\rho_0)) \partial_X \hat{v} \right. \\ \left. - V_e'(\rho_0) \hat{\rho}^2 - \frac{1}{2} \gamma V_e''(\rho_0) \hat{\rho}^2 \right] \\ + \varepsilon^3 \left[\partial_T \hat{v} + \hat{v} \partial_X \hat{v} - \frac{1}{6} \gamma V_e'''(\rho_0) \hat{\rho}^3 \right. \\ \left. + (\omega \rho_0^2 V_e''(\rho_0) + 2\omega \rho_0 V_e'(\rho_0)) \hat{\rho} \partial_X \hat{v} \right] = 0.\end{aligned}\quad (14)$$

From the coefficient of ε in (15), we have

$$\hat{v} = V_e'(\rho_0) \hat{\rho} + O(\varepsilon^2). \quad (16)$$

The relationship between the perturbation of density and velocity is given by (16) which is the basis of Burgers equation's derivation. According to (16), we have the value of b from the second term of ε in (14); that is,

$$b = \rho_0 V_e'(\rho_0) + v_0. \quad (17)$$

Then we have

$$\begin{aligned}\hat{q} = \hat{\rho} \hat{v} &= \left(V_e'(\rho_0) + \frac{1}{2} \gamma V_e''(\rho_0) \right) \hat{\rho}^2 \\ &+ (1 - \omega \rho_0) \rho_0 V_e'^2(\rho_0) \partial_X \hat{\rho}.\end{aligned}\quad (18)$$

Combining (14) and (18), the following equation is derived:

$$\begin{aligned}\partial_T \hat{\rho} + (2V_e'(\rho_0) + \gamma V_e''(\rho_0)) \hat{\rho} \partial_X \hat{\rho} \\ = (\omega \rho_0 - 1) \rho_0 V_e'^2(\rho_0) \partial_X^2 \hat{\rho}.\end{aligned}\quad (19)$$

In accordance with stability condition (11), the coefficient of the second derivative term on the right hand side of (19) is positive in the stable region. Thus, in the stable region, (19) is just the Burgers equation.

If $\hat{\rho}(X, 0)$ is of compact support, the solution $\hat{\rho}(X, T)$ of Burgers equation, (19), behaves like a train of N -triangular shock waves; that is,

$$\begin{aligned}\hat{\rho}(X, T) &= \frac{1}{|c_1| T} \left(X - \frac{\eta_n + \eta_{n+1}}{2} \right) - \frac{\eta_{n+1} - \eta_n}{2|c_1| T} \\ &* \tanh \left(\frac{c_2}{4|c_1| T} (\eta_{n+1} - \eta_n) (X - \xi_n) \right),\end{aligned}\quad (20)$$

where $c_1 = 2V_e'(\rho_0) + \gamma V_e''(\rho_0)$, $c_2 = (\omega \rho_0 - 1) \rho_0 V_e'^2(\rho_0)$, ξ_n are the coordinates of the shock fronts, and η_n are the coordinates of the intersections of the slopes with the x -axis ($n = 1, 2, \dots, N$). $\hat{\rho}(X, T)$ decays to 0 like $O(1/T)$ when $T \rightarrow +\infty$. That is to say, any shock wave expressed by (19) in stable traffic flow region will evolve to a uniform flow in the course of time. We see this phenomenon in Figures 1(a) and 1(e) of Section 4.

4. The Numerical Simulation

To check the theoretical results, we carry out numerical simulations for model (4) by using the numerical scheme in [21]. The difference equations are as follows:

$$\rho_i^{j+1} = \rho_i^j + \frac{\Delta t}{\Delta x} \rho_i^j (v_i^j - v_{i+1}^j) + \frac{\Delta t}{\Delta x} v_i^j (\rho_{i-1}^j - \rho_i^j), \quad (21)$$

$$(a) \text{ if } v_i^j < -\omega(\rho_i^j)^2 V_e'(\rho_i^j),$$

$$v_i^{j+1} = v_i^j + \frac{\Delta t}{\Delta x} \left(-\omega(\rho_i^j)^2 V_e'(\rho_i^j) - v_i^j \right) (v_{i+1}^j - v_i^j) \\ + \gamma \Delta t (V_e(\rho_i^j) - v_i^j), \quad (22)$$

$$(b) \text{ if } v_i^j \geq -\omega(\rho_i^j)^2 V_e'(\rho_i^j),$$

$$v_i^{j+1} = v_i^j + \frac{\Delta t}{\Delta x} \left(-\omega(\rho_i^j)^2 V_e'(\rho_i^j) - v_i^j \right) (v_i^j - v_{i-1}^j) \\ + \gamma \Delta t (V_e(\rho_i^j) - v_i^j), \quad (23)$$

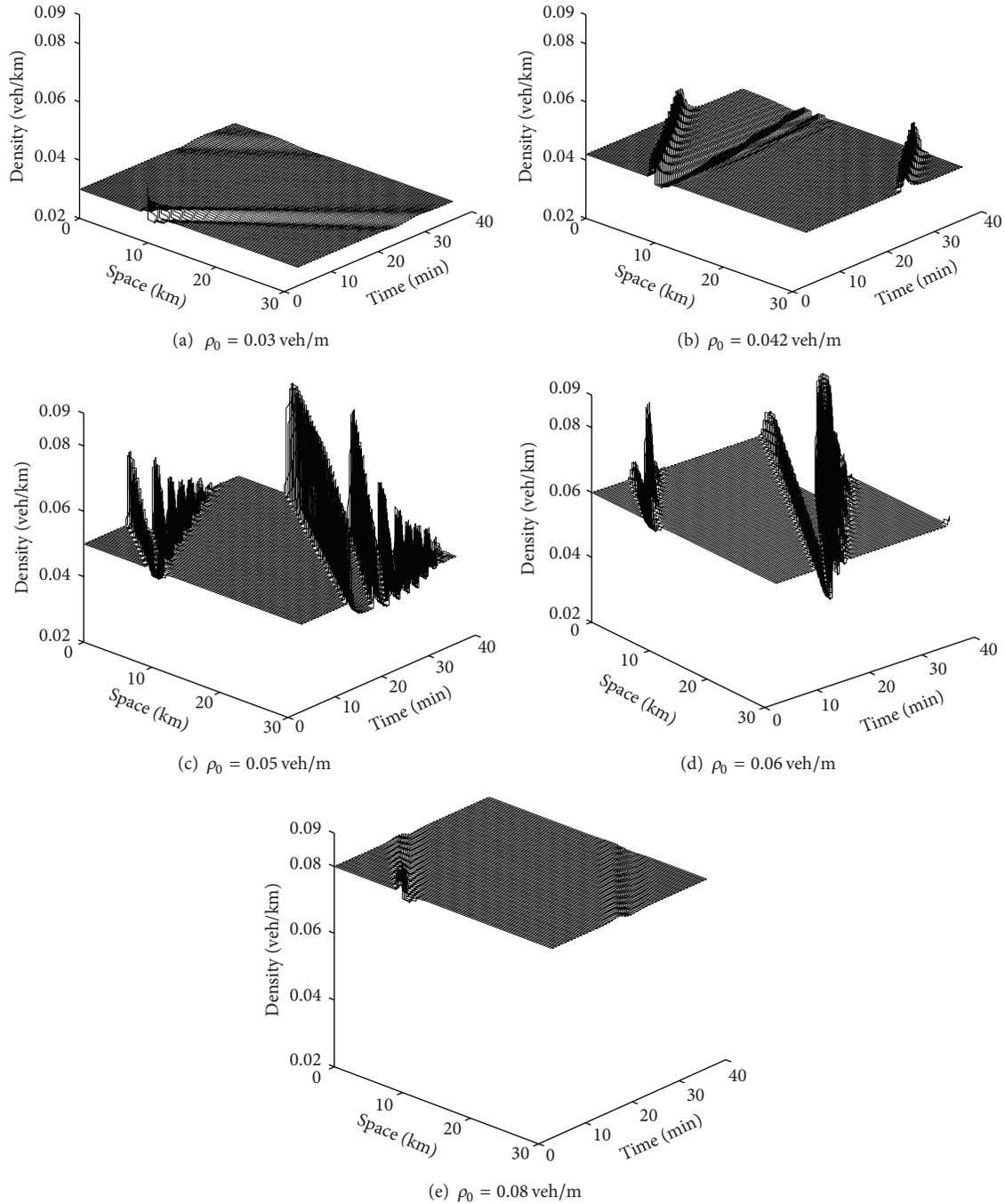


FIGURE 1: Temporal evolution of traffic flow before 40 minutes for different ρ_0 with $\beta = 0.2$.

where $i, j, \Delta x$, and Δt represent the road section, time, spatial step, and time step, respectively, $\rho_i^j \approx \rho(x_i, t_j)$, $v_i^j \approx v(x_i, t_j)$, $x_i = i\Delta x$, $t_j = j\Delta t$, and i, j are integers.

To study the local cluster effect of (4), we perform numerical simulations over a system of 32.2 km long highway using (21)–(23). The local cluster effect corresponds to the stop-and-go wave observed in the traffic flow due to the amplification of a small disturbance. We simulate the traffic

flow under the periodic boundary conditions. The following variation of the initial density ρ_0 is used as in [15]:

$$\rho(x, 0) = \rho_0 + \Delta\rho_0 \left\{ \cosh^{-2} \left[\frac{160}{L} \left(x - \frac{5L}{16} \right) \right] - \frac{1}{4} \cosh^{-2} \left[\frac{40}{L} \left(x - \frac{11L}{32} \right) \right] \right\}, \quad (24)$$

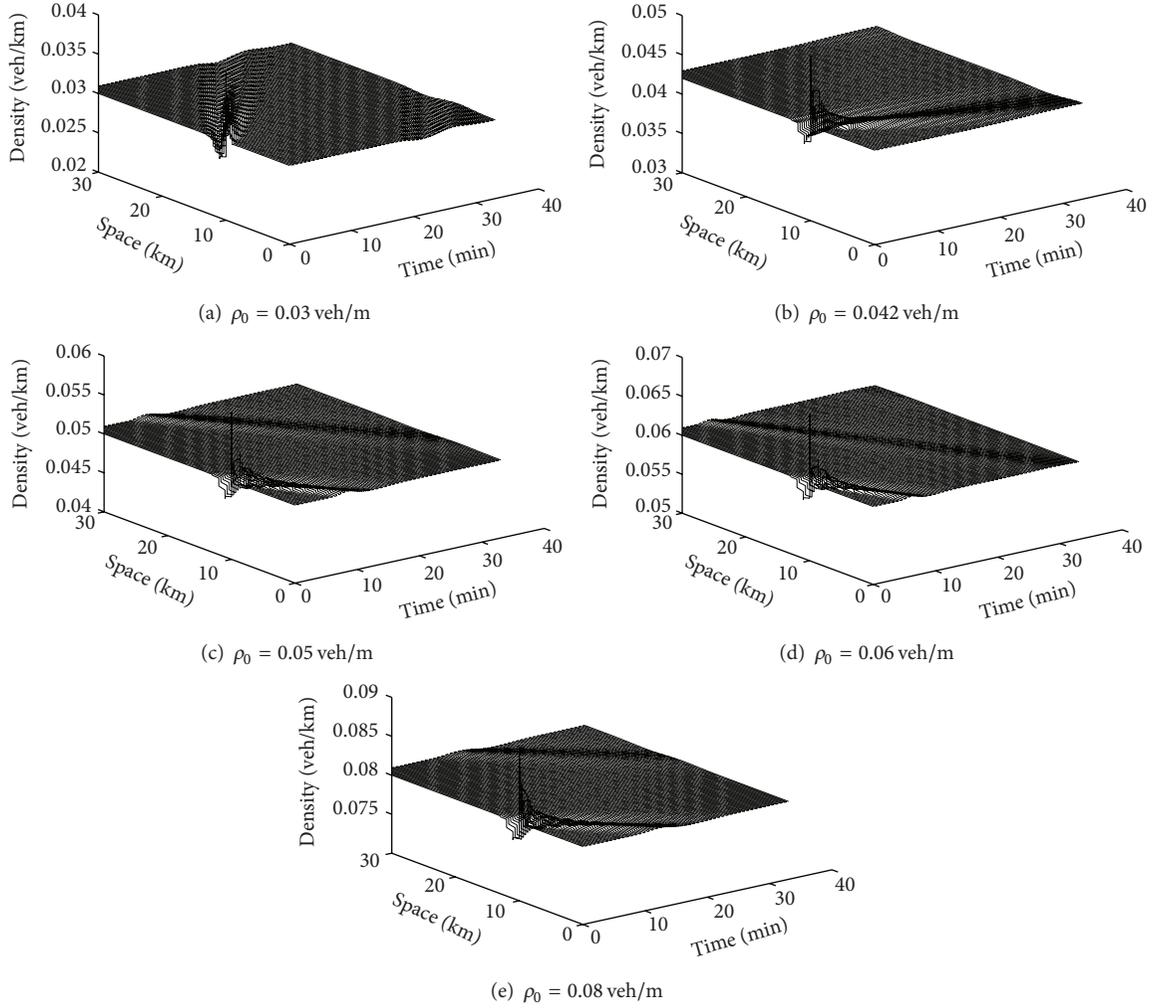


FIGURE 2: Temporal evolution of traffic flow before 40 minutes for different ρ_0 with $\beta = 0.4$.

where $L = 32.2$ km is the length of the road section under consideration. The periodic boundary conditions are given as follows:

$$\rho(L, t) = \rho(0, t), \quad v(L, t) = v(0, t). \quad (25)$$

Here we use the equilibrium speed-density relationship proposed in [22]:

$$V_e(\rho) = v_f \left\{ \left[1 + \exp\left(\frac{\rho/\rho_m - 0.25}{0.06}\right) \right]^{-1} - 3.72 \times 10^{-6} \right\}. \quad (26)$$

Assume the initial flow to be in local steady state; that is, $v(x, 0) = V_e(\rho(x, 0))$. Let $\Delta\rho_0 = 0.01$ veh/m, let the space interval Δx be 100 m, and let the time interval Δt be 1 s. The choice of Δx and Δt satisfies the Courant-Friedrichs-Levy (CFL) stability condition. The other parameter values we take are as follows:

$$\begin{aligned} \tau = 5 \text{ s}, \quad T = 10 \text{ s}, \quad v_f = 30 \text{ m/s}, \\ c_0 = 11 \text{ m/s}, \quad \rho_m = 0.2 \text{ veh/m}. \end{aligned} \quad (27)$$

Figures 1 and 2 show the evolution of initial uniform traffic flow under the small disturbance.

Figure 1 shows the temporal evolution of traffic flow before 40 minutes for different ρ_0 with $\beta = 0.2$. In pattern (a) of Figure 1, the initial density $\rho_0 = 0.03$ of traffic flow is so low that the disturbance dies out without any amplification with time. With the initial density ρ_0 increasing, small disturbance is amplified and so leads to the instability of traffic flow. The pattern (b) in Figure 1 shows that several local clusters form for the initial density $\rho_0 = 0.042$. In pattern (c) of Figure 1, the stop-and-go phenomenon, which is a complex local structure consisting of multiclusters, can be observed. Continuing to increase the initial density ρ_0 , we can see a dipole-like structure which is illustrated by Figure 1(d). In pattern (e) of Figure 1, a stable traffic flow is reached again with the initial density $\rho_0 = 0.08$.

Figure 2 shows the temporal evolution of traffic flow before 40 minutes for different ρ_0 with $\beta = 0.4$. In patterns of Figure 2, there are no clusters and the disturbances propagate backward without any amplification with different initial densities ρ_0 .

Comparing every subfigure of Figures 1 and 2, the value of β is the only different parameter. It is obvious that the increase of β leads to the stabilization of traffic flow. So the driver's forecast effect has the positive effect of reducing the local cluster. This result is coincident with the one in [21].

The above numerical results are not so perfect. The propagation speeds of the perturbation waves we get from our figures are about 40–60 km/h against the traffic direction, which are not very consistent with the ones from the real data. References [23, 24] point out that the propagation velocity which is obtained from the real data should be about –12–20 km/h. Many factors may lead to this result such as the option of equilibrium function, the difference scheme, or the parameters. The control and optimization problems of the parameters in our model can also be studied [25–30]. We will continue to study these problems in future.

5. Summary

In this paper, a new macroversion is proposed with the consideration of the driver's forecast effect proposed. The stability condition of the model is derived by using linear analysis. The triangular shock wave which is determined by Burgers equation in the stable region is discussed with reductive perturbation method. The stop-and-go phenomena appear when we carry out the numerical simulation for the model. The driver's forecast effect which has the positive effect of reducing the local cluster is proved.

Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

Acknowledgments

This work is supported by National Natural Science Foundation of China (Grants nos. 11102155 and 11102165), Natural Science Foundation of Shaanxi Province (2013JQ7014 and 2012JM1001), the new direction of the Young Scholar of Northwestern Polytechnical University, and Foundation for Fundamental Research of Northwestern Polytechnical University (JC20110265).

References

- [1] T. Ezaki and K. Nishinari, "Cluster size distribution in 1D-CA traffic models," *Journal of Physics A*, vol. 45, no. 4, Article ID 045101, 2012.
- [2] W.-X. Zhu and R.-L. Yu, "Nonlinear analysis of traffic flow on a gradient highway," *Physica A: Statistical Mechanics and Its Applications*, vol. 391, no. 4, pp. 954–965, 2012.
- [3] B. S. Kerner, "Physics of traffic gridlock in a city," *Physical Review E: Statistical, Nonlinear, and Soft Matter Physics*, vol. 84, Article ID 045102, 4 pages, 2011.
- [4] X. M. Zhao, D. F. Xie, B. Jia, R. Jiang, and Z. Y. Gao, "Disorder structure of free-flow and global jams in the extended BML model," *Physics Letters A*, vol. 375, no. 7, pp. 1142–1147, 2011.
- [5] D. Yang, J. Jin, B. Ran, Y. Pu, and F. Yang, "Modeling and analysis of car-truck heterogeneous traffic flow based on intelligent driver car-following model," in *Proceedings of the 92nd Annual Meeting on Transportation Research Board (TRB '13)*, Washington, Wash, USA, January 2013.
- [6] T. Nagatani, "Jamming and freezing transitions in CA model for facing pedestrian traffic with a soft boundary," *Physics Letters A*, vol. 374, no. 15-16, pp. 1686–1689, 2010.
- [7] S. Masukura, T. Nagatani, and K. Tanaka, "Jamming transitions induced by a slow vehicle in traffic flow on a multi-lane highway," *Journal of Statistical Mechanics: Theory and Experiment*, vol. 2009, no. 4, Article ID P04002, 2009.
- [8] T. Nagatani, "The physics of traffic jams," *Reports on Progress in Physics*, vol. 65, no. 9, pp. 1331–1386, 2002.
- [9] M. Bando, K. Hasebe, A. Nakayama, A. Shibata, and Y. Sugiyama, "Dynamical model of traffic congestion and numerical simulation," *Physical Review E*, vol. 51, no. 2, pp. 1035–1042, 1995.
- [10] D. Helbing and B. Tilch, "Generalized force model of traffic dynamics," *Physical Review E*, vol. 58, no. 1, pp. 133–138, 1998.
- [11] Y. Xue, L. Y. Dong, Y. W. Yuan, and S. Q. Dai, "The effect of the relative velocity on traffic flow," *Communications in Theoretical Physics*, vol. 38, no. 2, pp. 230–234, 2002.
- [12] M. J. Lighthill and G. B. Whitham, "On kinematic waves: II. A theory of traffic flow on long crowded roads," *Proceedings of the Royal Society A*, vol. 229, no. 1178, pp. 317–345, 1955.
- [13] P. I. Richards, "Shock waves on the highway," *Operations Research*, vol. 4, pp. 42–51, 1956.
- [14] H. J. Payne, "Models of freeway traffic and control," in *Mathematical Models of Public Systems*, G. A. Bekey, Ed., vol. 1 of *Simulation Councils Proceedings Series*, pp. 51–61, 1971.
- [15] R. Jiang, Q. Wu, and Z. Zhu, "A new continuum model for traffic flow and numerical tests," *Transportation Research Part B: Methodological*, vol. 36, no. 5, pp. 405–419, 2002.
- [16] D. C. Gazis, R. Herman, and R. B. Potts, "Car-following theory of steady-state traffic flow," *Operations Research*, vol. 7, pp. 499–505, 1959.
- [17] Z. Ou, "Density waves in the continuum analog of the full velocity difference model," *Physica A: Statistical Mechanics and its Applications*, vol. 387, no. 8-9, pp. 1799–1806, 2008.
- [18] D. A. Kurtze and D. C. Hong, "Traffic jams, granular flow, and soliton selection," *Physical Review E*, vol. 52, no. 1, pp. 218–221, 1995.
- [19] X. Zhou, Z. Liu, and J. Luo, "The density wave in a car-following model," *Journal of Physics A: Mathematical and General*, vol. 35, no. 20, pp. 4495–4500, 2002.
- [20] L. Yu, T. Li, and Z. K. Shi, "Density waves in a traffic flow model with reaction-time delay," *Physica A*, vol. 389, no. 13, pp. 2607–2616, 2010.
- [21] T. Q. Tang, H. J. Huang, and H. Y. Shang, "A new macro model for traffic flow with the consideration of the driver's forecast effect," *Physics Letters, Section A: General, Atomic and Solid State Physics*, vol. 374, no. 15-16, pp. 1668–1672, 2010.
- [22] B. S. Kerner and P. Konhäuser, "Cluster effect in initially homogeneous traffic flow," *Physical Review E*, vol. 48, no. 4, pp. R2335–R2338, 1993.
- [23] J. M. D. Castillo and F. G. Benítez, "On the functional form of the speed-density relationship-I: general theory," *Transportation Research Part B*, vol. 29, no. 5, pp. 373–389, 1995.
- [24] B. S. Kerner, "Synchronized flow as a new traffic phase and related problems for traffic flow modelling," *Mathematical and Computer Modelling*, vol. 35, no. 5-6, pp. 481–508, 2002.

- [25] W. C. Sun, H. J. Gao, and O. Kaynak, "Finite frequency H_∞ control for vehicle active suspension systems," *IEEE Transactions on Control Systems Technology*, vol. 19, no. 2, pp. 416–422, 2011.
- [26] W. Sun, Z. Zhao, and H. Gao, "Saturated adaptive robust control for active suspension systems," *IEEE Transactions on Industrial Electronics*, vol. 60, no. 9, pp. 3889–3896, 2013.
- [27] W. C. Sun, H. J. Gao, and O. Kaynak, "Adaptive backstepping control for active suspension systems with hard constraints," *IEEE/ASME Transactions on Mechatronics*, vol. 18, no. 3, pp. 1072–1079, 2013.
- [28] G. D. Tian, M. C. Zhou, J. W. Chu, and Y. M. Liu, "Probability evaluation models of product disassembly cost subject to random removal time and different removal labor cost," *IEEE Transactions on Automation Science and Engineering*, vol. 9, no. 2, pp. 288–295, 2012.
- [29] G. Tian, J. Chu, Y. Liu, H. Ke, X. Zhao, and G. Xu, "Expected energy analysis for industrial process planning problem with fuzzy time parameters," *Computers and Chemical Engineering*, vol. 35, no. 12, pp. 2905–2912, 2011.
- [30] B. Wang, G. D. Tian, Y. Liang, and T. G. Qiang, "Reliability modeling and evaluation of electric vehicle motor by using fault tree and extended stochastic Petri nets," *Journal of Applied Mathematics*, vol. 2014, Article ID 638013, 9 pages, 2014.

Research Article

3D Road Scene Monitoring Based on Real-Time Panorama

Yuezhou Wu,¹ Changjiang Liu,^{1,2} Shiyong Lan,¹ and Menglong Yang³

¹ National Key Laboratory of Fundamental Science on Synthetic Vision, Sichuan University, Chengdu 610065, China

² Key Laboratory of Higher Education of Sichuan Province for The Enterprise Informationalization and Internet of Things, Sichuan University of Science & Engineering, Zigong 643000, China

³ School of Aeronautics & Astronautics, Sichuan University, Chengdu 610065, China

Correspondence should be addressed to Changjiang Liu; liuchangjiang@189.cn

Received 12 May 2014; Revised 30 July 2014; Accepted 6 August 2014; Published 21 August 2014

Academic Editor: Yuxin Zhao

Copyright © 2014 Yuezhou Wu et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Road monitoring helps to control the regional traffic situation so as to adjust the traffic flow. Real-time panorama is conducive to timely treat traffic accidents and to greatly improve traffic capacity. This paper designs a 3D road scene monitoring framework based on real-time panorama. The system is the combination of large scale panorama, satellite map textures, and 3D scene model, in which users can ramble freely. This paper has the following contributions. Firstly, land-points were extracted followed by motion detection, then comotion algorithm was applied to land-points from adjacent cameras, and homography matrix was constructed. Secondly, reference camera was chosen and transformed to overhead viewpoint; subsequently multiviews were morphed to the same viewpoint and stitched to panorama. Finally, the registration based on high-precision GPS information between 2D road panorama and 3D scene model was also proposed. The proposed framework has been successfully applied to a large road intersection monitoring. Experimental results are furnished at the end of the paper.

1. Introduction

Independent camera only provides local information. As a result, it is difficult to make a reasonable judgment and immediate response to the global situation. Image mosaic can combine a set of images into a larger image with a wider field of view of the scene. It helps to improve observers' spatial awareness.

Over the years, numerous algorithms for image stitching have been developed. Image stitching is typically solved by finding global parametric warps to bring images into alignment [1]. 2D projective warps are parameterized by 3×3 homography, if the scene is planar or if the views differ purely by rotation [2, 3]. The warps of video images from multicameras were described in [4, 5]; however, they only focused on the images (infrared and visible images) of similar viewpoint. Lee et al. [6] introduced an automatic method to warp multiview images to a unified coordinate plane. They proposed time-space registration based on trajectory extraction and on the basis the homography matrices from images to the reference plane were estimated. However the trajectory

accuracy might influence fit precision. Reference [7] presented a method for ground plane estimation from image pairs and land-points, instead of static features such as color, shape, and contours, for image matching. Szlavik et al. [8] proposed a method for matching partially overlapping image pairs where the object of interest was in motion, namely, comotion algorithm. It was valid even if the motion is discontinuous and in an unstructured environment. Szlavik et al. [9] also applied comotion model to project a couple of camera images to the reference road plane.

Recently significant progress in the synthetic vision system (SVS) [10, 11], which is a computer-generated reality system, has been achieved. And that computer-generated image, namely, virtual image, is complementary to optical sensor-based vision. SVS supplies users with good immerse visualization. For example, virtual and reality technology has been studied in landmarks (horizon, runway) detection from an aircraft in low visibility conditions [12, 13].

In this paper, we propose a 3D road monitoring method based on real-time panorama. We have synthesized mosaicked images and static 3D scene model. Firstly, moving

targets were detected and land-points were extracted. Secondly, binary sequences in continuous time T were generated according to whether the pixel is land-point set or not. Subsequently land-points from adjacent cameras were matched according to similarity measure, resulting in homography transformation of adjacent cameras. Thirdly, choose a camera as a reference one and wrap the view to overhead view according to the world coordinates of ground control points (GCPs) and corresponding image coordinates. Thus with the transformation relationships between adjacent cameras, all the views were unified to overhead view. Finally, panorama was generated based on stitching lines. In the end of the paper, we proposed a novel panorama presentation of 3D road scene with real-time traffic situation. Specific process is shown in Figure 1.

2. Transformation of Multiview Based on Land-Points Comotion Method

The multiple views would be unified to the same viewpoint. In this paper 2D projective transformation brings all the images aligned to a common ground plane with overhead view. There are three main steps.

- (1) Homography relations are implemented via land-points comotion statistical maps between adjacent cameras.
- (2) Referenced image is projected to the ground plane on the basis of the world coordinate of GCPs and the corresponding image coordinates.
- (3) From steps (1) and (2), all the views are converted to the overhead view.

2.1. Calculation of Land-Points. Assuming the moving blobs has been detected before the calculation of land-points, land-point [7] is the point of central axis projected to the ground. The calculation of land-point is as follows.

- (1) Binarization of video images using hybrid Gauss background algorithm: moving blobs are denoted as

$$O = \{(x, y) \mid f(x, y) = 255\}. \quad (1)$$

- (2) Trace the contour of O , name the contour C , and only the blobs whose area of external polygon is greater than T_a (≥ 10) pixel² are counted.
- (3) The external rectangle of C is denoted as

$$R = \{(x, y) \mid x_{\min} \leq x \leq x_{\max}, y_{\min} \leq y \leq y_{\max}\}. \quad (2)$$

- (4) Line l_j perpendicular to the x -axis $x = x_j$, $x_{\min} \leq x_j \leq x_{\max}$; the distance from point $P_i(x_i, y_i) \in O$ to line l_j is

$$D(P_i, l_j) = |x_i - x_j|. \quad (3)$$

- (5) Calculate the abscissa of land-point

$$x_{\text{mid}} = \arg \min_j \left\{ \text{median}_{P_i \in O} (D(P_i, l_j)) \right\}. \quad (4)$$

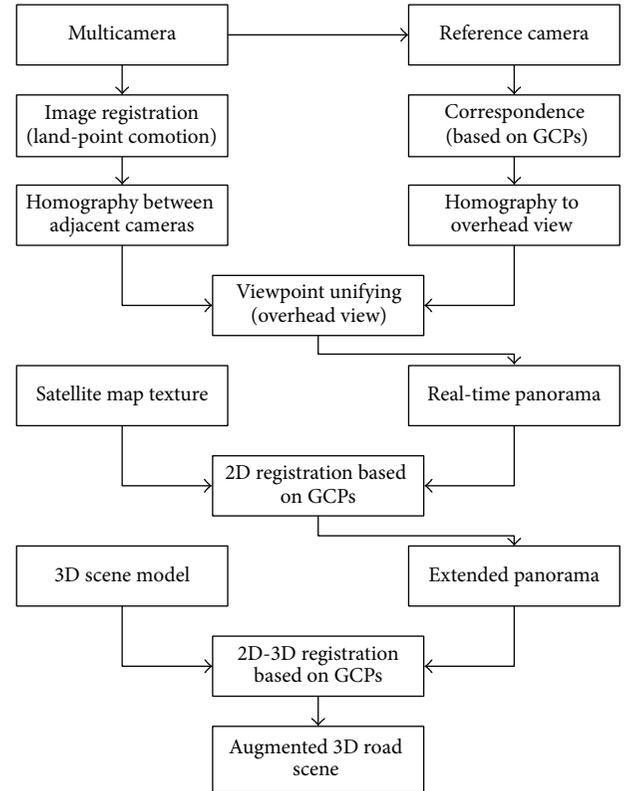


FIGURE 1: Flowchart of 3D road scene monitoring based on real-time panorama.

If there are more than one x_j meeting

$$x_j = \min_j \left\{ \text{median}_{P_i \in O} (D(P_i, l_j)) \right\}, \quad (5)$$

then x_{mid} is the average of these quantities.

- (6) Get land-point $(x_{\text{mid}}, y_{\text{max}})$.

The land-point of moving blobs is illustrated in Figure 2.

2.2. Matching of Land-Points Based on Comotion Method.

In our approach no prior knowledge is needed, and the method also works well in images of randomly scrambled motion. It is difficult to realize the registration among images with different position and angle of view. In this paper, the road surface served the reference plane so as to solve the image registration with comotion statistics maps [4, 5]. The hybrid Gauss background model with a fine performance in shadow removal was constructed to detect the foreground of camera video. For any pixel at t th frame, a binary variable $\text{seq}_\Lambda(i, j, t)$ indicates whether the pixel belongs to foreground. Continuous $\text{seq}_\Lambda(i, j, t)$ ($1 \leq t \leq T$) construct sequence vector $\text{seq}_\Lambda(i, j)$ [6],

$$\text{seq}_\Lambda(i, j) = [\text{seq}_\Lambda(i, j, 1), \dots, \text{seq}_\Lambda(i, j, T)], \quad (6)$$

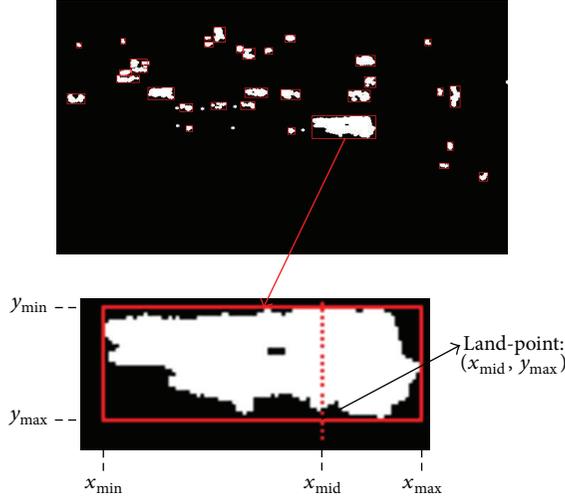


FIGURE 2: Land-point of moving blob.

where

$$\text{seq}_\Lambda(i, j, t) = \begin{cases} 0, & \text{if } (i, j) \text{ is background at } t\text{th frame} \\ 1, & \text{if } (i, j) \text{ is foreground at } t\text{th frame,} \end{cases}$$

$$1 \leq i \leq W_\Lambda, 1 \leq j \leq H_\Lambda, \Lambda \in \{A, B\}. \quad (7)$$

The aforementioned is illustrated in (6). Among them, subscript B is a reference image and A is a matched one. Let W and H denote the width and height of image, respectively. And T is the length of the video to participate in the registration.

Hence coded sequence seq can be divided into two categories: valid motion point and noise point with almost no or frequent movement. Let t_s be static point threshold and let t_n be the noise point threshold. Subsequently characteristic function $\alpha_\Lambda(i, j)$ is defined as follows:

$$\alpha_\Lambda(i, j) = \begin{cases} 0, & \text{if } \|\text{seq}_\Lambda(i, j)\|_1 < t_s \text{ or} \\ & \|\text{diff}(\text{seq}_\Lambda(i, j))\|_1 > t_n \\ 1, & \text{else,} \end{cases} \quad (8)$$

where $\|\cdot\|_1$ is the L_1 -norm and $\text{diff}(\cdot)$ is the difference of vector, for a vector \mathbf{v} is $[v(2) - v(1), \dots, v(n) - v(n-1)]$.

Point (i, j) will be removed from candidate matching pairs if $\alpha_\Lambda(i, j) = 0$. Here similarity measure of the binary sequence is depicted as L_1 -norm. Thus from (8) perfect matching objective function can be written as follows:

$$\text{map}(A(i, j)) = \arg \min_{(p, q) \in U_B} \|\text{seq}_A(i, j) - \text{seq}_B(p, q)\|_1, \quad (9)$$

$$(i, j) \in U_A, (p, q) \in U_B,$$

with

$$U_\Lambda = \{(i, j) \mid \alpha_\Lambda(i, j) = 1, 1 \leq i \leq W_\Lambda, 1 \leq j \leq H_\Lambda\},$$

$$\Lambda \in \{A, B\}. \quad (10)$$

In general, pixel (i, j) which is valid moving point in image A based on (9) matches the moving point (p^*, q^*) in image B if and only if the L_1 -norm of $\text{seq}_A(i, j) - \text{seq}_B(p^*, q^*)$ is minimum.

In (9) the nearest neighbor searching is implemented with ANN (approximate nearest neighbor) library [14] developed by Mount and Arya. Define the dimension of searching space as T and Minkowski norm to be L_1 -norm. ANN supports kd-trees and box-decomposition data structures, which greatly improves the efficiency of high dimensional feature matching.

2.3. *Homography between Adjacent Cameras.* According to the camera model, the relationship between world coordinate and image pixel coordinate is

$$Z_c \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \begin{bmatrix} \alpha_x & 0 & u_0 \\ 0 & \alpha_y & v_0 \\ 0 & 0 & 1 \end{bmatrix} [\mathbf{R} \mid \mathbf{t}] \begin{bmatrix} X_w \\ Y_w \\ Z_w \\ 1 \end{bmatrix} \quad (11)$$

$$= \begin{bmatrix} m_{11} & m_{12} & m_{13} & m_{14} \\ m_{21} & m_{22} & m_{23} & m_{24} \\ m_{31} & m_{32} & m_{33} & m_{34} \end{bmatrix} \begin{bmatrix} X_w \\ Y_w \\ Z_w \\ 1 \end{bmatrix},$$

where α_x , α_y , u_0 , and v_0 are called the internal camera parameters, \mathbf{R} and \mathbf{t} are called the external parameters, and Z_c is depth coefficient.

When the world point falls in the $z = 0$ plane, (11) is cast as

$$Z_c \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \begin{bmatrix} m_{11} & m_{12} & m_{13} \\ m_{21} & m_{22} & m_{23} \\ m_{31} & m_{32} & m_{33} \end{bmatrix} \begin{bmatrix} x_w \\ y_w \\ 1 \end{bmatrix}. \quad (12)$$

Customarily (12) is rewritten as

$$\lambda_i \begin{bmatrix} x'_i \\ y'_i \\ 1 \end{bmatrix} = \mathbf{H} \begin{bmatrix} x_i \\ y_i \\ 1 \end{bmatrix}, \quad (13)$$

where $\mathbf{H} = (h_{ij})_{3 \times 3}$ is called a homography.

Based on (9), we get the corresponding land-point pairs between images A and B to estimate 8 parameters of the projective transformation model; pixel (x'_i, y'_i) is the destination pixel in image B corresponding to the source pixel (x_i, y_i) in image A . Substitute $\{x_i, y_i\}$, $\{x'_i, y'_i\}$ in (13) for the land-point pairs; then the homography \mathbf{H} is computed such that

$$\sum_i \left[\left(x'_i - \frac{h_{11}x_i + h_{12}y_i + h_{13}}{h_{31}x_i + h_{32}y_i + h_{33}} \right)^2 + \left(y'_i - \frac{h_{21}x_i + h_{22}y_i + h_{23}}{h_{31}x_i + h_{32}y_i + h_{33}} \right)^2 \right]. \quad (14)$$

In many practical situations this assumption is not valid because land-points with noise are mismatched [15]. To improve the robustness of the transformation model, random sample consensus (RANSAC) procedure is applied to the estimation of homography matrix [16].

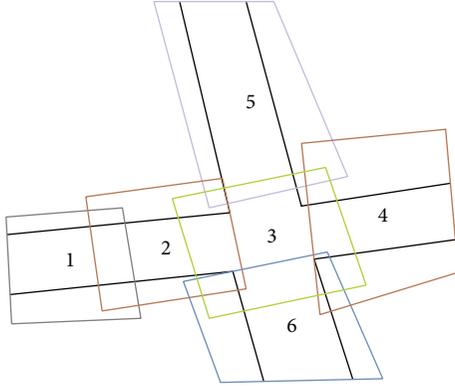


FIGURE 3: Visual field of each camera.

2.4. Projection of the Reference Camera Based on GCPs. Given the center of visual range of camera i C_i ($i = 1, 2, \dots$) and the center of the panorama C_0 , then the reference camera i^* is chosen via

$$i^* = \arg \min_i |C_i - C_0|. \quad (15)$$

As shown in Figure 3, camera 3 is the reference camera satisfied with (15).

The view of reference camera will be projected to the ground plane based on GCPs by means of measurement. That is to say, the view has been transformed to overhead view till now. For any quadrangle $ABCD$ on the ground, the intersection of edges AC and BD is the origin of the world coordinate and x -axis parallels AB pointing right. The coordinate system and length of each edge are as shown in Figure 4.

Given $|AB| = a$, $|BC| = b$, $|CD| = c$, $|DA| = d$, $|BD| = e$, and $|AC| = f$, according to simple geometric relations, the vertices of quadrilateral $ABCD$ are as below [17]:

$$\begin{aligned} x_A &= -\frac{\alpha \tan \beta}{\tan \alpha + \tan \beta}, \\ y_A &= -x_A \tan \alpha, \\ x_B &= a - x_A, \\ y_B &= y_A, \\ x_C &= x_B - b \cos \angle B, \\ y_C &= y_B - b \sin \angle B, \\ x_D &= x_A + d \cos \angle A, \\ y_D &= y_A - d \sin \angle A. \end{aligned} \quad (16)$$

In engineering application, so as to get better visual effect, similarity transformation is necessary:

$$x'_i = s_x x_i + \Delta x, \quad y'_i = s_y y_i + \Delta y, \quad i \in \{A, B, C, D\}, \quad (17)$$

where proper translation Δx and Δy makes the coordinate nonnegative, and scaling coefficient determines the height from viewpoint to the ground.

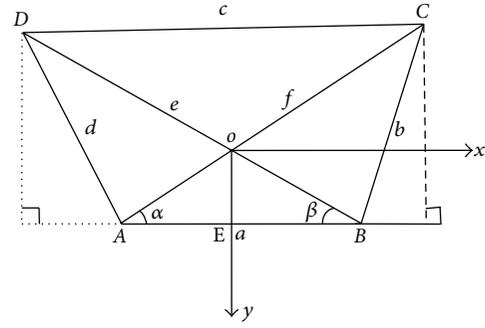


FIGURE 4: World coordinate system.

TABLE I: Parameters of GPS.

Item	Property value
Heading	0.05 degrees (4 m baseline)
Pitch	0.10 degrees (4 m baseline)
Position accuracy	0.3 m
Speed accuracy	0.02 m/s

Suppose the coordinates of 4 points from (16), (17), and the corresponding image points substitute $\{x'_i, y'_i\}$ and $\{x_i, y_i\}$ in (13); then the view of reference camera has been transformed to overhead view.

2.5. Viewpoints Unifying. Now suppose there are three cameras. Let G_i be the homography transforming the current viewpoint to overhead one. Suppose camera 2 is the reference camera; then

$$\begin{aligned} G_1 &\cong G_2 H_{12}, & H_{12} &\cong H_{21}^{-1}, \\ G_3 &\cong G_2 H_{32}, & H_{32} &\cong H_{23}^{-1}, \end{aligned} \quad (18)$$

where H_{21} and H_{23} are obtained in Section 2.3 and G_2 in Section 2.4.

3. 3D Road Scene Based on Multiview Panorama

2D multiview panoramas from multiple cameras installed top near the road were generated in real time. 2D panoramic road texture, satellite map texture, and 3D scene model were fused. The process is referred to as 2D-3D image registration. The whole data architecture has been shown in Figure 5.

3.1. Registration of Panorama and Satellite Map Texture. GPS geodetic coordinate transformation was introduced in [18, 19]. Suppose point of panorama corresponds to GPS geodetic coordinate and the position precision (Table 1) meets the demand of measurement.

The integrated navigation system XW-ADU5630 was used to obtain the GPS information of panorama's key-points, and the pixels of satellite map texture corresponding to the key-points also have GPS information. Therefore it is easy to

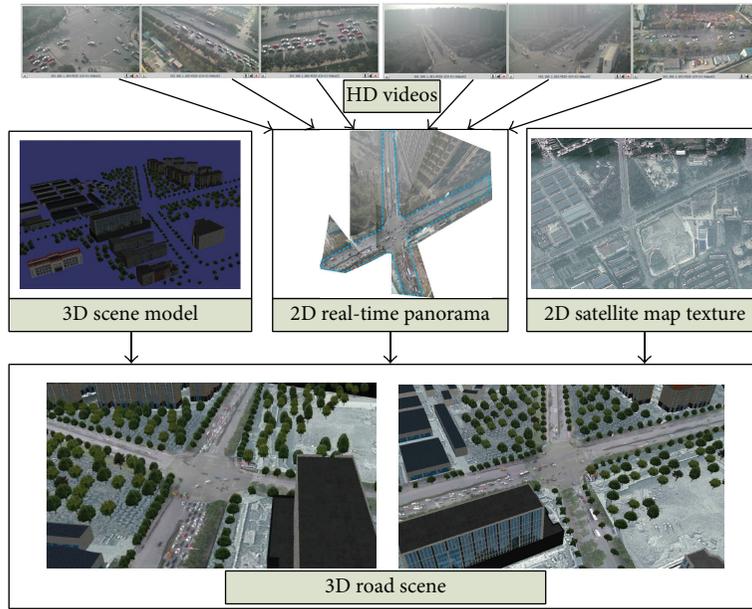


FIGURE 5: Data architecture of 2D and 3D image registration.

match the panorama with satellite map texture based on GPS information.

Applying the set of correspondences (GCPs) to (13), satellite map texture is registered to panorama. Therefore fusion of panorama and satellite map texture can be easily achieved. Panorama with satellite map texture is called extended panorama.

3.2. 2D-3D Image Registration. In our system, all the views are projected to the common road plane, which is the region of interest (ROI) of extended panorama. As a result there is much perspective deformation in top buildings above the ground. Therefore valid road area needs to be clipped out and registered to 3D model (Figure 12). Therefore observers can toggle 3D rendering engine to view the scene freely.

When we align the points on ground in 3D model and corresponding points in extended panorama, 2D-3D image registration is completed. A similarity transformation (or more simply similarity) is conducted so as to align the 3D model and 2D panorama. The similarity can be written more concisely in block form as

$$\begin{bmatrix} x' \\ y' \\ z' \\ 1 \end{bmatrix} \cong \begin{bmatrix} s\mathbf{R} & \mathbf{t} \\ \mathbf{0}^T & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix}. \quad (19)$$

A similarity in (19) has seven degrees of freedom and can be computed from four point pairs, that is, four corresponding GCPs. The GCPs are manually picked from ground plane of 3D scene model and 2D extended panorama.

4. Results and Discussion

Our experiments are conducted on the workstation with an Intel Xeon E5-2650 CPU and two graphic cards (NVIDIA GTX Titan Black). Six Samsung web cameras with 2.0-Megapixel are set up at the top of a building next to the road, monitoring a large crossroad. In the experiment, six cameras with three types of focal length (4 mm, 6 mm, and 8 mm) were used.

4.1. Abridged General View of Generation of Extended Panorama

(1) *Motion Detection.* The hybrid Gauss background model was applied to detect moving blobs (foreground). The results are demonstrated in Figure 6. The moving blobs, especially cars, have high accuracy to generate nice matching pairs.

After comotion process, land-points comotion statistics map [9] is shown in Figure 7. The higher the statistics value at a given position the brighter the corresponding pixel in the image.

(2) *Overhead View.* Four points easily to be measured on the ground were chosen, and the corresponding image coordinates were picked manually as shown in Figure 8.

(3) *Panorama and Satellite Map Texture Registration.* Figure 9 is the result of merging ROI of real-time traffic panorama and 2D satellite map texture. It is shown in Figure 9 on the left. On the right there are 6-channel real HD video images.

(4) *Viewpoint Roam.* The 3D model is driven based on OSG. Users can observe the road at any viewpoint. In Figure 10, we



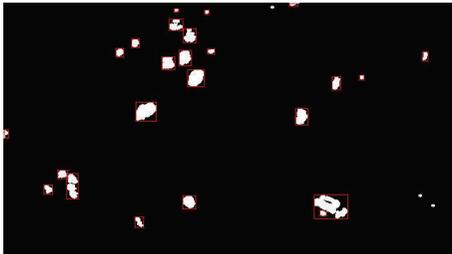
(a)



(b)



(c)



(d)

FIGURE 6: Moving blobs detection: (a) visible image from camera 2; (b) the moving blobs of Figure (a); (c) visible image from camera 3; (d) the moving blobs of Figure (c).

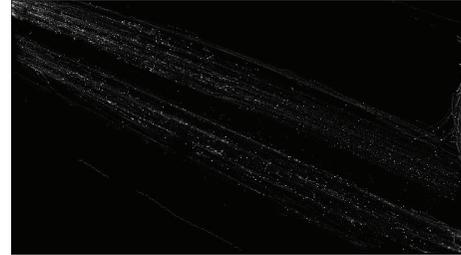
give two screenshots from two viewpoints. It helps observers to inspect the traffic situation in any direction and react promptly in emergency.

4.2. Experiments on Other Roads. Figure 11 shows the experiments on other two roads based on the method proposed. One is the panorama of civic center from 6-channel videos. The other is the panorama of a plaza from 2-channel videos. The polygon region in red dotted lines is mosaicked image.

4.3. Comparative Experiments. A multisource images (visible image, satellite map texture, and 3D scene model) fusion



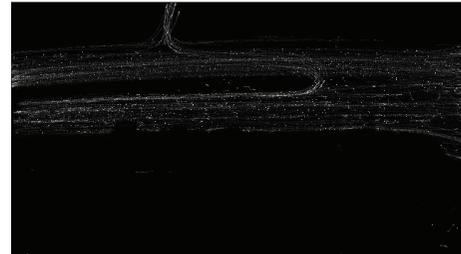
(a)



(b)



(c)



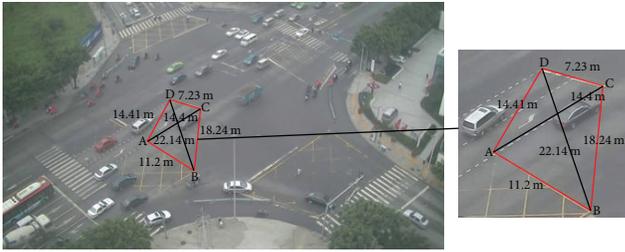
(d)

FIGURE 7: Comotion statistics maps for land-points: (a) visible image from camera 1; (b) comotion statistics map of camera 1; (c) visible image from camera 2; (d) comotion statistics map of camera 2.

method was proposed. By applying the method in [6], we have found the regions in the neighborhood of tall buildings have much projective deformation, as shown in Figure 12. The polygon region in red dotted lines is mosaicked image from real-time video images. And the polygon region in sky blue dotted lines is ROI of panorama, which is clipped out in our method as shown in Figure 12.

5. Conclusion

This paper implements a reasonable integration of 2D multiview overhead panorama, 2D satellite map texture, and



(a)



(b)

FIGURE 8: Overhead view: (a) visible image from camera 3; (b) overhead view of Figure (a).

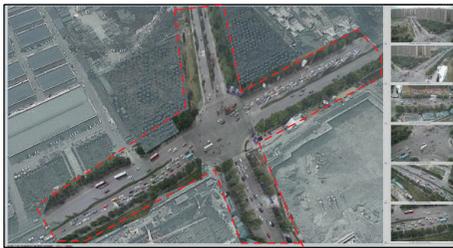
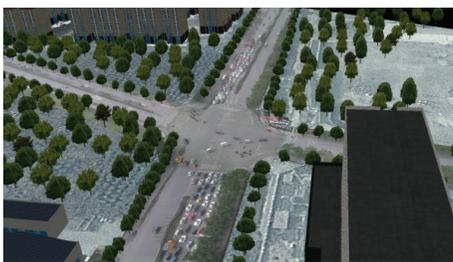


FIGURE 9: Fusion of multiview panorama and satellite map texture.



(a)

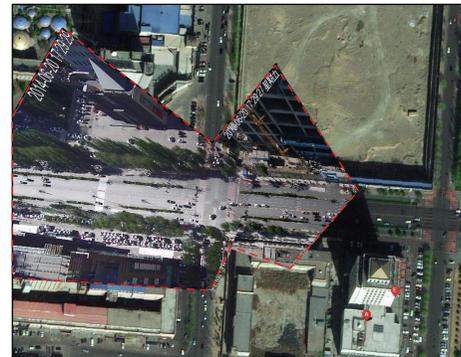


(b)

FIGURE 10: Effect of augmented reality 3D road scene: (a) from viewpoint 1, (b) from viewpoint 2.



(a)



(b)

FIGURE 11: Mosaicked images of two arterial roads.



FIGURE 12: Mosaicked images in [6].

3D model to synthesize 3D road scene. The 3D road scene monitoring system has been successfully applied to traffic supervision. The fusion method is innovative and enables the observers to have the sense of immersion. Besides, 3D road scene OSG driven can be translated, rotated, and scaled as one wishes.

In this paper, we focused mainly on the generation of 2D panorama and the presentation of 3D road scene. In future work, we will research on intelligent video analysis and traffic situation analysis.

Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

Acknowledgments

This work was supported in part by the National High Technology Research and Development Program (863 Program) of China under Grants 2012AA011804-1 and 2013AA013802, in part by the Open Research Fund of Key Laboratory of Higher Education of Sichuan Province through the Enterprise Informationalization and Internet of Things under Grant 2013WZJ01, and in part by the Scientific Research Project of Sichuan University of Science and Engineering under Grants 2014PY08 and 2014RC02.

References

- [1] C.-H. Chang, Y. Sato, and Y.-Y. Chuang, "Shape preserving half projective warps for image stitching," June 2014.
- [2] R. Hartley and A. Zisserman, *Multiple view geometry in computer vision*, Cambridge University Press, 2nd edition, 2003.
- [3] J. Zaragoza, T.-J. Chin, M. S. Brown, and D. Suter, "As-projective-as-possible image stitching with moving dlt," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR '13)*, pp. 2339–2346, IEEE, 2013.
- [4] R. Ai, Z. Shi, D. Xu, and C. Zhang, "A line mapping based automatic registration algorithm of infrared and visible images," in *Proceedings of the 5th International Symposium on Photoelectronic Detection and Imaging (ISPDI '13)*, p. 89072, International Society for Optics and Photonics, 2013.
- [5] X. W. Zhang, Y. N. Zhang, T. Yang, X. G. Zhang, and D. P. Shao, "Automatic visual-thermal image sequence registration based on co-motion," *Acta Automatica Sinica*, vol. 36, no. 9, pp. 1220–1231, 2010.
- [6] L. Lee, R. Romano, and G. Stein, "Monitoring activities from multiple video streams: establishing a common coordinate frame," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 22, no. 8, pp. 758–767, 2000.
- [7] M. Hu, J. Lou, W. Hu, and T. Tan, "Multi-camera correspondence based on principal axis of human body," in *Proceedings of the International Conference on Image Processing (ICIP '04)*, pp. 1057–1060, October 2004.
- [8] Z. Szlávik, L. Havasi, and T. Szirányi, "Estimation of common groundplane based on co-motion statistics," in *Image Analysis and Recognition*, vol. 3212 of *Lecture Notes in Computer Science*, pp. 347–354, Springer, 2004.
- [9] Z. Szlávik, T. Szirányi, and L. Havasi, "Video camera registration using accumulated co-motion maps," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 61, no. 5, pp. 298–306, 2007.
- [10] T. Schnell, Y. Kwon, S. Merchant, and T. Etherington, "Improved flight technical performance in flight decks equipped with synthetic vision information system displays," *International Journal of Aviation Psychology*, vol. 14, no. 1, pp. 79–102, 2004.
- [11] A. L. Alexander, C. D. Wickens, and T. J. Hardy, "Synthetic vision systems: the effects of guidance symbology, display size, and field of view," *Human Factors*, vol. 47, no. 4, pp. 693–707, 2005.
- [12] C. J. Liu, Y. Zhang, K. K. Tan, and H. Y. Yang, "Sensor fusion method for horizon detection from an aircraft in low visibility conditions," *IEEE Transactions On Instrumentation and Measurement*, vol. 63, no. 3, pp. 620–627, 2014.
- [13] C. J. Liu, Q. J. Zhao, Y. Zhang, and K. K. Tan, "Runway extraction in low visibility conditions based on sensor fusion method," *IEEE Sensors Journal*, vol. 14, no. 6, pp. 1980–1987, 2014.
- [14] D. M. Mount and S. Arya, ANN: library for approximate nearest neighbour searching, 1998.
- [15] M. Brown and D. G. Lowe, "Automatic panoramic image stitching using invariant features," *International Journal of Computer Vision*, vol. 74, no. 1, pp. 59–73, 2007.
- [16] S. Choi, T. Kim, and W. Yu, "Performance evaluation of ransac family," *Journal of Computer Vision*, vol. 24, no. 3, pp. 271–300, 1997.
- [17] T. K. Koo and Y. B. Aw, "A three-dimensional visualization approach to traffic accident mapping," *Photogrammetric Engineering & Remote Sensing*, vol. 57, no. 7, pp. 921–925, 1991.
- [18] D. H. Maling, "Coordinate systems and map projections for gis," in *Geographical Information Systems: Principles and Applications*, pp. 135–146, John Wiley & Sons, 1991.
- [19] R. Reulke, S. Bauer, T. Dring, and F. Meysel, "Traffic surveillance using multi-camera detection and multi-target tracking," in *Image and Vision Computing New Zealand*, pp. 175–180, 2007.

Research Article

An Aeromagnetic Compensation Algorithm for Aircraft Based on Fuzzy Adaptive Kalman Filter

Yao Zhang, Yu-Xin Zhao, and Shuai Chang

College of Automation, Harbin Engineering University, Harbin 150001, China

Correspondence should be addressed to Yu-Xin Zhao; zhaoyuxin@hrbeu.edu.cn

Received 13 June 2014; Accepted 4 August 2014; Published 17 August 2014

Academic Editor: Weichao Sun

Copyright © 2014 Yao Zhang et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

In order to ensure the effectiveness of geomagnetic navigation, as the foundation, the precise measurement of geomagnetic field must be guaranteed; namely, aircraft aeromagnetic compensation is worthy of being further studied. In this paper, the classical aircraft aeromagnetic compensation algorithm based on Leliak Model is analyzed and an aircraft aeromagnetic compensation algorithm based on fuzzy adaptive Kalman filter is proposed, which is a new approach for aircraft to achieve aeromagnetic compensation. Simulation results show that it has better compensation performance without relying on the aircraft attitude.

1. Introduction

Measurement of geomagnetic field refers to obtaining the field strength of carrier location accurately and in a real-time manner, and to improve the accuracy of measurement, there are mainly two ways: on one hand, improving the measurement precision of the sensor and, on the other hand, overcoming the interference caused by carrier's own magnetic field [1]. At present, the research on geomagnetic navigation and measurement of geomagnetic field generally uses orthotropic three-axis magnetic sensor as the basic magnetic components, because of its unique advantages that it can measure the overall field strength under any angle conveniently and need not consider the aircraft attitude and local magnetic field vector in an ideal world [2]. However, in practical application, due to the limitation of processing and assembling accuracy level, low cost of three-axis magnetic sensor cannot ensure that three-axis is orthogonal strictly, and the electric performance of the three axis sensitivity cannot completely be symmetrical [3]. The method of eliminating the magnetic field interference is generally neglected in the current research, especially, the interference caused by carriers themselves, and this kind of interference which is relative to carrier attitude is called aeromagnetic interference [4]. Tolles and Lawson reported the existence of interference

when aircraft was engine-driven in 1950, and consider it as a kind of constant noise [5]. The amplitude of aeromagnetic interference is small, which is composed by constant magnetic interference, excitation magnetic interference, and eddy current interference; however, it cannot be ignored relative to the weak magnetic field, and cannot simply it into a white noise because different aircraft attitude will cause different amplitude [6]. Therefore, it is necessary to consider the noise as a kind of complex form.

Leliak had a further study on this problem and found a 16th-order linear model, called "Leliak Model", to achieve the aeromagnetic compensation, which is used as the standard method in practical application [7]. Many scholars have studied the "Leliak Model" and improved it on certain aspect [8–11]. Bickel analyzed the magnetic interference under the small signal compensation and proposed a calculating method which is proved to be effective [8]; however, this research has bigger limitations and it cannot work when the signal is not small enough. Vijay Gopal et al. had a research on the high resolution airborne magnetic surveys and paid more attention to the real time compensation, which has better measurability and high value of application [9], but when the aircraft's location is less than 1000 km, this method will lose precision more or less. Jia and EiKon proposed a new aircraft compensation system for magnetic

terrains, which is based on the Ontario mineral exploration program, and its research object is mainly about Ontario mineral [10]; Above all, it can be seen that, by using this method to solve other field's magnetic interference problems, it has shown limitations. Liu had a comprehensive study on aeromagnetic compensation technology including setting up interference model, proposing compensation algorithm, and providing measuring method in practical experiment [11]; however, there are some aspects to be improved, such that the measure accuracy is not high enough and the method is effective when the geomagnetic field strength can be considered as a constant only. Therefore, it is necessary to provide a new aeromagnetic compensation algorithm which has no limitation in signal's value, aircraft flying height, and the change rate of geomagnetic field strength.

Kalman filter has been widely used to forecast system state vectors and estimate system parameters since it was invented by Stanley Schmidt because of its high accuracy with strong ability of interference suppression [12]. Under the situation that the calculation errors are ignored, classical Kalman filter can obtain the state vector value accurately; however, this situation is too hard to be satisfied, and it will cause divergent phenomenon sometime [13]. To solve the problem above, use measurement information constantly online to estimate the statistical feature of measurement noise in order to get accurate estimates of system state variables when filtering. At present, there are a lot of adaptive filtering methods, and the combination of Kalman filtering and fuzzy logic method is an ideal approach [14]. Bai et al. proposed an adaptive fuzzy Kalman filter to estimate system state vector, which adjusts estimation results based on observing noise in every step [15]; however, this method will cause oscillation of the covariance matrix.

Due to the fact that the existing works on aeromagnetic compensation were all found to rely on the aircraft attitude angle, the accuracy and singularity will be influenced. Therefore, it is necessary to make researches on aeromagnetic compensation algorithm using Kalman filter, which has better compensation performance without relying on the aircraft attitude.

At present, the research on aeromagnetic compensation algorithm using Kalman filter is seldom. In this paper, we proposed a new way to achieve aircraft aeromagnetic compensation, which can avoid the classical method limitation of accuracy and singularity. The rest of this paper consists of four sections. Leliak Model is analyzed in Section 2, and its shortcomings are pointed out. Section 3 improves the Leliak and gives the aircraft aeromagnetic compensation algorithm based on Kalman filter and adaptive fuzzy Kalman filter. Section 4 validates the effectiveness of the proposed method. Finally, we conclude this paper in Section 5.

2. Problem Description and Preliminaries

The classical aircraft aeromagnetic compensation algorithm based on Leliak Model can be expressed as follows:

$$\Delta \vec{T} = \vec{T}_p + \vec{T}_i + \vec{T}_e, \quad (1)$$

where \vec{T}_p , \vec{T}_i , and \vec{T}_e are constant magnetic interference, excitation magnetic interference, and eddy current interference, and $\Delta \vec{T}$ is the total geomagnetic interference caused by carrier moving.

Constant magnetic interference is caused by permanent magnet and ferromagnetic materials magnetized permanently, which cannot change according to the carrier attitude and can be considered as a constant, and it can be divided as follows:

$$T_{pd} = T_{px} \cos X + T_{py} \cos Y + T_{pz} \cos Z, \quad (2)$$

where T_{pd} is the sensor effect caused by constant magnetic interference, T_{px} , T_{py} , and T_{pz} are the projection weight of T_{pd} in the carrier coordinate system, and X , Y , and Z are the angle between magnetic field direction and Ox , Oy , and Oz of carrier coordinate system.

Excitation magnetic interference is caused by soft iron material magnetized by earth's magnetic field temporarily, which is proportional with the earth's magnetic field, and its expression is

$$\begin{aligned} T_{id} = T & [(a_{11} \cos X + a_{12} \cos Y + a_{13} \cos Z) \cos X \\ & + (a_{21} \cos X + a_{22} \cos Y + a_{23} \cos Z) \cos Y \\ & + (a_{31} \cos X + a_{32} \cos Y + a_{33} \cos Z) \cos Z], \end{aligned} \quad (3)$$

where T_{id} is the sensor effect caused by excitation magnetic interference, T is magnetic field strength, and a_{ij} ($i, j = 1, 2, 3$) is the proportionality coefficient, which is only related to carrier's material structure and material and can be considered as constants.

Eddy current interference is caused by the changes of magnetic flux when the carrier's attitude is changing, and there is eddy current among the conductor materials, which will bring interference to the magnetic field. This kind of interference is relative to the change rate of magnetic field, and its expression is

$$\begin{aligned} T_{ed} = T & [(b_{11} - b_{22}) \cos X (-\sin X \dot{X}) + b_{12} \cos X (-\sin Y \dot{Y}) \\ & + b_{13} \cos X (-\sin Z \dot{Z})] \\ & + b_{21} \cos Y (-\sin X \dot{X}) + b_{23} \cos Y (-\sin Z \dot{Z}) \\ & + b_{32} \cos Z (-\sin Y \dot{Y}) + (b_{33} - b_{22}) \cos Z (-\sin Z \dot{Z}), \end{aligned} \quad (4)$$

where T_{ed} is the sensor effect caused by eddy current interference and b_{ij} ($i, j = 1, 2, 3$) is the proportionality coefficient, which is only related to carrier's material structure and material and can be considered as constants.

The schematic diagram of the Leliak Model is shown in Figure 1.

From (1), (2), (3), and (4), if X , Y , Z , and T are known, we can calculate the sum interference caused by carrier moving by obtaining 16 parameters including T_{px} , T_{py} , T_{pz} , $(a_{11} - a_{22})$, $(a_{33} - a_{22})$, $(a_{21} + a_{12})$, $(a_{31} + a_{13})$, $(a_{32} + a_{23})$, $(b_{11} - b_{22})$, $(b_{33} - b_{22})$, b_{12} , b_{13} , b_{21} , b_{23} , b_{31} , and b_{32} . Then, we can achieve

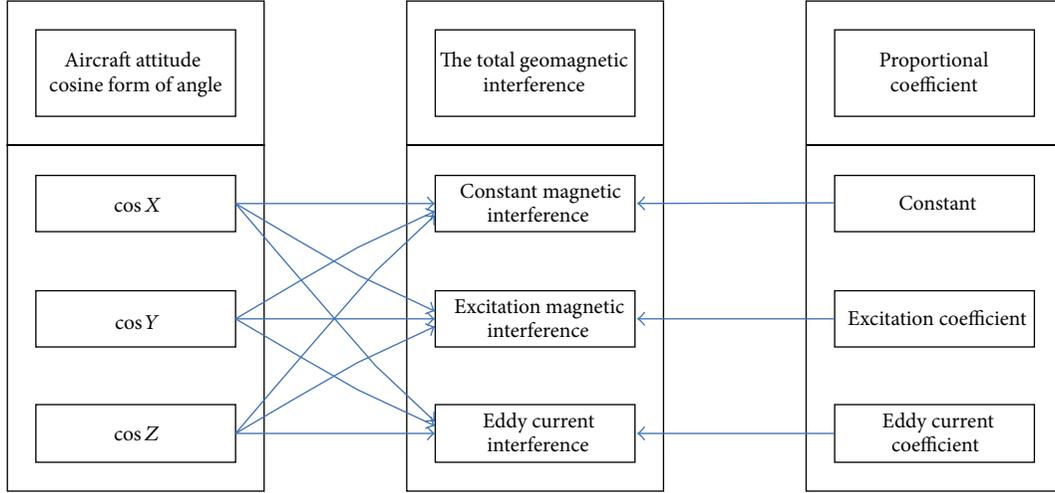


FIGURE 1: The schematic diagram of the Leliak Model.

the reverse compensation of the sensor measuring value according the calculation of interference value. In order to simplify the calculating process, Leliak Model uses reducing variable number method, whose specific approach is letting the aircraft do the single shaft motion to ignore the change rate of the other two attitudes.

Above all, classical aircraft aeromagnetic compensation algorithm based on Leliak Model is effective to achieve the aircraft aeromagnetic compensation; however, there is coupling among the aircraft three channel motions, which cannot be ignored, and it will influence the measure precision.

3. Aeromagnetic Compensation Algorithm Based on Fuzzy Adaptive Kalman Filter

3.1. Geomagnetic Interference Model Improving. In order to take the coupling among the aircraft three channels into consideration, we can establish the mathematical model in the carrier coordinate system according to the interference magnetic field generation mechanism in the following form [16]:

$$T_m = T + T_p + AT + B\dot{T}, \quad (5)$$

where $T_m = [T_{mx} \ T_{my} \ T_{mz}]^T$ is the sensor measurement, $T = [T_x \ T_y \ T_z]^T$ is the actual magnetic field strength, $T_p = [T_{px} \ T_{py} \ T_{pz}]^T$ is the constant magnetic interference, and $A = \{a_{ij}\}_{3 \times 3}$ ($i, j = 1, 2, 3$) and $B = \{b_{ij}\}_{3 \times 3}$ ($i, j = 1, 2, 3$) are the proportionality coefficient matrices.

We can obtain formula (6) by transferring carrier coordinate system to the geographic coordinate system, and $A(t)$ is the direction cosine matrix,

$$\begin{bmatrix} T_x \\ T_y \\ T_z \end{bmatrix} = A(t) \begin{bmatrix} T_{x0} \\ T_{y0} \\ T_{z0} \end{bmatrix}, \quad (6)$$

where T_{x0} , T_{y0} , and T_{z0} are the magnetic field strength of the shafts Ox , Oy , and Oz projection in the geographic

coordinate system. Combined with (6), (5) can be written in the following form:

$$\begin{bmatrix} T_{mx} \\ T_{my} \\ T_{mz} \end{bmatrix} = A(t) \begin{bmatrix} T_{x0} \\ T_{y0} \\ T_{z0} \end{bmatrix} + \begin{bmatrix} T_{px} \\ T_{py} \\ T_{pz} \end{bmatrix} + \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} A(t) \begin{bmatrix} T_{x0} \\ T_{y0} \\ T_{z0} \end{bmatrix} + \begin{bmatrix} b_{11} & b_{12} & b_{13} \\ b_{21} & b_{22} & b_{23} \\ b_{31} & b_{32} & b_{33} \end{bmatrix} \frac{dA(t)}{dt} \begin{bmatrix} T_{x0} \\ T_{y0} \\ T_{z0} \end{bmatrix}, \quad (7)$$

where

$$\frac{dA(t)}{dt} = \begin{bmatrix} 0 & \omega_z & -\omega_y \\ -\omega_z & 0 & \omega_x \\ \omega_y & -\omega_x & 0 \end{bmatrix} A(t), \quad (8)$$

ω_x , ω_y , and ω_z are the attitude angular velocity around the shafts Ox , Oy , and Oz in the inertial coordinate system.

3.2. Kalman Filter Model Establishing. From formula (7) above, it can be seen that the geomagnetic interference model can be written as

$$\begin{bmatrix} T_{mx} \\ T_{my} \\ T_{mz} \end{bmatrix} - \begin{bmatrix} T_x \\ T_y \\ T_z \end{bmatrix} = \begin{bmatrix} T_{px} \\ T_{py} \\ T_{pz} \end{bmatrix} + \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} \begin{bmatrix} T_x \\ T_y \\ T_z \end{bmatrix} + \begin{bmatrix} b_{11} & b_{12} & b_{13} \\ b_{21} & b_{22} & b_{23} \\ b_{31} & b_{32} & b_{33} \end{bmatrix} \begin{bmatrix} \dot{T}_x \\ \dot{T}_y \\ \dot{T}_z \end{bmatrix}; \quad (9)$$

its physical meaning is that the sum interference for sensor measurement consists of constant magnetic interference, excitation magnetic interference, and eddy current interference.

Let

$$x = [T_{px} \ T_{py} \ T_{pz} \ [a_{11} \cdots a_{33}]_{1 \times 9} \ [b_{11} \cdots b_{33}]_{1 \times 9}]^T,$$

$$y = [T_{mx} - T_x \ T_{my} - T_y \ T_{mz} - T_z]^T, \quad (10)$$

and formula (9) can be written as

$$y = H(t)x, \quad (11)$$

where $H(t)$ can be obtained as follows from formula (9):

$$H(t) = \begin{bmatrix} 1 & 0 & 0 & T_x & T_y & T_z & 0 & 0 & 0 & 0 & 0 & 0 & \dot{T}_x & \dot{T}_y & \dot{T}_z & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & T_x & T_y & T_z & 0 & 0 & 0 & 0 & 0 & 0 & \dot{T}_x & \dot{T}_y & \dot{T}_z & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & T_x & T_y & T_z & 0 & 0 & 0 & 0 & 0 & 0 & \dot{T}_x & \dot{T}_y & \dot{T}_z \end{bmatrix}. \quad (12)$$

Kalman filter can be expressed as formula (13) with formula (11) discretization as follows:

$$y(k) = H(k)x(k), \quad (13)$$

where $H(k)$ is in the following form:

$$H(k) = [\mathbf{E}_{3 \times 3} \ \mathbf{T}_{3 \times 9}(k) \ \mathbf{T}_{3 \times 9}(k+1)],$$

$$\mathbf{E}_{3 \times 3} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix},$$

$$\mathbf{T}_{3 \times 9}(k) = \begin{bmatrix} T_x(k) & T_y(k) & T_z(k) & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & T_x(k) & T_y(k) & T_z(k) & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & T_x(k) & T_y(k) & T_z(k) \end{bmatrix}, \quad (14)$$

$$\mathbf{T}_{3 \times 9}(k+1) = \begin{bmatrix} T_x(k+1) & T_y(k+1) & T_z(k+1) & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & T_x(k+1) & T_y(k+1) & T_z(k+1) & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & T_x(k+1) & T_y(k+1) & T_z(k+1) \end{bmatrix}.$$

It is obvious that $H(t)$ can be treated as the observe matrix of the Kalman filter, the state vector is $x(k)$, and the observe vector is $y(k)$.

Since the elements of state vector are constants, the state matrix of the Kalman filter is unit matrix; namely, the state equation is

$$x(k) = Ex(k-1) + W(k), \quad (15)$$

where E is an unit matrix, $u(k)$ is controller, and $W(k)$ is the process noise, which is white Gaussian noise and its covariance is Q .

Take observe error into consideration, the observe equation is

$$y(k) = H(k)x(k) + V(k), \quad (16)$$

where $V(k)$ is the observe noise, which is white Gaussian noise and its covariance is R .

Above all, Kalman filter for aeromagnetic compensation has been established which is expressed by state equation (15) and observe equation (16).

3.3. Aeromagnetic Compensation Algorithm Based on Kalman Filter. Aeromagnetic compensation algorithm proposed here uses Kalman filter method to solve the problem that classical aeromagnetic compensation algorithm cannot ensure the geomagnetic interference being measured accurately [17, 18], and the aeromagnetic compensation algorithm based on Kalman filter is introduced in this section. For a linear stochastic differential system, if process noise and observe noise are both white Gaussian noise, Kalman filter is the optimal method to estimate state vector. Aeromagnetic compensation algorithm based on Kalman filter is shown as Figure 2.

Firstly, the state equation of Kalman filter is used to achieve the state predicting, and it is in the following form:

$$x(k | k-1) = \Phi x(k-1 | k-1) + Bu(k), \quad (17)$$

where $x(k | k-1)$ is the predicting result at the previous state, $x(k-1 | k-1)$ is the optimal result at the previous state, Φ is state transition matrix, which is treated as unit matrix, and $u(k)$ is the controller at current state, which is treated

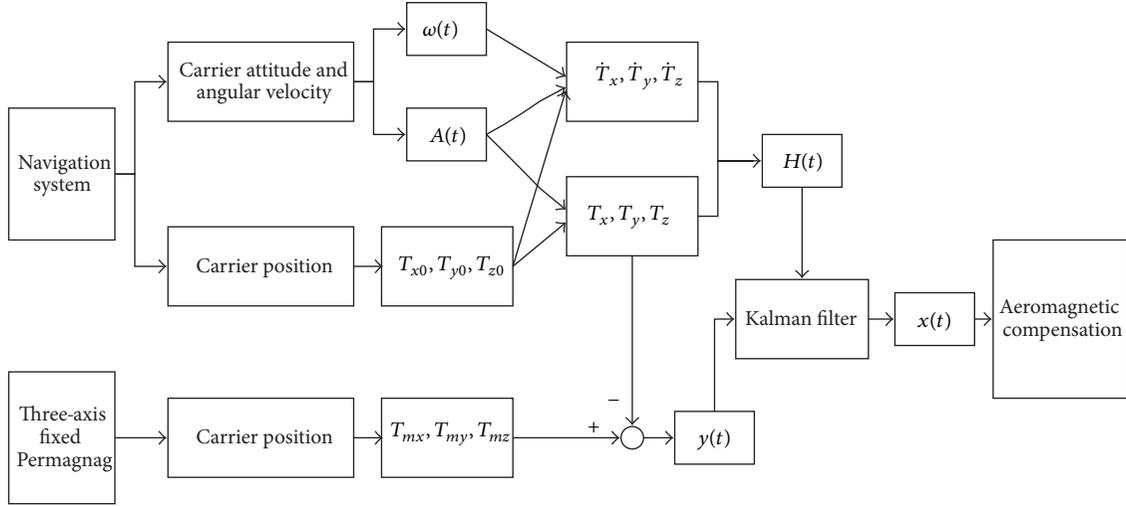


FIGURE 2: Aeromagnetic compensation algorithm based on Kalman filter.

as zero in aeromagnetic compensation algorithm combined with formula (15). Namely, formula (17) can be written as $x(k | k - 1) = x(k - 1 | k - 1)$. The current state has been updated by formula (17), but the covariance of $x(k | k - 1)$ has not been updated.

Secondly, we calculate the covariance as follows:

$$P(k | k - 1) = \Phi P(k - 1 | k - 1) \Phi^T + Q, \quad (18)$$

where $P(k | k - 1)$ is the covariance of $x(k | k - 1)$ and $P(k - 1 | k - 1)$ is the covariance of $x(k - 1 | k - 1)$. As $\Phi = E_{21 \times 21}$, formula (18) can be written as $P(k | k - 1) = P(k - 1 | k - 1) + Q$. Formula (17) and formula (18) bring out the predicting result of system state.

Thirdly, combined with the predicting results and measuring results, the optimal value of $x(k)$ can be obtained as

$$x(k | k) = x(k | k - 1) + K_g(k) (y(k) - H(k)x(k | k - 1)), \quad (19)$$

where $K_g = P(k | k - 1)H^T(k)/(H(k)P(k | k - 1)H^T(k) + R)$ is the Kalman gain.

Fourthly, in order to ensure that Kalman filter keeps working until system process is finished, the covariance of $x(k | k)$ must be updated by calculating the following formula:

$$P(kk) = (E - K_g(k)H(k))P(kk - 1). \quad (20)$$

By repeating iteration process above, we can get the estimated $y(k)$ when the filter converges under the situation that the initial $x(k)$ is small enough.

Above all, the aeromagnetic compensation algorithm estimates the elements of $y(k)$ and there is no need to ask aircraft to do single shaft motion with no errors during this process. After obtaining the elements of $y(k)$, we can bring them into aircraft magnetic field model to compensate.

However, there will exist divergent phenomenon if the noise cannot be estimated accurately or the aircraft geomagnetic interference is too big to measure.

3.4. Aeromagnetic Compensation Algorithm Based on Fuzzy Adaptive Kalman Filter. In order to ensure the effectiveness of kalman filter under the noise unmeasured, fuzzy adaptive method is considered to estimate the process noise and observe noise, which can avoid divergent phenomenon caused by noise uncertainty [19–23].

Using adaptive method to adjust process noise Q and observe noise R , we can get the following formulas:

$$\begin{aligned} Q(kk) &= s_a(b_a)Q(k - 1k - 1), \\ R(k | k) &= s_r(b_r)R(k - 1 | k - 1), \end{aligned} \quad (21)$$

where $s_a(b_a)$ is the adjustment parameter of Q , $s_r(b_r)$ is the adjustment parameter of R , and $b_a > 0$, $b_r > 0$ are constants standing for zoom level of Q and R , which are calculated by fuzzy inference system (FIS). Then, analysis of b_a and b_r is given as follows [24].

(i) Consider $b_a, b_r > 1$; magnify the effect on Q and R by $s_a(b_a)$ and $s_r(b_r)$. In this case, Q and R can approach the real process noise and observe noise in fewer steps. If the values of b_a, b_r are too big, the approaching rate is fast enough; however, it may cause small amplitude oscillation around the real value of noise.

(ii) Consider $0 < b_a, b_r < 1$; shrink the effect on Q and R by $s_a(b_a)$ and $s_r(b_r)$. In this case, there can be no oscillation and stability at the real noise value; however, it will cause more approach steps; namely, the adjustment time will be too long.

(iii) Consider $b_a, b_r = 0$; there is no effect on Q and R by $s_a(b_a)$ and $s_r(b_r)$. In this case, this method is equal to the method proposed in Section 3.3.

In order to establish the FIS of aeromagnetic compensation algorithm, performance indicators used to evaluate

estimation precision are introduced here with its definition as

$$\begin{aligned} ROR(k) &= \frac{T_r(c_r)}{T_r(p_r)}, \\ QOQ(k) &= \frac{T_r(c_a)}{T_r(p_a)}, \end{aligned} \quad (22)$$

where $T_r(\cdot)$ is the trace of the matrix, p_r and p_a are residual variance of process noise and observe noise, and c_r and c_a are variance of the real value of process noise and observe noise, and their definitions are as follows:

$$\begin{aligned} p_r &= H(k)(P(k-1|k-1) + Q(k|k))H^T(k) \\ &\quad + R(k-1|k-1), \end{aligned} \quad (23)$$

$$p_q = P(k-1|k-1) + R(k|k) + Q(k-1|k-1),$$

$$c_r = c_q = \frac{1}{M} \sum_{i=i_0}^k r_i r_i^T, \quad (24)$$

$$r_k = y(k) - H(k)x(k|k-1), \quad (25)$$

where r_k is residual error sequence.

From formulas (22) and (23), it can be seen that if the noise model is accurate enough, $ROR(k) = 1$, $QOQ(k) = 1$ can be established, and we use $R(k-1|k-1)$ to ensure $ROR(k)$ near to 1 and use $Q(k-1|k-1)$ to ensure $QOQ(k)$ near to 1.

Then, using FIS to calculate $s_a(b_a)$ and $s_r(b_r)$ with the following FIS rules.

(i) For observe noise, the FIS rules are as follows.

If $ROR(k) \in \{|ROR(k)-1| \leq \varepsilon\}$, then $s_r(b_r) \in \{|s_r(b_r)-1| \leq \varepsilon\}$.

If $ROR(k) \in \{ROR(k) > 1 + \varepsilon\}$, then $s_r(b_r) \in \{s_r(b_r) > 1 + \varepsilon\}$.

If $ROR(k) \in \{ROR(k) < 1 - \varepsilon\}$, then $s_r(b_r) \in \{s_r(b_r) < 1 - \varepsilon\}$,

where $0 < \varepsilon < 1$ and ε stands for boundary range of observe noise FIS.

(ii) For process noise, the FIS rules are as follows.

If $QOQ(k) \in \{|QOQ(k) - 1| \leq \eta\}$, then $s_a(b_a) \in \{|s_a(b_a) - 1| \leq \eta\}$.

If $QOQ(k) \in \{QOQ(k) > 1 + \eta\}$, then $s_a(b_a) \in \{s_a(b_a) > 1 + \eta\}$.

If $QOQ(k) \in \{QOQ(k) < 1 - \eta\}$, then $s_a(b_a) \in \{s_a(b_a) < 1 - \eta\}$,

where $0 < \eta < 1$ and η stands for boundary range of process noise FIS.

The output of FIS can be plugged in formula (21) to obtain the present noise value of $R(k|k)$ and $Q(k|k)$, and we

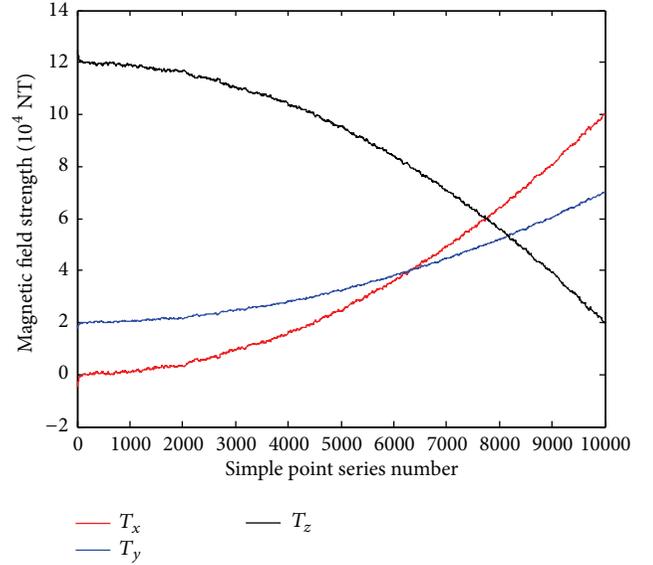


FIGURE 3: Magnetic field calculated by proposed aeromagnetic compensation algorithm.

achieve the process noise and observe noise estimation by FIS, which can avoid divergent phenomenon in theory.

4. Simulation and Analysis

In this section, experimental simulations will be carried out to evaluate the effectiveness of the proposed aeromagnetic compensation algorithm based on adaptive fuzzy Kalman filter.

Detailed parameters of this algorithm are presented as follows. We put the actual magnetic field data into adaptive Kalman filter proposed in this paper, and the observe noise and process noise are both white noise with 20 nT standard deviation in each axis. Initial condition is $Q(0) = 10^{-3} \times E_{21 \times 21}$, $R(0) = 2 \times 10^{-3} \times E_{21 \times 21}$, and $x(0) = [0.1 \ 0.1 \ 0.1 \ [10 \ \dots \ 10]_{1 \times 9} \ [15 \ \dots \ 15]_{1 \times 9}]^T$. System parameters are $\varepsilon = 0.05$ and $\eta = 0.025$. Simulation diagrams are as shown in Figures 3, 4, and 5.

From Figure 3, we can see that the calculated magnetic field strength conforms to the distribution law of actual magnetic field, and the 60 nT total noise is measured by Kalman filter. Figure 4 shows the Kalman filter output of system state vector elements, and it can be seen that these elements convergence to constants in a finite time, which stand for the compensation result calculated by Kalman filter. From Figure 5 above, it can be seen that the noise proposed algorithm estimated is nearly equal to the actual noise in each axis; namely, the proposed aircraft aeromagnetic compensation algorithm can calculate magnetic field effectively. Furthermore, the performance indicators used to evaluate estimation precision $ROR(k)$ and $QOQ(k)$ are around 1, which is to say that the adaptive FIS is effective in adjusting $s_a(b_a)$ and $s_r(b_r)$; namely, the variance of process noise and observe noise is estimated accurately.

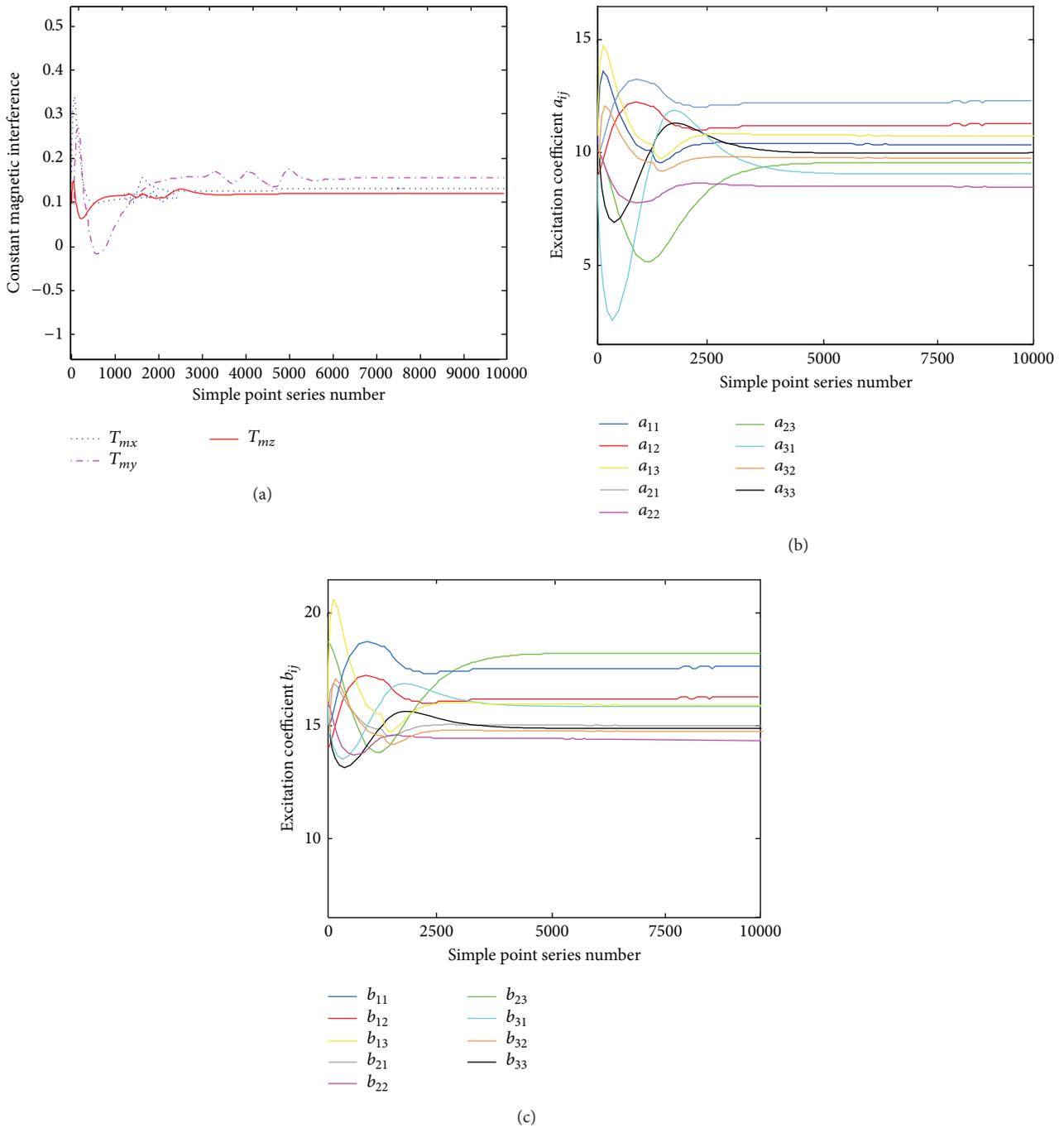


FIGURE 4: Kalman filter output of system state vector elements.

5. Conclusions

The limitations of classical aircraft aeromagnetic compensation algorithm based on Leliak Model have been analyzed in this paper and an aircraft aeromagnetic compensation algorithm based on fuzzy adaptive Kalman filter was proposed, which provided a new approach for aircraft to achieve aeromagnetic compensation. Simulation results confirmed the compensation performance of proposed algorithm and

proved that FIS is effective in estimating process noise and observe noise. Furthermore, this algorithm does not rely on the aircraft attitude, which has high application value.

Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

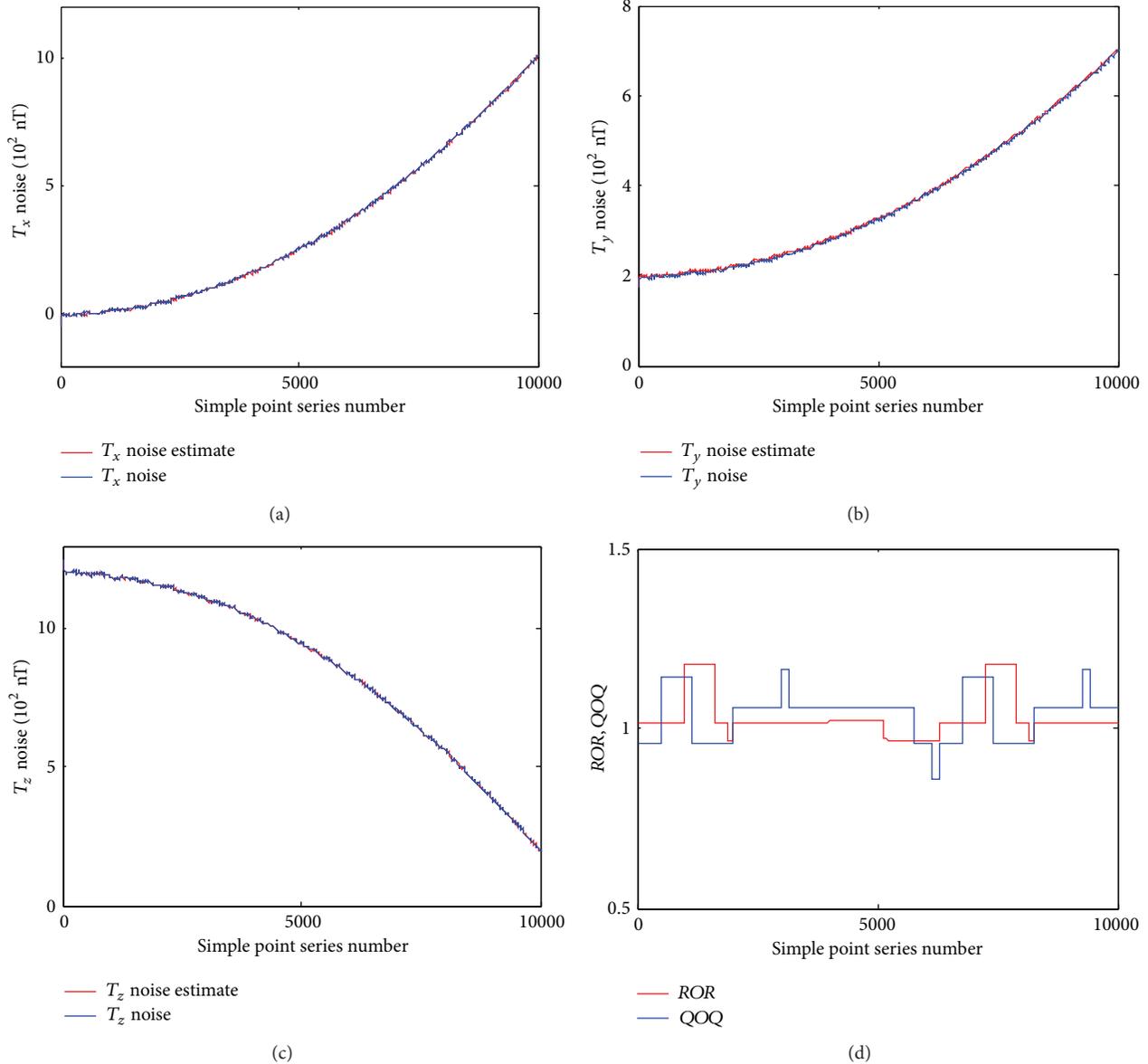


FIGURE 5: Compensation results of three-axis magnetic field and ROR, QOQ.

Acknowledgments

This work was supported by the National Natural Science Foundation of China (no. 51379049) and the Fundamental Research Funds for the Central Universities of China (nos. HEUCF110419 and HEUCFX41302).

References

- [1] M. P. Wu, Y. Liu, and X. P. Hu, "ICP algorithm application in the geomagnetic aided navigation," *Aerospace Control*, vol. 25, no. 6, pp. 17–21, 2007.
- [2] Y. R. You, "Rapid development of magnetic sensor," *Sensor Technology*, vol. 18, no. 4, pp. 5–8, 1999.
- [3] C. S. Lin, Q. Xiang, and S. G. Gong, "Three axis magnetometer orthogonal analysis and correction," *Journal of Detection & Control*, vol. 27, no. 2, pp. 9–12, 2005.
- [4] D. H. Wu, "Three axis flux-gate sensor error correction based on SVR," *Sensor & Micro System*, vol. 27, no. 6, pp. 43–46, 2008.
- [5] J. L. He, "Aircraft the automatic compensation method of the magnetic field," *Geophysical and Geochemical Exploration*, vol. 9, no. 6, pp. 464–469, 1993.
- [6] X. L. Li and W. L. Cai, "Airlines plane magnetic gradient measurement system installation and compensation," *Geophysical and Geochemical Exploration*, vol. 30, no. 3, pp. 224–228, 2006.
- [7] P. Leliak, "Identification and evaluation of magnetic field sources of magnetic airborne detector equipped aircraft," *IRE Transactions on Aerospace and Navigational Electronics*, vol. 8, pp. 95–105, 1961.
- [8] S. H. Bickel, "Small signal compensation of magnetic fields resulting from aircraft maneuvers," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 15, no. 4, pp. 518–525, 1979.

- [9] B. Vijay Gopal, V. N. Sarma, and H. V. Rambabu, "Real time compensation for aircraft induced noise during high resolution airborne magnetic surveys," *Journal of Indian Geophysical Union*, vol. 8, no. 3, pp. 185–189, 2004.
- [10] J. Jia and P. Eikon, "A New Aircraft Compensation System for Magnetic Terrains," pp. 23–50, Ontario Mineral Exploration Technology Program, Project #P02-03-043, 2007.
- [11] X. J. Liu, *Aeromagnetic Compensation Technology Research*, JiLin University, 2005.
- [12] P. J. Escamilla-Ambrosio and N. Mort, "Multi-sensor data fusion architecture based on adaptive Kalman filters and fuzzy logic performance assessment," in *Proceedings of the 5th International Conference on Information Fusion (FUSION '02)*, pp. 1542–1549, Annapolis, Md, USA, July 2002.
- [13] J. Z. Sasiak, Q. Wang, and M. B. Zeremba, "Fuzzy adaptive kalman filtering for INSPGPS data fusion," in *Proceedings of the 15th IEEE Intelligent Control*, pp. 181–186, Patras, Greece, July 2000.
- [14] X. R. Huang, *Development of Vehicle Integrated Navigation System Based on Information Fusion*, Harbin Institute of Technology, 2002.
- [15] J. Bai, J. Y. Liu, and X. Yuan, "Study of fuzzy adaptive kalman filtering technique," *Journal of Information and Control*, vol. 31, no. 3, pp. 193–197, 2002.
- [16] Y. Kou, H. Xia, R. Liu, and C. Wang, "Method for separation of low-frequency electromagnetic interference in geomagnetic navigation systems," *Journal of Harbin Institute of Technology*, vol. 43, no. 7, pp. 32–37, 2011.
- [17] W. Sun, Z. Zhao, and H. Gao, "Saturated adaptive robust control for active suspension systems," *IEEE Transactions on Industrial Electronics*, vol. 60, no. 9, pp. 3889–3896, 2013.
- [18] W. Sun, H. Gao, and B. Yao, "Adaptive robust vibration control of full-car active suspensions with electrohydraulic actuators," *IEEE Transactions on Control Systems Technology*, vol. 21, no. 6, pp. 2417–2422, 2013.
- [19] H. Zhang, X. Liu, J. Wang, and H. R. Karimi, "Robust H_∞ sliding mode control with pole placement for a fluid power electrohydraulic actuator (EHA) system," *The International Journal of Advanced Manufacturing Technology*, 2014.
- [20] H. Zhang and J. Wang, "Combined feedback-feedforward tracking control for networked control systems with probabilistic delays," *Journal of the Franklin Institute*, vol. 351, no. 6, pp. 3477–3489, 2014.
- [21] H. Zhang, X. Zhang, and J. Wang, "Robust gain-scheduling energy-to-peak control of vehicle lateral dynamics stabilisation," *Vehicle System Dynamics*, vol. 52, no. 3, pp. 309–340, 2014.
- [22] H. Zhang, Y. Shi, and J. Wang, "On energy-to-peak filtering for nonuniformly sampled nonlinear systems: a Markovian jump system approach," *IEEE Transactions on Fuzzy Systems*, vol. 22, no. 1, pp. 212–222, 2014.
- [23] H. Zhang, Y. Shi, and J. Wang, "Observer-based tracking controller design for networked predictive control systems with uncertain Markov delays," *International Journal of Control*, vol. 86, no. 10, pp. 1824–1836, 2013.
- [24] H. Zhang, Y. Shi, and M. Liu, " H_∞ step tracking control for networked discrete-time nonlinear systems with integral and predictive actions," *IEEE Transactions on Industrial Informatics*, vol. 9, no. 1, pp. 337–345, 2013.

Research Article

Research on Wavelet Singularity Detection Based Fault-Tolerant Federated Filtering Algorithm for INS/GPS/DVL Integrated Navigation System

Jianhua Cheng,^{1,2} Daidai Chen,¹ Rene Landry Jr.,² Lin Zhao,¹ and Dongxue Guan^{1,2}

¹ Marine Navigation Research Institute, College of Automation, Harbin Engineering University, Harbin 150001, China

² LASSENA Laboratory, École de Technologie Supérieure, Université du Québec, Montreal, QC, Canada H3C 1K3

Correspondence should be addressed to Daidai Chen; ins.dai@163.com

Received 4 April 2014; Revised 21 June 2014; Accepted 29 July 2014; Published 17 August 2014

Academic Editor: H. R. Karimi

Copyright © 2014 Jianhua Cheng et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Soft faults in navigation sensors will lead to the degradation of the accuracy and reliability of integrated navigation system. To solve this problem, a wavelet analysis and signal singularities based soft fault detection method are given out. To find signal singularities and detect the faults, the modulus maxima values are calculated after the wavelet transform of original signal. By calculating the Lipschitz exponent using the modulus maxima value at the fault point, the fault types are distinguished. Then, a fault-tolerant federated filtering algorithm for the calibration of INS/GPS/DVL integrated navigation system is proposed. Simulations are conducted and results show that sensor soft faults can be detected accurately. By effectively isolating the fault and refactoring information, the accuracy and reliability of navigation system are improved.

1. Introduction

Higher standards for the accuracy and reliability of navigation systems are required with the development of modern ship, aircraft, and land vehicle. Device redundancy and estimator based analytical redundancy are the two most effective methods of improving the accuracy and reliability of system [1].

Device redundancy appears to be a simple and efficient approach to improve the performance of navigation system [2]. Both system level and sensor level redundancy are considerable, and most of the system redundancies are in the form of integrated navigation system. And INS/GPS/DVL integrated navigation system is widely used in vehicles. Inertial navigation system (INS) has been usually selected as the core of integrated systems for its outstanding characters, such as high-accuracy, self-contained, and jam-proof. Meanwhile, global position system (GPS) and Doppler velocity log (DVL) commonly play the external auxiliary roles, because their navigation errors do not accumulate with time [3]. They also provide the position and velocity for the calibration and reset of INS. That means that the performance of GPS and DVL

will affect the calibration accuracy of INS directly as well as the entire INS/GPS/DVL integrated navigation system. As for the entire calibration process for INS, it is necessary to guarantee the accuracy and reliability of external information [4]. Otherwise, the gyroscope drifts cannot be estimated and the error of external information will be introduced.

Estimator based analytical redundancy method is another effective solution [5, 6]. The estimated information predicted by using system mathematical model is introduced, then conduct data fusion by estimator (i.e., Kalman filter) so that almost all outputs of different kinds of sensors can be fully used. However, the fault of sensors might happen, which must be detected and processed. Kalman filter based fault detection methods are commonly used in integrated navigation system, especially χ^2 detection method.

In χ^2 detection method, faults are detected by comparing the assumptive mean and variance with those of constructed n -dimension Gaussian variables, which is easy to realize. However, this method insensitive and large-delay for the detection of soft faults, and the type of fault cannot be accurately determined. In the integrated system, there are two most common sensor faults: (1) step error caused by clock

jumping and power fluctuation in GPS; (2) slope error caused by clock drift and the inaccuracy of rail parameter model [7].

To solve the problems above, a sensor fault detection method based on wavelet analysis and signal singularities is put forward, and a fault-tolerant federated filter for INS/GPS/DVL integrated navigation system is designed. In Section 2, signal fault detection method based on wavelet modulus maxima is introduced, and the relationships between signal singularity and modulus maxima are given out, and by searching modulus maxima the fault can be accurately positioned. In Section 3, fault type detection method based on the calculation of signal Lipschitz exponent is proposed. In Section 4, by utilizing the proposed fault detection method, a fault-tolerant federated filter for the calibration progress in INS/GPS/DVL integrated navigation system is designed. And in Section 5, simulations for fault-tolerant federated filter in calibration application are conducted. Finally, the conclusions are given in Section 6.

2. Signal Singularity Detection Method Based on Wavelet Analysis

2.1. Wavelet Transform and Modulus Maxima. Wavelet analysis is an efficient signal analysis tool and has been widely used in signal process [8, 9]. Wavelet transform appears to be extremely sensitive for the signal singularities, which makes it highly appropriate for the fault signal detection. In this transform, the original signal is processed by using wavelet as follows:

$$\psi_{s,u}(t) = \frac{1}{\sqrt{|s|}} \psi\left(\frac{t-u}{s}\right) \quad s, u \in \mathbf{R}, \quad s \neq 0, \quad (1)$$

where the $\psi(t)$ is mother wavelet, s is scale parameter, and u is translation parameter.

By (1), with the decrease of $|s|$, the locality and resolution in time domain of wavelet function are improved, and resolution in frequency domain is reduced.

Assuming that wavelet $\psi(t)$ and the sensor output signal $f(t) \in L^2(\mathbf{R})$ are both real functions, scale parameter $s > 0$. The continuous wavelet transform for $f(t)$ can be calculated as follows:

$$W_f(s, u) = s^{1/2} (f * \bar{\theta}_s)(u) = s^{1/2} \frac{d}{du} (f * \bar{\theta}_s)(u), \quad (2)$$

where “ $*$ ” denotes the convolution operation and $\bar{\theta}_s$ is the complex conjugate of θ_s .

As known from (3), under the determined scale s_0 , we could obtain the local maxima of $W_f(s, u)$ when the derivative $\partial W_f(s_0, u)/\partial u = 0$ at $u = u_0$. If any points u within the neighbourhood of u_0 satisfy equation $W_f(s_0, u) \leq W_f(s_0, u_0)$ and any points u within the left or right neighbourhood of u_0 satisfy equation $W_f(s_0, u) < W_f(s_0, u_0)$, we call the point (s_0, u_0) as the extremum of the wavelet transform modulus $W_f(s, u)$ on the scale s_0 and $|W_f(s_0, u_0)|$ as the modulus maxima for the wavelet transform for the original $f(t)$.

Due to the high-sampling ratio and nature discrete characters for the output signals of navigation systems, the

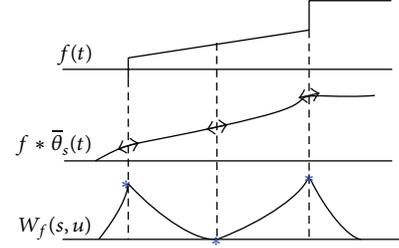


FIGURE 1: Signal singularities detection based on wavelet modulus maxima.

continuous wavelet transform leads to extremely heavy calculation burden. It seems impossible to implement continuous wavelet transform in practical signal processing. Therefore, binary wavelet transform is selected in this paper, which can be seen as a compromise between continuous and discrete wavelet transforms. And let $s = 2^j$; the binary wavelet can be expressed as

$$\psi_{2^j, u}(t) = 2^{-j/2} \psi \left[2^{-j} (t - u) \right] \quad j \in \mathbf{Z}, \quad u \in \mathbf{R}. \quad (3)$$

Accordingly, the binary wavelet transform is

$$W_f(2^j, u) = 2^{-j/2} \int_{-\infty}^{+\infty} f(t) \psi \left[2^{-j} \left(\frac{t-u}{s} \right) \right] dt. \quad (4)$$

Signal singularities, which are the most important local information of signals, can be inferred by searching the modulus maxima. By calculating the modulus maxima on different scales, the entire modulus maxima curve can be obtained, and then we can observe the local signal singularities. The relationships among the original signal and the modulus maxima are shown in Figure 1.

As shown in Figure 1, if the characteristics of signal changes, their corresponding wavelet modulus appears as maxima. This character can be utilized to detect and position the sensor faults.

2.2. Fault Positioning Based on Modulus Maxima Searching. To detect the singularity of fault signal, we have to obtain the curves of modulus maxima. And ad hoc algorithm is commonly used in searching the points of modulus maxima [10].

The specific steps of the algorithm are shown as follows.

- (1) Choose Morlet wavelet basis and scale 2^j and process the sensor output signal by using binary discrete wavelet transform. The chosen scale should guarantee that the signal extrema are dominated on the largest scale and keep all the important singularities of signal. In order to make the output signal meet the requirement of real time for the integrated navigation system, $j = 5$ is set.
- (2) Assuming that the largest amplitude of extrema is M on the largest scale, the extrema whose amplitudes are smaller than M/j should be eliminated, that is, because these extrema are mainly caused by noise.

- (3) For arbitrary extremum x_0 on scale 2^j , its neighbour points are x_1 and x_2 , and the manifold-points of x_1 are x'_1 should be in the interval $L = [\max(x_1, x'_1), x_2]$. Among the points (a_1, a_2, \dots, a_n) have the same sign with x_0 in the interval L , if a_k satisfies $|a_k - x_0| \leq |a_i - x_0|/3, i = 1, 2, \dots, n, k \neq i, a_k$ is the manifold-point of x_0 . Otherwise, the point with the largest amplitude is the manifold point of x_0 . However, once the amplitude of the manifold point is twice that of x_0 , this point will be treated as the extrema of noise and have to be eliminated [11].
- (4) Repeat the process above until the scale is 2^2 .
- (5) Eliminate the maximum points on scale 2^1 , and use the extrema on scale 2^2 . So far, the search for the modulus maxima on scales $2^j \sim 2^1$ is completed.

Through the steps above, the modulus maxima on each scale can be obtained as well as the modulus maxima curve. According to the modulus maxima on fine scale, we can accurately position the singular points and the moments of sensor faults that occurred. However, only with these wavelet modulus maxima, we still could not identify the types of sensor faults.

3. Fault Type Detection Based on Lipschitz Exponent

3.1. Lipschitz Exponent Characteristics of Signal Faults. Singularity refers to the signal with discontinuous points or its derivative function is not continuous. Mathematically, signal singularity is described by Lipschitz exponent [12]. The type of sensor faults could be identified based on Lipschitz exponent.

The Lipschitz exponent α of signal $f(t)$ at $t = h_0$ is defined as follows. There are two constants $M, h_0 (M > 0, h_0 > 0)$ and a polynomial $g_n(h)$. When $h < h_0$, inequation is established as

$$|f(h_0 + h) - g_n(h_0)| \leq M|h|^\alpha. \quad (5)$$

For any point h_0 , the polynomial $g_n(h)$ is uniquely determined. If $f(t)$ is $n(n = \lfloor \alpha \rfloor)$ order continuously differentiable at point h_0 and its neighborhood, the $g_n(h)$ is equal to the Taylor expansion of $f(t)$ at point h_0 . If function $f(t)$ has uniform Lipschitz exponent in any neighborhood of h_0 point and $\alpha < n$, $f(t)$ is n order differentiable in this neighborhood of point h_0 . Therefore, a great Lipschitz exponent α means better smoothness of signal and the smaller α characterizes the singularities of $f(t)$.

As for the most typical sensor errors, the relationships between slope error, step error, random noise error, and Lipschitz exponents are given out.

- (1) Assuming that $f(t)$ is a step signal, we have

$$|f(h_0 + h) - g_n(h_0)| \leq O|h|^0. \quad (6)$$

According to (5), the Lipschitz exponent α of step signal at h_0 point is 0.

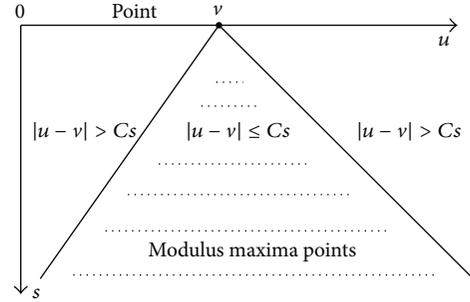


FIGURE 2: The conical surface composed of modulus maxima points.

- (2) Assuming that $f(t)$ is a slope signal, we have

$$|f(h_0 + h) - g_n(h_0)| \leq O|h|^1. \quad (7)$$

According to (5), the Lipschitz exponent α of slope signal at h_0 point is 1.

However, the Lipschitz exponents related to random noise are usually less than 0, because Gaussian white noise is a random distribution whose singularities are almost everywhere. On this condition, the Lipschitz exponents are negative and could be denoted as $\alpha = -0.5 - \epsilon, \forall \epsilon > 0$.

3.2. Calculation of Signal Lipschitz Exponent. By analyzing the relationship between the modulus maxima and signal singularity, it is found that the attenuation of $|W_f(s, u)|$ could be controlled by its local maxima. Therefore, according to the relationship between wavelet modulus maxima and signal Lipschitz exponent, the calculation of signal Lipschitz exponent can be realized [13].

Assuming that the compactly supported set of wavelet function ψ is $[-C, C] (C > 0)$ and when $s < s_0$, all the modulus maxima points which converge to the point v are located in the conical surface, which is denoted as $|u - v| \leq Cs$, as shown in Figure 2.

Assuming the signal $f(t) \in L^2(\mathbf{R})$ and its modulus maximum values distribute within a conical surface. That is

$$|W_f(s, u)| \leq Aa^{\alpha+0.5} \cdot \left| 1 + \left| \frac{t - t_0}{s} \right|^\alpha \right|, \quad (8)$$

where A is a constant.

$$|W_f(s, u)| \leq Aa^{\alpha+0.5}. \quad (9)$$

Calculating the logarithm for both sides of (9) and using binary wavelet transform, it can be obtained that

$$\log_2 |W_f(2^j, t_j)| \leq \log_2 A + j(\alpha + 0.5). \quad (10)$$

Equation (10) shows that the Lipschitz exponent value at point t_0 of $f(t)$ depends on the attenuation of $|W_f(s, u)|$ under fine scales. And when $t = t_0$, we get the following conclusions.

- (a) If Lipschitz exponent $\alpha > 0$, the modulus maxima is proportional to the scale.

- (b) If Lipschitz exponent $\alpha < 0$, the modulus maxima is inversely proportional to the scale.
- (c) If Lipschitz exponent $\alpha = 0$, the modulus maxima is uncorrelated with the scale.

And for the next scale $j + 1$, we have

$$\log_2 |W_f(2^{j+1}, t_{j+1})| \leq \log_2 A + (j + 1)(\alpha + 0.5). \quad (11)$$

Assuming that the equalities hold up in both (10) and (11), the formula for the calculation of the Lipschitz exponents is as follows:

$$\alpha = \log_2 |W_f(2^{j+1}, t_{j+1})| - \log_2 |W_f(2^j, t_j)| - 0.5. \quad (12)$$

Through the steps above, calculate the modulus maxima on every scale and draw the modulus maxima curve. According to the modulus maxima on fine scale, we can find the singular point position, namely, the moment of fault occurrence.

Select the modulus maxima point on every scale calculated in step (6) and the logarithm of decomposition scale as ordinate and abscissa, respectively. The slope of curve is estimated by linear least squares, and Lipschitz exponent can be obtained by subtracting 0.5 from the curve slope. According to the Lipschitz exponent, we can distinguish the fault type, and the fault information is provided for fault detection and isolation module to dispose of the sensors with faults.

4. Design for Fault-Tolerant Filtering Scheme

To maintain the high accuracy of INS/GPS/DVL integrated navigation system for long term operation, comprehensive calibration is one of the key technologies [14]. External information of INS such as position provided by GPS is introduced to estimate the gyroscope drifts and reset the navigation information. Usually, this progress might last for hours. Moreover, during the entire calibration phase, the accuracy of external information must be maintained at a stable level. However, the GPS signals appear to be easily influenced by dynamic environment or human disturbances.

In order to guarantee the accuracy and reliability of INS/GPS/DVL integrated navigation system, a fault detection method based fault-tolerant filtering scheme for the calibration phase is presented. In this scheme, decentralized filter is adopted; the fault detection method mentioned above is used to process the outputs of the two subintegrated systems INS/GPS and INS/DVL. And the diagram of this fault-tolerant filtering algorithm is given in Figure 3.

As shown in Figure 3, position measurement Z_{IG} is the position difference between GPS and INS outputs, and velocity measurement Z_{ID} is the velocity difference between DVL and INS outputs. Then, Z_{IG} and Z_{ID} are processed through the fault detection method based on binary wavelet transform and Lipschitz exponent. Once faults are detected, the faults can be quickly and accurately positioned and then the type of fault would be obtained. Consequently, the fault sensor would be isolated, and the corresponding output would be

reconstructed to form measurement information Z_{IG} and Z_{ID} which are sent to subfilters. That means that the integrated navigation system exhibits better fault-tolerant ability. With more precise external navigation information, the integrated system could normally output precise navigation information and even sensor soft fault occurs.

4.1. State and Measurement Equations for Subfilter. Considering that the outputs from the sensors are different, we use indirect method to estimate the navigation parameters. The state vector of INS is selected as

$$\mathbf{X}_I = [\delta\varphi_I \ \delta\lambda_I \ \delta V_x \ \delta V_y \ \alpha \ \beta \ \gamma \ \varepsilon_x \ \varepsilon_y \ \varepsilon_z \ \varepsilon_{rx} \ \varepsilon_{ry} \ \varepsilon_{rz}]^T, \quad (13)$$

where $\delta\varphi$, $\delta\lambda$ are the position errors of INS; δV_x , δV_y are the velocity errors of INS; α , β and γ are attitude errors of INS; ε_x , ε_y and ε_z are the gyro constant drifts; ε_{rx} , ε_{ry} and ε_{rz} are the gyro random drifts.

The state vector of GPS is selected as

$$\mathbf{X}_G = [\delta\varphi_G \ \delta\lambda_G]^T, \quad (14)$$

where $\delta\varphi_G$, $\delta\lambda_G$ are the position error of GPS.

Position integrated mode is used for INS/GPS subfilter, so the state equation and measurement equation are given by

$$\begin{aligned} \begin{bmatrix} \dot{\mathbf{X}}_I \\ \dot{\mathbf{X}}_G \end{bmatrix} &= \begin{bmatrix} \mathbf{F}_I & \mathbf{0} \\ \mathbf{0} & \mathbf{F}_G \end{bmatrix} \begin{bmatrix} \mathbf{X}_I \\ \mathbf{X}_G \end{bmatrix} + \begin{bmatrix} \mathbf{W}_I \\ \mathbf{W}_G \end{bmatrix}, \\ \mathbf{Z}_{IG} &= \begin{bmatrix} \varphi_I - \varphi_G \\ \lambda_I - \lambda_G \end{bmatrix} = \mathbf{H}_{IG} \begin{bmatrix} \mathbf{X}_I \\ \mathbf{X}_G \end{bmatrix} + \mathbf{V}_{IG}, \end{aligned} \quad (15)$$

where \mathbf{F}_I is INS state transition matrix in [15]; \mathbf{W}_I , \mathbf{W}_G are the system noise matrixes for INS and GPS, respectively; \mathbf{V}_{IG} is measurement noise matrix; and \mathbf{H}_{IG} is state transition matrix for INS/GPS measurement equation. Consider

$$\begin{aligned} \mathbf{W}_I &= [\omega_{gx} \ \omega_{gy} \ \omega_{gz}]^T, \\ \mathbf{W}_G &= [\omega_\varphi \ \omega_\lambda]^T, \end{aligned} \quad (16)$$

where ω_{gx} , ω_{gy} and ω_{gz} are system white noises of INS and ω_φ , ω_λ are system white noise of GPS. Consider

$$\mathbf{H}_{IG} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -1 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -1 \end{bmatrix}. \quad (17)$$

Using first order Markov process to complete the approximate fitting, GPS state transition matrix \mathbf{F}_G can be expressed as follows:

$$\mathbf{F}_G = \text{diag} \left(\frac{1}{\tau_{\text{GPS}\varphi}}, \frac{1}{\tau_{\text{GPS}\lambda}} \right), \quad (18)$$

where $\tau_{\text{GPS}\varphi}$ and $\tau_{\text{GPS}\lambda}$ are the correlation times.

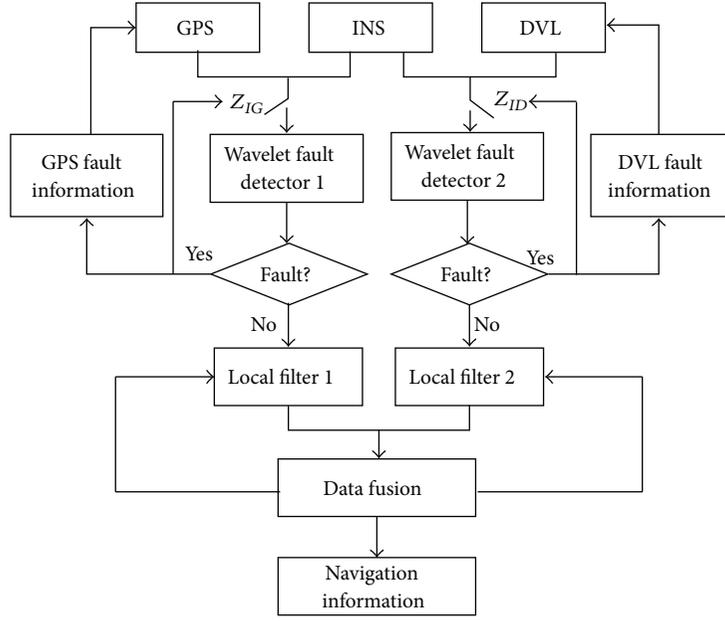


FIGURE 3: The fault-tolerant federated filter based on fault detection method for INS/GPS/DVL integrated navigation system.

Velocity matching method is used in INS/DVL subfilter, and the state equation and measurement equation are given by

$$\begin{aligned} \begin{bmatrix} \dot{\mathbf{X}}_I \\ \dot{\mathbf{X}}_D \end{bmatrix} &= \begin{bmatrix} \mathbf{F}_I & \mathbf{0} \\ \mathbf{0} & \mathbf{F}_D \end{bmatrix} \begin{bmatrix} \mathbf{X}_I \\ \mathbf{X}_D \end{bmatrix} + \begin{bmatrix} \mathbf{W}_I \\ \mathbf{W}_D \end{bmatrix}, \\ \mathbf{Z}_{ID} &= \begin{bmatrix} \mathbf{V}_{Ix} - \mathbf{V}_{Dx} \\ \mathbf{V}_{Iy} - \mathbf{V}_{Dy} \end{bmatrix} = \mathbf{H}_{ID} \begin{bmatrix} \mathbf{X}_I \\ \mathbf{X}_D \end{bmatrix} + \mathbf{V}_{ID}, \end{aligned} \quad (19)$$

where $\mathbf{X}_D = [\delta V_{Dx} \ \delta V_{Dy} \ \delta k]^T$ is state vector of DVL velocity estimation error, δV_{Dx} , δV_{Dy} are velocity error, and δk is scale error. \mathbf{F}_D is DVL state transition matrix:

$$\mathbf{F}_D = \text{diag} \left(\frac{1}{\tau_{Dx}}, \frac{1}{\tau_{Dy}}, \frac{1}{\tau_k} \right), \quad (20)$$

where τ_{Dx} , τ_{Dy} and τ_k are Markov correlation time, \mathbf{W}_D is system noise matrix, ω_{Dx} , ω_{Dy} are driven white noise, and \mathbf{V}_{ID} is measure noise matrix. Consider

$$\begin{aligned} \mathbf{W}_G &= [\omega_{Dx} \ \omega_{Dy}]^T, \\ \mathbf{H}_{ID} &= \begin{bmatrix} 0 & 0 & 1 & 0 & 0 & 0 & -V_y & 0 & 0 & 0 & 0 & 0 & -1 & 0 & -V_x \\ 0 & 0 & 0 & 1 & 0 & 0 & V_x & 0 & 0 & 0 & 0 & 0 & 0 & -1 & -V_y \end{bmatrix}. \end{aligned} \quad (21)$$

4.2. *INS/GPS/DVL Global Filter.* To make the integrated filter realizable for computer calculation, it is necessary to discretize the state equation and measurement equation mentioned above and then we have

$$\begin{aligned} \mathbf{X}_k &= \Phi_{k,k-1} \mathbf{X}_{k-1} + \Gamma_{k-1} \mathbf{W}_{k-1}, \\ \mathbf{Z}_k &= \mathbf{H}_k \mathbf{X}_k + \mathbf{V}_k. \end{aligned} \quad (22)$$

Assuming that the local estimates are unrelated, the global estimate can be expressed as

$$\begin{aligned} \hat{\mathbf{X}}_g &= \mathbf{P}_g (\mathbf{P}_{11}^{-1} \hat{\mathbf{X}}_1 + \mathbf{P}_{22}^{-1} \hat{\mathbf{X}}_2), \\ \mathbf{P}_g &= (\mathbf{P}_{11}^{-1} + \mathbf{P}_{22}^{-1})^{-1}. \end{aligned} \quad (23)$$

Local estimates ($\hat{\mathbf{X}}_1, \hat{\mathbf{X}}_2$) from the subfilters and their covariance matrix ($\mathbf{P}_1, \mathbf{P}_2$) are sent to master filter, and the obtained results would integrate with master filter estimates to get the global optimal estimates. Besides, global estimate $\hat{\mathbf{X}}_g$ and its covariance matrix \mathbf{P}_g , which are amplified to be $\beta_i^{-1} \mathbf{P}_g$ ($\beta_i \leq 1$), are feed back to sub filters, and estimates of subfilter are reset as

$$\begin{aligned} \hat{\mathbf{X}}_i &= \hat{\mathbf{X}}_g, \\ \mathbf{P}_{ii} &= \beta_i^{-1} \mathbf{P}_g. \end{aligned} \quad (24)$$

5. Simulation

5.1. *Simulation for Fault Positioning.* Compared with the detection step error, slope error is more difficult to be detected. Consequently, the slope error is selected to be researched to verify the efficiency of the proposed fault detection method.

(1) Simulation conditions:

During 100~200 sampling points, the slope error is added to the sensor signal.

(2) Simulation results:

As a comparison, the fault detection method based on Gaussian wavelet transform is applied to faults detection; its fault detection result is shown in Figure 4.

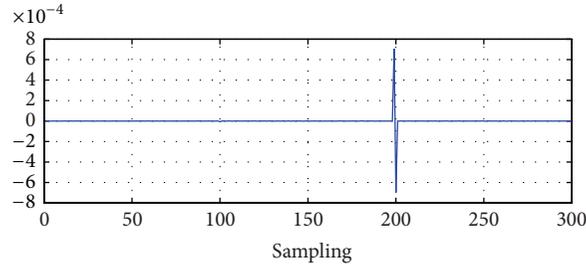


FIGURE 4: Slope error detection based on Gaussian wavelet transform.

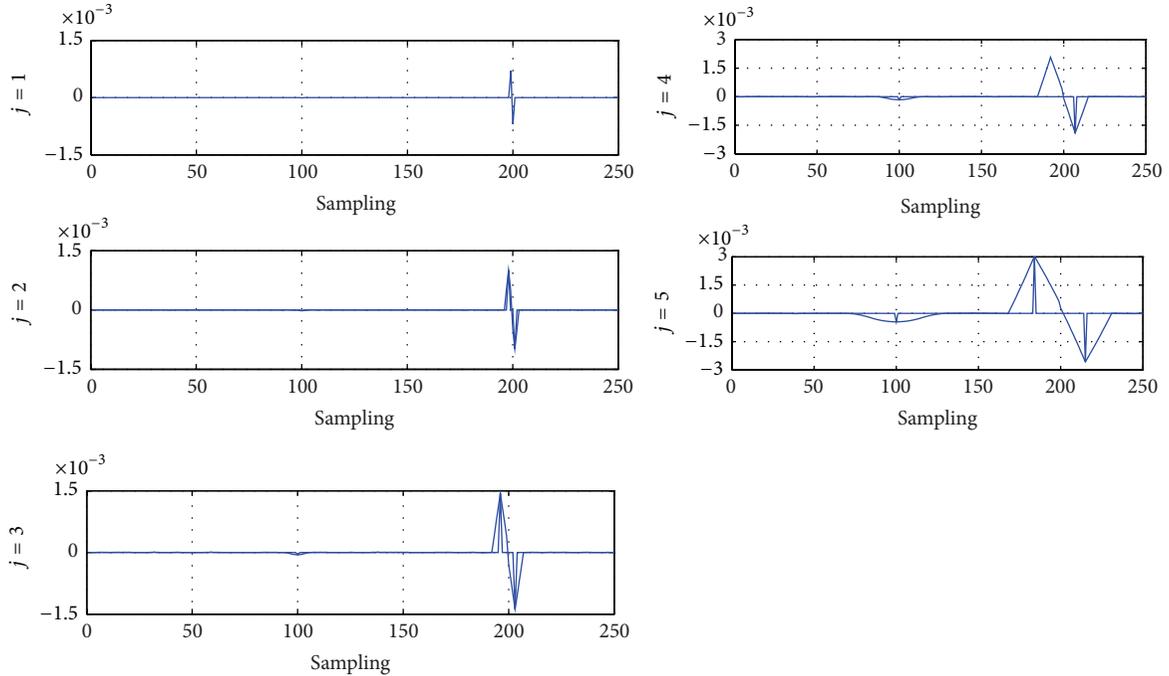


FIGURE 5: Slope error detection by binary wavelet transform.

It is obvious that there is a time delay for the fault detection algorithm based on Gaussian wavelet transform in Figure 4. And the sensor faults cannot be efficiently detected until the soft faults accumulate to a certain degree.

And the simulation result of the proposed fault detection method based on binary wavelet transform and modulus maxima is shown in Figure 5.

As known from Figure 5, the proposed fault detection method can efficiently detect the slope error of sensor fault and exhibits excellent accuracy for the fault positioning and sensational resolution in time domain. However, with the increase of the scale, multiple extrema appear in the signal which are caused by noise, and some of extrema would attenuate sharply or even disappear. However, those extrema which have appeared in all scales would not attenuate. And they are corresponding to the real singularities of signal.

5.2. Simulation for Fault Type Detection. The step fault and slope fault are selected to analyze the fault type detection performance of the proposed fault detection method.

During 100~200 sampling points, the constant error (to form the step fault) and slope error are added in the sensor

TABLE 1: Lipschitz exponent α .

	Step fault	Slope fault
Theoretical value of α	0	1
Modulus maxima attenuation ratio	0.4737	1.5143
Measured value of α	-0.0263	1.0143

signals, respectively. The logarithm of modulus maxima is calculated at $u = 100$, as shown in Figure 6.

In Figure 6, the logarithm curves of modulus maxima exhibit obvious attenuation ratio at $u = 100$. In order to identify the fault type, the attenuate ratio for logarithm curves of modulus maxima with respect to the scales at $u = 100$ should be obtained. By using linear least square method, the slope of curve on each scale of the curve can be estimated, and the Lipschitz exponent can be obtained by subtracting 0.5 from the slope. As for the two kinds of faults, the scopes of modulus maxima curves and Lipschitz exponents at $u = 100$ are shown in Table 1.

According to Table 1, the measured Lipschitz exponents with respect to step fault and slope fault approximate to

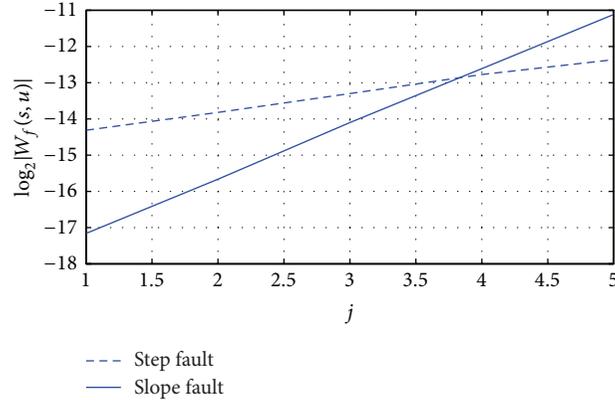


FIGURE 6: Logarithm curves of modulus maxima at $u = 100$.

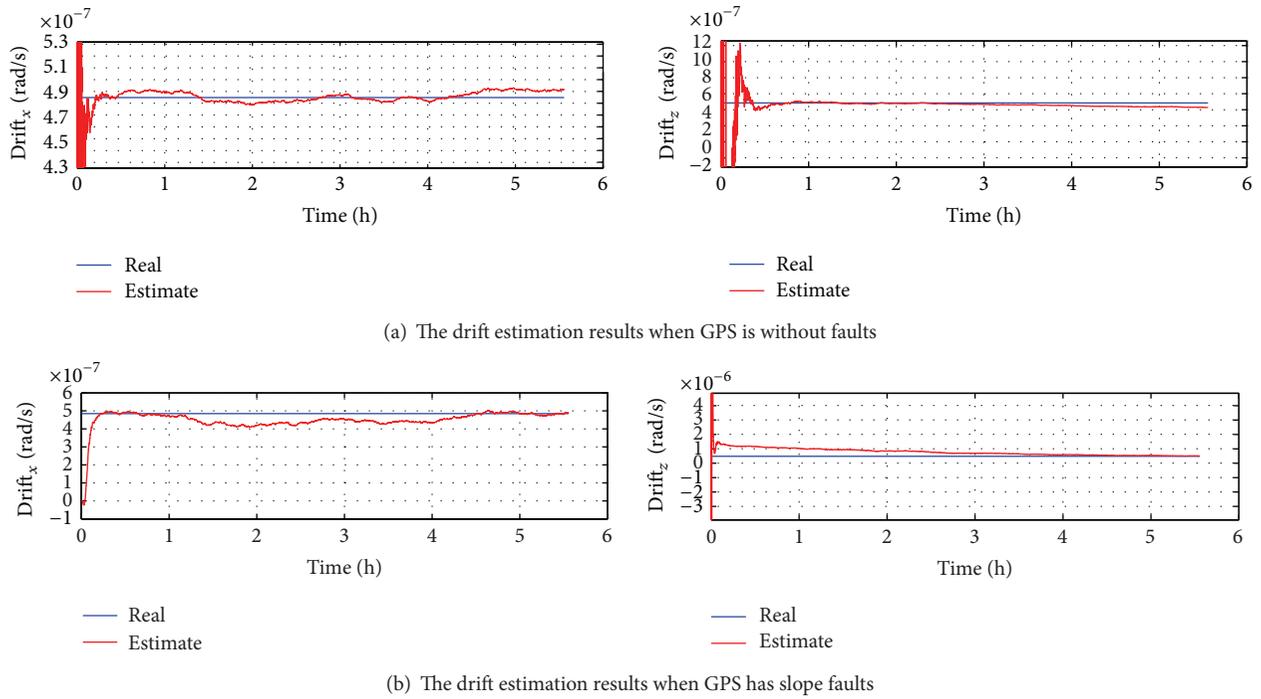


FIGURE 7: Gyro drifts estimation result.

their ideal values, respectively. By calculating the Lipschitz exponent of signal, the fault type can be detected.

In conclusion, by searching wavelet modulus maxima on different scales, the fault detection method based on modulus maxima can accurately position fault points and distinguish the fault type through Lipschitz exponent which is calculated by the attenuation of wavelet modulus maxima.

5.3. Simulation for Fault-Tolerant Filter

- (1) Assuming that initial latitude is $\varphi_0 = 30^\circ$ and initial longitude is $\lambda_0 = 120^\circ$, gyro constant drift is $0.01^\circ/\text{h}$; heading angel is 45° and east velocity and north velocity are 5 kn; simulation time is 5.5 h.
- (2) INS/DVL/GPS integrated system is applied to calibration. After 4-hour system work, inertial system is

calibrated by the gyro drift estimated and system error information.

- (3) Two conditions are involved in the calibration simulation: (a) GPS without any fault; (b) GPS with slope fault.

For the two conditions, the drift estimated results for east and azimuth gyroscopes are given in Figure 7.

It is obvious from Figure 7 that on the no-faults condition the gyro drifts and position information can be well estimated, especially since the estimation error for east gyro drift is less than 10%. On the other hand, when slope fault occurs, we cannot obtain the position measurement, and the gyro drift estimation performance decreases, especially since the estimate accuracy of the azimuth gyro drift is seriously restricted by limited available information. Through

the detection and isolation of slope faults in INS/GPS subintegrated system, such a single system fault will have no effect on the output of SINS/DVL system but slightly decrease for the position estimation comparing with the situation without GPS faults.

6. Conclusion

This current work has presented a sensor fault detection method based on binary wavelet transform, modulus maxima, and Lipschitz exponent. In this method, after binary wavelet transform of original signal, sensor faults can be accurately positioned by searching modulus maxima, and the type of fault can be obtained by calculation of the Lipschitz exponents. To guarantee the performance during calibration phase for INS navigation errors, a developed fault-tolerant federated filter for INS/GPS/DVL integrated navigation system is proposed. Simulation shows that even when GPS slope faults occur, INS/GPS/DVL system could keep on providing reliable position information only with a slight drop of positioning accuracy.

Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

Acknowledgments

The authors would like to thank all the editors and anonymous reviewers for improving this paper. Funding for this work was provided by the National Nature Science Foundation of China under Grants no. 61374007 and 61104036 and the Fundamental Research Funds for the Central Universities under the Grant of HEUCFX41309.

References

- [1] J. Bokor and Z. Szabó, "Fault detection and isolation in nonlinear systems," *Annual Reviews in Control*, vol. 33, no. 2, pp. 113–123, 2009.
- [2] M. Jafari and J. Roshanian, "Optimal redundant sensor configuration for accuracy and reliability increasing in space inertial navigation systems," *Journal of Navigation*, vol. 66, no. 2, pp. 199–208, 2013.
- [3] G. T. Schmidt, "INS/GPS technology trends," *Advances in Navigation Sensors and Integration Technology Technical Report*, Massachusetts Institute of Technology, Lexington, Mass, USA, 2010.
- [4] U. I. Bhatti, W. Y. Ochieng, and S. Feng, "Integrity of an integrated GPS/INS system in the presence of slowly growing errors. Part II. Analysis," *GPS Solutions*, vol. 11, no. 3, pp. 183–192, 2007.
- [5] A. Korff, C. Kellermann, M. Mueller, S. C. Kunze, K. Schmieder, and K. Radermacher, "Using analytical redundancy to increase safety of a synergistic manually guided instrument for craniotomy," in *Proceedings of the 34th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, pp. 912–915, September 2012.
- [6] M. Carminati, G. Ferrari, R. Grassetto, and M. Sampietro, "Real-time data fusion and MEMS sensors fault detection in an aircraft emergency attitude unit based on kalman filtering," *IEEE Sensors Journal*, vol. 12, no. 10, pp. 2984–2992, 2012.
- [7] M. I. Cao and P. Y. Cui, "Information fusion technology of INS/GPS navigation system based on wavelet fault detection," *Journal of Astronautics*, vol. 30, no. 5, pp. 1885–1887, 2009.
- [8] A. Phinyomark, C. Limsakul, and P. Phukpattaranont, "Application of wavelet analysis in EMG feature extraction for pattern classification," *Measurement Science Review*, vol. 11, no. 2, pp. 45–52, 2011.
- [9] J. A. Fairley, G. Georgoulas, O. L. Smart et al., "Wavelet analysis for detection of phasic electromyographic activity in sleep: influence of mother wavelet and dimensionality reduction," *Computers in Biology and Medicine*, vol. 48, pp. 77–84, 2014.
- [10] J. Cheng and M. Li, "Research of fault detection algorithm based on wavelet analysis in integrated navigation system," *Journal of Astronautics*, vol. 33, no. 4, pp. 419–425, 2012.
- [11] Y. Liu, Y. Li, and J. T. Xu, "Application of a multi-algorithm fusion real time filter in FOGs," *Acta Photonica Sinica*, vol. 39, no. 6, pp. 1116–1119, 2010.
- [12] Y. K. Sun, *Wavelet Analysis and Its Application*, China Machine Press, 2004.
- [13] S. Mallat and S. F. Zhong, "Characterization of signals from multi-scale edges," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 14, no. 7, pp. 710–713, 1992.
- [14] P. D. Groves, *Principles of GNSS, Inertial, and Multi-Sensor Integrated Navigation Systems*, Artech House, London, UK, 2013.
- [15] Y. N. Wang, X. Deng, and W. Zhao, "Application of fuzzy adaptive Kalman filter based on weighted matrix in integrated navigation system," *Journal of Chinese Inertial Technology*, vol. 16, no. 3, pp. 334–338, 2008.

Research Article

Probabilistic Harmonic Calculation in Distribution Networks with Electric Vehicle Charging Stations

Jianxue Wang,¹ Yanlin Cui,² and Minghui Zhu¹

¹ Department of Electrical Engineering, Xi'an Jiaotong University, Xi'an, Shaanxi 710049, China

² Power Dispatch and Control Center of Guangdong Power Grid Corporation, Guangzhou, Guangdong 510080, China

Correspondence should be addressed to Jianxue Wang; jxwang@mail.xjtu.edu.cn

Received 25 April 2014; Accepted 28 July 2014; Published 14 August 2014

Academic Editor: Weichao Sun

Copyright © 2014 Jianxue Wang et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Integrating EV charging station into power grid will bring impacts on power system, among which the most significant one is the harmonic pollution on distribution networks. Due to the uncertainty of the EV charging process, the harmonic currents brought by EV charging stations have a random nature. This paper proposed a mathematical simulation method for studying the working status of charging stations, which considers influencing factors including random leaving factor, electricity price, and waiting time. Based on the proposed simulation method, the probability distribution of the harmonic currents of EV charging stations is obtained and used in the calculation of the probability harmonic power flow. Then the impacts of EVs and EV charging stations on distribution networks can be analyzed. In the case study, the proposed simulation and analysis method is implemented on the IEEE-34 distribution network. The influences of EV arrival rates, the penetration rate, and the accessing location of EV charging station are also investigated. Results show that this research has good potential in guiding the planning and construction of charging station.

1. Introduction

Car industry has been developing for decades at a cracking pace. In order to pursue better performance and higher utility of vehicles, many valuable researches have been done to address the issues like adaptive control, mathematical modeling, and fault diagnosis [1–3]. In recent years, the pressure of fossil fuel shortage and other environmental problems leads to high-speed development of electric vehicles (EVs) due to their emission-free characteristic and effective utilization of renewable energy [4, 5]. The promising trend of EVs attracts plenty of researchers to concentrate on the functioning and the influence of electrical vehicles. Unlike conventional vehicles, the power supply of EVs mainly comes from local power distribution system through EV charging stations [6, 7] and that will lead to impacts on distribution network. In order to study the impacts, the EV charging station should be modeled properly and many factors need to be taken into consideration.

The impacts of EV charging station integration can be divided into two categories basically. The first category

contains the influences induced by uncoordinated charging of large amounts of EVs such as voltage deviations, thermal overloads, increase in losses, and decrease in voltage stability [6, 8–10]. Researches have investigated aspects including the different EV penetrations on distribution network investment [7], the long-term operating capacity reserve [11], and transformer life [12]. Coordinated charging strategies are proposed [12–15] to promote the distribution grid's ability to hold more EVs without reinforcements, improve voltage profile, and reduce power losses.

The other category is the impacts brought by high rating nonlinear switching devices in EV chargers [16] and the typical one in this category is harmonic pollution. Harmonic pollution will shorten the life span of transformers, reduce the utilization of transmission lines, and interfere with the signals of measuring and protection equipment. Therefore, studying the harmonic pollution brought by EV charging stations is of great importance. Paper [17] establishes a stochastic aggregate harmonic load model to study the harmonic characteristics of EV charger. Paper [18] presents a harmonic simulation method to study the harmonic impact of EVs. Paper [19]

develops an approach to decrease the harmonics in a high EV penetration grid, and the harmonic load flow is used to model the distorted grid and identify the worst bus based on the typical daily load curve.

In most of the calculation, the worst situation of harmonic of charging station was used. However, the probability of the worst situation is rare in daily life. In most of these studies the harmonic currents are generated by individual EV charger. The random characteristic of charging behavior was not considered in the calculation of charging station harmonic and was hardly consistent with the actual situation in our daily life. Besides, there are limited studies focused on the system level to evaluate the harmonic distortion of charging stations.

In this paper, this problem is investigated from single equipment to the entire system. First, a simplified simulation model of EV charger is used to analyze the harmonic characteristics. Second, based on the queuing theory and charger model, a mathematical simulation method to get the probability distribution of charging station state is proposed, considering the influences of random factors, electricity price, and waiting time. The statistics of the simulation results are performed to obtain the probability distribution of total harmonics. Third, through the probability harmonic power flow, the probability characteristics of probability distribution of charging stations' influences on system harmonics are analyzed. At last, the key factors, including the arrival rate, penetration rate, and the access location of EV charging station, were analyzed based on a modified IEEE-34 system.

2. The Working State Model of EV Charging Station

2.1. Model the EV Charger. In order to analyze the harmonic characteristic of EV charging station, the basic working model of a single EV charger needs to be built first. In this paper, we consider a kind of currently the most widely used EV charger. This typical EV charger consists of a three-phase uncontrolled rectifier and a high frequency DC/DC converter which are coupled with a high frequency transformer [4]. The output voltage and current of the charger are controlled by feedback control equipment in the high frequency DC/DC converter. Since elaborately modeling the EV charger is too complicated, it consumes much time for the simulation.

To reduce the complexity of the EV charger model, a simplification proposed in [20] is adopted here, where the high frequency DC/DC converter and battery load are replaced by an equivalent resistance. The simplified equivalent circuit is shown in Figure 1.

Based on Ohm's law, the equivalent resistance can be calculated as

$$R(t) = \frac{U_I}{I_I(t)} = \frac{U_I^2}{P_I(t)} = \frac{\eta U_I^2}{P_O(t)}, \quad (1)$$

where R is the equivalent resistance of high frequency DC/DC converter, P_I is the input power of the high frequency DC/DC converter, P_O is the output power of EV charger, η is the efficiency of high frequency DC/DC converter, U_I is

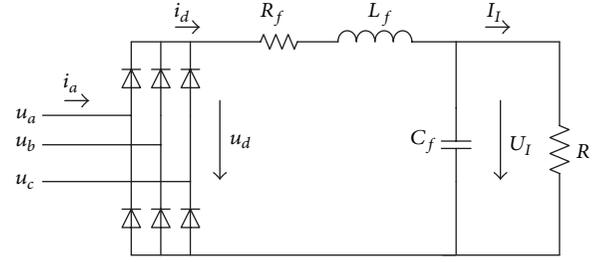


FIGURE 1: The structure diagram of equivalent simplified mode for EV charger [20].

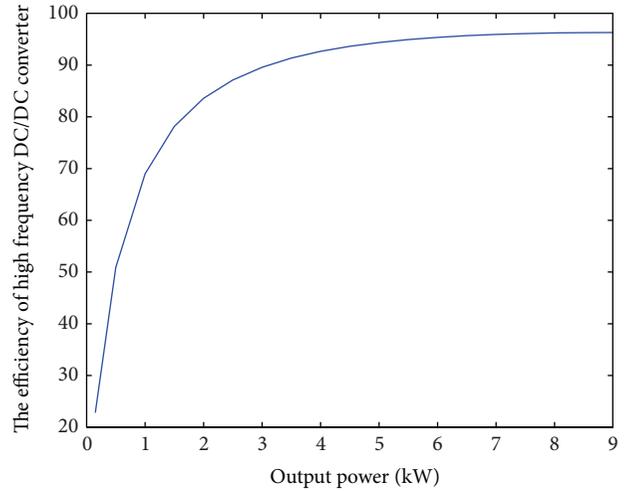


FIGURE 2: The relationship between efficiency and output power.

the output voltage of three-phase uncontrolled rectifier, and I_I is the output current of three-phase uncontrolled rectifier.

According to (1), the equivalent resistance should be a function of the charging time to reflect the different output power in different charging stages. The empirical formula of P_O during the charging process was obtained through fitting the actual-measured data [20], which can be put as

$$P_O(t) = \begin{cases} 0.79P_{O\max}t^{0.048}, & 0 < t < t', \\ P_{O\max}e^{-0.021(t-t')}, & t' < t < t_{\max}, \end{cases} \quad (2)$$

where P_O is the output power of EV charger, t is charging time, t' is the time spot when charging mode turns from constant current charging to constant voltage charging, t_{\max} is the maximum charging time, and $P_{O\max}$ is the rated output power.

In [20], η was treated as a constant; however, according to the simulation result in MATLAB/Simulink environment, the efficiency of the high frequency DC/DC converter η varies with the output power of the EV charger P_O . Their relationship is presented in Figure 2.

Using the curve-fitting technique, the expression can be obtained as

$$\eta(P_O) = \frac{-0.011 + 193P_O}{1 + 1.785P_O + 0.012P_O^2}. \quad (3)$$

Through (1)–(3), we can get the value of equivalent resistance at any time spot during charging.

To illustrate the effectiveness of the simplified model in harmonic analysis, a comparison between the original and the simplified model (considering the varying efficiency during charging) has been made through simulating in MATLAB/Simulink environment. Simulation results show that the maximum error of the harmonic between the original and the simplified model is merely 0.45%, while the equivalent simplified model costs 40% less time than the original one. Therefore, the simplified model appears to be more effective and suitable, especially when repetitive calculations are required.

2.2. Build the Working State Model for an EV Charging Station. Analysis of the harmonic impact of EV charging stations requires knowing the harmonic currents injection of the charging stations to distribution network. With the EV charger simulation model mentioned in Section 2.1, the harmonic currents that one charger brings to the grid can be obtained in simulation software. The harmonic currents of the entire charging station can be calculated by adding up the harmonic currents generated by every charger in work. Therefore, it is necessary to build the working state model for the EV charging station to acquire the number and the status of working chargers.

To build the working state model, the charging process in the charging station, which depends on EVs' charging behaviors, needs to be studied. Charging behaviors of EVs have much to do with the power conditions of their batteries. Normally, the power condition of a battery is measured by its SOC, which is defined as the available capacity in percentage of full charge capacity [21]. The battery power needs to be adequate to ensure the normal operation of electric vehicles. With the battery power consuming, the SOC will keep dropping and may get down to an overdischarging state [21]. When the battery power drops below a threshold value, the EV cannot work normally and thus needs to get charged at a nearby charging station. The charging process of an EV (as shown in Figure 3) can be modeled as a queuing process based on queuing theory (introduced in [22]), which can be summarized as "first come, first served." The EV that needs charging adds to the tail of the queue when it enters the charging station. If there is no available EV charger, the EV will be waiting for an opening; if there is an available EV charger, it will get charged immediately. The EV will leave when the SOC reaches 100% or when there are some factors to make it leave early.

In [23, 24], the queuing theory is used to describe the charging process and it is proved feasible. In [24], a typical queuing model, M/M/c, is used to portray the customer streams of EV charging stations. The typical queuing theory assumes that the inter-arrival and service time are generally subject to exponential distribution. As for EV charging, it is feasible to assume that the interarrival time of two consecutive EVs is exponentially distributed but it will lead to large error to consider that the charging time also obeys exponential distribution. The charging time is a random

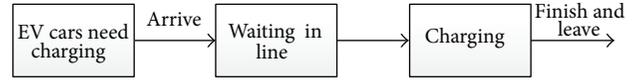


FIGURE 3: The diagram of EVs' charging process.

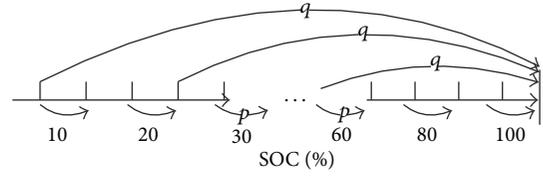


FIGURE 4: The diagram of charging state transition.

variable influenced mainly by the initial SOC, the random leaving factor, the electricity price, and the waiting time, so it does not seem to follow any already known probability distributions. Therefore, the typical queuing model may not describe the charging behavior effectively.

In order to make the model more realistic, this paper adopts a numerical simulating method to study the EV charging process. First, the states of an EV charger should be modeled. The state of charger is determined by the SOC of the battery connected to it, where the range of SOC is from 0 to 100 in percentage. The battery charging process can be treated as a Markov process since the SOC of the next time spot is only related to the current SOC and has nothing to do with the SOC before, which can be put as follows:

$$\begin{aligned}
 & P \{ \text{SOC}(t_{m-1}) \in A \mid \text{SOC}(t_1) = \text{SOC}_1, \\
 & \quad \text{SOC}(t_2) = \text{SOC}_2, \dots, \text{SOC}(t_m) = \text{SOC}_m \} \quad (4) \\
 & = P \{ \text{SOC}(t_{m+1}) \in A \mid \text{SOC}(t_m) = \text{SOC}_m \},
 \end{aligned}$$

where t_m is the current time and $\text{SOC}(t_m)$ is the current state of charge.

In order to make it easier for study, we discretize the domain of SOC's possible values. Figure 4 shows the transition of the states during charging. The SOC of the battery goes towards a higher level over the charging time if the EV does not quit at the middle of charging. If it does quit, the SOC of the battery goes to 100% directly. It is worth noting that 100% of SOC here means that the EV leaves and the charger is free. So there are two conditions that a state is treated as a 100% state. One is that the EV gets fully charged and leaves and the other is that the EV leaves in the middle of charging.

The diagram can be described in

$$\begin{aligned}
 P_{m,m+\Delta\text{SOC}} &= p(\text{soc}) = 1 - q(\text{soc}) \\
 P_{m,m+\Delta\text{SOC}} &= \text{SOC}_{m,m} = 0 \\
 P_{m,100} &= q(\text{soc}), \\
 & (0 < m < 100).
 \end{aligned} \quad (5)$$

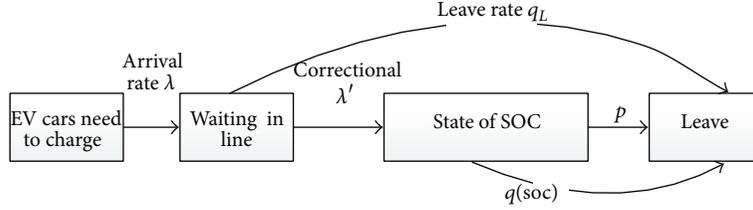


FIGURE 5: The correctional diagram of EVs' charging process.

$P_{i,j}$ means the transition probability from state i to j , ΔSOC is the incremental power in a simulation time unit, and $q(\text{soc})$ is the probability of quitting before the EV is fully charged, which will vary due to the current value of SOC.

To start the simulation, the initial SOC should be given. Setting it to zero or some fixed low value will be inappropriate because the remaining SOC of last time is random. In fact, according to [25], the battery's initial SOC obeys two-parameter β distribution, whose probability density function is

$$\begin{aligned} f(x; \alpha, \beta) &= \frac{x^{\alpha-1}(1-x)^{\beta-1}}{\int_0^1 u^{\alpha-1}(1-u)^{\beta-1} du} \\ &= \frac{\Gamma(\alpha + \beta)}{\Gamma(\alpha)\Gamma(\beta)} x^{\alpha-1}(1-x)^{\beta-1} \quad (6) \\ &= \frac{1}{B(\alpha, \beta)} x^{\alpha-1}(1-x)^{\beta-1}. \end{aligned}$$

Quitting probability $q(\text{soc})$ needs to be pinned down to complete the model for EV chargers. Generally, it contains 3 components, including random leaving factor, waiting time q_{time} , and electricity price q_{price} .

The typical threshold value SOC' is set to 80%. Leaving earlier before SOC reaches SOC' will do some harm to the battery, so people usually do not leave earlier no matter how long they will wait or how much the electricity costs. When SOC is equal to or greater than the threshold SOC' , waiting time factor and electricity cost factor begin to work. The whole process can be illustrated by

$$q(\text{soc}) = \begin{cases} q', & \text{soc} < \text{SOC}', \\ q' + q_{\text{time}} + q_{\text{price}}, & \text{soc} \geq \text{SOC}'. \end{cases} \quad (7)$$

Furthermore, q_{time} is determined not only by the waiting time but also by the difference between the current SOC and the threshold SOC' . With the increasing of SOC, both the waiting time and the difference to SOC' will increase, which leads to a higher waiting time-related quitting probability. In (8), q_{time} can be thought to be in a linear relationship with the product of the charging time t and the difference between the battery's SOC and SOC' in this study, where time coefficient k' is a constant:

$$q_{\text{time}} = \begin{cases} k't(\text{soc} - \text{SOC}'), & \text{soc} \geq \text{SOC}', \\ 0, & \text{soc} < \text{SOC}'. \end{cases} \quad (8)$$

q_{price} is also decided by two components. One is the difference between the real-time electricity price (p_{now}) and

the expected electricity price (P_E), and the other is the difference between SOC and SOC' . The expression is shown in (9), where price coefficient k'' is a constant:

$$\begin{aligned} q_{\text{price}} &= \begin{cases} k''(p_{\text{now}} - P_E)(\text{soc} - \text{SOC}'), & \text{soc} \geq \text{SOC}', \quad p_{\text{now}} > P_E, \\ 0, & \text{soc} < \text{SOC}'. \end{cases} \quad (9) \end{aligned}$$

Finally, a correction should be made to improve the model of charging process. It is noted that an EV driver may be scared off by a long queue or get impatient waiting in line and then decide to leave before getting the car charged, so the effective arrival rate is a little bit less than the arrival rate. The correctional charging process is demonstrated in Figure 5.

Considering the waiting time, the arrival rate is fixed in (10), where λ is the real arrival rate and λ' is the fixed arrival rate. The directly leaving rate (q_L) is the product of the leaving factor (k_{wait}) and the average waiting time (T_{wait}) which can be acquired in the simulation of queuing process:

$$\lambda' = \lambda - q_L = \lambda - k_{\text{wait}} T_{\text{wait}}. \quad (10)$$

Using the model given above to simulate the activity of each EV charger, the state of the whole charging station can be simulated by the sequential Monte Carlo method. This makes the preparation for the calculation of harmonic distribution.

2.3. Calculate the Harmonic Distribution of Charging Station.

The harmonic currents of the whole charging station can be calculated by adding up the harmonic currents of all EV chargers. However, the states of chargers are uncertain so the harmonics of the charging station cannot be calculated directly. Using the working state model built in Section 2.2, the states of chargers in an EV charging station can be determined at any time spot; then the simulation model of the entire charging station can be built in simulators to calculate the total harmonics. After simulation, the probability distribution of the charging station can be obtained through the statistical analysis of simulation results.

The simulation model of the charging station is built by the simplified models of EV chargers mentioned in Section 2.1. In each step of the EV charging station's mathematical simulation, all the chargers' output power is determined; thus, the equivalent resistance can be calculated through (1). After that, the value of equivalent resistance can be set in MATLAB/Simulink and then the harmonics of the charging station can be simulated for statistical analysis to acquire

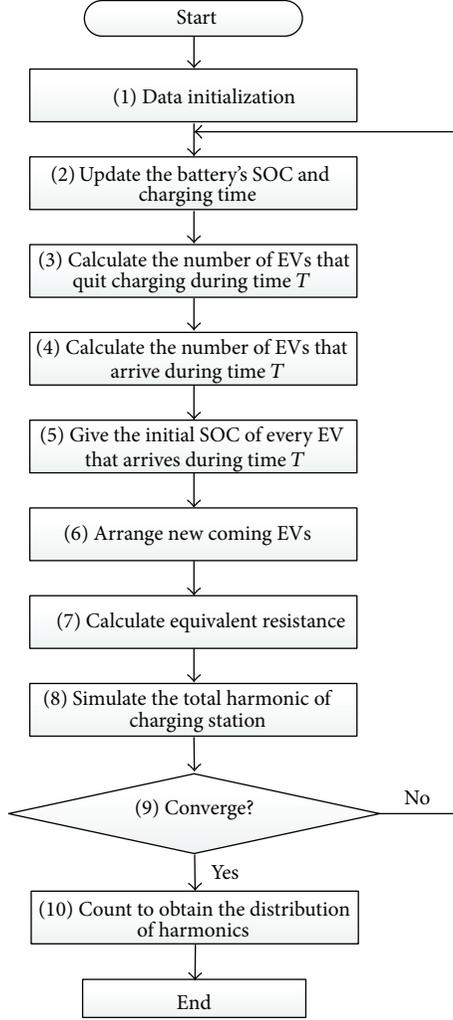


FIGURE 6: Calculating the EV charging station's harmonic distribution.

the probability distribution. Figure 6 illustrates the steps in a flowchart.

Details of the steps in Figure 6 are given as follows.

- (1) Initialize the valid data, for example, the simulation step time (T) and the convergence criterion. Then start the mathematical simulation process.
- (2) Update the battery's SOC in time T through

$$\Delta\text{SOC}(t) = \frac{\int_t^{t+T} P_O(t) dt}{\int_0^{t_{\text{all}}} P_O(t) dt} \times 100\%, \quad (11)$$
 where T is the time of simulation step, t is the battery's charging time, and t_{all} is the total time of a battery's SOC changing from 0 to 100 in percent.
- (3) Calculate the number of EVs which quit charging in time T , considering the random leaving factor, the price of electricity, and the waiting time.
- (4) Calculate the number of EVs which arrive in time T using the correctional arrival rate to increase the calculation accuracy. Then update the number of EV cars in charging station.

- (5) Give the initial SOC of each EV which arrives in time T with (1). Then find the corresponding charging time (t) and output power (P_O).
- (6) Arrange the EV cars in the queue for charging according to the "first come, first served" principle.
- (7) Calculate the output power by (2) and (3), and then the value of equivalent resistance can be obtained by (1).
- (8) Set the resistance parameters of each charger in the simulation model of EV charging station. Then start the simulation and record results of harmonic and total power in an array.
- (9) Keep simulation until the probability distribution of charging state meets the convergence criterion that Euclidean distance of probability distributions obtained using different simulation time must be within the range of allowable error, which can be put as

$$d_{ij} = \sqrt{\sum_{i=1}^M (x_{ik} - x_{jk})^2} < \epsilon, \quad (12)$$

where x_{ik}, x_{jk} stands for the difference of the k th component of random variable x in the i th and j th distributions obtained in different simulation time, M stands for the number of variables, and ϵ stands for the allowable error.

- (10) Obtain the probability distribution of the charging state and the harmonics.

3. Probabilistic Harmonic Power Flow of Distribution System Integrating EVs

The influence of integrating EVs into the distribution network has a probabilistic nature due to the random and uncertain characteristic of EVs. Based on the probability distribution of total harmonic of charging stations, the probability characteristics of charging stations' influences on system harmonics can be analyzed through the probabilistic harmonic power flow [10, 26]. From the simulation results of harmonics, we find that there is no typical curve that can describe the probability characteristics of charging stations' harmonic. Thus in this research, the Monte Carlo sampling is used to calculate the probabilistic harmonic power flow.

Because the harmonic power brought by EV charging station is much less than the fundamental power in distribution network, the fundamental power flow and the harmonic power flow can be decoupled and calculated, respectively. Newton-Raphson method was used to calculate the fundamental power flow.

The harmonic models of load, transmission line, and transformer adopt the typical harmonic models in [27]. Based on the harmonic model of electrical components and the result of fundamental power flow, n time harmonic admittance matrix (\mathbf{Y}_n) can be obtained according to the harmonic analysis theory. Harmonic sources are treated as an

TABLE 1: The detailed information of the charging station.

Voltage of network/kV	The number of transformers	The ratio of transformation	The number of chargers	The total charging power/kW	The number of parking spaces
10	1	10 kV/400 V	35	315	50

TABLE 2: Parameters of the EV charger.

Voltage/V	Rated output power/kW	Output voltage/V	Output current/A	L_f (mH)	C_f (uF)
400	9	300	30	1.5	2115

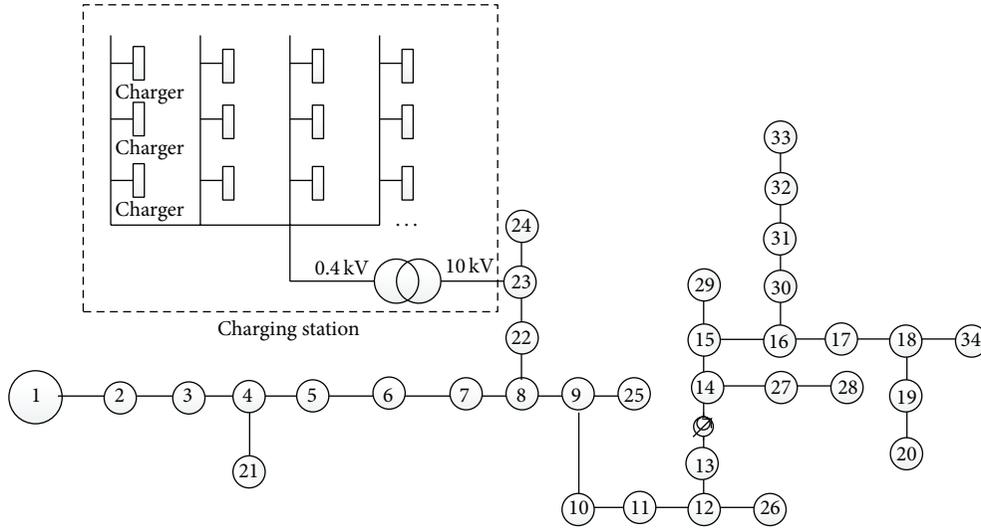


FIGURE 7: The diagram of the modified IEEE-34 test system.

ideal current source, so the n time injection current matrix (\mathbf{I}_n) can be obtained. The n time harmonic voltage matrix (\mathbf{U}_n) can be obtained by solving

$$\mathbf{Y}_n \mathbf{U}_n = \mathbf{I}_n. \quad (13)$$

After solving harmonic power flow at each time, the effective voltage (U_i) and total harmonic distortion of voltage (THD_v) can be calculated through

$$U_i = \sqrt{\sum_{n=1}^H |U_i^n|^2}, \quad (14)$$

$$\text{THD}_v = \left[\frac{(\sum_{n=2}^H |U_i^n|^2)^{0.5}}{|U_i^1|} \right] \times 100\%.$$

After the calculation of probabilistic harmonic power flow based on Monte Carlo, the probability distribution of effective voltage and THD_v can be acquired. Then the probability characteristics of charging stations' influences on system harmonics under different conditions can be analyzed.

4. Case Study

The case study is implemented on a modified IEEE-34 test system [28], which is shown in Figure 7.

TABLE 3: Parameters of the working state model.

Parameters	α	β	q	$k/10^{-4}$	k''	$k_{\text{wait}}/10^{-4}$
Value	2.7	5.5	0.004	1.25	0.2	1

An EV charging station is connected to the original IEEE-34 test system at Node 23 through a transformer. The penetration power of the EV charging station is set to 15% in this case. Detailed information of the charging station is listed in Table 1.

The charging time of the EV battery from out-of-power to fully charged is 270 minutes. The parameters of the EV charger are listed in Table 2.

The arrival rate is set to 0.1 during the day and 0.2 at night since fewer people use cars at night and the electricity price is lower. According to the load curve, the load level during the night is lower than other times. The random leaving factor, the waiting time, and electricity price are considered during the daytime. The parameters of the working state model are listed in Table 3.

4.1. The Efficiency Analysis of Working State Model. In order to prove the effectiveness of the model proposed in this paper, the probability distribution of charging station's state is calculated using two different models. The first is the conventional queuing theory model, in which the average service time of

TABLE 4: Parameters of the probability distribution of station state.

Model	Expected number of chargers in work	The number of chargers with highest possibility	The probability of all chargers in work	The average length of the queue
Typical	20.3	20	0.002	20.3
Proposed model	20.3	20	0.003	20.4

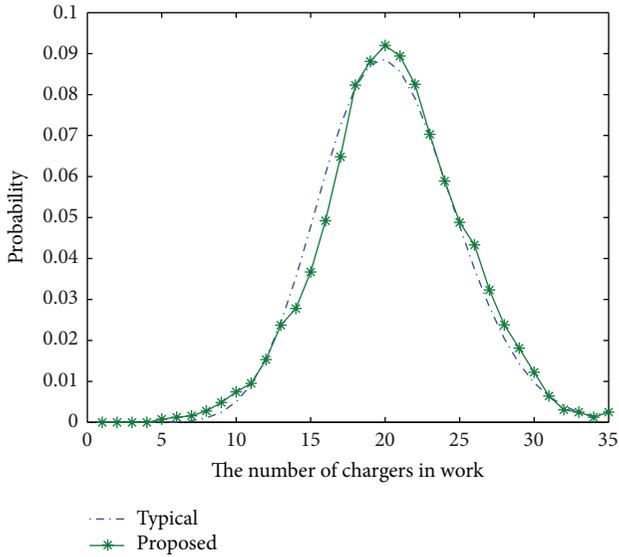


FIGURE 8: The comparison of probability distribution of EV charger in work using typical and proposed model.

charger is 203 minutes. For in the typical model, the random leaving factor, the waiting time, and the price of electricity are not reflected. So in order to prove the effectiveness of the method proposed in this paper, these factors are ignored. The probability distribution of EV charging station's state obtained using two models is shown in Figure 8.

From Figure 8, it is obvious that the probability distribution obtained by the typical model is similar to that by the proposed model.

The parameters of probability distribution of station state are listed in Table 4.

Among the parameters listed in Table 4, the expected number of working chargers and the number of chargers with the highest probability are the same, while the other two parameters are similar in two models. The simulation step time (T) is set to be 3 minutes in this case. If it is set smaller, the value of those similar parameters will get closer. Therefore, the method proposed in this paper can be used to simulate the charging process and calculate the probability distribution of charging station's harmonic.

4.2. Result of Harmonic Probability Power Flow. Based on the working state model of charging station, the distribution of charging station's total power and harmonics can be acquired by doing statistics on simulation results. The probability distributions of charging station's total power during daytime and night are shown in Figures 9 and 10, respectively.

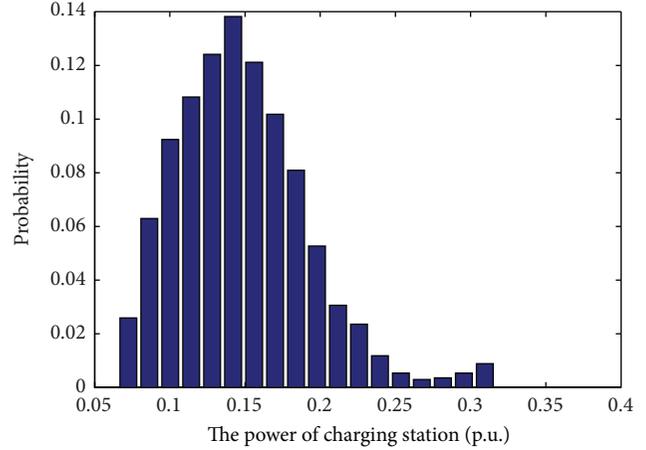


FIGURE 9: The probability distribution of charging station's power during daytime.

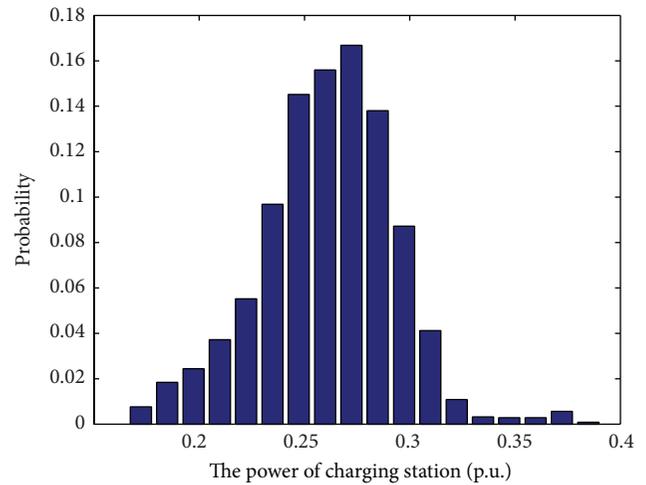


FIGURE 10: The probability distribution of charging station's power during night.

From Figures 9 and 10, it can be noted that the EV charging station is more heavily utilized during night than during daytime.

The distribution of harmonics of charging station can also be acquired. Based on the distribution, the expectation of the current amplitude, the harmonic content, and the expectation of the current angle during day and night are calculated, which are listed in Table 5. It is worth noting that only $6k \pm 1$ th ($k = 1, 2, \dots$) harmonics are listed because they are the major harmonics of EV charging station.

TABLE 5: The detailed information of charging station's harmonics.

Time	Parameters	Order					
		5	7	11	13	17	19
Daytime	The expectation of current amplitude/ 10^{-2} p.u.	3.77	1.61	0.82	0.44	0.41	0.25
	The harmonic content/%	34.2	14.6	7.5	4.0	3.7	2.3
Night	The expectation of current amplitude/ 10^{-2} p.u.	6.45	2.13	1.47	0.74	0.66	0.45
	The harmonic content/%	31.5	10.4	7.2	3.6	3.3	2.2

TABLE 6: The detailed information of voltage RMS over the range.

Node	8	9	10	11	12	13	22	23	24	25	26
The probability of voltage RMS over the range/ 10^{-2}	0.01	0.27	4.88	5.50	6.73	7.70	0.03	0.03	2.07	0.27	13.69

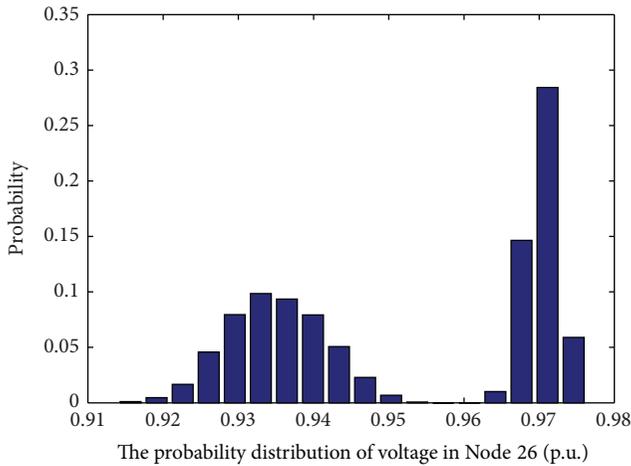


FIGURE 11: The probability distribution of voltage in node 26.

The expectation of current amplitude decreases with the increase of the harmonic order. It is easy to see that the low-order harmonic current accounts for larger proportion than the high-order harmonic. The harmonics with orders higher than 19 are not listed because the amplitudes are too small; thus, the influences are neglectable.

Based on the result of the probabilistic harmonic power flow, the probability distributions of root mean square (RMS) voltages of each node are obtained. The overranging nodes are listed in Table 6.

According to Table 6, there are 11 nodes whose RMS voltages are overrange. Among these nodes, Node 26 has the highest overranging probability. The probability distribution of voltage RMS of Node 26 is shown in Figure 11.

It can be seen that the distribution has a bimodal characteristic, which is due to the peak-valley load during 24 hours.

Based on the result of the probabilistic harmonic power flow, the expected THD_v of each node is calculated, which is shown in Figure 12.

From Figure 12, it can be observed that Node 23, the EV charging station node, and its neighbor node, Node 24, have very high expected THD_v. Generally, it is obvious that the expected THD_v of the node near Node 23 is high while the expected THD_v of the node far from Node 23 is low.

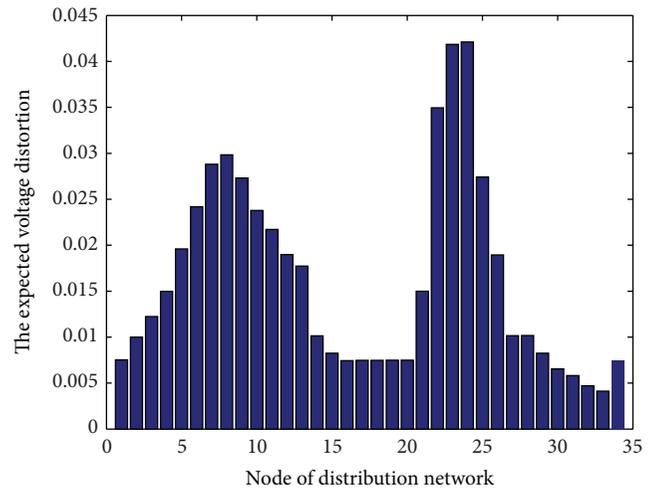


FIGURE 12: The expected voltage distortion of each node.

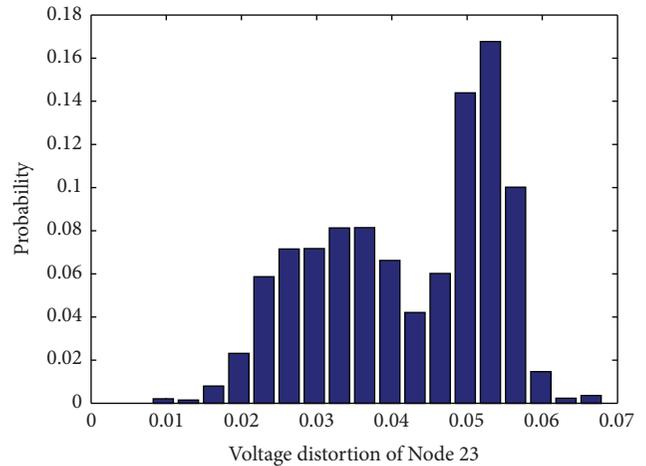


FIGURE 13: The probability distribution of voltage distortion in node 23.

The probability distribution of THD_v of Node 23 is shown in Figure 13.

Table 7 lists the nodes whose THD_v is over range. There are 11 nodes whose probability of THD_v is over the range; 7 of them have an overranging probability that is higher than

TABLE 7: The detailed information of THD_v over the range.

Node	5	6	7	8	9	10	11	22	23	24	25
The probability of THD _v over the range/%	0.4	29.2	49.2	53.6	49.2	29.6	3.8	68.1	76.1	76.1	49.2

TABLE 8: The detailed information of THD_v over the range in different arrival rate.

Arrival rate	Node	6	7	8	9	10	22	23	24	25		
0.1	The probability of THD _v over the range/%	0.9	5.5	9.3	3.6	0.9	25.8	53.9	53.9	3.7		
Arrival rate	Node	5	6	7	8	9	10	11	22	23	24	25
0.2	The probability of THD _v over the range/%	0.8	43.8	93.8	96.6	89.9	45.6	3.4	100	100	100	89.9

TABLE 9: The detailed information of THD_v over the range when penetration rate is 30%.

Node	4	5	6	7	8	9	10	11
The probability of THD _v over the range	0.048	0.454	0.700	0.889	0.889	0.889	0.747	0.668
Node	12	13	21	22	23	24	25	
The probability of THD _v over the range	0.239	0.187	0.048	0.982	0.994	0.994	0.889	

TABLE 10: The detailed information of THD_v over the range in different accessing location.

Accessing location	Node	7	8	9	10	11	12	13	22	23	24	25	26
10	The probability of THD _v over the range/%	1.4	1.4	3.2	8	5.1	1.4	0.6	1.4	1.6	1.6	3.2	1.4
Accessing location	Node	6	7	8	9	10	22	23	24	25			
23	The probability of THD _v over the range/%	0.9	5.5	9.3	3.6	0.9	25.8	53.9	53.9	3.7			
Accessing location	Node	18	19	20	34								
34	The probability of THD _v over the range/%	1.4	1.4	1.4	23.7								

49%. The result shows that the integration of EV charging station significantly affects the safety and power quality of the distribution network. So filter devices need to be installed to diminish the influence.

4.3. Influence Factors of Charging Station. In this part of research, the influences of the arrival rate, the penetration rate, and the accessing location of EV charging station on the probabilistic harmonic power flow are analyzed as follows.

(1) *The Arrival Rate.* First, the influence of arrival rate is studied. Calculations have been made when arrival rate is set to different values, 0.1 and 0.2. The results of the probabilistic harmonic power flow during daytime are listed in Table 8.

With the increase of the arrival rate, the harmonic distortion of the network becomes worse: two new nodes, Node 5 and Node 11, become THD_v overranging nodes. The influenced area becomes larger due to the increase of the arrival rate and so does the THD_v overranging probability. In fact, the overranging probability even reaches 100% at Nodes 22, 23, and 24. Therefore, it is easy to conclude that the harmonic pollution is more significant when the arrival rate increases.

(2) *The Penetration Rate.* The harmonic impact on the distribution network is studied during 24 hours when the penetration rate increases to 30%. The results of probabilistic harmonic power flow are listed in Table 9.

Compared to the results in Table 7, the number of THD_v overranging nodes has increased by 4, the influence area is

larger than the original result, and the probability of THD_v over the range changes from 0.761 to 0.994 at Node 23. It can be seen that the impact on the power quality of the distribution network will be more significant when the penetration rate increases.

(3) *Accessing Location.* Situations where the EV charging station accesses Nodes 10, 23, and 34 are compared. The results are listed in Table 10.

From the results, two conclusions which are useful in EV charging station planning can be drawn. The influenced area when the EV charging station accesses to Node 10 is larger than that of Node 23. It can be interpreted in a way that Node 10 is the main feeder of distribution network so accessing it will result in a larger influenced area compared with accessing a branch feeder. When accessing location is Node 34, the influenced area and the degree of influence are smaller than that when accessing the other two nodes. That is because Node 34 is close to Node 15, which is connected with reactive power compensator that has an inhibitory effect on harmonics. Therefore, it can be concluded that it is better to set the charging station close to reactive power compensation devices or away from the main feeder.

5. Conclusions

A new method for studying the working status of EV charging stations is proposed in this paper, while considering the random leaving factor, the electricity price, and the waiting time. This method is proved to be effective through

the comparison with the conventional queuing theory model. In this paper, this method is used to analyze the characteristics of harmonics from the charging station. The probabilistic harmonic power flow is used to analyze the impact of charging station's harmonics on the distribution network. The proposed model can be directly applied in some other fields that involve charging station's state simulation. In addition, other factors can also be considered in the simulation process. Since the method proposed in this paper is a simulation process, further study will focus on the mathematical theory of the charging behavior.

Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

Acknowledgment

This work was supported by National Natural Science Foundation of China (no. 51277141).

References

- [1] W. Sun, H. Gao, and O. Kaynak, "Adaptive backstepping control for active suspension systems with hard constraints," *IEEE/ASME Transactions on Mechatronics*, vol. 18, no. 3, pp. 1072–1079, 2013.
- [2] W. Sun, H. Gao, and B. Yao, "Adaptive robust vibration control of full-car active suspensions with electrohydraulic actuators," *IEEE Transactions on Control Systems Technology*, vol. 21, no. 6, pp. 2417–2422, 2013.
- [3] W. Sun, Z. Zhao, and H. Gao, "Saturated adaptive robust control for active suspension systems," *IEEE Transactions on Industrial Electronics*, vol. 60, no. 9, pp. 3889–3896, 2013.
- [4] M. Yilmaz and P. T. Krein, "Review of battery charger topologies, charging power levels, and infrastructure for plug-in electric and hybrid vehicles," *IEEE Transactions on Power Electronics*, vol. 28, no. 5, pp. 2151–2169, 2013.
- [5] C. Shen, P. Shan, and T. Gao, "A comprehensive overview of hybrid electric vehicles," *International Journal of Vehicular Technology*, vol. 2011, Article ID 571683, 7 pages, 2011.
- [6] K. Clement-Nyns, E. Haesen, and J. Driesen, "The impact of charging plug-in hybrid electric vehicles on a residential distribution grid," *IEEE Transactions on Power Systems*, vol. 25, no. 1, pp. 371–380, 2010.
- [7] L. P. Fernández, T. G. S. Román, R. Cossent, C. Domingo, and P. Frias, "Assessment of the impact of plug-in electric vehicles on distribution networks," *IEEE Transactions on Power Systems*, vol. 26, no. 1, pp. 206–213, 2011.
- [8] P. Papadopoulos, S. Skarvelis-Kazakos, I. Grau, L. M. Cipcigan, and N. Jenkins, "Electric vehicles' impact on British distribution networks," *IET Electrical Systems in Transportation*, vol. 2, no. 3, pp. 91–102, 2012.
- [9] C. H. Dharmakeerthi, N. Mithulananthan, and T. K. Saha, "Impact of electric vehicle fast charging on power system voltage stability," *International Journal of Electrical Power & Energy Systems*, vol. 57, pp. 241–249, 2014.
- [10] S. Deilami, A. S. Masoum, P. S. Moses, and M. A. S. Masoum, "Voltage profile and THD distortion of residential network with high penetration of plug-in electrical vehicles," in *Proceedings of the IEEE PES Innovative Smart Grid Technologies Conference Europe (ISGT Europe '10)*, pp. 1–6, IEEE, October 2010.
- [11] L. E. Bremermann, M. Matos, J. A. Peças Lopes et al., "Electric vehicle models for evaluating the security of supply," *Electric Power Systems Research*, vol. 111, pp. 32–39, 2014.
- [12] J. C. Gómez and M. M. Morcos, "Impact of EV battery chargers on the power quality of distribution systems," *IEEE Transactions on Power Delivery*, vol. 18, no. 3, pp. 975–981, 2003.
- [13] S. Deilami, A. S. Masoum, P. S. Moses, and M. A. S. Masoum, "Real-time coordination of plug-in electric vehicle charging in smart grids to minimize power losses and improve voltage profile," *IEEE Transactions on Smart Grid*, vol. 2, no. 3, pp. 456–467, 2011.
- [14] P. Richardson, D. Flynn, and A. Keane, "Optimal charging of electric vehicles in low-voltage distribution systems," *IEEE Transactions on Power Systems*, vol. 27, no. 1, pp. 268–279, 2012.
- [15] P. Han, J. Wang, Y. Han, and Y. Li, "Resident plug-in electric vehicle charging modeling and scheduling mechanism in the smart grid," *Mathematical Problems in Engineering*, vol. 2014, Article ID 540624, 8 pages, 2014.
- [16] M. A. S. Masoum, P. S. Moses, and S. Deilami, "Load management in smart grids considering harmonic distortion and transformer derating," in *Proceedings of the Innovative Smart Grid Technologies Conference (ISGT '10)*, pp. 1–7, IEEE, January 2010.
- [17] M. T. Au and J. V. Milanović, "Establishing harmonic distortion level of distribution network based on stochastic aggregate harmonic load models," *IEEE Transactions on Power Delivery*, vol. 22, no. 2, pp. 1086–1092, 2007.
- [18] C. Jiang, R. Torquato, D. Salles et al., "Method to assess the power-quality impact of plug-in electric vehicles," *IEEE Transactions on Power Delivery*, vol. 29, no. 2, pp. 958–965, 2014.
- [19] M. A. S. Masoum, S. Deilami, and S. Islam, "Mitigation of harmonics in smart grids with high penetration of plug-in electric vehicles," in *Proceedings of the IEEE PES General Meeting*, pp. 1–6, IEEE, July 2010.
- [20] M. Huang, S. F. Huang, and J. C. Jiang, "Harmonic study of electric vehicle chargers," *Journal of Beijing Jiaotong University*, vol. 32, no. 5, pp. 85–88, 2008.
- [21] A. Baba and S. Adachi, "State of charge estimation of HEV/EV battery with series Kalman filter," in *Proceedings of the 51st Annual Conference on the Society of Instrument and Control Engineers of Japan (SICE '12)*, pp. 845–850, August 2012.
- [22] U. N. Bhat, *An Introduction to Queueing Theory: Modeling and Analysis in Applications*, Springer, New York, NY, USA, 2008.
- [23] R. Garcia-Valle and J. G. Vlachogiannis, "Letter to the editor: electric vehicle demand model for load flow studies," *Electric Power Components and Systems*, vol. 37, no. 5, pp. 577–582, 2009.
- [24] G. Li and X. Zhang, "Modeling of plug-in hybrid electric vehicle charging demand in probabilistic power flow calculations," *IEEE Transactions on Smart Grid*, vol. 3, no. 1, pp. 492–499, 2012.
- [25] W. Su and M. Chow, "Performance evaluation of an EDA-based large-scale plug-in hybrid electric vehicle charging algorithm," *IEEE Transactions on Smart Grid*, vol. 3, no. 1, pp. 308–315, 2012.
- [26] A. A. Romero, H. C. Zini, G. Rattá, and R. Dib, "Harmonic load-flow approach based on the possibility theory," *IET Generation, Transmission and Distribution*, vol. 5, no. 4, pp. 393–404, 2011.

- [27] P. F. Ribeiro, *Time-Varying Waveform Distortions in Power Systems*, Wiley, 2009.
- [28] W. H. Kersting, "Radial distribution test feeders," in *Proceedings of the IEEE Power Engineering Society Winter Meeting*, pp. 908–912, February 2001.

Research Article

Multifeature Fusion Vehicle Detection Algorithm Based on Choquet Integral

Wenhui Li,^{1,2,3} Peixun Liu,¹ Ying Wang,^{1,2} and Hongyin Ni¹

¹ College of Computer Science and Technology, Jilin University, Changchun 130012, China

² State Key Laboratory of Automotive Simulation and Control, Jilin University, Changchun 130022, China

³ Key Laboratory of Symbolic Computation and Knowledge Engineering of Ministry of Education, Jilin University, Changchun 130012, China

Correspondence should be addressed to Ying Wang; wangying_jlu@163.com

Received 13 May 2014; Accepted 25 June 2014; Published 24 July 2014

Academic Editor: Weichao Sun

Copyright © 2014 Wenhui Li et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Vision-based multivehicle detection plays an important role in Forward Collision Warning Systems (FCWS) and Blind Spot Detection Systems (BSDS). The performance of these systems depends on the real-time capability, accuracy, and robustness of vehicle detection methods. To improve the accuracy of vehicle detection algorithm, we propose a multifeature fusion vehicle detection algorithm based on Choquet integral. This algorithm divides the vehicle detection problem into two phases: feature similarity measure and multifeature fusion. In the feature similarity measure phase, we first propose a taillight-based vehicle detection method, and then vehicle taillight feature similarity measure is defined. Second, combining with the definition of Choquet integral, the vehicle symmetry similarity measure and the HOG + AdaBoost feature similarity measure are defined. Finally, these three features are fused together by Choquet integral. Being evaluated on public test collections and our own test images, the experimental results show that our method has achieved effective and robust multivehicle detection in complicated environments. Our method can not only improve the detection rate but also reduce the false alarm rate, which meets the engineering requirements of Advanced Driving Assistance Systems (ADAS).

1. Introduction

As an important part of the intelligent transportation system (ITS), the Advanced Driving Assistance Systems (ADAS) can significantly improve the driving safety. Forward Collision Warning Systems (FCWS) and Blind Spot Detection Systems (BSDS) are principal portions of ADAS, and their performance depends on the real-time capability, accuracy, and robustness of the vehicle detection method. Recently, with the increasing maturity of visual sensors, vision-based vehicle detection has become a hot topic in the field of intelligent vehicle. There are plenty of approaches proposed for the day time vehicle detection. These methods can be divided into the following categories: methods based on prior knowledge, such as shadow-based [1, 2], taillight-based [1, 2], horizontal (vertical) edge-based [2–4], and symmetry-based vehicle detection method [2]; methods based on stereo vision; this type of method detects vehicles by using

the three-dimensional information. The most widely used methods are inverse perspective transformation (IPM) based method and disparity map based method [1]; template-based methods use predefined patterns of vehicle class and perform correlation between the image and the template [1]; the main detection steps of appearance-based methods are as follows: the appropriate descriptors are first used for representing vehicles in the image; then the machine learning methods are used to train these descriptors. Much processes have been made in appearance-based vehicle detection, such as algorithm based on HOG + AdaBoost [5], Haar + HMM [6], Haar + AdaBoost [7–9], HOG + SVM [10], PCA – ICA + GMM [11], and minimum Mahalanobis distance classifier [12]. The method based on the motion information detects vehicles by using the motion information between vehicles and scenes, such as finding out vehicles by calculating the change of optical flow information which is caused by the relative motion of vehicles or scenes [13].

To improve the accuracy of vehicle detection methods, several of above methods are combined together to detect vehicles. Lin et al. [3] applied the SURF and edge features to represent the vehicle and, combining with probabilistic methods, their methods have achieved vehicle detection in the blind spot area. Chen et al. [6] first used a road modelling method to confine detection regions, and then Haar-like features and eigencolours were used for detecting vehicles. Finally, a tackling method was used. Tehrani Niknejad et al. [10] proposed a deformable vehicle model based on HOG feature; the method can achieve the adaptive threshold vehicle detection under urban roads. Wang and Lien [11] proposed a vehicle detection method based on a statistical model of local feature. They applied the public dataset Caltech Cars (Rear) [16] to test their method. Alonso et al. [12] proposed a vehicle detection method based on multidimensional classification. They represented vehicles in form of rectangular subregions based on the robust classification of features vectors result of a combination of multiple morphological vehicle features. Their method can detect vehicles with very different aspect-ratio, color, and size. Chang and Cho [8] presented a vehicle detection algorithm based on combination of Haar feature and online boosting. Their algorithm has realized vehicle detection in various environments. Sivaraman and Trivedi [9] proposed an active-learning framework based on Haar feature and AdaBoost for vehicle detection on the highway. Jazayeri et al. [13] proposed an optical flow and hidden Markov model-based vehicle detection method which mode the locations and motion information of vehicles in the image level. Their method can deal with the vehicle identification problem under the scene of changing illumination and environment well.

Vehicle detection method based on a single feature can quickly detect vehicles in images. However, using single feature method may cause a lot of false alarms, because single feature only describes one certain characteristic of vehicles. Most of the appearance-based vehicle detection methods' performance excessively depends on the number and scale of training samples. Various samples in different situations are needed to generate more powerful classifiers. In addition, detecting vehicles in images using appearance-based methods which has to scan the whole image requires excessive calculation and cannot meet the real-time requirement of FCW. To solve the above-mentioned problem, a widely used method is multifeature fusion which combines several single feature-based algorithms together by using voting method. This can significantly reduce the false alarm rate, but the detection rate is reduced either. In recent years, mathematical theory has been widely used for improving the performance of complex vehicle systems. Much process has been made in the field of mathematical modeling and control methods [17–23], such as adaptive back stepping control for active suspension systems with hard constraints [17], saturated adaptive robust control for active suspension systems [18], and adaptive robust vibration control of full-car active suspensions with electrohydraulic actuators [19]. Choquet integral is a widely used method in data fusion [24–26]; it can seek the maximum consistency of decision from the consistency and conflict detection results of multiple features. To improve

the performance of vehicle detection algorithm and to solve problems above, we propose a multifeature fusion vehicle detection algorithm based on Choquet integral. Experiment results show that our multifeature fusion method will not only improve the detection rate but also reduce the false alarm rate.

Figure 1 illustrates the workflow of our approach. The rest of the paper is organized as follows. Section 2 briefly introduces the shadow-based vehicle region of interest (ROI) detection method. In Section 3, vehicle taillight feature similarity measure, vehicle symmetry feature similarity measure, and HOG + AdaBoost feature similarity measure are presented, respectively. Then our multifeature fusion vehicle detection algorithm based on Choquet integral is introduced in Section 4. Experiment results for the proposed method are shown in Section 5; finally Section 6 draws conclusions.

2. Shadow-Based Vehicle ROI Detection

The shadow-based vehicle detection algorithm is usually applied to extract the vehicle ROIs in the whole images for reducing computation complexity [1]. We have developed a shadow-based vehicle detection method, and the basic principle of the method is that regions underneath vehicles are distinctly darker than any other regions on an asphalt road. The grayscale of pixels in shadow regions is much lower than that in any other regions in the same image. Grayscale histogram (GH) can reflect the whole image grayscale distribution well. The grayscale of vehicle shadow pixels belongs to the lower parts of GH. So we can detect the shadow regions underneath vehicles by segmenting GH with a threshold th_BW . Figure 2(a) is a vehicle image from Caltech Cars (Rear) [27]. Black regions in Figure 2(b) are shadow regions segmented by setting th_BW to 0.1. The green lines in Figure 2(c) are vehicle shadow lines detected by shadow-based vehicle detection method.

3. Feature Similarity Measure

To make full use of the Choquet integral in our multifeature fusion vehicle detection framework, each single feature should be first represented in form of fuzzification before calculating the Choquet integral. After this phase, the algorithm can fuzz the output of each single feature; then the result can be determined by using the fuzzy judgment instead of direct judgment. Therefore, in this section, we detailedly introduce three feature-based vehicle detection methods and their feature similarity measure functions.

3.1. Vehicle Taillight Feature Similarity Measure. The red taillights and braking lights are important features for detecting the rear-view vehicle. Taillight-based feature provides an important criterion for our multifeature fusion vehicle detection framework. The RGB components of pixels in taillight regions are obviously different from the other parts of vehicle ROI (except red cars). Therefore, by following this rule, we present a similarity measure method based on color feature of vehicle taillights. First, taillight regions in

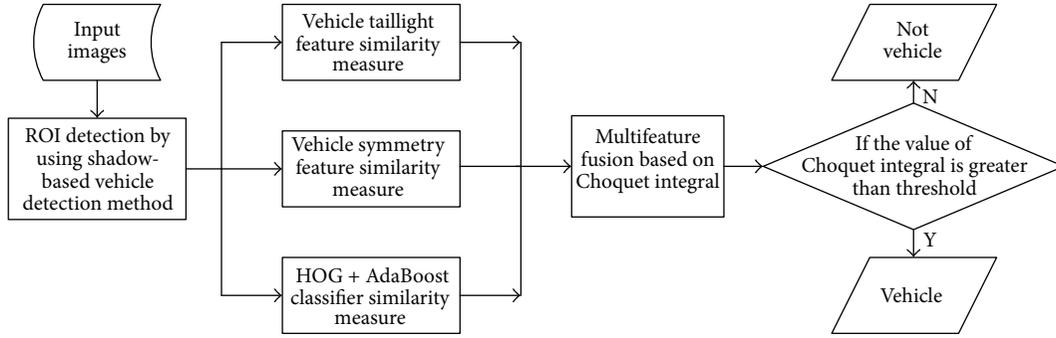


FIGURE 1: Framework of our approach.

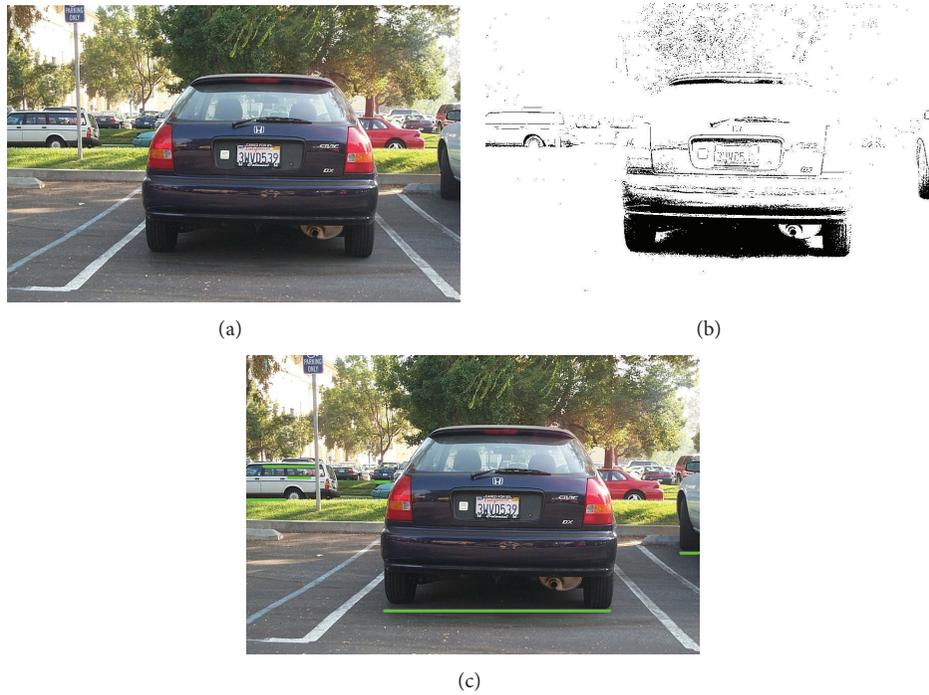


FIGURE 2: Result of shadow-based vehicle ROI detection method.

the vehicle ROI are detected by threshold value segmentation method. The key threshold of method can be acquired by analyzing RGB components from images of taillights and the other parts of vehicle. The collection of images for setting the threshold should be large enough and includes different vehicles and various scenes. We acquire the R , G , and B components distributions by analyzing the public collection [27] and images captured by our camera. As shown in Figure 3(b), differences between the R component and the G component of the other regions of vehicle are mainly distributed on the range of $[1, 31]$. It is different from the values of $|R - G|$ in taillight regions illustrated in Figure 3(a); therefore, the taillight regions of vehicle ROI can be detected by setting a certain threshold $Th_Taillight$:

$$I_{Taillight}(x, y) = \begin{cases} 255, & \text{if } |R(x, y) - G(x, y)| \geq Th_Taillight \\ 0, & \text{otherwise.} \end{cases} \quad (1)$$

Figure 4(a) is the vehicle ROI detected by the shadow-based vehicle detection method. Figure 4(b) is the binary image of taillights detected by employing (1) on the vehicle ROI. Canny-based edge detection method is used to detect the edges of taillights in Figure 4(b), and Figure 4(c) is the edge image of Figure 4(b). Then the connected domains in Figure 4(c) are extracted. The input images of connected domain extraction method are binary and edge image which are illustrated as Figures 4(b) and 4(c), respectively.

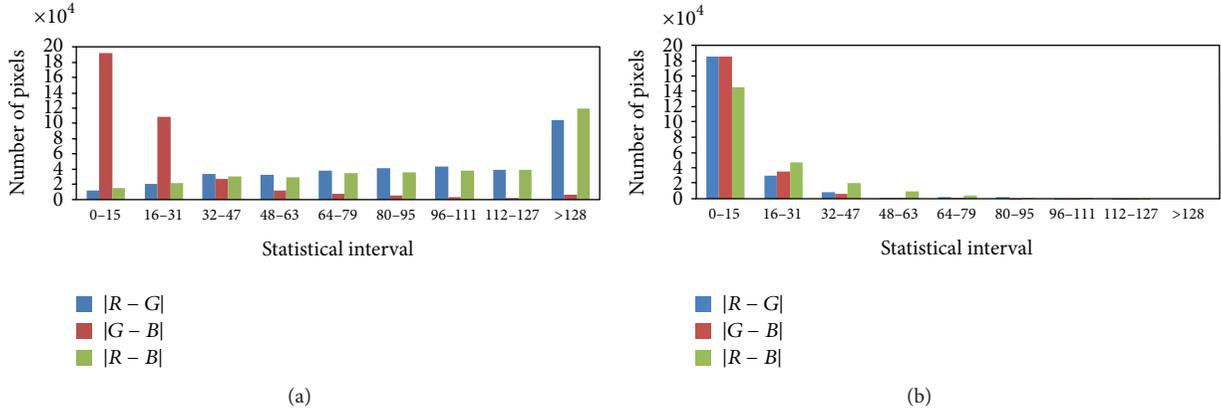


FIGURE 3: Comparison of RGB components between taillights and the other parts of vehicles.

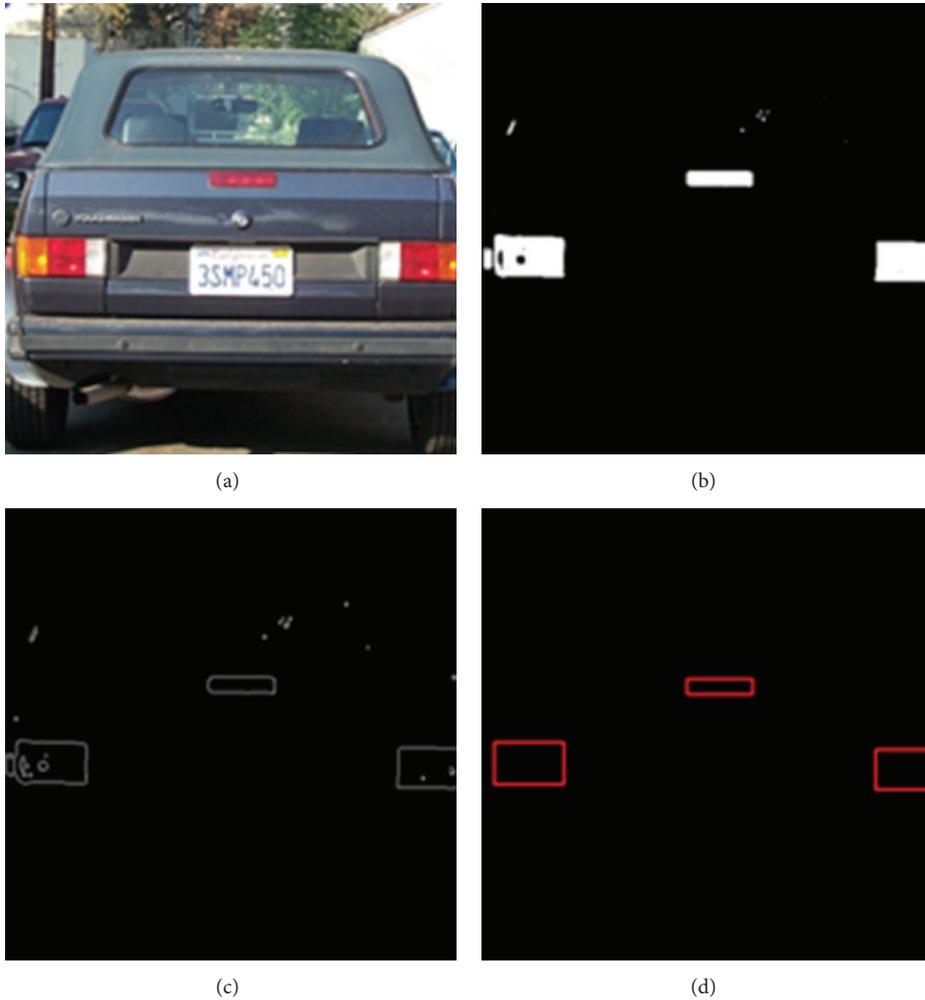


FIGURE 4: Extraction of taillight areas.

Finally, the minimum circumscribed rectangles (MCR) of connection domains are calculated. The detected MCRs are illustrated as the red rectangles in Figure 4(d).

Each MCR of connected domain is represented by the left top point $\text{MinPoint}_i(x, y)$ and the right top point

$\text{MaxPoint}_i(x, y)$ of MCR. Two left top points in vehicle ROI can form a straight line; the slope of straight line is defined as

$$K_{\text{condomains}}^i = \frac{y_{\text{minpoint}}^i - y_{\text{minpoint}}^{i-1}}{x_{\text{minpoint}}^i - x_{\text{minpoint}}^{i-1}}. \quad (2)$$

The distance between each MCR is represented as (3). The two taillights of vehicle are usually on a horizontal line, and the thresholds th_L and th_H can get rid of the straight lines that are not horizontal or almost horizontal:

$$\text{width_taillights} = \begin{cases} \text{Max}(x_{\text{maxpoint}}^i - x_{\text{minpoint}}^{i-1}), & \text{if } th_L \leq K_{\text{condomains}}^i \leq th_H, \\ 0, & \text{otherwise.} \end{cases} \quad (3)$$

Definition 1. The taillight feature similarity measure function $C_{\text{tailCoeff}}$ is defined as

$$C_{\text{tailCoeff}} = \frac{\text{width_taillights}}{\text{width_ROI}}. \quad (4)$$

3.2. Vehicle Symmetry Feature Similarity Measure. The symmetry measure is a statistic to describe the symmetry of target. Vehicles are obviously symmetrical objects; therefore, we use the symmetry feature as a similarity measure in our algorithm. According to the symmetry-based method described in [28], we use the symmetry measure method based on normalized entropy to calculate the symmetry value of each vehicle ROI. The symmetry measure is described as (5), where $S(x_s)$ is the symmetry measure of target. $E(l)$ is the information entropy, which is also the mathematical expectation of information content. E_m is the max value of information entropy. Consider

$$\begin{aligned} \bar{s}_g &= \frac{[(S(x_s) + 1) / 2 + E(l) / E_m]}{2} \\ &= \frac{S(x_s) \times E_m + 2 \times E(l) + E_m}{4 \times E_m}. \end{aligned} \quad (5)$$

Definition 2. The symmetry feature similarity measure function C_{symCoeff} is defined as

$$C_{\text{symCoeff}} = \begin{cases} \bar{s}_g, & 0 \leq \bar{s}_g \leq 1, \\ 1, & \bar{s}_g > 1. \end{cases} \quad (6)$$

3.3. HOG and AdaBoost Classifier Feature Similarity Measure. The histogram of oriented gradient (HOG) is a descriptor of feature which has been widely used in object detection. Zhu et al. [29] introduced an efficient pedestrian detection method based on HOG and AdaBoost. In our previous work, we use HOG feature to detect pedestrian [14]. The HOG feature is represented by calculating the histogram of oriented gradient of local region in the image. First, the image is divided into a plurality of grids according to a certain size; these grids are called BLOCK which are illustrated as in Figure 5(a). Then each BLOCK is divided into four regions which are called CELL. Each CELL projects an orientation-based histogram which includes nine bins. In this histogram, the horizontal ordinate is a range of direction angles which divide 180° into nine equal parts, and the vertical coordinates are an accumulation of each angle range. Finally, a 36D feature vector named BLOCK is formed. Due to the strong edge

feature of vehicles, we employ the HOG feature to represent vehicles; then the AdaBoost-based algorithm [30] is applied to generate weak classifiers.

In this paper, the training samples of generating HOG + AdaBoost classifiers are images captured from actual driving environments. Vehicle regions of these images are positive samples, and other regions of images are negative samples. The amount of positive samples and negative samples are both 10000. These samples are normalized to the same size (30×30). Screenshots of samples are shown in Figures 5(b) and 5(c). There are two phases to employ the HOG + AdaBoost classifier and the training and the detection phase. In the training phase, we extract HOG features by applying CELL size of 5×5 , 10×10 , and 15×15 , respectively; the scanning step size is three pixels, and the weak classifiers are selected by AdaBoost algorithm. After training, we use the samples (positive 10000 and negative 10000) which are different from that of the training phase to test the weak classifiers. The ROC curves of HOG + AdaBoost algorithm under three different CELL sizes are illustrated as in Figure 6; the performance of HOG + AdaBoost classifiers whose CELL size is 15×15 is the best among these three types; therefore, we set the CELL size to 15×15 in our further experiments.

To enhance the performance of HOG + AdaBoost classifiers, inspired by method in [9], the active-learning based HOG + AdaBoost framework is used by following the steps in the Active-Learning Framework. The advantage of this framework is that you are only adding negative samples that would otherwise be causing false positives. There is no point in adding more negative samples that are handled by the original training anyways.

Active-Learning Framework.

Step 1. Train HOG + AdaBoost classifiers using the 10000 positive samples and 30000 negative samples.

Step 2. Run the algorithm by using well-trained HOG + AdaBoost classifiers on a large video set (not the training set from Step 1).

Step 3. Any false positives from the run in Step 2 can be put in the negative set.

Step 4. Retrain the algorithm using the original true positive set and the updated negative set (negatives from both Step 1 and Step 3).

Step 5. This can be repeated as many times as appropriate, using new video on each iteration.

In detection phase, each vehicle ROI detected by shadow-based vehicle detection method is resized to the same size of the training sample; the HOG feature is extracted in the same way of training phase. Then use the well-trained classifiers to identify the vehicle ROI; the classification value of each vehicle ROI is calculated by

$$\text{hogadbCoeff} = \sum_{i=1}^T \alpha_i \cdot h_i^{\text{base}} \cdot th_strong. \quad (7)$$

Most of AdaBoost-based object detection methods decide whether the ROI is object or interference by judging whether

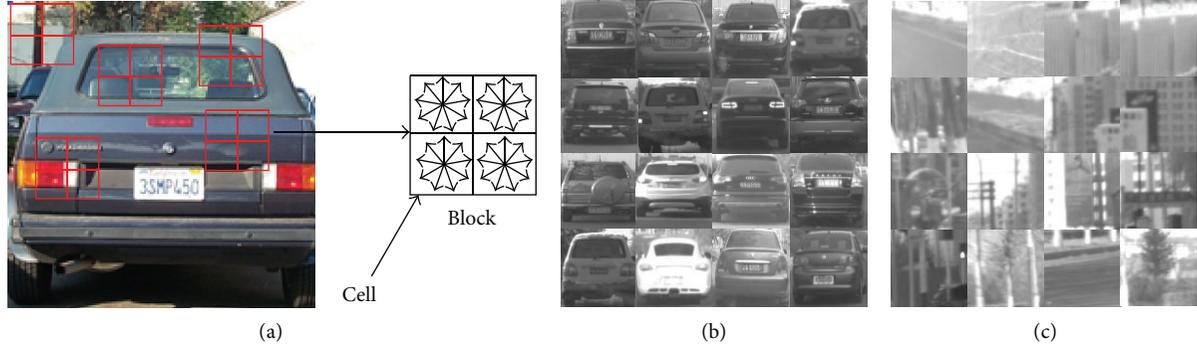


FIGURE 5: Some samples of training dataset.

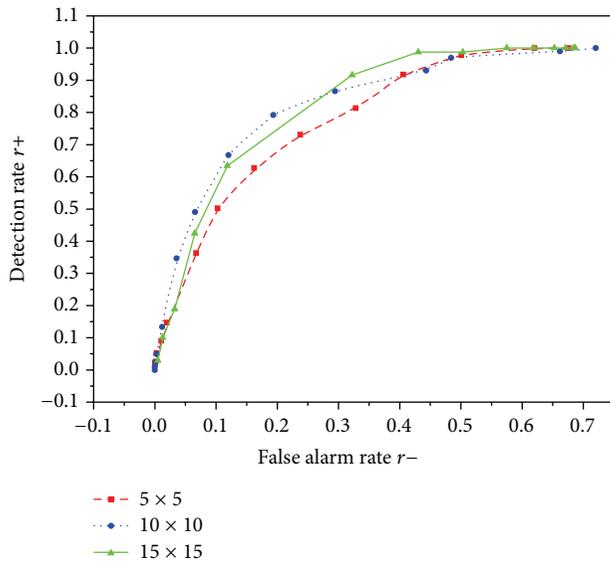


FIGURE 6: ROC curves of HOG + AdaBoost algorithm under three different HOG CELL sizes.

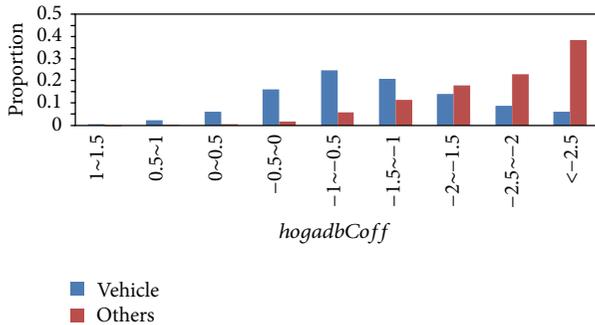


FIGURE 7: Statistic distribution of hogadbCoeff belonging to vehicles and interferences.

the symbol of the hogadbCoeff is positive or negative. This way is not suitable to employ the HOG and AdaBoost-based classifiers in our Choquet integral-based multifeature fusion vehicle detection framework. To represent the hogadbCoeff in form of probability, we first test the well-trained classifiers

TABLE 1: Mapping table between $C_{symCoeff}$ and $hogadbCoeff$.

$hogadbCoeff$	<-2.5	$-2.5\sim-2$	$-2\sim-1.5$	$-1.5\sim-1$	$-1\sim-0.5$
$C_{symCoeff}$	0	0.65	0.7	0.75	0.8
$hogadbCoeff$	$-0.5\sim 0$	$0\sim 0.5$	$0.5\sim 1$	$1\sim 1.5$	>1.5
$C_{symCoeff}$	0.85	0.9	0.95	0.99	1

by using the testing sample set which is different from the training sample set. And then the statistic distribution of hogadbCoeff is calculated. Finally the mapping table between the HOG and AdaBoost classifier feature similarity measure value $C_{hogadbCoeff}$ and the hogadbCoeff is formed. The statistic distributions of hogadbCoeff belonging to vehicles and interferences are illustrated in Figure 7; we use the algorithm precision $p+$ corresponding to interval of hogadbCoeff to be the $C_{hogadbCoeff}$; the precision is defined as (16) in this paper; the mapping table is created as Table 1.

Definition 3. The HOG + AdaBoost classifier feature similarity measure function $C_{hogadbCoeff}$ is defined as Table 1.

4. Multifeature Fusion Vehicle Detection Algorithm Based on Choquet Integral

In this paper, fuzzy integral theory is applied to vehicle detection in complex scenarios. First, the basic theory of Choquet integral is introduced here. And then the fuzzy measure of each feature is defined. Finally, the features of taillight, symmetry, and HOG + AdaBoost classifier are fused by Choquet integral of fuzzy theory. The brief concepts of Choquet integral and the fuzzy measure used in our algorithm are followed from the concepts in [24–26].

Definition 4. Let X be a finite set, and Y is a power set which is composed of subsets of X , $g : Y \rightarrow [0, \infty]$ is the mapping function from the power set Y to the range of $[0, \infty]$. If g satisfies the following three conditions, g is a fuzzy measure on Y .

- (1) Boundedness: $g(\emptyset) = 0$.
- (2) Monotonicity: $\forall A, B \in Y$, if $A \subseteq B$, then $g(A) \leq g(B)$.

- (3) Continuity: if $\forall A_n \in Y$, and $\{A_i \mid i \in [1, +\infty)\}$ is monotonous. This is also represented in the form of $A_1 \subseteq A_2 \subseteq \dots \subseteq A_n \dots$ or $A_1 \supseteq A_2 \supseteq \dots \supseteq A_n \dots$, then $\lim_{i \rightarrow \infty} g(A_i) = g(\lim_{i \rightarrow \infty} A_i)$.

The fuzzy measure which is widely applied in multifeature fusion is the regular fuzzy measure: if $X \in Y$ and $g(X) = 1$, the fuzzy measure g is regular.

Definition 5. If the fuzzy measure satisfies the following conditions: $\forall A, B \in Y, A \cap B = \emptyset$, if there exists a constant value $\lambda, \lambda > -1$ satisfying $g(A \cup B) = g(A) + g(B) + \lambda g(A)g(B)$, then g is a λ -fuzzy measure. λ can be calculated by (8), where $g^i = g(\{x_i\})$; it is used to indicate the importance of a single feature classifier for the final evaluation, where $x_i \in X = \{x_1, x_2, \dots, x_n\}$. Consider

$$1 + \lambda = \prod_{i=1}^n (1 + \lambda \times g^i). \quad (8)$$

Definition 6. $f : X \rightarrow [0, 1]$ is a nonnegative function defined on X , g is a fuzzy measure defined on power set Y , and then Choquet integral of function f on X with respect to fuzzy measure is defined by

$$\int f dg = \int_0^\infty g(Y_\mu) d\mu, \quad (9)$$

where $Y_\mu = \{x \mid f(x) \geq \mu, x \in X\}, \mu \in [0, \infty)$; the main idea of (9) is determining the value of Choquet integral using Riemann integral by an infinite approximation method. The definition of Choquet integral is as follows when X is a finite set:

$$\int f dg = \sum_{i=1}^n [f(x_{\theta(i)}) - f(x_{\theta(i-1)})] g(K_{\theta(i)}), \quad (10)$$

where θ is a permutation of the indices such that

$$0 = f(x_{\theta(0)}) \leq f(x_{\theta(1)}) \leq \dots \leq f(x_{\theta(n)}) \leq 1,$$

$$K_{\theta(i)} = \{x_{\theta(i)}, x_{\theta(i+1)}, x_{\theta(i+2)}, \dots, x_{\theta(i+n)}\}, \quad i = 1, 2, \dots, n. \quad (11)$$

When fuzzy measure g is a λ -fuzzy measure, any subset is defined by

$$\begin{aligned} g(K_{\theta(1)}) &= g(\{x_{\theta(1)}\}) = g_{\theta(1)}, \\ g(\{x_{\theta(i)}\}) &= g_{\theta(i)}, \\ g(K_{\theta(i)}) &= g_{\theta(i)} + g(K_{\theta(i-1)}) \\ &\quad + \lambda g_{\theta(i)} g(K_{\theta(i-1)}), \quad i = 2, \dots, n. \end{aligned} \quad (12)$$

To apply the Choquet integral to detect vehicles in complex environments, O is first initialized as the vehicle ROI detected by shadow-based vehicle detection algorithm. $F = \{\text{vehicle}, \text{interference}\}$ is a classification framework. $X = \{x_1, x_2, x_3\}$ is the feature set for detecting vehicle, where x_1, x_2 , and x_3 represent the vehicle symmetry feature, the vehicle

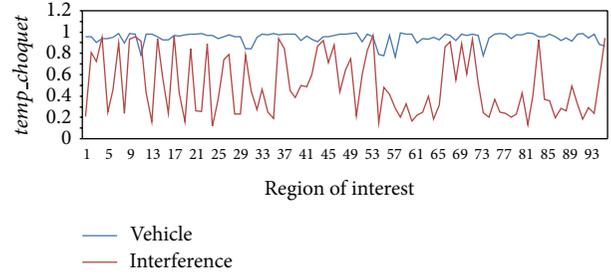


FIGURE 8: Comparison of temp_choquet between the vehicle and the interference.

taillight feature, and vehicle HOG + AdaBoost classifiers feature, respectively. Let $g : X \rightarrow [0, 1]$ be the fuzzy density of vehicle ROI O belonging to the class F_i ; define $g(x_i)$ to be the degree of importance of the feature x_i in deciding whether vehicle ROI is vehicle or interference. Define $g(x_1) = g(\{x_1\})$, $g(x_2) = g(\{x_2\})$, and $g(x_3) = g(\{x_3\})$; the higher the $g(x_i)$ is, the more important the feature x_i is. The fuzzy function f is defined in $[0, 1]$ so that $f(x_1) = C_{\text{tailCoeff}}$, $f(x_2) = C_{\text{symCoeff}}$, and $f(x_3) = C_{\text{hogadbCoeff}}$. To calculate the value of Choquet integral for each vehicle ROI, the features x_i in the set X are needed to be rearranged with respect to the order $f(x_1) \leq f(x_2) \leq f(x_3)$.

Main steps of our multifeature fusion vehicle detection algorithm based on Choquet integral are as follows.

Multifeature Fusion Vehicle Detection Algorithm Based on Choquet Integral.

Step 1. Calculate the fuzzy measure of each feature. We test each feature-based vehicle detection method on the same vehicle sample set, and, according to (16), the precision of each vehicle detection method can be acquired. Let the precision $p+$ be the fuzzy measure g corresponding to each feature-based method.

Step 2. Calculate λ by (8).

Step 3. Estimate the λ -fuzzy measure by (12).

Step 4. The Choquet integral value of each ROI temp_choquet can be calculated by (10) combining with three feature similarity measures.

Step 5. Decide whether the vehicle ROI is vehicle according to (13). As it is illustrated in Figure 8, the temp_choquet belonging to the vehicle and that belonging to the interference are much more different; the threshold Th_vehicle can be set according to Figure 8:

$$\text{isVehicle} = \begin{cases} 1, & \text{if temp_choquet} > \text{Th_vehicle}, \\ 0, & \text{otherwise.} \end{cases} \quad (13)$$

5. Experiment Results

To verify the performance of the algorithm, experimental platform has been built in c using OpenCV 1.0 library

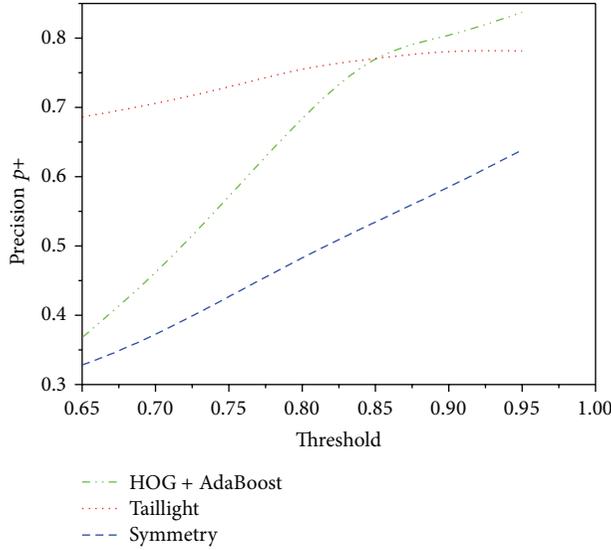


FIGURE 9: Algorithm precision under various thresholds.

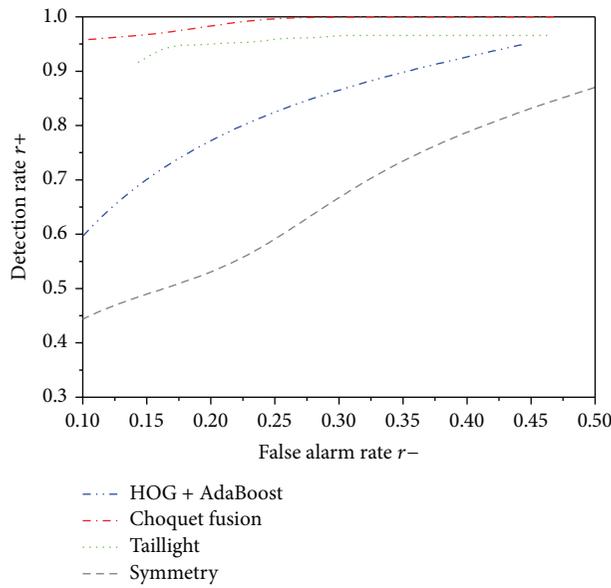


FIGURE 10: Algorithm ROC curves.

and Visual Studio 2010. The vehicle detection algorithm is performed on an Intel Core i7-3770GHZ PC. A part of vehicle images for testing are from the public test library Caltech Cars (Rear) [27]. The rest of vehicle images are captured in the real environments (parking lot and urban road) by using DEWETRON DEWE2-M4 (camera: DEWE-CAM-01, lens: computar M3Z1228C) and SAMSUNG GT-S7562 camera (5,000,000 pixels). There are 5 video sequences in our test datasets; the frames of our datasets are 1500, and the number of vehicles in datasets is 3219. The test images include single vehicle, multivehicle, and illumination changing in the scene. We use three indicators to measure the performance of algorithms: the detection rate r_+ , the false alarm rate r_- , and the algorithm precision p_+ . The criterion to determine

a “good” detection in this paper is the overlap of the detected bounding box versus the annotated bounding box. If the overlap is larger than a certain threshold, the detection is a “good” detection. Consider

$$r_+ = \frac{\text{Number of detected vehicles}}{\text{Total number of vehicles in testing data set}}, \quad (14)$$

$$r_- = \frac{\text{Number of false alarms}}{\text{Total number of vehicle ROI}}, \quad (15)$$

$$p_+ = \frac{\text{Number of detected vehicles}}{(\text{Number of detected vehicles} + \text{Number of false alarms})^{-1}}. \quad (16)$$

Experiment 1 (calculate the fuzzy measure of each algorithm). In our multifeature fusion vehicle detection algorithm, fuzzy measure of each feature-based algorithm is set according to the performance of its own. We test each feature-based vehicle detection method on the same vehicle sample set named JVTI. The images in JVTI are vehicle ROIs detected by shadow-based method which is introduced in Section 2. The positive samples of JVTI are vehicles, and the negative samples are interferences in JVTI. The numbers of positive and negative samples are 3219 and 6000. According to (16), the precision of each vehicle detection method can be acquired. Let the precision p_+ be the fuzzy measure g corresponding to each method. According to Figure 9, we can set the fuzzy measure of every algorithm.

Experiment 2 (performance of our multifeature fusion vehicle detection algorithm). After setting fuzzy measure of each feature-based algorithm, we apply the sample set JVTI to test our method and every feature-based algorithm. As shown in Figure 10, the single feature cannot meet the requirement of high detection rate and low false alarm rate. Our algorithm fuzzes the output of each single feature, and the result is determined by using the fuzzy judgment instead of direct judgment. At the same time, the use of fuzzy integral can give full consideration to the cooperation of multifeatures and the importance degree of each feature in the recognition phase. Therefore, our method outperforms each single feature. In our experiment, the average processing time (AVT) of our method can achieve 50 ms per frame when processing on the Caltech Rear public test images whose resolutions are 896×592 , which basically achieve real-time processing. And the processing time is 36 ms per frame on images whose resolutions are 640×480 . Part of results of our algorithm are shown in Figures 11 and 12. Figure 11 is the result of algorithm on Caltech Rear public vehicle images; we set the main thresholds as follows: $th_{BW} = 0.1$ and $Th_{vehicle} = 0.9$. Experimental results show that our method can detect well vehicles in different distances. The distances are different in Figures 12(a) and 12(b); the distances between vehicles and camera are from 3 m to 50 m. Figure 12(c) shows that our method can not only detect the single vehicle, but also handle the multivehicle detection problem. Figure 12(d) is the detection result on urban road.

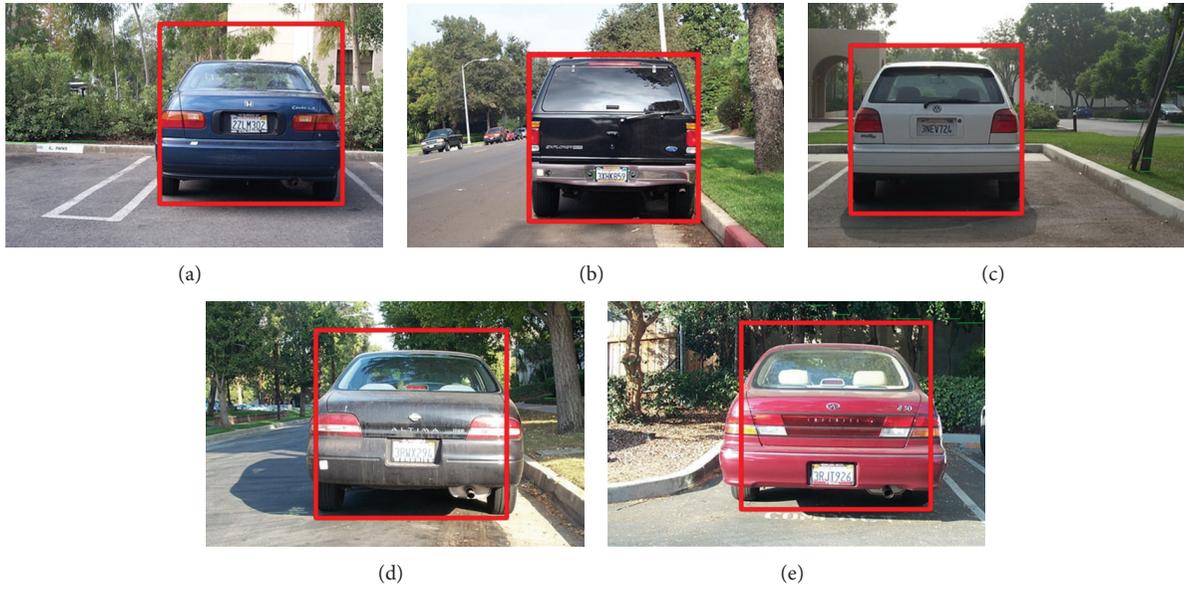


FIGURE 11: Detection results on Caltech Rear public vehicle images.

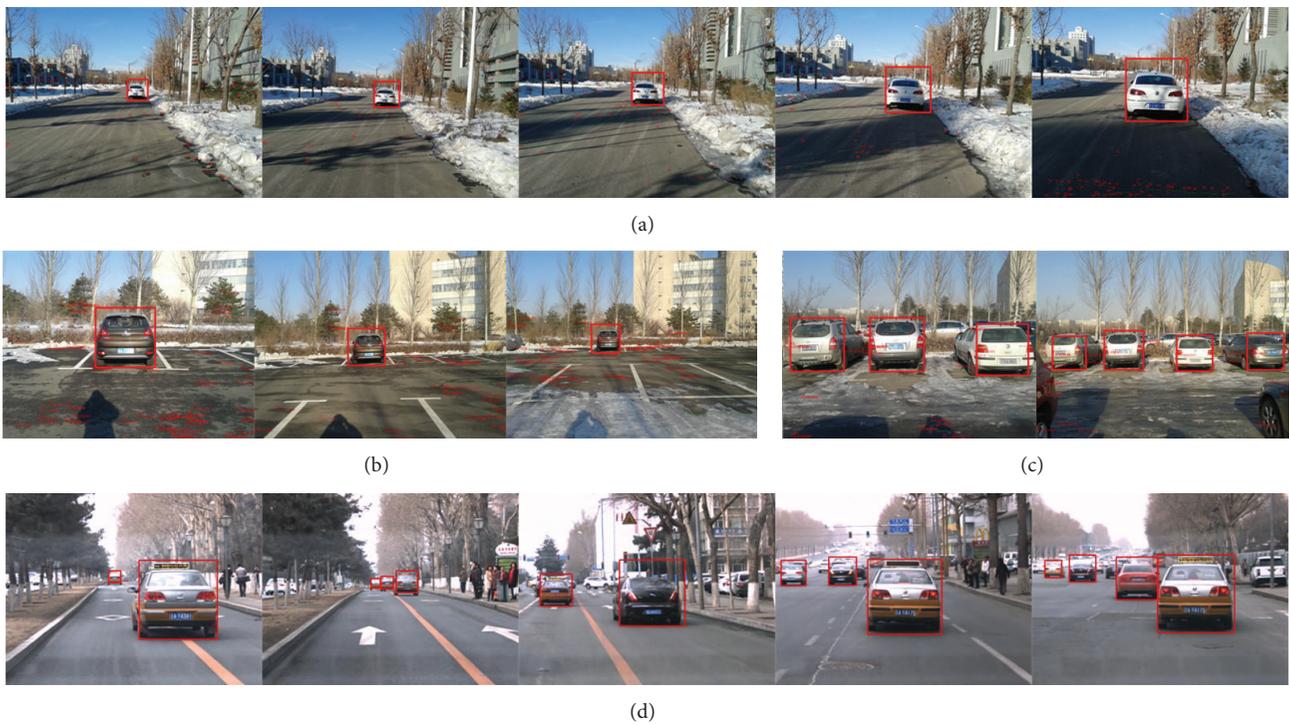


FIGURE 12: Detection results on our data set.

Experiment 3 (algorithm comparison). To verify the performance of our method, we compare our method to three feature-based methods, the voting method of these three feature-based methods, and vehicle detection methods in [11, 15, 16]. Algorithms used for comparison are all tested on the same collection (the public test library Caltech Cars (Rear) [27]). There are two ways to get the algorithms' results. On one hand, we download the source code from the websites

which have been provided in their articles to get the testing results. On the other hand, we directly use the testing results illustrated in the articles. Comparison result is shown in Table 2; it shows that the single feature-based methods can detect vehicle better, but the false alarm rate is also the highest. Although the voting method can reduce the false alarm rate, the detection rate is reduced either. Processing time is another indicator to measure the performance of

TABLE 2: Algorithm comparison.

Methods	Accuracy (DR/FAR)	AVT (ms/frame) (896 × 592)
Wang and Lien [11]	98%/0%	510
Li et al. [14]	98%/1%	500
Ali and Shah [15]	90.2%/0.6%	500
Taillight-based method	95.3%/23.4%	16
Symmetry-based method	86.1%/48%	15
HOG + AdaBoost	95.1%/44.8%	16
Voting method	83.3%/0%	45
Our method	95.5%/8.2%	50

algorithms; Ali and Wang's methods outperform our method in terms of accuracy, but the processing time of their methods is above 500 ms. Considering both the accuracy and the processing time of algorithms, our method outperforms the other methods.

6. Conclusions

In this paper, we propose a multifeature fusion vehicle detection algorithm based on Choquet integral. There are two major contributions in this paper. First, we propose a taillight-based vehicle detection method, and a vehicle taillight feature similarity measure is defined. In addition, the vehicle symmetry and HOG + AdaBoost feature similarity measures are introduced combining with the definition of Choquet integral. Second, these three feature similarity measures are fused by Choquet integral to detect vehicles in both static test images and videos. In experiment part, our algorithm has been evaluated by using public collections and our own test images, and the experiment results are encouraging. But, to generalize our algorithm, there are still several problems to solve, such as improving accuracy of HOG + AdaBoost feature. To improve the performance of vehicle detection methods, we will address these issues and improve the multivehicle detection to an upper level.

Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

Acknowledgments

The authors would like to thank the reviewers and editors for their comments regarding enhancing the quality of the paper. This work is supported by Grants from Jilin Planned Projects for Science Technology Development (Grant no. 20120305 and no. 20130522119JH) and Ph.D. Programs Foundation of Ministry of Education of China (Grant no. 20130061110054).

References

- [1] Z. Sun, G. Bebis, and R. Miller, "On-road vehicle detection: a review," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 28, no. 5, pp. 694–711, 2006.
- [2] Y. M. Chan, S. S. Huang, L. C. Fu, P. Y. Hsiao, and M. F. Lo, "Vehicle detection and tracking under various lighting conditions using a particle filter," *IET Intelligent Transport Systems*, vol. 6, no. 1, pp. 1–8, 2012.
- [3] B. Lin, Y. Lin, L. Fu et al., "Integrating appearance and edge features for sedan vehicle detection in the blind-spot area," *IEEE Transactions on Intelligent Transportation Systems*, vol. 13, no. 2, pp. 737–747, 2012.
- [4] J. Hwang, K. Huh, and D. Lee, "Vision-based vehicle detection and tracking algorithm design," *Optical Engineering*, vol. 48, no. 12, Article ID 127201, 2009.
- [5] B. Southall, M. Bansal, and J. Eledath, "Real-time vehicle detection for highway driving," in *Proceeding of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops (CVPR '09)*, pp. 541–548, Miami, Fla, USA, June 2009.
- [6] D. Y. Chen, G. R. Chen, and Y. W. Wang, "Real-time dynamic vehicle detection on resource-limited mobile platform," *IET Computer Vision*, vol. 7, no. 2, pp. 81–89, 2013.
- [7] Y. Tsai, K. Huang, C. Tsai, and L. Chen, "An exploration of on-road vehicle detection using hierarchical scaling schemes," in *Proceedings of the 17th IEEE International Conference on Image Processing (ICIP '10)*, pp. 3937–3940, Hong Kong, September 2010.
- [8] W. Chang and C. Cho, "Online boosting for vehicle detection," *IEEE Transactions on Systems, Man, and Cybernetics, Part B: Cybernetics*, vol. 40, no. 3, pp. 892–902, 2010.
- [9] S. Sivaraman and M. M. Trivedi, "A general active-learning framework for on-road vehicle recognition and tracking," *IEEE Transactions on Intelligent Transportation Systems*, vol. 11, no. 2, pp. 267–276, 2010.
- [10] H. Tehrani Niknejad, A. Takeuchi, S. Mita, and D. McAllester, "On-road multivehicle tracking using deformable object model and particle filter with improved likelihood estimation," *IEEE Transactions on Intelligent Transportation Systems*, vol. 13, no. 2, pp. 748–758, 2012.
- [11] C. R. Wang and J. J. Lien, "Automatic vehicle detection using local features—a statistical approach," *IEEE Transactions on Intelligent Transportation Systems*, vol. 9, no. 1, pp. 83–96, 2008.
- [12] D. Alonso, L. Salgado, and M. Nieto, "Robust vehicle detection through multidimensional classification for on board video based systems," in *Proceedings of the 14th IEEE International Conference on Image Processing (ICIP '07)*, pp. IV321–IV324, San Antonio, Tex, USA, September 2007.
- [13] A. Jazayeri, H. Cai, J. Y. Zheng, and M. Tuceryan, "Vehicle detection and tracking in car video based on motion model," *IEEE Transactions on Intelligent Transportation Systems*, vol. 12, no. 2, pp. 583–595, 2011.
- [14] W. H. Li, H. Y. Ni, Y. Wang, B. Fu, P. X. Liu, and S. J. Wang, "Detection of partially occluded pedestrians by an enhanced cascade detector," *IET Intelligent Transport Systems*, 2014.
- [15] S. Ali and M. Shah, "A supervised learning framework for generic object detection in images," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR '05)*, vol. 2, pp. 1347–1354, San Diego, Calif, USA, June 2005.

- [16] J. Gall and V. Lempitsky, "Class-specific hough forests for object detection," in *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops (CVPR '09)*, pp. 1022–1029, Miami, Fla, USA, June 2009.
- [17] W. Sun, H. Gao, and O. Kaynak, "Adaptive backstepping control for active suspension systems with hard constraints," *IEEE/ASME Transactions on Mechatronics*, vol. 18, no. 3, pp. 1072–1079, 2013.
- [18] W. Sun, Z. Zhao, and H. Gao, "Saturated adaptive robust control for active suspension systems," *IEEE Transactions on Industrial Electronics*, vol. 60, no. 9, pp. 3889–3896, 2013.
- [19] W. H. Sun, H. J. Gao, and B. Yao, "Adaptive robust vibration control of full-car active suspensions with electrohydraulic actuators," *IEEE Transactions on Control Systems Technology*, vol. 21, no. 6, pp. 2417–2422, 2013.
- [20] W. Sun, H. Gao Sr., and O. Kaynak, "Finite frequency H_∞ control for vehicle active suspension systems," *IEEE Transactions on Control Systems Technology*, vol. 19, no. 2, pp. 416–422, 2011.
- [21] G. D. Tian, M. C. Zhou, and J. W. Chu, "A chance constrained programming approach to determine the optimal disassembly sequence," *IEEE Transactions on Automation Science and Engineering*, vol. 10, no. 4, pp. 1004–1013, 2013.
- [22] G. D. Tian, M. C. Zhou, J. W. Chu, and Y. M. Liu, "Probability evaluation models of product disassembly cost subject to random removal time and different removal labor cost," *IEEE Transactions on Automation Science and Engineering*, vol. 9, no. 2, pp. 288–295, 2012.
- [23] G. Tian, J. Chu, Y. Liu, H. Ke, X. Zhao, and G. Xu, "Expected energy analysis for industrial process planning problem with fuzzy time parameters," *Computers and Chemical Engineering*, vol. 35, no. 12, pp. 2905–2912, 2011.
- [24] Z. Y. Wang and G. J. Klir, *Fuzzy Measure Theory*, Plenum Press, New York, NY, USA, 1992.
- [25] Y. Wang and W. Li, "High-precision video flame detection algorithm based on multi-feature fusion," *Journal of Jilin University*, vol. 40, no. 3, pp. 769–775, 2010.
- [26] Y. Ding, W. H. Li, J. T. Fan, and H. M. Yang, "A moving object detection algorithm base on choquet integrate," *Acta Electronica Sinica*, vol. 38, no. 2, pp. 263–268, 2010.
- [27] "Caltech Cars (Rear)," <http://www.vision.caltech.edu/html-files/archive.html>.
- [28] M. B. Qi, Y. Pan, and Y. X. Zhang, "Preceding moving vehicle detection based on shadow of chassis," *Journal of Electronic Measurement and Instrument*, vol. 26, no. 1, pp. 54–59, 2012.
- [29] Q. Zhu, S. Avidan, M. C. Ye, and K. T. Cheng, "Fast human detection using a cascade of Histograms of Oriented Gradients," in *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR '06)*, New York, NY, USA, 2006.
- [30] K. Tieu and P. Viola, "Boosting image retrieval," in *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition (CVPR '00)*, pp. 228–235, Hilton Head Island, SC, USA, June 2000.

Research Article

Quad-Rotor Helicopter Autonomous Navigation Based on Vanishing Point Algorithm

Jialiang Wang,¹ Hai Zhao,¹ Yuanguo Bi,¹ Xingchi Chen,¹ Ruofan Zeng,¹ and Shiliang Shao²

¹ College of Information Science & Engineering, Northeastern University, No. 11, Lane 3, Wenhua Road, Heping, Shenyang, Liaoning Province 110819, China

² Shenyang Institute of Automation Chinese Academy of Sciences, No. 114, Nanta Street, Shenhe District, Shenyang, Liaoning 110016, China

Correspondence should be addressed to Yuanguo Bi; biyuanguo@ise.neu.edu.cn

Received 14 April 2014; Accepted 22 June 2014; Published 22 July 2014

Academic Editor: Yuxin Zhao

Copyright © 2014 Jialiang Wang et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Quad-rotor helicopter is becoming popular increasingly as they can well implement many flight missions in more challenging environments, with lower risk of damaging itself and its surroundings. They are employed in many applications, from military operations to civilian tasks. Quad-rotor helicopter autonomous navigation based on the vanishing point fast estimation (VPFE) algorithm using clustering principle is implemented in this paper. For images collected by the camera of quad-rotor helicopter, the system executes the process of preprocessing of image, deleting noise interference, edge extracting using Canny operator, and extracting straight lines by randomized hough transformation (RHT) method. Then system obtains the position of vanishing point and regards it as destination point and finally controls the autonomous navigation of the quad-rotor helicopter by continuous modification according to the calculated navigation error. The experimental results show that the quad-rotor helicopter can implement the destination navigation well in the indoor environment.

1. Introduction

As things get further away from us, they seem smaller. When they get far enough away, distances become ever tinier and so form a single point called vanishing point. In one-point perspective, all the horizontal lines go straight across, while the lines that move away from us—the sides of boxes, the road we are on, or the railway lines in front of us—all converge towards the center of the picture [1].

So the vanishing point contains much three-dimension information of the real environment, which can help to understand the real environment especially containing many parallel lines. The position estimation of vanishing point is widely used in applications such as navigation of robot, three-dimensional reconstruction, and camera calibration, and it has already become the hot research topic nowadays.

In the recent years, quad-rotor helicopter received an increasing attention from the research community. Many authors focused on the autonomous navigation control of

them. Hoffmann et al. [2] presented a model-based algorithm for autonomous flying with their STARMAC-quad rotor. Their system flew outdoors and utilized GPS and inertial measurement. Achtelika et al. [3] developed an indoor autonomous quad rotor equipped with a laser range scanner and cameras enabling autonomous hovering in a constraint indoor environment. Kendoul et al. [4] developed vision system for autonomous navigation 3D localization and control of small aerial vehicles based on optic flow.

Vanishing point is very important because it contains abundant three-dimension information for the indoor environment, and the innovation for this paper is that it focuses on the implementation of autonomous navigation using vanishing point algorithm and VPFE algorithm based on clustering principle in the indoor environment.

This paper is organized as follows: Section 2 presents system model of quad-rotor helicopter and its physical parameters. Section 3 describes the principle of vanishing point algorithm. Section 4 presents the implementation process

of proposed VPEE algorithm in detail. Section 5 describes the navigation based on improved PID control. Section 6 presents experimental results in the indoor environment in terms of execution time and flight deviation. Conclusions are given at the end of the paper.

2. System Model of Quad-Rotor Helicopter

The quad rotor has the flight motions such as vertical take-off and landing, pitching, rolling, and yawing, which are implemented by motor speed control so as to make system generate different force on each propeller.

$$\begin{aligned} \begin{bmatrix} Xg \\ Yg \\ Zg \end{bmatrix} &= \begin{pmatrix} \cos \psi & \sin \psi & 0 \\ -\sin \psi & \cos \psi & 0 \\ 0 & 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} 1 & 0 & 0 \\ 0 & \sin \varphi & \cos \varphi \\ 0 & -\cos \varphi & \sin \varphi \end{pmatrix} \cdot \begin{pmatrix} -\sin \theta & \cos \theta & 0 \\ -\cos \theta & -\sin \theta & 0 \\ 0 & 0 & 1 \end{pmatrix} \cdot \begin{bmatrix} X \\ Y \\ Z \end{bmatrix} \\ &= \begin{pmatrix} -\cos \psi \sin \theta - \sin \psi \sin \varphi \cos \theta & \cos \psi \cos \theta - \sin \psi \sin \theta \sin \varphi & \sin \psi \cos \varphi \\ \sin \psi \sin \theta - \cos \psi \sin \varphi \cos \theta & -\sin \psi \cos \theta - \cos \psi \sin \theta \sin \varphi & \cos \psi \cos \varphi \\ \cos \theta \cos \varphi & \cos \varphi \sin \theta & \sin \varphi \end{pmatrix} \cdot \begin{bmatrix} X \\ Y \\ Z \end{bmatrix}, \end{aligned} \quad (1)$$

Where straight lines $L_1, L_2,$ and L_3 represent the parallel lines of $OX, OY,$ and $OgXg$ separately. Lines $O'P_1$ and $O'P_2$ are the projection lines while lines OX and OY mapped into the ground coordinate. $\theta, \varphi,$ and ψ mean included angles as Figure 1 shows.

Besides, for the whole quad-rotor helicopter, the input variable is rotate speed of 4 electric motors [5], and they can be expressed as

$$\begin{aligned} u_1 &= F_1 + F_2 + F_3 + F_4, \\ u_2 &= F_4 - F_2, \\ u_3 &= F_3 - F_1, \\ u_4 &= F_1 - F_2 + F_3 - F_4, \end{aligned} \quad (2)$$

where $u_1, u_2, u_3,$ and u_4 represent the lifting force, rolling force, pitching force, and drifting force, respectively. In the meanwhile, the input variable contains 3 other pieces of poseinformation of $\theta, \varphi,$ and ψ for the whole system [6].

3. Principle of Vanishing Point Algorithm

3.1. Model of Vanishing Point. A group of parallel lines in three-dimension (3D) space can be mapped into some intersection lines in two-dimension (2D) image, and the intersection point formed by these intersection lines is called vanishing point [7].

Figure 2 shows the mapping relationship between point $A(x_0, y_0, z_0)$ in the three-dimension space and point $B(x'_0, y'_0, z'_0)$ in the two-dimension image [8], and we can obtain

$$(x'_0, y'_0, z'_0) = \left(\frac{x_0 f}{y_0}, \frac{z_0 f}{y_0}, f \right), \quad (3)$$

where f represents focal length of visual system, and point $p(x_0, y_0, z_0)$ can also be expressed as the sum of several

(Og, Xg, Yg, Zg) is the ground coordinate, and (O, X, Y, Z) is the quad-rotor helicopter coordinate as Figure 1 shows. Quad-rotor helicopter coordinate is used to control the state of motion, while ground coordinate is used to adjust the position and posture of the quad-rotor helicopter. From the analysis about transformation of above two coordinates, we can get the transformation equation as formula (1) shows:

vectors in the directions of $x, y,$ and z coordinate axes [9]. Using the attributes of $i, j,$ and k to express the base vector of the three axes, in the meantime vector \bar{p} can be expressed as $\bar{p} = x_0 \bar{i} + y_0 \bar{j} + z_0 \bar{k}$, so the vanishing point v can be calculated by

$$\bar{v} = x_v \bar{i} + y_v \bar{j} + f \bar{k}. \quad (4)$$

The indoor flight environment for quad-rotor helicopter navigation is shown as Figure 3(a), and the vanishing point can be obtained by the above calculation, and then it is set as the destination point navigation for quad-rotor helicopter in the indoor environment.

3.2. Straight Line Extraction of Destination Region. Detecting the parallel lines existing in the test environment is the premise for calculating the vanishing point, so this procedure is very important. It contains the below process: preprocessing of image, deleting noise interference, edge extracting using Canny operator, and extracting lines by RHT method.

3.2.1. Preprocessing of Image. While using histogram equalization method to improve contrast of overall situation, the luminance information can well be reflected in the histogram, and it is also effective for foreground and background no matter it is too bright or too dark. The most important thing is that histogram equalization is a reversible operation, which means only knowing the function expression of equalization, and then system can recover the original histogram with lower computational complexity.

3.2.2. Canny Operator. Canny operator is the optimization approximation operator obtained by measuring signal to noise ratio (SNR) and location product. Its principle in nature is to execute smooth image operation by using Gaussian

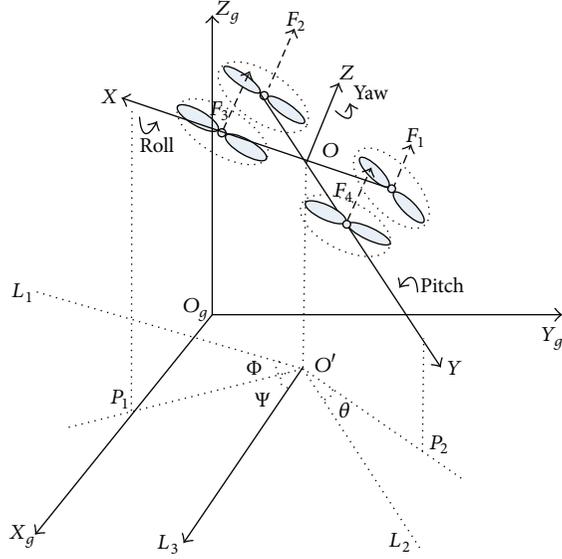


FIGURE 1: Transformation relationship between quad-rotor helicopter and ground coordinate.

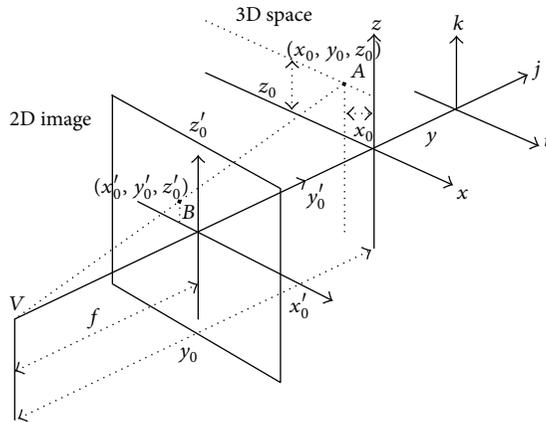


FIGURE 2: A point in 3D space mapped into 2D image.

function and calculating max derivative through first order differential operator. The process of edge detection is used to delete much useless information. In fact, it is used to detect points having obvious brightness variation, because these points include much information such as discontinuous depth and surface direction, change of material attribute, and scene illumination. Canny algorithm includes the below 4 procedures.

(1) *Gaussian Filter Denoising*. Gaussian blur is a filter for image penumbra, and it calculates the conversion for each pixel in the image by using below normal distribution:

$$G(r) = \frac{1}{\sqrt{2\pi\sigma^2}N} e^{-r^2/2\sigma^2}. \quad (5)$$

In the two-dimensional space, it can be expressed as

$$G(u, v) = \frac{1}{2\pi\sigma^2} e^{-(u^2+v^2)/2\sigma^2}, \quad (6)$$

where r is the blur radius which can be calculated by $r^2 = u^2 + v^2$; besides σ is the standard deviation of the above normal distribution, and it is used to control the smoothness of Gaussian filter parameters.

System executes convolution operation for original image and Gaussian smoothing template and then obtains an image having the slightest obscure. So image noise is reduced and the level of detail is also decreased after this process.

(2) *Gradient Operator Sobel*. The edges in image may point to different directions, so the Canny algorithm uses 4 masks to detect the margin of horizontal verticality and diagonals lines. The convolution value for original image and any mask is calculated and stored.

The first derivative of one two-dimension equation can be regarded as the gradient, and its vector can be expressed as

$$\bar{L}(x, y) = \begin{bmatrix} \bar{L}_x \\ \bar{L}_y \end{bmatrix} = \begin{bmatrix} \partial f / \partial x \\ \partial f / \partial y \end{bmatrix}. \quad (7)$$

The direction of vector $\bar{L}(x, y)$ stands for the direction having max changing rate, and the modulus of the gradient can be calculated by

$$|\bar{L}(x, y)| = \sqrt{\bar{L}_x^2 + \bar{L}_y^2}. \quad (8)$$

The direction of the gradient is

$$\delta(x, y) = \arctan\left(\frac{\bar{L}_y}{\bar{L}_x}\right), \quad (9)$$

where δ is the included angle formed by gradient and x -axis.

So the sample gradient model can be calculated by

$$\begin{aligned} \bar{L}_x &= f[i, j+1] - f[i, j], \\ \bar{L}_y &= f[i, j] - f[i+1, j], \end{aligned} \quad (10)$$

where i and j stand for the base vectors about y -axis negative direction and x -axis positive direction, respectively.

Canny edge detection function adopts a kind of gradient operator Sobel. Each pixel in the image uses this operator, and then the corresponding gradient vector or normal vector will be generated. The operator contains two three-dimensional matrices, and the modulus value of gradient can be expressed as

$$H = \sqrt{L_x^2 + L_y^2}. \quad (11)$$

Partial derivative can be calculated by

$$\begin{aligned} L_x &= (\beta \cdot n_3 + n_2 + n_4) - (\alpha \cdot n_7 + n_0 + n_6), \\ L_y &= (\beta \cdot n_1 + n_0 + n_2) - (\alpha \cdot n_5 + n_4 + n_6), \end{aligned} \quad (12)$$

where the constant values $\alpha = 2$ and $\beta = 2$ and then the above L_x and L_y can be expressed as the below convolution template:

$$L_x = \begin{bmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{bmatrix}, \quad L_y = \begin{bmatrix} 1 & 2 & 1 \\ 0 & 0 & 0 \\ -1 & -2 & -1 \end{bmatrix}. \quad (13)$$

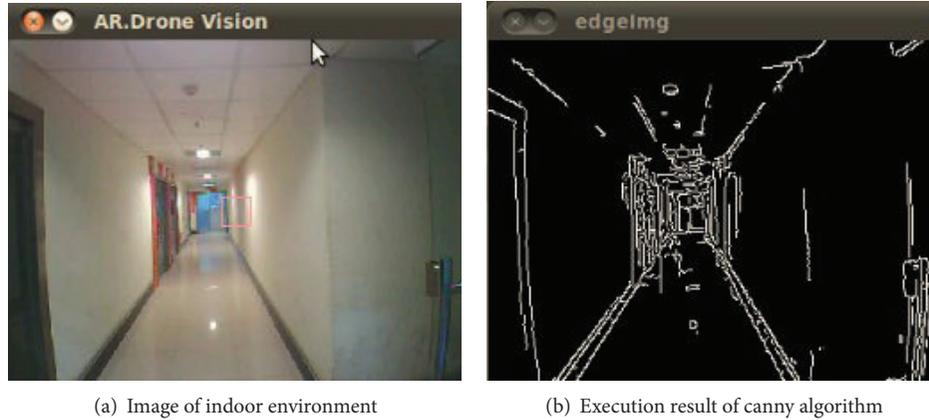


FIGURE 3: Edge detection of indoor flight environment.

Neighbourhood pixel point can be expressed as

$$L_{nb} = \begin{bmatrix} m_0 & m_1 & m_2 \\ m_7 & [i, j] & m_3 \\ m_6 & m_5 & m_4 \end{bmatrix}. \quad (14)$$

So the emphasis on calculating Sobel operator is to obtain the central pixel in the area whose size is similar to the template.

(3) *Control of Gradient Value.* The system compares the gray level and direction of current pixel with the pixel gradient value whose position is two pixels prior to current pixel, if it is smaller than the pixel gradient value of prior pixel, and then it would be set as 0, which means it is not an edge.

(4) *Partition of Lag Threshold.* The system could obtain the point whose gradient is changing by the above principle, but for those having high luminance gradient, there would exist some interference factors to judge whether they are edges, so a threshold is introduced to solve this problem. To be exact, Canny uses two lag thresholds of high and low threshold. If the important edge is continuous curve in the image, then the system can track the obscure part of the given curve and avoid to regard these noise pixels which actually do not form curve as edges.

The system begins this process with a large threshold, firstly marks these luminance gradients having high edge probability, and then tracks them in the whole image by the calculated direction information. For the process of edge tracking, the system uses the low threshold to control tracking path and makes the tracking path finally return the beginning position along with the obscure part.

Figure 3(a) shows the indoor environment captured by the front camera of quad-rotor helicopter. System stores the detecting results in the form of binary image, and Figure 3(b) is the edge effect picture using the above edge detecting process. Experimental results show the basic edge information can be successfully extracted.

3.2.3. *Randomized Hough Transformation (RHT) Method Used to Extract Straight Lines.* The main idea used to detect

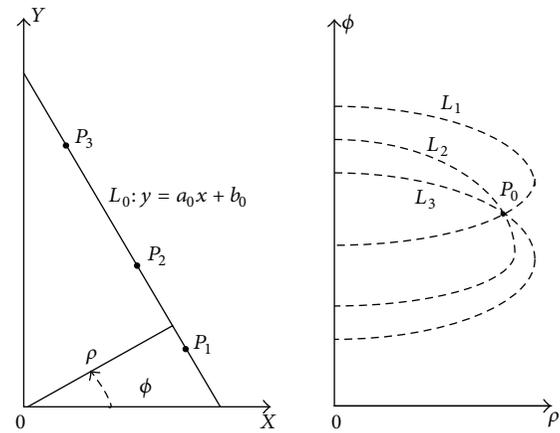


FIGURE 4: The mapping relation for a line in the image space and parameter space.

straight lines is to establish a mapping between image coordinate space and parameter coordinate space. Many pixels in the image space are mapped into the parameter space, if these pixels are mapped into the same point in the parameter coordinate, and then the counter will be increased by 1. The point's coordinate having largest counter value is corresponding to a straight line in the image space.

As Figure 4 shows, for any three points P_1 , P_2 , and P_3 in the straight line $Y = a_0x + b_0$ in the image space, they will form three sinusoidal curves L_1 , L_2 , and L_3 while they are mapped into the parameter space, and the intersection point P_0 of the three curve is corresponding to the straight line L_0 in the image space.

This paper uses improved randomized Hough transformation for the quad-rotor helicopter by using randomized sampling pixels in the image space and mapping them into parameter space. System firstly randomly picks two or n pixels and maps them into one point in the parameter space, then uses a set P with each element containing both a real-valued vector and an integer score to represent the parameter space, and updates the set P at each step by the point mapped from

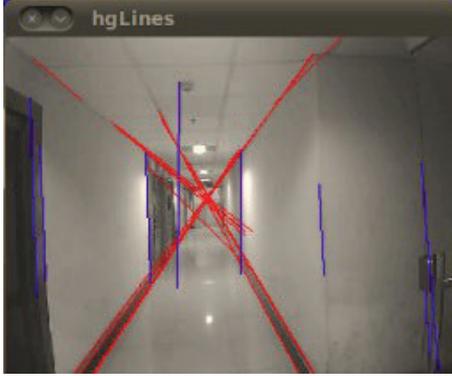


FIGURE 5: Straight lines extraction.

the randomly picked pixels. Finally, system finds accumulated cells of the detected curves [10].

RHT method uses many-to-one mapping so as to avoid huge amount of calculations in the traditional Hough transformation (HT) method using one-to-many mapping.

The system uses dynamic list structure and allocates corresponding memory space to these parameters only accumulated by the many-to-one mapping. The experimental result about using RHT method to extract straight lines is shown as Figure 5.

The interfering accumulations caused by disturbance pixels can be significantly improved in RTH method. For Figure 5, HT method uses a $320 * 240$ array while RHT method uses the storage of only around 36 cells. Besides, the computing time of HT method is about 2.4 times larger than that of RHT method.

Consequently, RTH method has the advantages of infinite parameter space, small storage, and high computation speed, which can better satisfy the real-time control of the quadrotor helicopter.

4. Implementation of Vanishing Point Algorithm

Vanishing point is obtained by evaluating the straight lines extracted from collected image. As in the indoor environment, there inevitably existed some disturbing lines, so the emphasis on this vanishing point algorithm is to reduce the effect of error component and estimate the exact position of vanishing point.

4.1. The Vanishing Point Algorithm.

(1) *Deleting Unreasonable Straight Line.* There is a straight line set as $C\{l_0, l_1, \dots, l_k \mid k = 0, \dots, M\}$, where M is the number of straight lines system that has been detected and k is integer. Each line is expressed by two endpoints $\{(x_{s0}, y_{s0}), (x_{e0}, y_{e0})\}, \{(x_{s1}, y_{s1}), (x_{e1}, y_{e1})\}, \dots, \{(x_{sM}, y_{sM}), (x_{eM}, y_{eM})\}$.

For endpoints set as C_p , the angle between each line and horizontal coordinate axis can be calculated by

$$\text{deg} = \arctan \frac{|y_1 - y_0|}{|x_1 - x_0|}. \quad (15)$$

There are some straight lines in the indoor scenario such as the edges of ceilings, the edges of floors, etc., and these lines are disturbing lines in the calculation of the vanishing point. Therefore, the removal of these lines is essential to accurately calculate the vanishing point and alleviate the error of the algorithm.

Unreasonable lines in the image are deleted by means of judging the deg angle. As OpenCV library function is used by this system, so the value ranges from 0 to 360 degrees after transformation. Judging about whether a line is close to horizontal or vertical straight lines depends on

$$\begin{aligned} f(\text{deg}) &= 1, \\ \text{VT} &> |90 - \text{deg}|, \\ \text{VT} &> |270 - \text{deg}|, \\ \text{HT} &> |\text{deg}|, \\ \text{HT} &> |180 - \text{deg}|, \end{aligned} \quad (16)$$

where VT (vertical line threshold) and HT (horizontal line threshold), respectively, represent the vertical and horizontal angle thresholds used to delete undesired lines.

In order to obtain the optimized value of VT and HT, we do the below test experiments so as to make the coverage rate to achieve the max value. Supposing the error radius is r for the vanishing point, so these straight lines will generate some points both inside and outside the circular error of vanishing point. The coverage rate is defined as $I_i/(I_i + O_k)$ used to mean the ratio of points inside the circular error of vanishing point accounts for the percentage of total points, where i and k are both positive integers.

As Figure 6(a) shows, there exist many intersections caused by straight lines, and these intersections are enlarged in order to be seen clearly; for example I_1 and I_2 (inside the circular error of vanishing point) and O_1, O_2, O_3 , and O_4 (outside the circular error of vanishing point) consist in straight line L_1 . If L_1 is deleted, then these six points are all deleted, which will affect the coverage rate. So using the threshold of horizontal line and vertical line is necessary so as to reduce the calculation work and improve coverage rate.

The values of VT and HT are both integers so as to lower the calculation. While VT and HT are both 0, which means no straight lines are deleted in this case, $\text{VT} + \text{VT} > 90$, which means all straight lines will be deleted. So choosing a suitable value to make the algorithm have the best accuracy is the key point. Figure 6(b) shows the test results about the relationship of VT and HT affecting the coverage rate. The corresponding value of VT and HT is 10 and 15, respectively, while the coverage rate reaches to its peak value.

Figure 6(b) shows the coverage rate affected by the thresholds of VT and HT. Because VT and HT both vary from 0 to 90 degrees, so there are total 810 image frames (while accuracy is 1 degree) in Figure 6(b). As the coverage rate will affect the outcome accuracy of vanishing point, which directly affects the flight deviation, the experiment of these 810 image frames well reflects the flight deviation according to

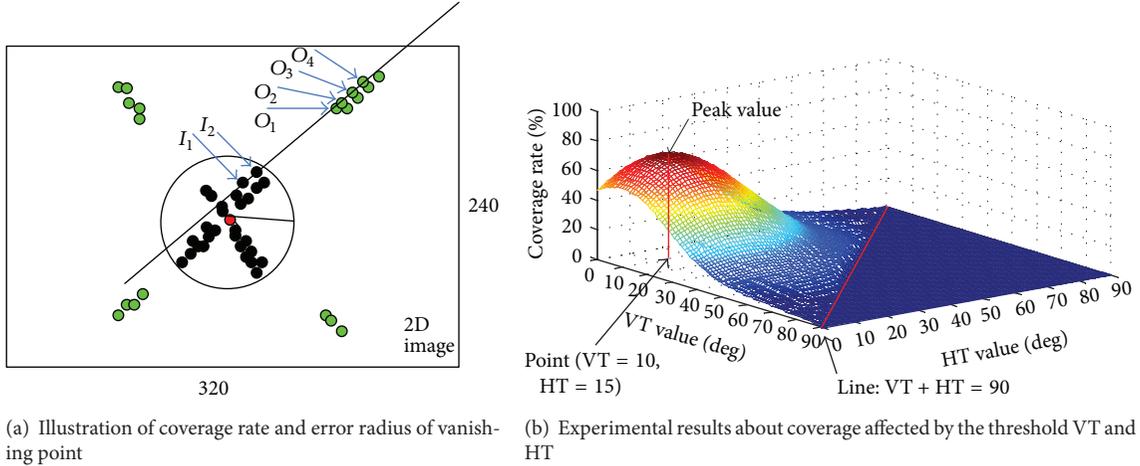


FIGURE 6: Error radius of vanishing point and coverage rate affected by the threshold VT and HT.

the thresholds VT and HT while the number of images frames increase.

For those lines making $f(\text{deg}) = 1$ is undesirable, so those lines should be deleted, and they are marked as blue in Figure 5.

(2) *Calculating the Intersection Point.* After deleting undesirable lines, the number of lines in the line set is M , so the number of their intersections is $N = M \cdot (M - 1)/2$, and it is stored as the initial value in the intersection array.

Any two straight lines in image can be expressed as

$$\begin{aligned} A_1x + B_1y + C_1 &= 0, \\ A_2x + B_2y + C_2 &= 0. \end{aligned} \quad (17)$$

Points $A(x_1, y_1)$ and $B(x_2, y_2)$ are on a straight line while points $C(x_3, y_3)$ and $D(x_4, y_4)$ are on another straight line. Then distance d between the two lines can be calculated by

$$d = (x_4 - x_3) \cdot (y_2 - y_1) - (x_2 - x_1) \cdot (y_4 - y_3). \quad (18)$$

While d is 0, it means straight lines AB and CD form parallel line; otherwise, they are intersection lines, and then the coordinate of their intersection point can be calculated by

$$\begin{aligned} x_0 &= x_1 \cdot (x_4 - x_3) \cdot (y_2 - y_1) \\ &+ (x_2 - x_1) \cdot (x_4 - x_3) \cdot (y_3 - y_1) \\ &- \frac{x_3}{d} \cdot (x_2 - x_1) \cdot (y_4 - y_3), \\ y_0 &= y_1 \cdot (x_2 - x_1) \cdot (y_4 - y_3) \\ &+ (x_3 - x_1) \cdot (y_2 - y_1) \cdot (y_4 - y_3) \\ &+ \frac{y_3}{d} \cdot (x_4 - x_3) \cdot (y_2 - y_1). \end{aligned} \quad (19)$$

Calculate all the intersections by the above equations, where n is the number of all intersection points, and their coordinates are expressed as $(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)$,

respectively. Those intersection points whose coordinate does not exist in the range of $(0, 0)$ to (h, w) are undesirable ones, and they should be deleted, where h and w are the height and weight of the image.

The system uses Canny algorithm to obtain the edge of the collected image and then adopts RHT method to extract straight lines. From the analysis of abundant practical flight experimental results, we can know that the successful rate about collecting elements of intersection point set can reach to above 92.5%.

4.2. Proposed VPFE Algorithm Based on Clustering Principle.

As the bottleneck for quad-rotor helicopter navigation is the speed of image processing, it includes the calculation of every pixel in the image. In order to improve the real-time ability of visual navigation, it is important to eliminate some unnecessary calculations.

A new vanishing point fast estimation (VPFE) algorithm is proposed to evaluate the vanishing points and obtain a navigation destination point finally. In fact, it is 1-mean cluster process. The system deals with all the points in the intersection set by the method of convergence. Specifically, the system firstly deletes those points whose distances to mass center are larger than cluster radius, and then adjusts value of the cluster radius by the method of convergence. The above procedure is executed repeatedly, and finally the mass center is obtained as the vanishing point. The total number of iterative procedure can be calculated by

$$R \cdot p^n < R_e. \quad (20)$$

As n is the number of iterative process, the value is

$$n = \left\lceil \log \frac{R_e}{R} \right\rceil + 1, \quad (21)$$

where $p(<1)$ is the convergence coefficient, R represents radius of convergence, and R_e is a dynamical value that varies from each cluster process, which stands for the deviation of mass center about each cluster process; finally, it is the



FIGURE 7: Clustering processing of VPFE algorithm.

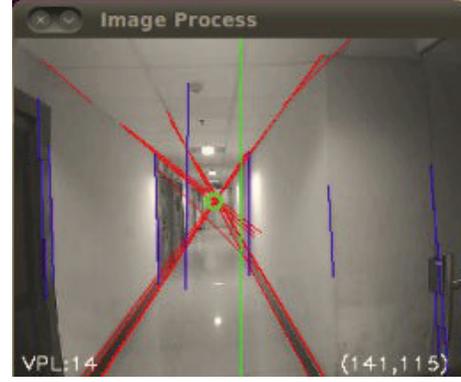


FIGURE 8: Collected image of the indoor environment.

radius of convergence used to describe the region how much vanishing point will be located. In order to facilitate the calculation, an element amount threshold NT is preset.

The procedure of algorithm is described as follows.

Step 1. Set the initialization values of R and R_e and calculate the n according to R and R_e .

Step 2. The new radius of convergence is $p \cdot R$.

Step 3. Calculate the mass center of the intersection points set as C and regard it as the vanishing point.

Step 4. For all the elements in the intersection point set as C , calculate the distance between intersection point and vanishing point.

Step 5. If the number of elements in the intersection points is smaller than NT or the amount of iterative process is larger than n , then go to Step 6; otherwise, return to Step 2.

Step 6. Output value of the mass center in the intersection point set is regarded as the position of the vanishing point.

The verification result of the above algorithm is shown as Figure 7. The green points are the undesirable ones who do not meet the requirement of threshold in the convergence process, and these points should be deleted. The red marker is the vanishing point obtained from the above VPFE algorithm, and the corresponding image in the indoor environment collected by the front camera of the quad-rotor helicopter is shown as Figure 8.

5. Navigation Based on PID Control

The implementation of quad-rotor helicopter navigation is composed of image data collection and receiving control command, which forms a closed-loop control system [11, 12]. Besides, the control command calculated by PID control forms another closed-loop control processing as Figure 9 shows.

In this experiment, the destination point is the vanishing point for the quad-rotor helicopter, and the image central

position collected by the quad rotor is its current position. The quad-rotor helicopter modifies the image central position according to the vanishing position. Suppose the image central position is $O(x_o, y_o)$ and the vanishing point coordinate is $VP(x_{vp}, y_{vp})$, so the error for the navigation can be expressed as

$$\begin{aligned} dx &= x_{vp} - x_o, \\ dy &= y_{vp} - y_o. \end{aligned} \quad (22)$$

The system adapts to PID to implement the process of navigation by close-loop control. In fact, the effect of PID control is to decrease the difference of current position and the vanishing point position till the difference reaches to a certain threshold [13]:

$$V = K_P \cdot e + K_I \cdot \sum e + K_D \cdot de, \quad (23)$$

where e_1 is the difference between current value and real value, e_2 represents accumulated error sum for the period ranges from beginning of current close-loop control to current control, e_3 stands for difference between the current control difference and last control difference in the process of the same close-loop control, and V is the real parameter to control the quad rotor. This navigation process is implemented by adjusting the three attitude angles of yaw, roll, and pitch.

The adjustment includes direction of three-dimension axis, dy is the position difference of y -axis used to control the vertical flight, which is corresponding the height control of quad-rotor helicopter, and dx is the position difference of x -axis used to control the horizontal flight, which is corresponding the yaw angle and roll angle control of quad-rotor helicopter [14]:

$$e_i = dx_i = x_{vi} - x_o \quad (24)$$

$$\text{out} = K_P \cdot e_i + \alpha \cdot K_I \cdot \sum e_i + K_D \cdot (e_i - e_{i-1}),$$

where α is defined by

$$\alpha = \begin{cases} 0, & |e_i| > \varepsilon, \\ 1, & |e_i| < \varepsilon, \end{cases} \quad (25)$$

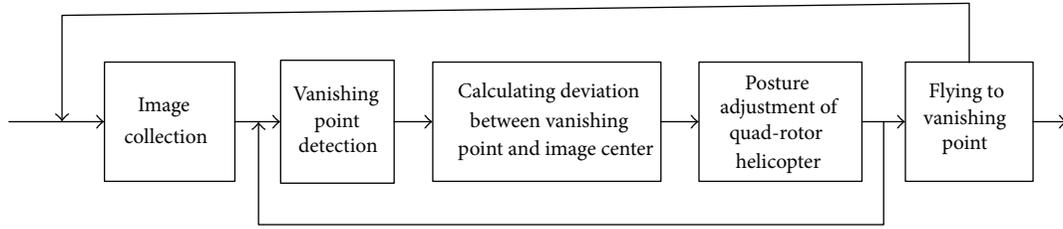


FIGURE 9: The workflow of quad-rotor helicopter navigation control.

where ε is the error threshold of integral separation while, under the condition of y -axis and z -axis, the e_i is $dy_i = y_{vi} - y_o$ and $dz_i = z_{vi} - z_o$, respectively.

Due to the fact that the control of quad-rotor helicopter belongs to complex nonlinear control, so in the process of PID control, while biased error became larger, the current integral separation parameter should be set as 0 to close integral action in order to eliminate shake, and while biased error became smaller, the current integral separation parameter should be set as 1 so as to open integral action.

In order to test the performance of improved PID control, indoor flight experiments about testing flight angle (roll, pitch, and yaw) and flight deviation (horizontal and vertical) are executed.

The quad-rotor helicopter is AR.Drone which is developed by Parrot Corporation in France, and it is designed based on four-axis intelligent controlling by WIFI, and it has the sensors such as three-axis accelerometer, two-axis gyrometer, one-axis yaw gyrometer, and ultrasonic wave height meter. Besides, it has a front camera (320 * 240 pix, 15 fps) and bottom camera (176 * 144 pix, 60 fps). AR.Drone has an ARM9 microcontroller which is based on Linux OS development platform.

Quad-rotor helicopter collects the sensors data such as gyroscope, accelerometer, and altimeter and system further calculates the navigation data such as posture information of pitch, roll, yaw, and height. The quad rotor sends these navigation data and video streaming to PC through WIFI [15]. Then PC refreshes the current navigation data and handles the video streaming including decoding, video navigation, and image handling algorithms. After PC calculating the navigation error according to the output of visual navigation, then system sends the navigation error to the PID controller and finally obtains the output of flight control command. The output is sent to quad rotor by WIFI to control its flight effectively [16].

Experiment 1. While the quad-rotor helicopter hovers in the indoor environment, disturbance is imposed to the three angles of roll angle, pitch angle, and yaw angle at the time 22 s. The response process was reflected as Figures 10(a), 10(b), and 10(c) show. The experimental results show that the improved PID control is better than original PID control in terms of overshoot and adjustment time.

Experiment 2. While the quad-rotor helicopter hovers in the indoor environment, disturbance is imposed in the direction of horizontal and vertical at the time 22 s. The response

process was reflected as Figures 10(d) and 10(e) show. The experimental results show that the improved PID control is also better than original PID control as to overshoot and adjustment time.

Experiment 3. The quad-rotor helicopter was flied to the coordinate of destination point (3,3,3) from the coordinate of original point (1,1,1) in order to show the performance test of flight stability for the improved PID. The experimental results about x , y , and z coordinate axes are shown in Figures 11(a), 11(b), and 11(c). The experimental results show that the improved PID control is also better than original PID control in terms of overshoot and adjustment time.

From the analysis of the above experimental results, no matter whether disturbance is imposed in the quad-rotor helicopter, we can know that the improved PID can well conduct the stability of flight for the quad-rotor helicopter in the indoor environment compared with the original PID.

6. Flight Experimental Results in the Indoor Environment

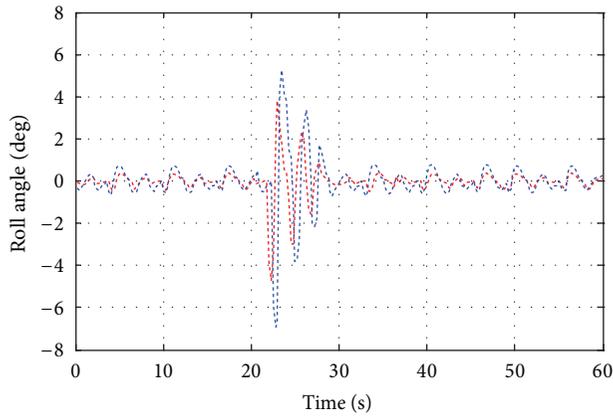
In order to test the performance of the VPFE algorithm, we test the flight of the quad-rotor helicopter in the indoor environment as Figure 3(a) shows. Figure 12 shows two frames of the software display image in PC. Each image describes the video image collected by the quad rotor after the corresponding handling by the above algorithm. A smaller picture in the right bottom of the image is used to show the flight scene of the quad-rotor helicopter at the corresponding time point.

In the below experiments, we measure the system efficiency in terms of algorithm execution time, and the experimental results about calculating vanishing point for these continuous 150 frames are shown in Figure 13.

As we can see from the results, the average execution time for the whole algorithm is 26.46 ms, which can meet the indoor flight environment for AR.Drone whose video flowing returning rate is 15 fps [17, 18].

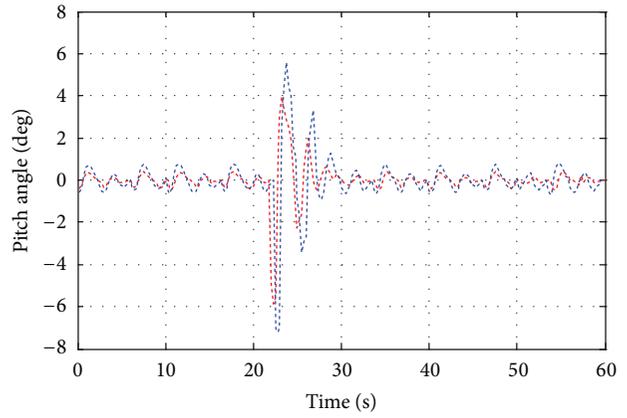
There are two other kinds of algorithm used to detect the vanishing point.

(1) *Spatial Alternation Technique.* Information of image coordinate space is mapped into another coordinate space, and then operations of deleting straight lines are executed. Gaussian ball transformation method is proposed by Tayebi and McGilvray [18]; after the process of mapping, the straight



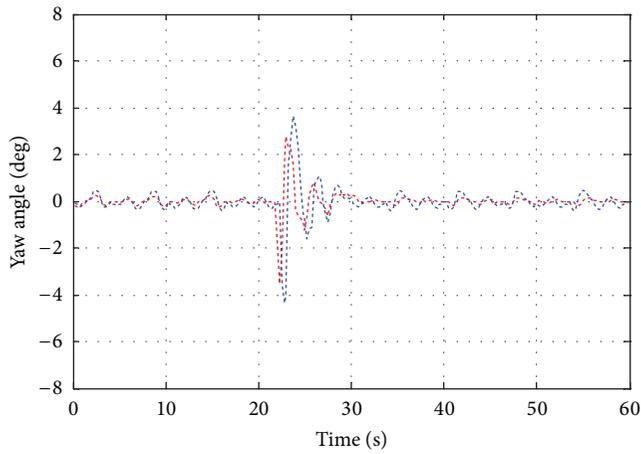
Improved PID
Original PID

(a) Flight test of roll angel



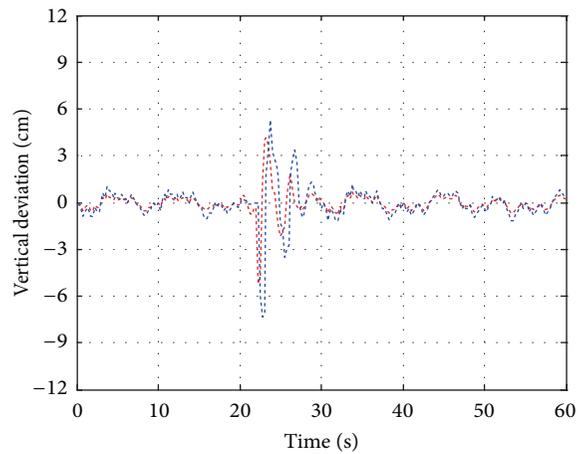
Improved PID
Original PID

(b) Flight test of pitch angel



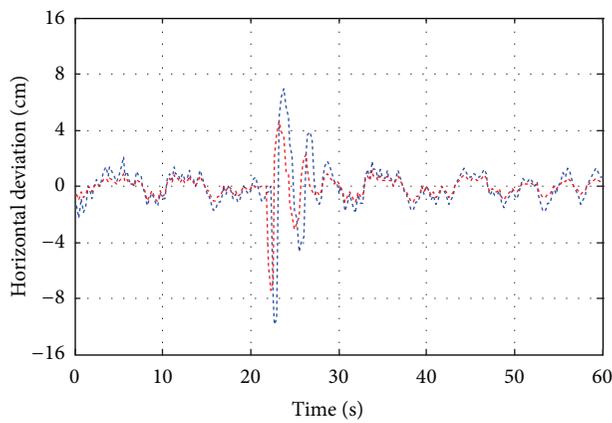
Improved PID
Original PID

(c) Flight test of yaw angle



Original PID
Improved PID

(d) Flight test of vertical deviation



Original PID
Improved PID

(e) Flight test of horizontal deviation

FIGURE 10: Performance comparison about flight angle and flight deviation.

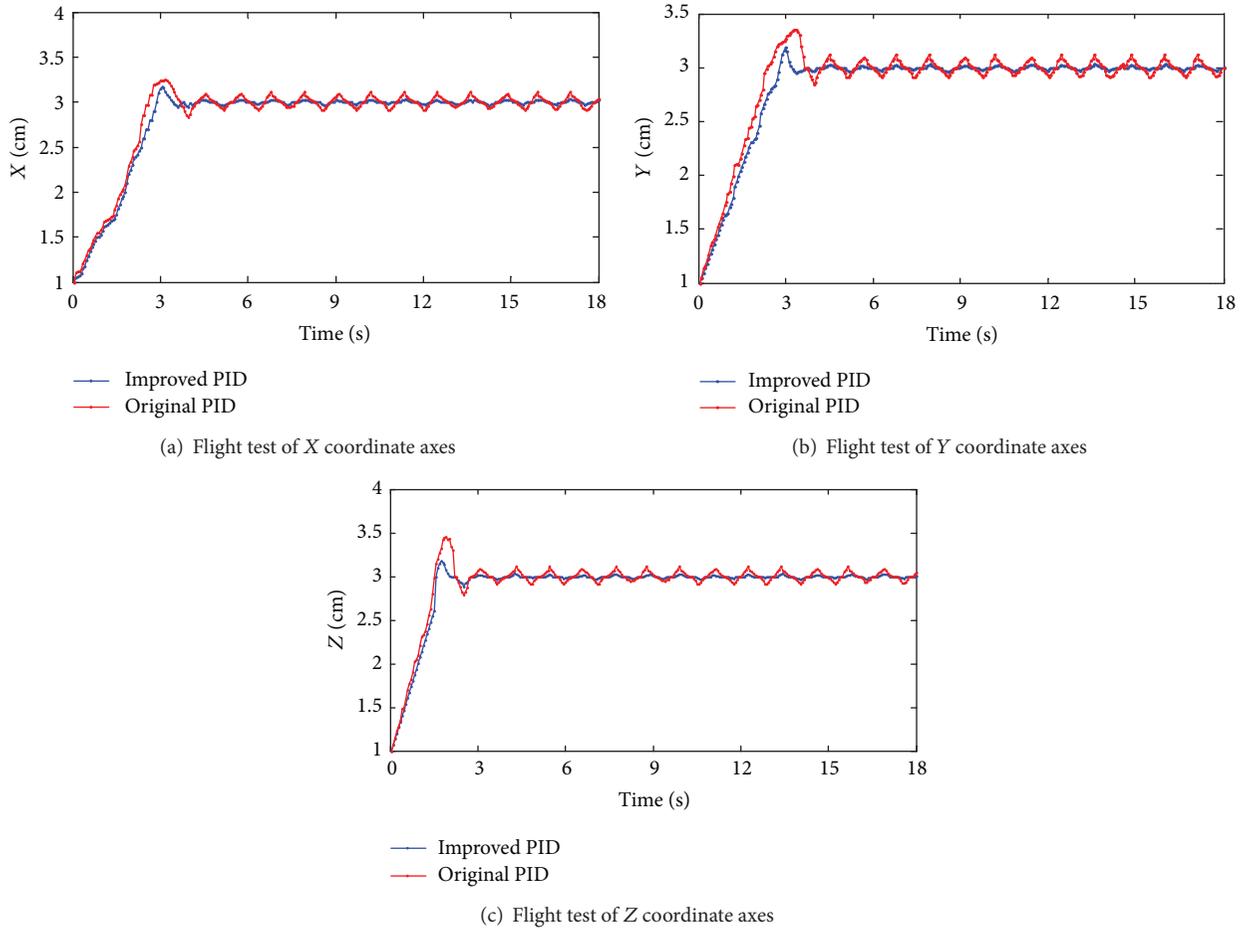


FIGURE 11: Performance test of flight stability.

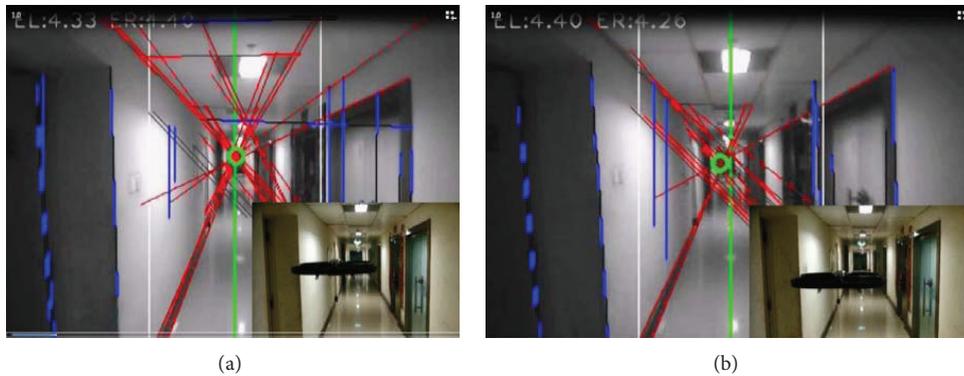


FIGURE 12: Flight test in the indoor environment.

lines and spatial location information of vanishing point is decreased at the cost of missing the information of straight line length and the distance between straight line and vanishing point. In the meantime, the calculation accuracy is affected by the accuracy of accumulate units.

(2) *Statistical Estimation Method.* Parameter of lines is estimated by the feature points in the edge of image, and these

parameters are used to calculate the vanishing point, or cost function is constructed using vanishing point and feature points of edge [19], but in the practical application, the algorithm is having higher computation complexity and lower computational efficiency.

This paper proposes algorithm mainly using straight line information. Straight line information is directly used to obtain the vanishing point in the image, and then the

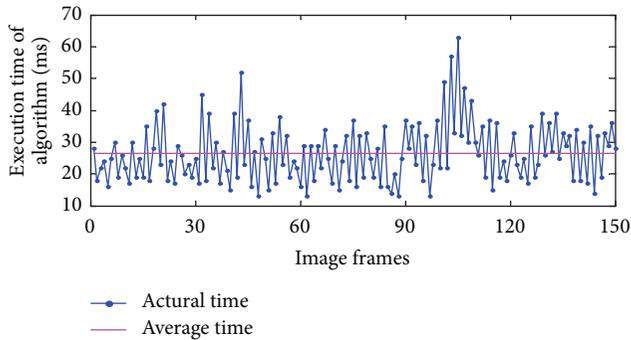


FIGURE 13: The efficiency of VPFE algorithm.

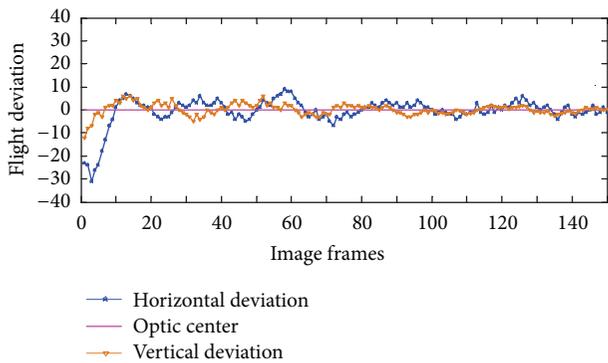


FIGURE 14: The horizontal deviation of autonomous navigation.

intersection of these straight lines is calculated, and this process is implemented by calling the matrix operations function in the OpenCV. Finally clustering principle is used to obtain the vanishing point so as to reduce the computation complexity of the algorithm.

The average execution time of proposed algorithm is about 26.46 ms. Compared with the other two vanishing point detection algorithms, the experimental results show that the execution time is separately about 13.7% and 21.3% shorter than Gaussian ball transformation method and statistical estimation method [20]. Consequently, the proposed algorithm has better real-time ability than the other two vanishing point detection algorithms.

Figure 14 shows the horizontal and vertical deviations between vanishing point and optic center for these continuous 150 image frames. In this experiment, the flight speed for the quad-rotor helicopter is constant value (0.4 m/s). As we can see from the experimental results, the quad-rotor helicopter can be well controlled during the whole process of autonomous navigation flight in the indoor environment [21, 22].

7. Conclusions

After the implementation of preprocessing of image, deleting noise interference, edge extracting using Canny operator, and extracting straight lines by RHT, the system finally obtains the position of vanishing point through the proposed VPFE

algorithm and effectively operates the autonomous navigation of the quad-rotor helicopter by continuous calculation about the navigation error. The experiment shows that it can implement the destination navigation toward the calculated vanishing point well in the indoor environment, and the experimental results show that its average execution time for the whole algorithm is 26.46 ms per image frame, besides the flight deviation error can well be controlled into the acceptable range, so it can meet the requirements of real-time and good stability in the indoor environment.

Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

Acknowledgments

This work is partly supported by the National Natural Science Foundation of China (no. 61101121), the National High Technology Research and Development Program (863 Program) (no. 2013AA102505), the Key Laboratory Project Funds of Shenyang ligong University under Grant no. 4771004kfs03, and the Educational Committee of Liaoning Province Science and Technology Research Projects under Grant no. L2013096. The authors thank Jan Vitek, Ales Plsek, and Lei Zhao for their help while the first author studied at the Purdue University as a Visiting Scholar from September, 2010, to September, 2012. The authors also thank the anonymous reviewers for their valuable comments.

References

- [1] <http://drawsketch.about.com/od/drawingglossary/g/vanishingpoint.htm>.
- [2] G. Hoffmann, D. G. Rajnarayan, S. L. Waslander, D. Dostal, J. S. Jang, and C. J. Tomlin, "The Stanford testbed of autonomous rotorcraft for multiagent control," in *Proceedings of the 23rd Digital Avionics Systems Conference*, vol. 2, October 2004.
- [3] M. Achtelika, A. Bachrach, R. He, S. Prentice, and N. Roy, "Autonomous navigation and exploration of a quadrotor helicopter in GPS-denied indoor environments," in *Proceedings of the Robotics: Science and Systems*, 2008.
- [4] F. Kendoul, I. Fantoni, and K. Nonami, "Optic flow-based vision system for autonomous 3D localization and control of small aerial vehicles," *Robotics and Autonomous Systems*, vol. 57, no. 6-7, pp. 591–602, 2009.
- [5] S. Grzonka, G. Grisetti, and W. Burgard, "A fully autonomous indoor quadrotor," *IEEE Transactions on Robotics*, vol. 28, no. 1, pp. 90–100, 2012.
- [6] P. Pounds, R. Mahony, and P. Corke, "Modelling and control of a large quadrotor robot," *Control Engineering Practice*, vol. 18, no. 7, pp. 691–699, 2010.
- [7] F. Schaffalitzky and A. Zisserman, "Planar grouping for automatic detection of vanishing lines and points," *Image and Vision Computing*, vol. 18, no. 9, pp. 647–658, 2000.
- [8] J. A. Shufelt, "Performance evaluation and analysis of vanishing point detection techniques," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 21, no. 3, pp. 282–288, 1999.

- [9] K. Seo, J. Lee, and H. Choi, "An efficient detection of vanishing points using inverted coordinates image space," *Pattern Recognition Letters*, vol. 27, no. 2, pp. 102–108, 2006.
- [10] L. Xu, E. Oja, and P. Kultanen, "A new curve detection method: randomized Hough transform (RHT)," *Pattern Recognition Letters*, vol. 11, no. 5, pp. 331–338, 1990.
- [11] G. M. Hoffmann, H. Huang, S. L. Waslander, and C. J. Tomlin, "Precision flight control for a multi-vehicle quadrotor helicopter testbed," *Control Engineering Practice*, vol. 19, no. 9, pp. 1023–1036, 2011.
- [12] C. Nicol, C. J. B. MacNab, and A. Ramirez-Serrano, "Robust adaptive control of a quadrotor helicopter," *Mechatronics*, vol. 21, no. 6, pp. 927–938, 2011.
- [13] T. Dierks and S. Jagannathan, "Output feedback control of a quadrotor UAV using neural networks," *IEEE Transactions on Neural Networks*, vol. 21, no. 1, pp. 50–66, 2010.
- [14] G. V. Raffo, M. G. Ortega, and F. R. Rubio, "An integral predictive/nonlinear H_∞ control structure for a quadrotor helicopter," *Automatica*, vol. 46, no. 1, pp. 29–39, 2010.
- [15] G. R. Rodríguez-Canosa, S. Thomas, J. del Cerro, A. Barrientos, and B. MacDonald, "A real-time method to detect and track moving objects (DATMO) from unmanned aerial vehicles (UAVs) using a single camera," *Remote Sensing*, vol. 4, no. 4, pp. 1090–1111, 2012.
- [16] M. Bonak, D. Matko, and S. Blai, "Quadrocopter control using an on-board video system with off-board processing," *Robotics and Autonomous Systems*, vol. 60, no. 4, pp. 657–667, 2012.
- [17] J. F. Guerrero-Castellanos, N. Marchand, A. Hably, S. Lesecq, and J. Delamare, "Bounded attitude control of rigid bodies: real-time experimentation to a quadrotor mini-helicopter," *Control Engineering Practice*, vol. 19, no. 8, pp. 790–797, 2011.
- [18] A. Tayebi and S. McGilvray, "Attitude stabilization of a VTOL quadrotor aircraft," *IEEE Transactions on Control Systems Technology*, vol. 14, no. 3, pp. 562–571, 2006.
- [19] R. Hartley and A. Zisserman, *Multiple View Geometry in Computer Vision*, Cambridge University Press, Cambridge, UK, 2nd edition, 2003.
- [20] A. Minagawa and N. Tagawa, "Vanishing point and vanishing line estimation with line clustering," *IEICE Transactions on Information & Systems*, vol. 7, pp. 1574–1582, 2000.
- [21] W. Sun, H. Gao, and B. Yao, "Adaptive robust vibration control of full-car active suspensions with electrohydraulic actuators," *IEEE Transactions on Control Systems Technology*, vol. 21, no. 6, pp. 2417–2422, 2013.
- [22] W. Sun, Y. Zhao, J. Li, L. Zhang, and H. Gao, "Active suspension control with frequency band constraints and actuator input delay," *IEEE Transactions on Industrial Electronics*, vol. 59, no. 1, pp. 530–537, 2012.

Research Article

Stepped Fault Line Selection Method Based on Spectral Kurtosis and Relative Energy Entropy of Small Current to Ground System

Xiaowei Wang, Xiangxiang Wei, Jie Gao, Yaxiao Hou, and Yanfang Wei

School of Electrical Engineering and Automation, Henan Polytechnic University, Jiaozuo 454000, China

Correspondence should be addressed to Xiaowei Wang; proceedings@126.com

Received 9 April 2014; Revised 8 June 2014; Accepted 8 June 2014; Published 22 July 2014

Academic Editor: Hamid R. Karimi

Copyright © 2014 Xiaowei Wang et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

This paper proposes a stepped selection method based on spectral kurtosis relative energy entropy. Firstly, the length and type of window function are set; then when fault occurs, enter step 1: the polarity of first half-wave extremes is analyzed; if the ratios of extremes between neighboring lines are positive, the bus bar is the fault line, else, the SK relative energy entropies are calculated, and then enter step 2: if the obtained entropy multiple is bigger than the threshold or equal to the threshold, the overhead line of max entropy corresponding is the fault line, if not, enter step 3: the line of max entropy corresponding is the fault line. At last, the applicability of the proposed algorithm is presented, and the comparison results are discussed.

1. Introduction

Most of 3~66 kV distribution networks are so-called small current to ground system in china, which includes unearthed neutral system, arc suppression coil compensated neutral system, and high resistance-grounded neutral system. Since the fault current is small and the arc is unstable instability, suitable fault line selection methods are lacking [1, 2].

When fault occurs, the researchers are attracted by the rich transient fault features. At present, the main research methods are based on the transient fault signal, such as Wavelet transform [3, 4], S transform [5, 6], Prony algorithm [7], Hilbert-Huang transform (HHT) [8, 9], and correlation analysis [10, 11]. In [3, 4], the transient zero-sequence current (TZSC) is decomposed by wavelet, and the transient energy is calculated; according to the differences, the fault line is selected, when the high grounding resistance fault occurs, which can lead to misjudgment because of the compensation of arc suppression coil to the TZSC. Good time frequency of S transform can be used to select fault line; in [5, 6], the dominant frequencies of TZSC are obtained by S transform, which are used as the selection criterions; however, there is much information after S transform, and what should be further studied is how to take advantage of the phase angles. Prony algorithm can fit the low frequency transient

signal well; in [7], Prony is used to fit the TZSC to select the dominant components, and then the relative entropy of dominant components is obtained to select the fault line, but what should be further studied is how to determine the model orders. In [8, 9], HHT is used to decompose the TZSC, and then the most high-frequency component of the intrinsic mode functions (IMF) can be obtained, and based on this, the selection criterion is built; however, the decomposition process may cause modal aliasing. In [10, 11], the poor waveform similarity of the TZSC between the fault line and the healthy line is noticed, and the coefficients are obtained to select the fault line; however, when the small-angle fault occurs, the similarity of the TZSC that flows the healthy cable line and the healthy overhead lines gets worse.

Aiming at few features that are available in arc suppression coil to ground system, single fault line selection method is not very reliable, and this paper regards the ratios of the first half-wave extremes between neighboring lines and spectral kurtosis (SK) relative energy entropy as features to select the fault line, a stepped fault line selection method is proposed by utilizing two kinds of features. When fault occurs, enter step 1, and if the ratios of first half-wave extremes between neighboring lines are positive, the bus bar is selected as the fault line, and the line selection process is done. Otherwise, the SK relative energy entropies of branch lines are computed,

then enter step 2, if the obtained entropy multiple is greater than or equal to the threshold, the overhead line is judged as the fault line, and the overhead line of max entropy corresponding is selected as the fault line, and the fault line selection process is done. If the entropy multiple is less than the threshold, enter step 3, and the line of max entropy corresponding is selected as the fault line.

The remaining of this paper is organized as follows. In Section 2, characteristics of TZSC are presented. In Section 3, basic theories about spectral kurtosis and wavelet denoising are presented. In Section 4, feasibility analysis of spectral kurtosis algorithm is presented. In Section 5, basic theories about fault line selection are presented. In Section 6, fault line selection criteria are proposed. In Section 7, simulation and verification are presented. In Section 8, applicability analysis is presented. In Section 9, two kinds of comparison results are discussed. In Section 10, the paper is completed with conclusions and future directions.

2. Characteristics of TZSC

Zero sequence equivalent circuit of single phase to ground is shown in Figure 1, C_0 and L_0 are the zero-sequence equivalent capacitance and inductance, respectively, R_g is the ground transition resistance, R_p and L_p are the equivalent resistance and inductance of arc suppression coil, respectively, and $G(t)$ is the zero-sequence voltage.

From Figure 1, the TZSC $i_0(t)$ through the fault point is

$$i_0(t) = i_{0L}(t) + i_{0C}(t) = I_{Lm} \cos \varphi e^{-t/\tau_L} + I_{Cm} \left(\frac{\omega_f}{\omega} \sin \varphi \sin \omega t - \cos \varphi \cos \omega_f t \right) e^{-\delta t}, \quad (1)$$

$i_{0L}(t)$ and $i_{0C}(t)$ are the inductor and capacitive current, respectively, I_{Lm} and I_{Cm} are the initial value of $i_{0L}(t)$ and $i_{0C}(t)$, respectively, ω is the angular frequency, ω_f and δ are the angular frequency of free oscillations and the attenuation coefficient, respectively, and $\delta = 1/\tau_c$, τ_c is the decay time constant of $i_{0C}(t)$, τ_L is the decay time constant of $i_{0L}(t)$, and φ is the initial phase of the fault line [12].

From (1), $i_{0L}(t)$ and $i_{0C}(t)$ make up the fault TZSC, both of which include nonperiodic damped component and stable periodic component. At beginning, τ_L is greater than τ_c ; therefore, $i_{0L}(t)$ is deteriorated slower than $i_{0C}(t)$.

It can be known from the above analysis that the TZSC is mutant signal when fault occurs, and as time goes on, it would decrease gradually. Considering the good frequency-domain effect for unstable, nonperiodic, mutant, and damped signal of SK, SK algorithm is introduced to extract the feature information of TZSC in this paper.

3. Basic Theories

3.1. Definition of SK. Kurtosis is random 4-order cumulative amount and cannot reflect specific signal changes as a global index; therefore, kurtosis is improper to detect the signal with strong noise. In order to overcome the weakness of kurtosis, SK is proposed by Immovilli et al. [13–15].

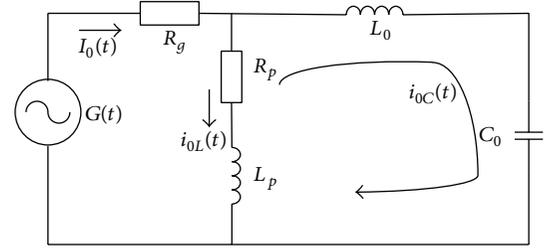


FIGURE 1: Zero sequence equivalent circuit.

$x(t)$ is decomposed by Wold-Cramer under the nonstationary condition, and the result is as follows:

$$Y(t) = \int_{-\infty}^{+\infty} e^{j2\pi ft} H(t, f) dX(f), \quad (2)$$

$H(t, f)$ is the complex envelope at frequency f of $Y(t)$, and also $H(t, f)$ changes with time. In a practical system, $H(t, f)$ is random, $X(f)$ is the strict white noise spectrum, and ω is random with time varying of filter.

Fourth-order spectrum cumulant of $Y(t)$ can be defined as follows:

$$C_{4Y}(f) = S_{4Y}(f) - S_{2Y}^2(f) \quad (f \neq 0), \quad (3)$$

$S_{2nY}(f)$ is the $2n$ moments, which can be defined as follows:

$$S_{2nY}(f) \triangleq \frac{E \{ |H(t, f) dX(f)|^{2n} \}}{df}. \quad (4)$$

Therefore SK can be defined as follows:

$$K(f) \triangleq \frac{C_{4Y}(f)}{S_{2Y}^2(f)} = \frac{S_{4Y}(f)}{S_{2Y}^2(f)} - 2, \quad (f \neq 0). \quad (5)$$

The important properties of SK are as follows.

- (1) The SK of pure stationary process is -1 , $f \neq 0$.
- (2) The SK of a stationary Gaussian process is 0 .

The TZSC $i_0(t)$ can be expressed as follows:

$$i_0(t) = Y(t) + b(t), \quad (6)$$

$Y(t)$ is the actual TZSC and $b(t)$ is the noise. $Y(t)$ and $b(t)$ are independent; therefore, the SK of $i_0(t)$ is

$$K_{(x+b)}(f) = \frac{K_x(f)}{[1 + \rho(f)]^2}, \quad (7)$$

$\rho(f) = (S_b(f)/S_{i_0}(f))$ is the ratio of noise to signal, and $S_b(f)$ and $S_{i_0}(f)$ are power spectral densities of $b(t)$ and $Y(t)$, respectively.

It can be known from (7) that $K_{(x+b)}(f) \approx K_{i_0}(f)$ when $\rho(f)$ is very small and $K_{(x+b)}(f) \approx 0$ when $\rho(f)$ is very

large. So the SK can search the whole frequency band, and the characteristic frequency bands can be detected easily.

3.2. *SK Based on Short-Time Fourier Transforms (STFT).* $i_0(n)$ is the discrete form of $i_0(t)$, $n = 1, 2, \dots, N$, and the calculation of SK based on STFT is as follows:

$$Y_w(kP, f) = \sum_{n=1}^N i_0(n) w(n - kP) e^{-j2\pi n f}, \quad (8)$$

$w(n)$ is the window function, whose length is N_w , P is the time step, and k is the number of time steps.

The $2n$ -order spectrum distance of $Y_w(kP, f)$ is defined as follows:

$$S_{2nY} = \langle |Y_w(kP, f)|^{2n} \rangle_k, \quad (9)$$

$\langle \cdot \rangle_k$ is k -order time average.

Considering the definition of SK in Section 3.1, when $n = 1$ and $n = 2$, based on STFT, and SK can be described as follows [16]:

$$K(f) = \frac{S_{4Y}(f)}{S_{2Y}^2(f)} - 2. \quad (10)$$

It can be seen that the $i_0(t)$ can be characterized by SK based on STFT, and the frequency band of the max SK corresponding can be obtained finally.

3.3. Basic Principle of Wavelet Denoising

3.3.1. *Basic Principle of Wavelet Transform.* The basis function of the wavelet transform can be defined as follows [17, 18]:

$$\psi_{u,v}(t) = \frac{1}{\sqrt{u}} \psi\left(\frac{t-v}{u}\right) \quad u > 0, v \in R, \quad (11)$$

where u is the scale parameters and v is the translation parameter.

The continuous wavelet transform of $\psi(t)$ is a function defined as

$$\text{WT}_f(u, v) = \langle f, \psi_{u,v} \rangle = \frac{1}{\sqrt{u}} \int_R f(t) \bar{\psi}\left(\frac{t-v}{u}\right) dt. \quad (12)$$

In order for an inverse wavelet transform to exist, the mother wavelet $\psi(t)$ should satisfy the admissibility condition: $C_\psi = \int_R (|\psi(\hat{\omega})|^2 / |\omega|) d\omega < +\infty$ ($\psi(\hat{\omega})$ is the complex conjugate of $\psi(t)$). So $\psi(t)$ is an admissible wavelet. Choosing admissible wavelet, the original signals are reconstructed by

$$f(t) = \frac{1}{C_\psi} \iint_R \text{WT}_f(u, v) \cdot \left[\frac{1}{\sqrt{u}} \psi\left(\frac{t-v}{u}\right) \right] \frac{1}{u^2} du dv. \quad (13)$$

In practice, u and v must be discrete. Let $u = u_0^m$ and $v = nv_0 u_0^m$, and (13) can be presented as

$$\psi_{m,n}(t) = u_0^{-m/2} \psi(u_0^{-m} t - nv_0). \quad (14)$$

Function family $\{\psi_{m,n}(t), m, n \in Z\}$ is called discrete wavelet. When $u_0 = 2$ and $v_0 = 1$, function family $\{\psi_{m,n}(t)\}$

constructs a group orthonormal basis on which signals can be reconstructed.

3.3.2. *Process of Wavelet Denoising.* The basic principle of wavelet denoising is based on the spectrum distribution characteristics of signals and noise. According to the principle, noise of each decomposition level is eliminated, and useful signal is preserved [19, 20].

The substance of denoising is restraining the unwanted signal and boosting the useful signal. The one-dimensional K_i ($i = 1, 2, \dots, N_w$) can be expressed as follows:

$$K_i = f_i + \varepsilon e_i, \quad (15)$$

K_i is the observable SK with noise, f_i is the real SK without noise, e_i is the gauss random signal, and ε is the standard deviation of noise.

The denoising process of one-dimensional signals can be divided into three steps.

Step I. Mother wavelet function and the levels of decomposition are determined, and then K_i can be decomposed by wavelet transform.

Step II. High frequency coefficients are processed by the threshold.

The threshold thr is

$$\text{thr} = \sigma \sqrt{2 \log_e(N_w)}, \quad (16)$$

where σ is the standard deviation of K_i and N_w is the length of K_i .

Step III. Signals can be reconstructed by adding the low-frequency coefficients and high frequency coefficients, and the result is K'_i .

4. Feasibility Analysis of Spectral Kurtosis Algorithm

The ideal signal is defined to show the SK algorithm characteristics, and the details are as follows:

$$\begin{aligned} i_0(n) &= i_{01}(n) + i_{02}(n) + i_{03}(n) + i_{04}(n) + \varepsilon(n), \\ i_{01}(n) &= 5.6 \cos\left(2\pi \times 50t + \frac{\pi}{3}\right), \\ i_{02}(n) &= 40e^{-56t} \cos\left(2\pi \times 250t + \frac{\pi}{6}\right), \\ i_{03}(n) &= 72e^{-102t} \cos(2\pi \times 315t), \\ i_{04}(n) &= 10e^{-5.5t}, \end{aligned} \quad (17)$$

where $i_{01}(n)$ is the fundamental frequency component, $i_{02}(n)$ is the 5th harmonic component, $i_{03}(n)$ is noninteger harmonic component, $i_{04}(n)$ is the decaying direct current (DC) component, $\varepsilon(n)$ is gauss white noise signal, and $\text{SNR} = -13.5$ dB.

The ideal TZSC $i_0(n)$ is a nonperiodic signal, added by $i_{01}(n)$, $i_{02}(n)$, $i_{03}(n)$, $i_{04}(n)$, and $\varepsilon(n)$; what should be noted is

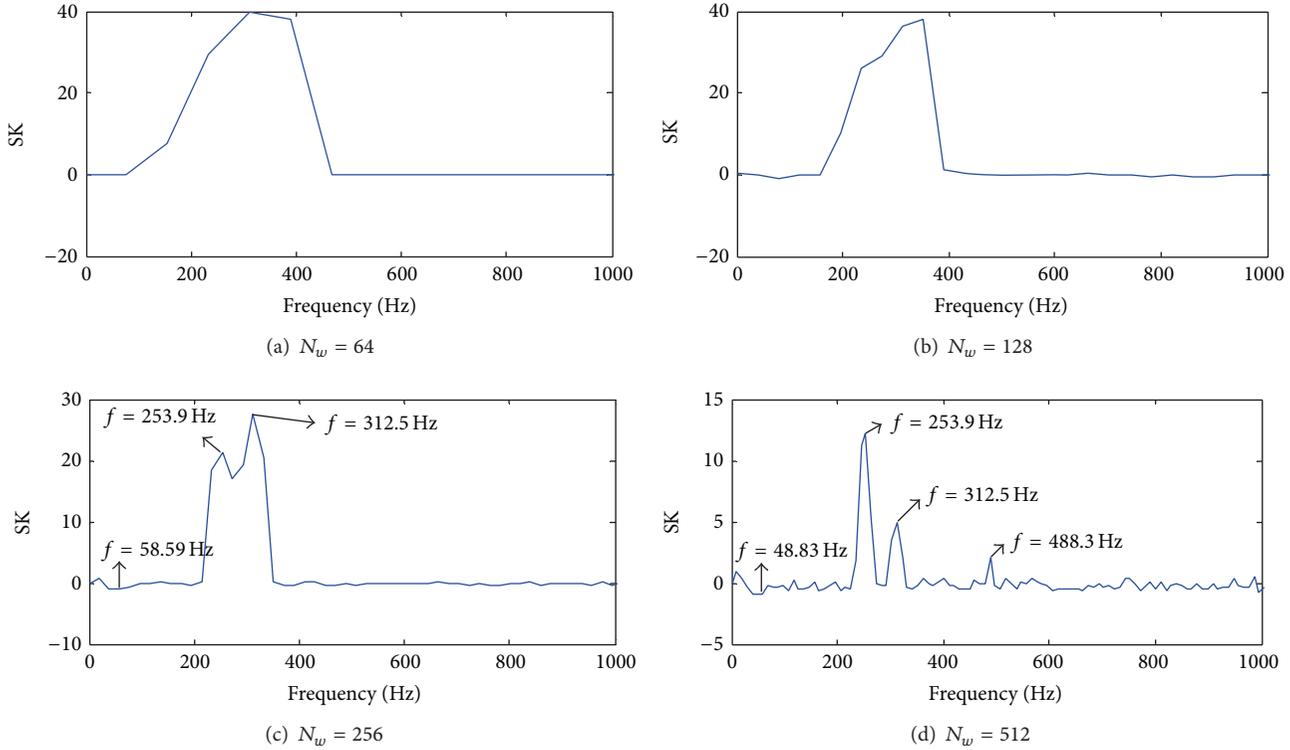


FIGURE 2: Length of window function.

that the ideal signal is generated according to the features of the TZSC [21, 22].

4.1. Set the Length of Window Function. The sampling frequency $f_c = 10000$ Hz, the simulation time is 0.6 s, and SK algorithm is used to analyze the ideal TZSC. The adjacent frequency interval is $f_c/2N_w$, as N_w are 64, 128, 256, and 512, and the frequency intervals are 78.13 Hz, 39.06 Hz, 19.53 Hz, and 9.77 Hz, respectively. The results of SK based on STFT are shown in Figure 2.

As Figure 2 shows, different lengths of window function make a difference for SK. When N_w is 64 or 128, to a certain extent, the ideal TZSC can be characterized; however, the frequency resolution is lower, fundamental frequency is not imprecise, and the 5th harmonic signal cannot be characterized. When N_w is 256, frequency resolution meets the requirements, and three types of signal frequency can be distinguished, because of the interval between adjacent SK. The frequency f causes errors, and the results are $f = 58.59$ Hz, $f = 253.9$ Hz, and $f = 312.5$ Hz. When N_w is 512, frequency resolution is higher, and fundamental frequency is more accurate; however, SK contains much noise; for example, as Figure 2(d) shows, $f = 488.3$ Hz belongs to spurious frequency components.

4.2. Set the Type of Window Function. After the length of window function is determined, the window functions are selected, such as Gaussian window (gausswin), Hamming

window, Hanning window, and Kaiser window. Based on the SK algorithm, the results are shown in Figure 3.

It is known from Figure 3 that different window functions can also characterize the TZSC's frequency, and the SK of white noise is close to 0. Therefore, the types of window function have a little impact on the SK. Because of being less effective than other window functions, Kaiser should avoid being used.

In conclusion, SK algorithm can distinguish the frequency components of the TZSC. In order to extract the features more perfectly, Hanning window is selected in this paper, whose length is 256.

5. Basic Theories about Fault Line Selection

5.1. Steps of Denoising. The calculation process of SK can be interfered by various noise; firstly, wavelet analysis is introduced to eliminate the noises, whose frequency is approximate to the frequency of TZSC's SK K_i ($i = 1, 2, \dots, N_w$), since symlets wavelet are limited compactly supported wavelets and whose local capabilities in the time domain and frequency domain are good, considering SK algorithm has good local capabilities in the frequency domain, and the frequency characteristic of sym6 is good, whose frequency resolution can also meet the requirements. White noise belongs to a stationary random interference signal; according to the properties (2) in Section 3.1, the SK of white noise is 0, but the defects from the outside condition make the SK of white noise fluctuate around 0. Finally, the threshold method is proposed to correct SK, which can also eliminate the effects

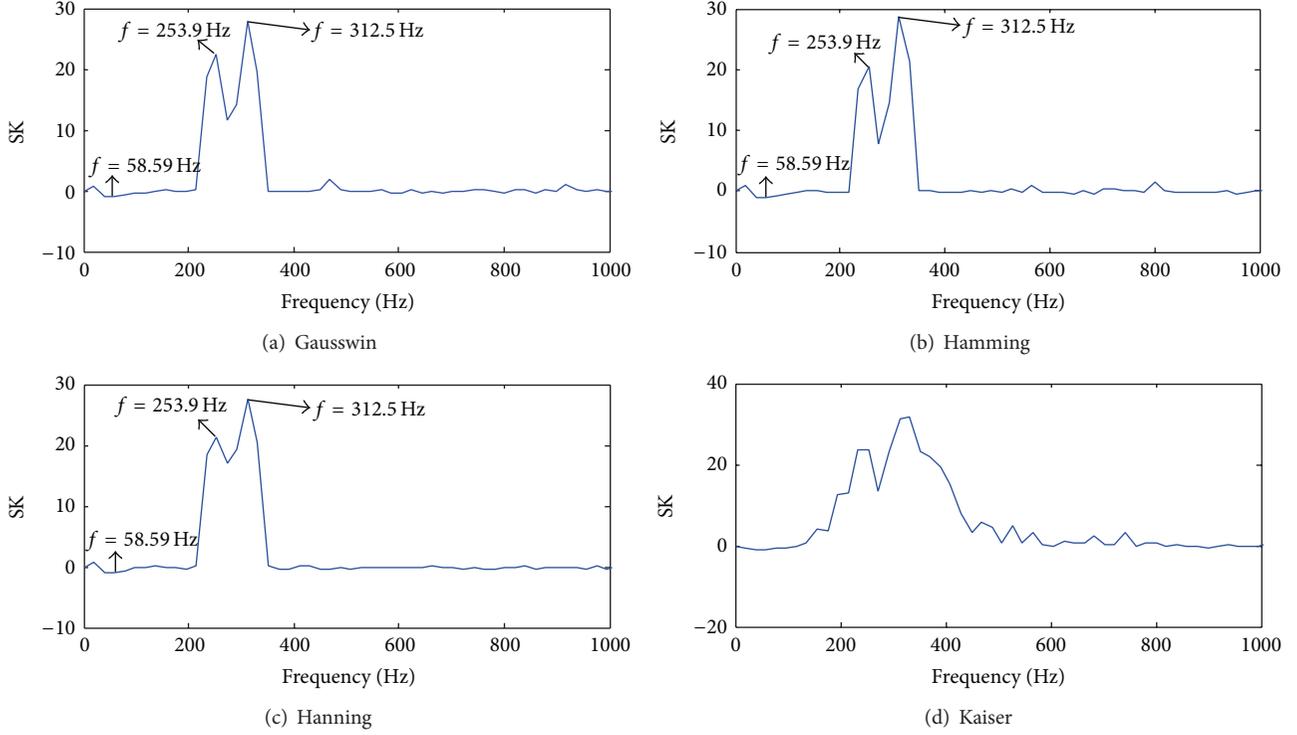


FIGURE 3: Types of window function.

on the SK relative energy entropy caused by the white noise. Specific processes are as follows.

- (1) K_i is decomposed into 4 layers by sym6.
- (2) Soft threshold process the high frequency coefficients of each decomposition scale in Section 3.3.2 Step II.
- (3) Reconstruct one-dimensional wavelet by adding the bottom of the low frequency coefficients and high frequency coefficients, and the result is K'_i .
- (4) The waveform correction coefficient λ is introduced, and the equation is

$$\xi = \lambda \cdot K'_{\max}, \quad 0 \leq \lambda \leq 1, \quad (18)$$

where K'_{\max} is the max K'_i , and considering the value relation between K'_i and K'_{\max} , the coefficient λ is 20% in the paper.

- (5) K'_i is retained the same value as long as $K'_i \geq \xi$; otherwise, $K'_i = 0$, the result is \widehat{K}_i .

After wavelet denoising and threshold correcting, the noise in \widehat{K}_i decreases significantly, and a premise of calculating the SK relative energy entropy is accurately provided.

5.2. The SK Relative Energy Entropy. In information theory, entropy represents uncertainty of average information and can also be used to estimate the complexity of random signals, it provides useful information about underlying the dynamic processes of signal, SK algorithm can magnify a

local frequency characteristic of the TZSC, and small and short abnormalities are found in the frequency domain by SK energy entropy. The impact of fault conditions mainly reflected on time-frequency distribution of the TZSC [23–25], in order to reflect the changes of SK in the frequency domain, based on the relative entropy and energy entropy, and the SK relative energy entropy C_j is defined in the paper:

$$C_j = -\sum_{i=1}^{N_w} u_{ji} \log_2 u_{ji}, \quad j = 1, 2, \dots, l, \quad (19)$$

where u_{ji} is the SK relative energy probability distribution, $u_{ji} = \widehat{K}_{ji}^2 / E_{\max}$, $E_{\max} = \max(E_j)$, $E_j = \sum_{i=1}^M \widehat{K}_{ji}^2$, the number of overhead lines is l_1 , and the SK relative energy entropies are C_a ($a = 1, 2, \dots, l_1$). The number of hybrid lines and cable lines is l_2 , and the SK relative energy entropies are C_b ($a = 1, 2, \dots, l_2$).

It can be known from the above definition that C_j is the information entropy that the SK energy of line j is relative to the max energy of all lines, the SK relative energy probability distributions are based on the sum of the largest SK energy, the probability space of SK energy is united, u_{ji} is a relative quantity, which has a unique sensitivity for the change of output frequency, and C_j can describe the relative uniformity for SK energy distribution in the frequency domain.

6. Fault Line Selection Criteria

In [26], it can be known from the simulation waveforms that the sign of first half-wave extremes between the fault branch

line and the healthy branch lines is opposite. However, the signs of first half-wave extremes among the branch lines are the same when the bus bar fault occurs. According to this, the criterion of bus bar fault is obtained.

6.1. Step 1—The Bus Bar Fault Criterion

- (1) The first half-wave extremes i_{0zj} of branch lines are recorded when fault occurs, and then the ratios G_{0zm} of i_{0zj} are calculated:

$$G_{0zm} = \frac{i_{0zj}}{i_{0z(j+1)}}, \quad m = 1, 2, \dots, l-1. \quad (20)$$

- (2) If $G_{0zm} > 0$, the bus bar is selected as the fault line, and the selection process is done; otherwise, the following steps should be executed.

6.2. Step 2 and Step 3—The Branch Lines Fault Criteria.

The distributed complexity and uncertainty of the TZSC are reflected by the energy entropy in the frequency domain [27]; when fault occurs, the SK relative energy entropy reflects the amount of information that can be used to select the fault line. Because of the interaction between the overhead lines and cable lines, the line with the max relative energy entropy may be not the fault line; in order to avoid the influence, the relative energy of the overhead line and cable line, hybrid line would be compared in this paper, respectively.

The entropy multiple β is introduced to judge the overhead line fault, and the calculation process is as follows:

$$\beta = \frac{C_{a \max}}{C_{as \max}}, \quad \beta \geq 1, \quad (21)$$

where $C_{a \max}$ is the max of C_a and $C_{as \max}$ is the second biggest of C_a . After a number of simulations, the threshold of entropy multiple β is set to 2 in this paper.

When the branch line fault occurs, firstly, enter step 2, judge whether the overhead line is fault or not, according to (21), calculate β , if $\beta \geq 2$ is satisfied, the overhead line of $C_{a \max}$ corresponding is selected when the fault line is selected, and the selection process is done. Otherwise, enter step 3, the hybrid line or cable line of $C_{b \max}$ corresponding is selected as the fault line, and the selection process is done.

The specific selection flowchart is shown in Figure 4.

7. Simulation and Verification

In this paper, the ATP-EMTP is used to simulate a single phase-to-ground fault, and the simulation model is shown in Figure 5, there are 4 lines (S_1 , S_2 , S_3 and S_4).

The parameters of simulation model are as follows.

Overhead line positive-sequence parameters are $R_1 = 0.17 \Omega/\text{km}$, $L_1 = 1.2 \text{ mH}/\text{km}$, and $C_1 = 9.697 \text{ nF}/\text{km}$, and zero-sequence parameters are $R_0 = 0.23 \Omega/\text{km}$, $L_0 = 5.48 \text{ mH}/\text{km}$, and $C_0 = 6 \text{ nF}/\text{km}$.

Cable line positive-sequence parameters are $R_{11} = 0.193 \Omega/\text{km}$, $L_{11} = 0.442 \text{ mH}/\text{km}$, $C_{11} = 143 \text{ nF}/\text{km}$,

TABLE 1: First half-wave extremes.

Overhead line S_1	Overhead line S_2	Hybrid line S_3	Cable line S_4
178.7320	-15.0541	-72.0865	-149.6746

and zero-sequence parameters are $R_{00} = 1.93 \Omega/\text{km}$, $L_{00} = 5.48 \text{ mH}/\text{km}$, $C_{00} = 143 \text{ nF}/\text{km}$.

Transformer is 110/10.5 kV, the leakage impedance of high voltage is $(0.40 + j12.20) \Omega$, the leakage impedance of low voltage is $(0.006 + j0.183) \Omega$, excitation current is 0.672 A, magnetizing flux is 202.2 Wb, and magnetic resistance is 400 k Ω . Load is all are delta, $Z_L = 400 + j20 \Omega$, and arc suppression coil is $L_N = 1281.9 \text{ mH}$, $R_N = 40.2517 \Omega$.

When the overhead line S_1 fault occurs, the initial fault angle is 0° and the grounding resistance is 50 Ω , and the TZSC waveforms are shown in Figure 6.

As Figure 6 shows, the TZSC is mutant when fault occurs, then they are damped. The conclusions are consistent with the analysis about the characteristics of the TZSC in Section 2; therefore, SK algorithm is introduced to analyse the fault information contained in the TZSC, and the characteristic frequency bands can be selected.

7.1. Select Based on Step 1. The first half-wave extremes i_{0zj} of $i_{0j}(t)$ ($j = 1, 2, 3, 4$) are obtained when fault occurs, and the results are in Table 1.

Because $i_{0z1}/i_{0z2} = 178.7320/(-15.0541) = -11.8726 < 0$, according to step 1, it can be known that the bus bar is healthy in the power distribution system. In order to select the fault line, step 2 is needed.

7.2. Select Based on Steps 2 and 3. The definition of characteristic frequency f_T is the frequency f that the max SK K_{\max} is corresponding in the specific condition.

According to the above analysis of SK algorithm in Section 4, Hanning window is selected, whose length is 256. And then K_i are calculated, the results are in Figure 7 (SNR = -5.5623 dB), considering too high frequency without actual physical meaning, and only the frequency of SK below 10000 Hz is listed in this paper.

The following conclusions are obtained from Figure 7.

- (1) Several different frequency components in the TZSC can be contained. For example, two kinds of different frequency within 2000 Hz are contained in S_2 at least.
- (2) Because of the flaws of SK algorithm and the impact of auxiliary white noise, the values of SK are fluctuant around 0, which are not equal to the theoretical value 0.
- (3) Different frequency components of the transient zero sequence current signals are amplified by SK; among them, the most prominent frequency components are $f_T = 781.3 \text{ Hz}$ for S_1 , because the amplitude of SK in 781.3 Hz is larger than any other amplitude of SK. Similarly, the most prominent frequency components are $f_T = 390.6 \text{ Hz}$ for S_2 , the most prominent frequency components is $f_T = 585.9 \text{ Hz}$ for S_3 , and

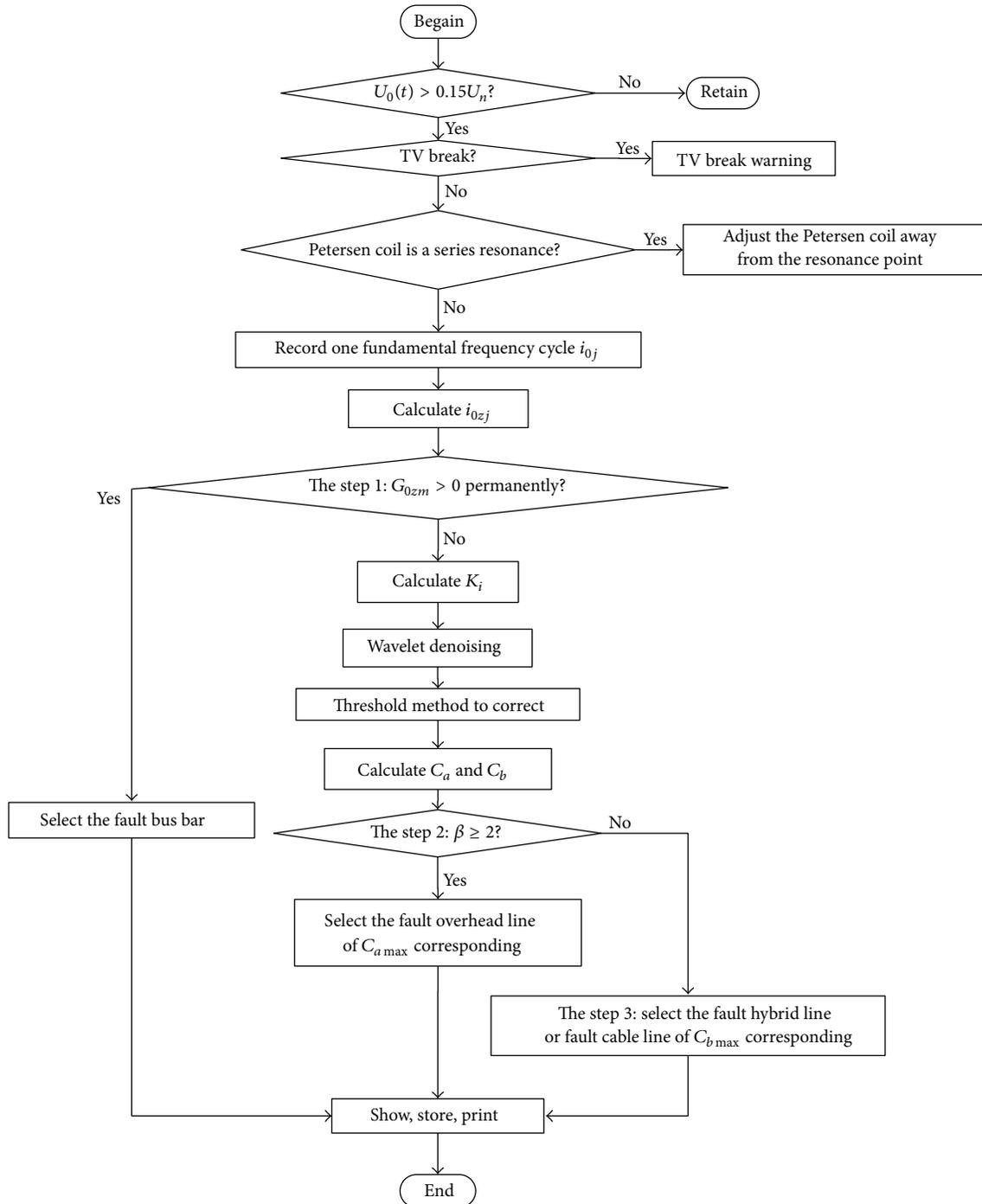


FIGURE 4: Flowchart of fault line selection.

the most prominent frequency components is $f_T = 585.9$ Hz for S_4 .

The conclusion (3) shows that the characteristic frequency f_T of fault line is larger than the f_T of healthy line in this fault condition. Therefore calculate SK of each branch line, and try to select the fault line by observing the characteristic frequency f_T directly [2]. According to the SK algorithm, f_T can be obtained with different initial angles θ

and grounding resistances R , Hanning window is selected, whose length is 256, SNR = -5.5623 dB, and the results are shown in Table 2.

The following conclusions can be obtained from Table 2.

- (1) S_1 is the fault line, and f_T of S_1 is larger than the f_T of S_2 , and f_T between S_3 and S_4 are either equal or close.
- (2) Since the inductance of the cable lines is less far than the inductance of overhead lines, the capacitance to

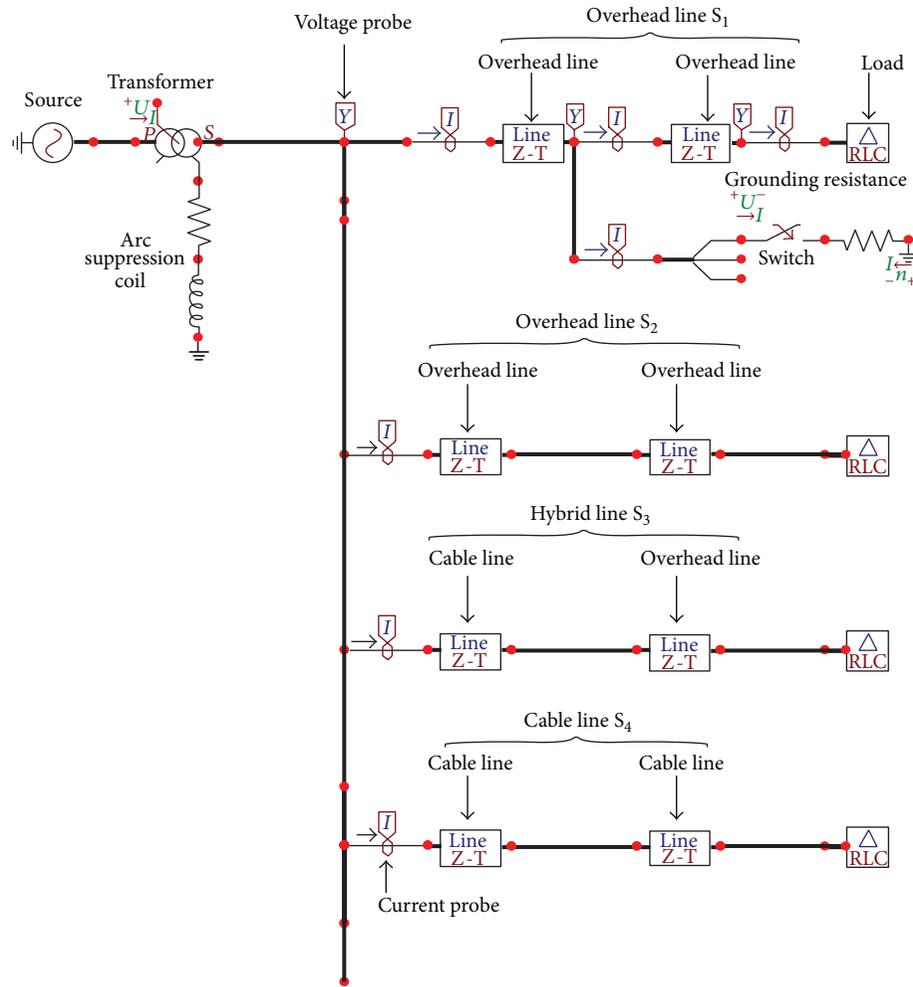


FIGURE 5: Arc suppression coil grounding system.

TABLE 2: f_T of each line with S_1 fault.

Initial angle θ	Grounding resistance		Characteristic frequency f_T (Hz)			
	R/Ω	Overhead line S_1	Overhead line S_2	Hybrid line S_3	Cable line S_4	
0°	20	585.9	390.6	585.9	585.9	
	200	3320	2344	3320	3125	
150°	20	585.9	390.6	585.9	585.9	
	200	3320	2344	3320	3125	

ground of the cable lines is larger several times than the latter, the parameter of cable lines can make a difference in the spectral characteristics of TZSC [28], and it shows that the f_T of fault overhead line S_1 is either equal or close to S_3 and S_4 .

- (3) When S_1 is fault, the initial phase angle θ is from 0° up to 150° and the grounding resistance R is fixed. The f_T of same line is unchanged; however, if the grounding resistance R is from 20Ω up to 200Ω and the initial phase angle θ fixed, the f_T of S_1 is from 585.9 Hz up to 3320 Hz , and the f_T of S_2 is from 390.6 Hz up to 2344 Hz . In conclusion, the initial phase angle θ has

little effect on f_T , but the grounding resistance has great effect on the f_T , because of the effect of the grounding resistance making the f_T nonunique.

It can be seen that the f_T would be nonunique with different resistances; therefore, the healthy line and the fault line cannot be distinguished only by f_T . Considering that the more obvious of the fault features are, the larger difference of entropy between the healthy line and fault lines is, so the SK relative energy entropy is introduced to distinguish the fault line and the healthy lines in this paper.

To eliminate the noises, whose frequency is approximate to the frequency of K_i , wavelet analysis is introduced to

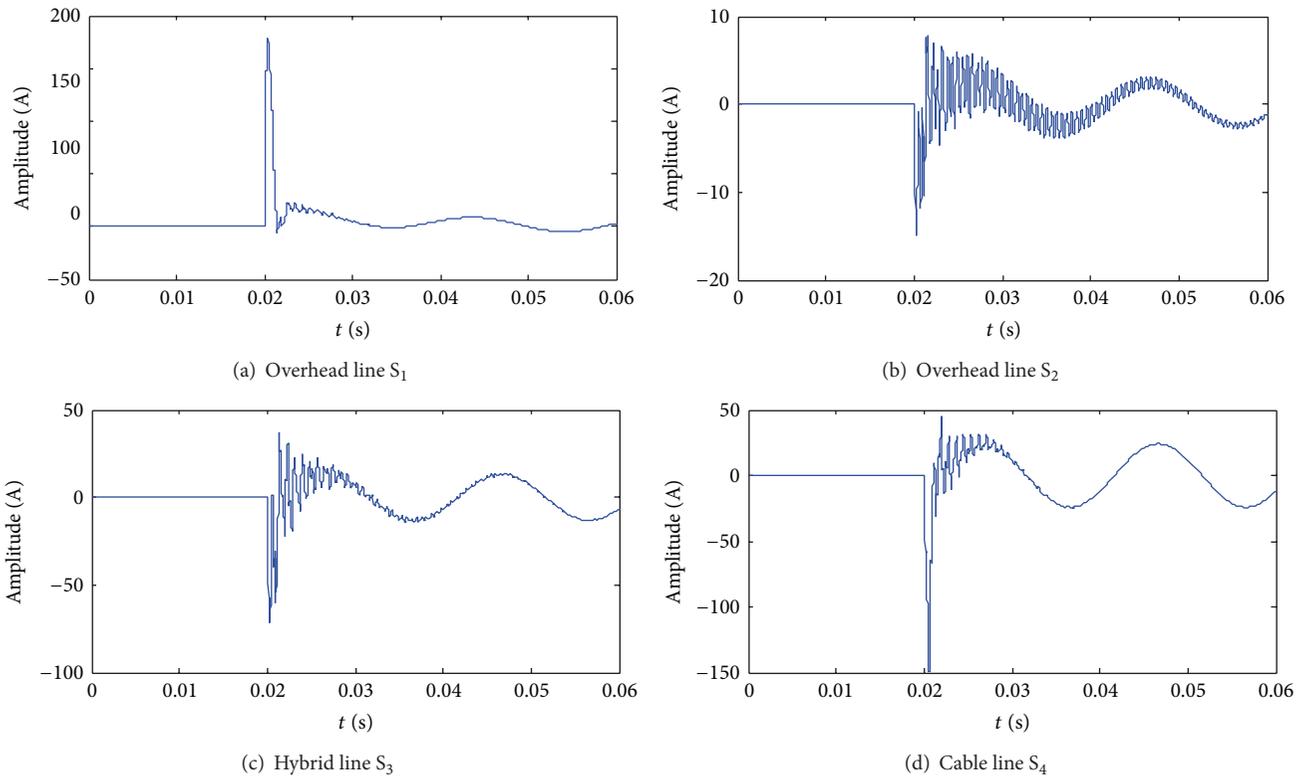


FIGURE 6: Transient zero-sequence current.

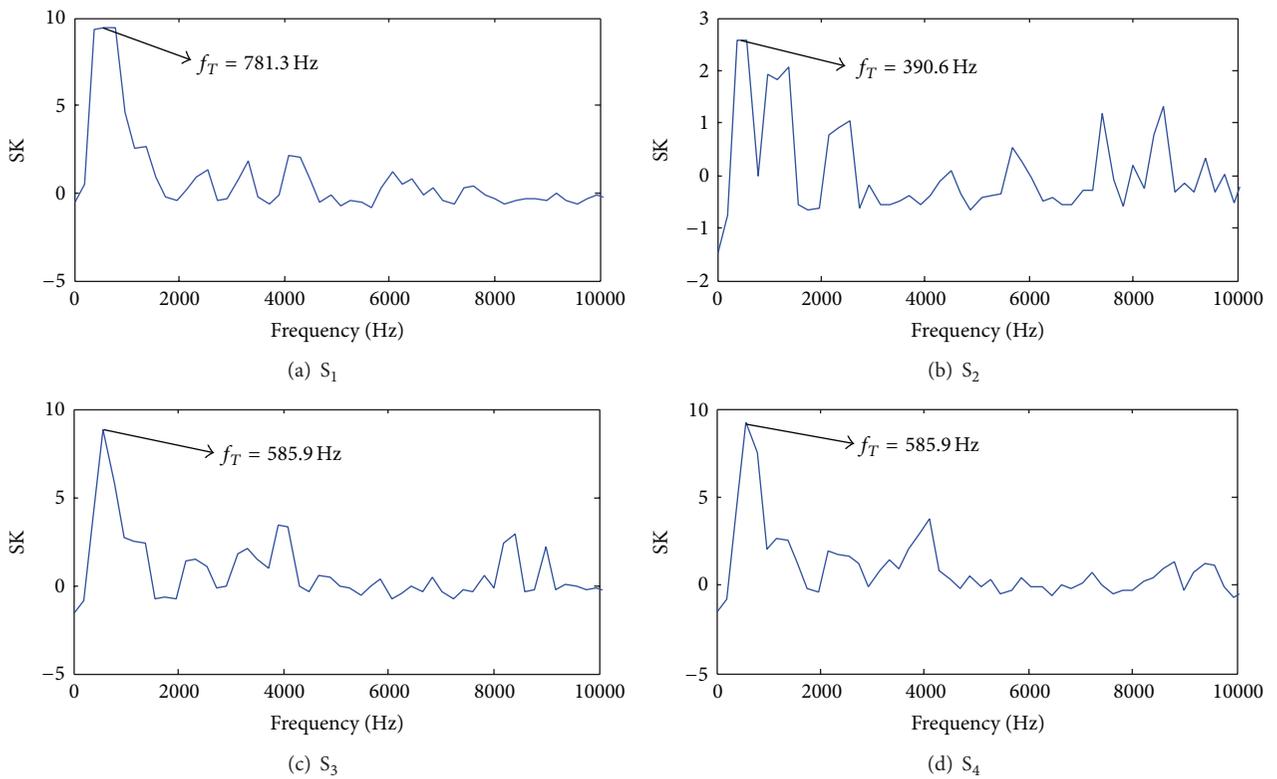


FIGURE 7: Corresponding f_T of SK.

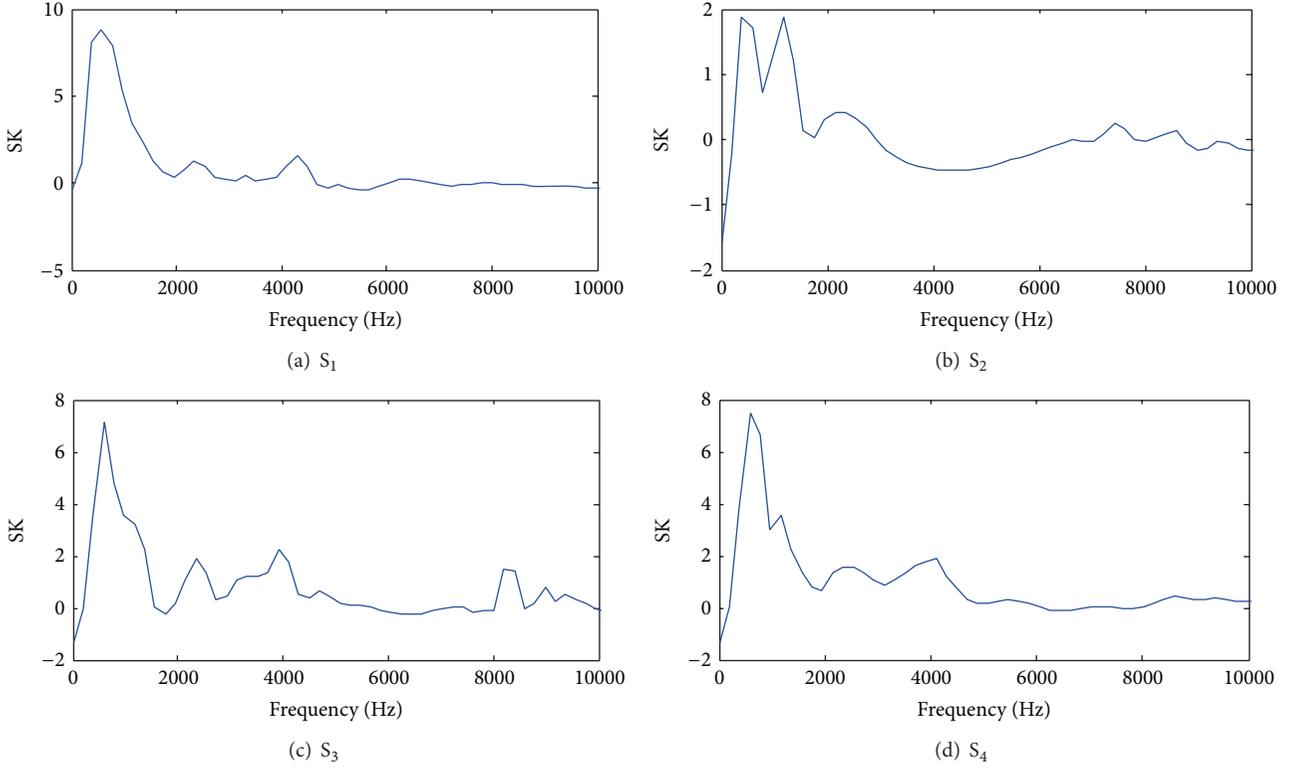


FIGURE 8: Denoising spectral kurtosis.

TABLE 3: SK relative energy entropy.

Branch line	S_1	S_2	S_3	S_4
C_j	2.2788	0.9179	1.6239	2.2834

denoising before calculating the relative energy entropy, and the results are in Figure 8.

From Figure 8, the noise signals decrease obviously after denoising, the values of SK fluctuate around 0 in the high frequency (over 5000 Hz), the waveforms are more smoothly, and the result is K'_i .

To decrease the interference from the SK of white noise to the SK relative energy entropy of TZSC, the threshold method is introduced to correct K'_i in this paper, and the finally results \widehat{K}_i are shown in Figure 9.

From Figure 9, most of the frequencies are below 5000 Hz in \widehat{K}_j , and the SK of white noise is set to 0 artificially by the threshold method actually. The results show that the effect of white noise is decreased, and a premise that calculates C_j accurately is provided. The results are in Table 3.

Based on the fault line selection method of step 2, according to (21), calculate β , $\beta = C_{a \max} / C_{as \max} = 2.2788 / 0.9179 = 2.48 > 2$ is satisfied, although C_4 is larger than C_1 , C_2 , and C_3 , in view of $\beta > 2$, and S_1 is selected as the fault line.

For S_1 belongs to the overhead line, it can be known that the bus bar is healthy after step 1, then enter step 2, calculate

β , $\beta > 2$ is satisfied, and therefore S_1 is judged as the fault line without step 3.

7.3. Results of Different Fault Types. To verify the accuracy of the proposed stepped fault line selection method further, the results are given when the bus bar fault occurs, the overhead line is S_1 , the hybrid line is S_3 , and the cable line is S_4 with different θ and R .

7.3.1. Bus Bar Fault. When the bus bar fault occurs with different initial phase angles θ and grounding resistances R , according to step 1, the selection results are in Table 4.

When the bus bar fault occurs, it can be known from Table 4 that R , θ would make a difference for i_{0zj} , which can be described as the larger R is, the lower i_{0zj} is, i_{0zj} is from negative to positive when θ is from 0° up to 150° , and the absolute value of i_{0zj} decreases. When $\theta = 0^\circ$, $i_{0zj} < 0$ is satisfied; however, when $\theta = 150^\circ$, $i_{0zj} > 0$ is satisfied, the conclusion that whatever R and θ change, $G_{0zm} > 0$ is right permanently, the criterion of bus bar fault are satisfied. Therefore, the fault bus bar can be selected by step 1, the theory of bus bar fault selection method is simple, the results are correct and credible, and the influence from R and θ is avoided.

7.3.2. Overhead Line Fault in S_2 . When the overhead line S_2 fault occurs with different θ and R , according to step 1 and step 2, the selection results are in Table 5, phase A to grounding

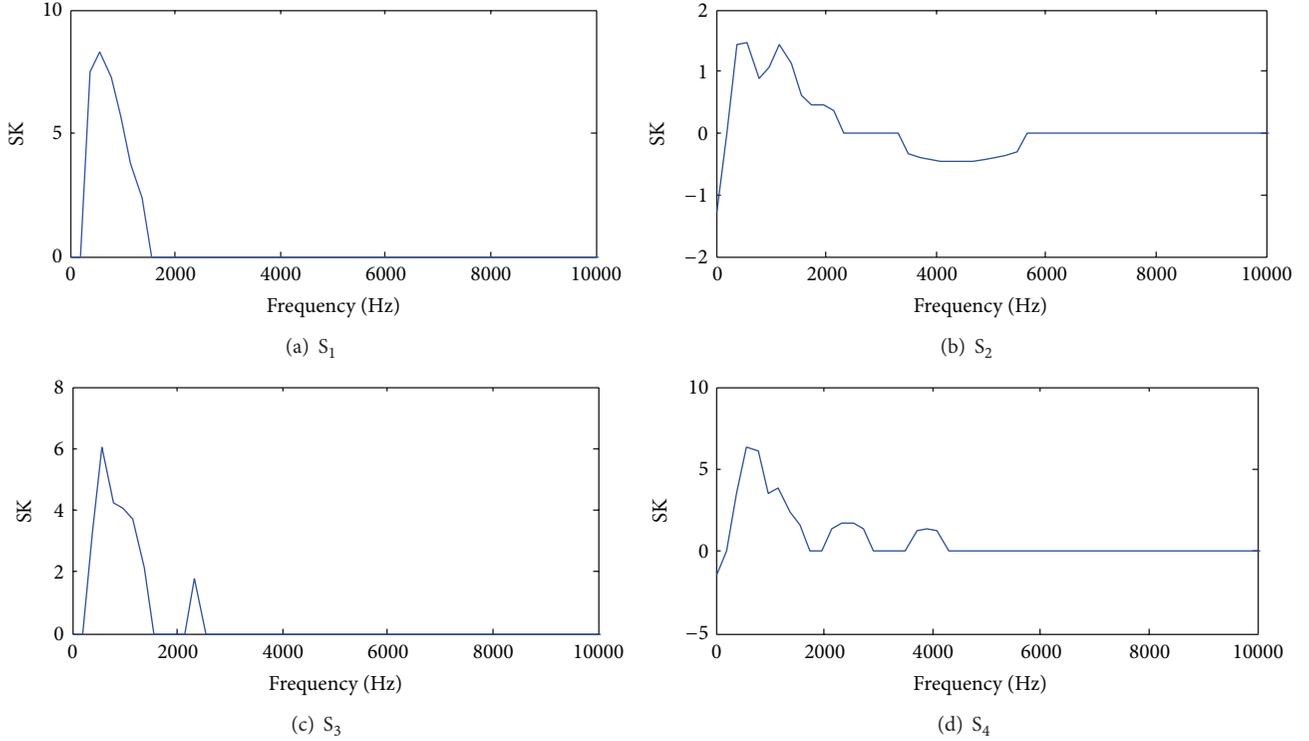


FIGURE 9: Revising spectral kurtosis.

TABLE 4: First half extremes of the TZSC with bus bar faults.

Fault line	θ	R/Ω	First half extremes of TZSC i_{0zj}				$G_{0zm} > 0?$	Results
			S_1	S_2	S_3	S_4		
Bus bar	0°	20	-36.7519	-23.2945	-157.7988	-229.0609	Y	Bus bar
		200	-9.7630	-8.8730	-35.6166	-68.9971	Y	Bus bar
	150°	20	22.1604	20.3579	98.0850	150.7914	Y	Bus bar
		200	6.0437	5.8571	23.0452	44.5919	Y	Bus bar

TABLE 5: SK relative energy entropy with S_2 fault.

Fault line	θ	R/Ω	SK relative energy entropy C_j				Entropy multiple β	Results
			S_1	S_2	S_3	S_4		
Overhead line S_2	0°	20	0.5324	3.9787	1.8911	2.4950	7.47 > 2	S_2
		200	1.1521	4.4164	3.5851	3.2250	3.83 > 2	S_2
	150°	20	0.5761	2.9607	2.2864	2.2134	5.14 > 2	S_2
		200	2.9918	7.0234	7.8605	5.0687	2.35 > 2	S_2

fault occurs, and the fault location is 5 km away from the bus bar.

From Table 5, the SK relative energy entropy of fault line is larger than the healthy lines in most cases. According to step 2, calculate β , because of $\beta = C_{a \max}/C_{as \max} \geq 2$, the overhead line is judged as fault line preliminary, because C_2 belongs to the $C_{a \max}$, S_2 is selected as the fault line finally, and the selection results are accurate without step 3. When the overhead line fault occurs, the results show that utilizing the stepped fault line selection method to select the fault line only

to enter step 1 and step 2, and without step 3. The selection process is simplified.

From the results of Tables 4 and 5, the fault line selection method not only ensures the accuracy of fault line selection but also can save the time of fault line selection by steps 1 and 2. Besides, the fault line selection method cannot be influenced by R and θ .

7.3.3. Hybrid Line S_3 and Cable Line S_4 Fault. When the hybrid line S_3 or cable line S_4 fault occurs with different θ

TABLE 6: SK relative energy entropy with S_3 and S_4 faults.

Fault line	θ	R/Ω	SK relative energy entropy C_j				Entropy multiple β	Results
			S_1	S_2	S_3	S_4		
Hybrid line S_3	0°	20	1.4287	1.5408	4.9297	4.8318	$1.08 < 2$	S_3
		200	1.0894	1.1655	3.5634	2.8949	$1.07 < 2$	S_3
	150°	20	1.1355	1.1165	4.4099	3.8247	$1.02 < 2$	S_3
		200	0.7395	0.4809	3.5649	2.1245	$1.54 < 2$	S_3
Cable line S_4	0°	20	0.8222	0.7218	3.7140	5.1669	$1.14 < 2$	S_4
		200	1.7250	1.1495	2.5230	4.0801	$1.50 < 2$	S_4
	150°	20	0.4447	0.4522	3.3416	4.1139	$1.02 < 2$	S_4
		200	2.8497	4.0614	3.1244	3.6334	$1.43 < 2$	S_4

TABLE 7: SK relative energy entropy with different fault distances ($\theta = 0^\circ$, $R = 20 \Omega$).

Fault line	SK relative energy entropy C_j				Entropy multiple β	Results
	S_1	S_2	S_3	S_4		
Overhead line S_1	2.7267	1.2616	2.1665	2.8215	$2.16 > 2$	S_1
Cable line S_4	1.6140	1.1455	3.9845	4.3558	$1.41 < 2$	S_4

TABLE 8: SK relative energy entropy with different fault phases ($\theta = 0^\circ$, $R = 20 \Omega$).

Fault line	Fault phase	SK relative energy entropy C_j				Entropy multiple β	Results
		S_1	S_2	S_3	S_4		
Overhead line S_1	B	3.0161	1.3442	1.7646	2.9313	$2.24 > 2$	S_1
	C	2.1819	0.4247	1.0312	1.4207	$5.14 > 2$	S_1
Cable line S_4	B	0.6462	0.7357	3.5225	5.2236	$1.14 < 2$	S_4
	C	2.0681	3.6244	1.2483	1.4300	$1.75 < 2$	S_4

and R , respectively, according to steps 1, 2, and 3, phase A to ground fault occurs, the fault location is 5 km away from the bus bar, and the stepped fault line selection results are in Table 6.

From Table 6, according to steps 1 and 2, calculate β , because of $\beta = C_{a \max}/C_{as \max} < 2$, the fault line belongs to S_3 or S_4 preliminary, then step 3 is needed, the line of $C_{b \max}$ corresponding is selected as the fault line, the fault line S_3 or S_4 can be selected after step 3 finally. From the above, the fault hybrid line or the fault cable line can be selected accurately after steps 1, 2, and 3.

8. Applicability Analyses

8.1. Different Fault Distances. To verify the proposed method adaptability with different fault distances, θ and R are fixed, and the simulations are as follows. Taking the fault line S_1 and S_4 as examples, the fault locations are 10 km and 2 km away from the bus bar, respectively; the results are in Table 7.

From Table 7, when S_1 fault occurs with different fault distances, based on step 2, because of $\beta_1 = 2.7267/1.2616 = 2.16 > 2$, step 3 is not needed, and S_1 is selected as the fault line finally. Similarly, based on step 2 when S_4 fault occurs, because of $\beta_2 = 1.6140/1.1455 = 1.41 < 2$, step 3 is needed, since $4.3558 > 3.9845$, and S_4 is selected as the fault line finally. The results show the stepped fault line selection method still applies to the changing fault distances.

8.2. Different Fault Phases. To verify the proposed method adaptability about different fault phases, taking the fault line S_1 and S_4 as examples, the results are as shown in Table 8.

From Table 8, when S_1 fault occurs with different fault phases, based on step 2, because of $\beta_1 = 2.7267/1.2616 = 2.16 > 2$ and $\beta_2 = 2.1819/0.4247 = 5.15 > 2$, step 3 is not needed, and S_1 is selected as the fault line finally. Similarly, based on step 2 when S_4 fault with the changing fault phases, since $\beta_3 = 0.7357/0.6462 = 1.14 < 2$ and $\beta_4 = 3.6244/2.0681 = 1.75 < 2$, step 3 is needed, because of $5.2236 > 3.5225$ and $1.4300 > 1.2483$, and S_4 is selected as the fault line finally. The results show the stepped fault line selection method still applies to the changing fault phases.

8.3. Simulate the Practical Data. Because the conditions of practical environment are poor when fault occurs, therefore, the practical fault data contains much noise. In this paper, in order to analyze the differences between the practical data and the simulation data intuitively, the different intensity noises are injected into the TZSC to simulate the practical data, when SNR is -1.5678 dB and -13.5678 dB, and the results are as shown in Figure 10.

From Figure 10, it can be known that the simulation data can simulate the practical data with the noise background. That is to say, it is possible to verify the proposed algorithm using the practical data. Taking the fault lines S_1 and S_4 as examples, the results are shown in Table 9.

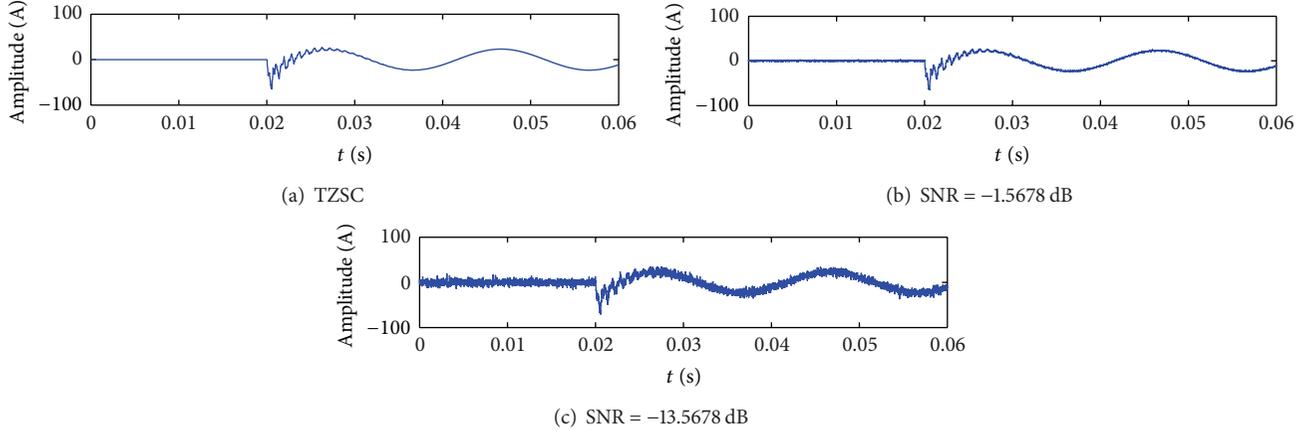


FIGURE 10: Simulate the practical fault TZSC.

TABLE 9: SK relative energy entropy with different intensity noise ($\theta = 0^\circ$, $R = 20 \Omega$).

SNR/dB	Fault line	SK relative energy entropy C_j				Entropy multiple β	Results
		S_1	S_2	S_3	S_4		
-1.5678	Overhead line S_1	3.0665	1.2711	1.3520	2.7091	$2.41 > 2$	S_1
	Cable line S_4	0.9898	0.8089	4.5347	5.6922	$1.22 < 2$	S_4
-5.5678	Overhead line S_1	2.9061	1.1725	1.8192	2.7349	$2.48 > 2$	S_1
	Cable line S_4	0.9375	0.5388	3.7573	5.2826	$1.74 < 2$	S_4
-10.5678	Overhead line S_1	2.8449	1.0359	1.7698	2.4071	$2.75 > 2$	S_1
	Cable line S_4	0.8267	0.4481	3.4586	3.9673	$1.84 < 2$	S_4
-13.5678	Overhead line S_1	2.6814	0.6951	2.0311	2.1538	$3.86 > 2$	S_1
	Cable line S_4	1.1979	0.7646	4.5726	5.8559	$1.57 < 2$	S_4

From Table 9, when overhead line S_1 fault occurs, no matter what SNR is equal to, according to steps 1 and 2, $\beta > 2$ establish permanently, step 3 is not needed, and S_1 is selected as the fault line finally. Similarly, when S_4 fault occurs, no matter what SNR is equal to, based on steps 1, 2, and 3, $\beta < 2$ establish all the time, step 3 is needed, since $C_4 > C_3$, and S_4 is selected as fault line finally. It can be seen that the different fault lines can be selected correctly with different intensity noise, and the results show the stepped fault line selection method can apply to the simulative practical fault data.

8.4. Incipient Faults. Compared with the overhead line, the cable line is safer and more economic, so the cable lines are widely used in modern cities. However, the cables are more easily in the state of incipient faults before they fail into permanent faults. Such as the cable line S_4 is in the state of incipient faults in this paper. Usually, incipient faults in power cables gradually result from the aging process, where the localized deterioration in insulation exists. Incipient faults are normally characterized as the faulty phenomena with the relatively small fault currents and the relatively short duration ranging from one-quarter cycle to multicycle. These short lasting current variations cannot be detected by the traditional distribution protection schemes because of their short duration and low increment in magnitude.

However, such faults must be detected at the early stage to avoid the consequent catastrophe induced by the degradation themselves [29].

The incipient faults of underground cables are similar to the intermittent arc faults; therefore, the arc model is essential to the analysis of the incipient faults. Paper [30] introduces the arc model; after theoretical analysis and practical verification, arc has nonlinear and time-varying characteristics, and the high-frequency components can be produced. Due to the fact that the arc model is easy, the nonlinear and time-varying characteristics can be expressed, and the arc model is introduced in this paper.

The arc model is built by the energy balance of arc column, and the differential equations of conductivity are used to express

$$\frac{dp}{dt} = \frac{1}{\tau} (P - p), \quad (22)$$

where τ is the time constant of arc, p is the instantaneous conductivity of arc, and P is stable conductivity of arc.

Stable conductivity can be defined as

$$P = \frac{|i_{\text{arc}}|}{u_{\text{st}}}, \quad u_{\text{st}} = u_0 + r_0 \cdot |i_{\text{arc}}|, \quad (23)$$

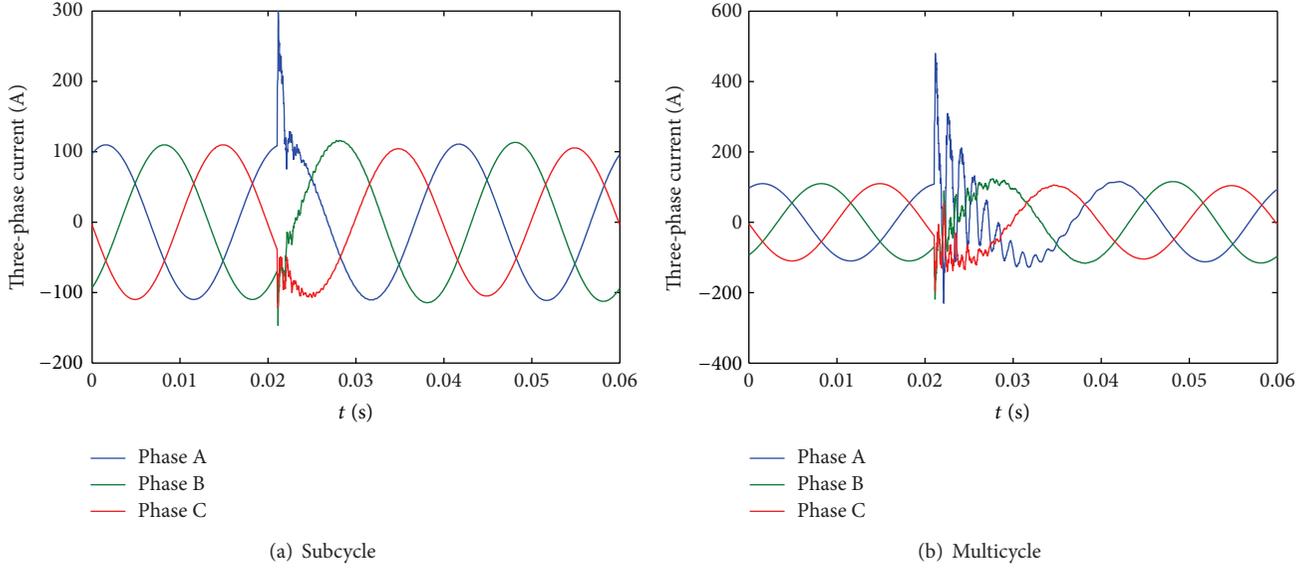


FIGURE 11: Feeder currents of incipient faults.

TABLE 10: SK relative energy entropy with incipient faults.

Fault line	θ	R/Ω	SK relative energy entropy C_j				Entropy multiple β	Results
			S_1	S_2	S_3	S_4		
Cable line S_4	0°	1	1.1021	0.6623	4.7252	5.6418	$1.66 < 2$	S_4
		50	0.3872	0.3602	3.5047	4.1438	$1.07 < 2$	

where i_{arc} is the instantaneous current of arc, u_{st} is static voltage of arc, u_0 is the voltage characteristics of arc, r_0 is the arc resistance, and the characteristic parameters u_0 and r_0 depend on the length l_{arc} of arc, which can be calculated by

$$u_0 = \left(\frac{0.9l_{\text{arc}}}{m} + 0.4 \right) \text{kV} \quad (24)$$

$$r_0 = \left(\frac{40l_{\text{arc}}}{m} + 8 \right) \text{m}\Omega. \quad (25)$$

Equation (25) is generalized arc equation. In the resonance grounding system, the parameters of the small current arc change depend on the length of arc, which is the arc elongation elongspd . elongspd can be defined as

$$\text{elongspd} = \frac{dl_{\text{arc}}}{dt} = \frac{7l_0}{(0.2v_{\text{th}}/v_{\text{max}}) + 0.2}, \quad (26)$$

where v_{th} is the instantaneous value of transient initial voltage and v_{max} is the max normal voltage.

The length of arc is defined as a time constant before simulating. The time constant of arc is defined as

$$\tau = \tau_0 \cdot \left(\frac{l_{\text{arc}}}{l_0} \right)^\alpha, \quad (27)$$

where τ_0 is the initial time constant, l_0 is the initial length of arc, α is the negative coefficient, and the general value of α is -4 .

The length of arc is

$$l_{\text{arc}} = (\text{elongspd} \cdot (t - t_f) + 1) \cdot l_0, \quad (28)$$

where t is the simulation time and t_f is the moment that the arc occurs.

In underground cables, the incipient fault is one type of transient, which is prone to an intermittent arc fault. The typical incipient faults are composed of two types: subcycle and multicycle incipient fault. The subcycle incipient fault always occurs near a voltage peak where the arc is ignited, lasts around one-quarter cycle, and self-clears when the current crosses zero. Figure 11(a) shows the three-phase feeder currents when a subcycle incipient fault occurs between phase A to ground. The multicycle incipient fault also likely occurs near a voltage peak, lasts 1–4 cycles, and self-clears when the arc is quenched. The waveforms of the currents are shown in Figure 11(b).

To verify the adaptability of the proposed method when the incipient faults occur, the TZSC of incipient faults is obtained, and then, according to the steps 1, 2, and 3, the selection results are as shown in Table 10.

From Table 10, based on step 2 when S_4 occur incipient faults, it can be known that $\beta_1 = 1.1021/0.6623 = 1.66 < 2$ and $\beta_2 = 0.3872/0.3602 = 1.07 < 2$, step 3 is needed, and since $5.6418 > 4.7252$ and $4.1438 > 3.5047$, S_4 is selected as the fault line finally. The results show that the stepped fault line selection method can also adapt to the incipient faults.

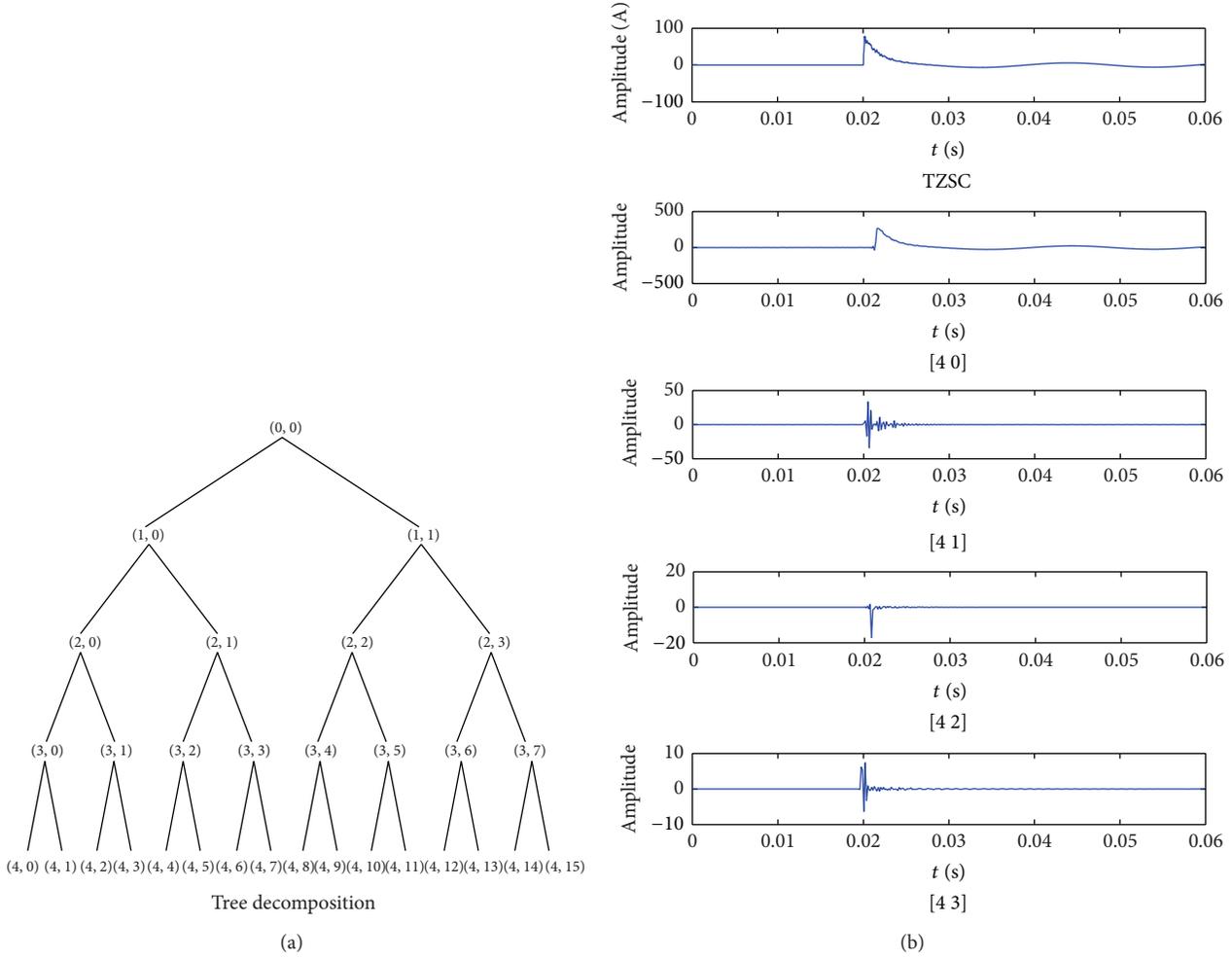


FIGURE 12: Wavelet decompose the TZSC.

TABLE 11: Wavelet relative energy entropy.

Fault line	θ	R/Ω	Wavelet relative energy entropy C_j				Entropy multiple β	Results	Right?
			S_1	S_2	S_3	S_4			
Overhead line S_2	0°	20	2.1015	2.4464	2.6801	4.1830	$1.16 < 2$	S_4	N
		200	1.2593	2.1661	3.6776	3.3619	$1.72 < 2$	S_3	N
	150°	20	2.1669	2.5476	2.5867	4.4014	$1.18 < 2$	S_4	N
		200	1.1975	1.9841	3.6561	3.6506	$1.66 < 2$	S_3	N

9. Compared with Other Fault Line Selection Methods

9.1. *Wavelet Relative Energy Entropy.* The wavelet mother function “db10” is used to decompose the TZSC [12], and the decomposed layers are 4. Select the frequency bands [4 0], [4 1], [4 2], and [4 3], and the results are shown in Figure 12.

From Figure 12, the lowest frequency band [4 0] is removed, the biggest energy band [4 1] is selected as the characteristic band, because the characteristic band contains the main features of TZSC. And then wavelet denoising and the threshold are used to correct the characteristic band coefficients, successively. Finally, the wavelet relative energy

entropies C_j can be obtained, and the selection results are as shown in Table 11.

From Table 11, the method based on wavelet relative energy entropy cannot select the fault line accurately. When the overhead line S_2 fault occurs, according to (21), the entropy multiple β can be calculated, $C_1 > C_2$, while $\beta < 2$ is established, the selection results are incorrectly, because of the big differences on the distribution parameters between the cable line and the overhead line, which make the healthy cable line’s entropies bigger than the fault overhead line. Moreover, before the wavelet transform, the appropriate mother wavelet function should be selected, while the selection process is difficult. It can be seen that the accuracy of

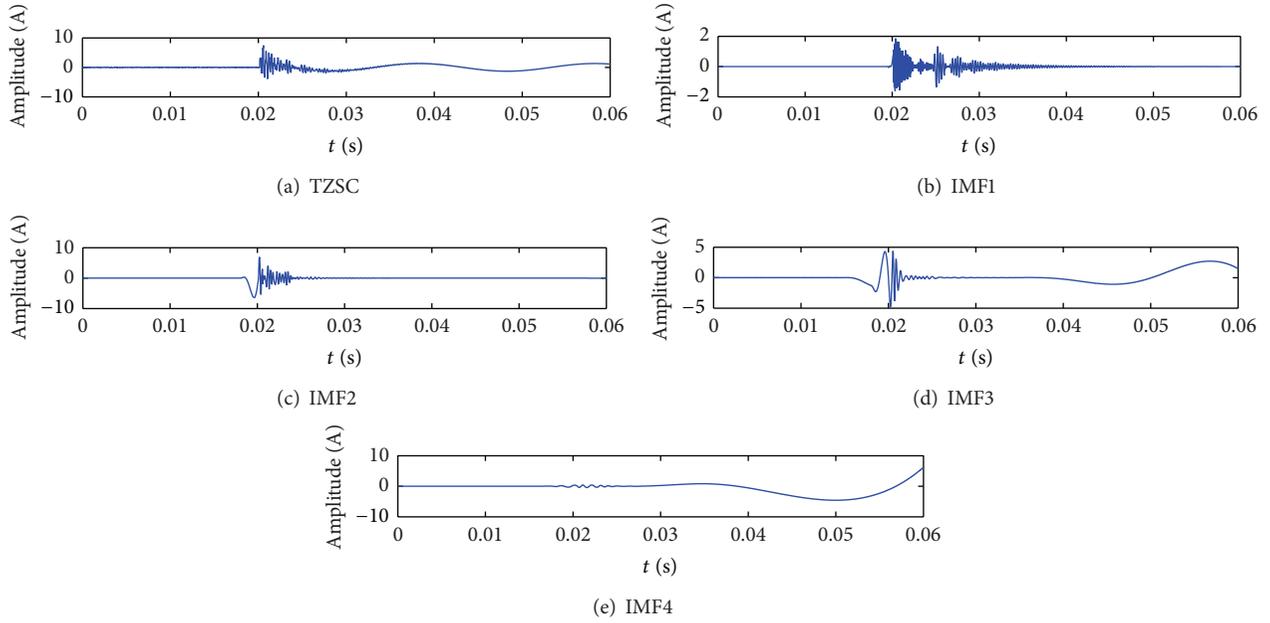


FIGURE 13: EMD decompose the TZSC.

TABLE 12: EMD relative energy entropy.

Fault line	θ	R/Ω	EMD relative energy entropy C_j				Entropy multiple β	Results	Right?
			S_1	S_2	S_3	S_4			
Overhead line S_2	0°	20	1.9112	4.5143	5.1857	4.9335	$2.36 > 2$	S_2	Y
		200	0.8032	3.8554	5.9764	4.8981	$4.80 > 2$	S_2	Y
	150°	20	2.0273	4.4264	5.7486	4.3812	$2.18 > 2$	S_2	Y
		200	0.9719	4.1269	6.8780	6.2598	$4.25 > 2$	S_2	Y
Hybrid line S_3	0°	20	0.7617	1.1209	4.8099	2.8481	$1.47 < 2$	S_3	Y
		200	0.6714	0.9501	4.7789	2.9695	$1.12 < 2$	S_3	Y
	150°	20	1.0806	1.5968	5.3013	4.0003	$1.48 < 2$	S_3	Y
		200	0.7927	1.3501	5.0642	2.7225	$1.70 < 2$	S_3	Y
Cable line S_4	0°	20	0.5502	0.9428	5.4257	4.2543	$1.71 < 2$	S_4	Y
		200	0.2029	0.7532	2.0443	3.7942	$3.71 > 2$	S_2	N
	150°	20	0.2706	0.3861	4.8885	4.0831	$1.43 < 2$	S_3	N
200		0.3037	1.5814	4.2905	5.5391	$5.21 > 2$	S_2	N	

the method based on wavelet relative energy entropy is not guaranteed.

9.2. EMD Relative Energy Entropy. The local characteristics of the TZSC can be represented by empirical mode decomposition (EMD); after decomposition, IMF components can be obtained [8], and the waveforms of the mode function IMF1~IMF4 are as shown in Figure 13.

In view of IMF1 component belonging the highest frequency component of the TZSC, containing rich transient information, so IMF1 is selected as the characteristic component, and then wavelet denoising and the threshold are used to correct the characteristic component coefficients. Finally, the relative energy entropies of IMF1 can be obtained, and the selection results are as shown in Table 12.

From Table 12, when the cable line S_4 fault occurs, after decomposition by EMD, the EMD relative energy entropy of each branch line can be obtained, and then the entropy multiple β can be calculated, β is bigger than the threshold, according to steps 1 and 2, the cable line S_4 belongs to the healthy line, and the selection results are incorrect. Moreover, the decomposition process may cause modal aliasing, which can make a big difference on IMF1 component to represent the TZSC. Therefore, it can be seen that the accuracy of the method based on EMD relative energy entropy is not guaranteed.

From Tables 5, 6, 11, and 12, when the branch line fault occurs, because of the big difference of distribution parameters between the cable line and overhead line, the relative energy entropy of each branch line can be obtained, according to the fault line selection steps 1, 2, and 3, and the accuracy

of the method based on wavelet relative energy entropy and EMD relative energy entropy is not guaranteed. Moreover, wavelet transform needs to select the appropriate mother wavelet function, while the selection process is difficult and the modal aliasing may be caused by EMD decomposition process, while SK algorithm can decompose the TZSC in time-frequency domain adaptively, which contributes to the accuracy of the selection results. Therefore, the accuracy of selection method based on SK relative energy entropy is higher than the methods based on the wavelet relative energy entropy and EMD relative energy entropy.

10. Conclusions

This paper proposes a stepped fault line selection method based on SK relative energy entropy, and the conclusions are as follows.

- (1) SK belongs to higher order statistics with good ability to restrain the noise and can characterize the TZSC in frequency domain as well. The relative homogeneity of SK energy distribution is described quantitatively by the SK relative energy entropy. In addition, utilizing the max SK relative energy entropy as the criterion to select the fault line, the transient universal characteristic quantity is selected in fact. However, the SK is limited by the compromise problem of time-frequency domain, and the type of window function and the length of window should be determined at the beginning.
- (2) The ratios of the first half-wave extremes between neighboring lines and SK relative energy entropy are regarded as fault features to select the fault line. When the bus bar occur fault, the ratios of the first half-wave extreme between neighboring lines are positive. If the entropy multiple is greater than or equal to the threshold, judge the overhead line fault, if not, judge the hybrid line or the cable line fault.

Future work should concentrate on how to determine the appropriate auxiliary noise intensity according to the characteristics of TZSC. Moreover, higher adaptability of the method based on SK relative energy entropy should be advanced with the high resistance.

Notations

HHT: Hilbert-Huang Transform
 IMF: Intrinsic mode functions
 EMD: Empirical mode decomposition
 SK: Spectral kurtosis
 TZSC: Transient zero-sequence current
 DC: Direct current
 STFT: Short-time Fourier transform.

Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

Acknowledgments

This work was supported by the Science and Technology Research (12B470003, 14A470004), the Control Engineering Lab Project (KG2011-15) of Henan Province, China, and the Youth Foundation (Q2012-28, Q2012-43A) of Henan Polytechnic University, China.

References

- [1] Z. X. Zhang, X. Liu, and L. Z. Piao, "Fault line detection in neutral point ineffectively grounding power system based on phase-locked loop," *IET Generation, Transmission & Distribution*, vol. 8, no. 2, pp. 273–280, 2014.
- [2] M. F. Guo, G. J. Yang, and S. Y. Huang, "Adaptive faulty line selection based on transient characteristics for resonant earthing system," *Electric Power Automation Equipment*, vol. 32, no. 12, pp. 25–31, 2012.
- [3] L. Wu, C. Huang, D. Lin, Z. Zhu, and H. Jiang, "Faulty line selection based on transient wavelet energy for non-solid-earthed network," *Electric Power Automation Equipment*, vol. 33, no. 5, pp. 70–75, 2013.
- [4] D. Zhou, Z. Liu, F. Hu, and W. Bai, "A new method for fault line selection based on wavelet de-noising and transient current energy grouping comparison in ineffective grounding system," *Power System Protection and Control*, vol. 38, no. 7, pp. 22–28, 2010.
- [5] J. Zhang, Z. Y. He, and Y. Jia, "Fault line identification approach based on S-transform," *Proceedings of the CSEE*, vol. 31, no. 10, pp. 109–155, 2011.
- [6] H. C. Shu and S. X. Peng, "A fault line detection algorithm for distribution network of overhead line and underground cable mixed lines using S-transform energy from short window data," *Transactions of China Electrotechnical Society*, vol. 24, no. 10, pp. 152–158, 2009.
- [7] X. W. Wang, J. W. Wu, and R. Y. Li, "A novel method of fault selection based on voting mechanism of Prony relative entropy theory," *Electric Power*, vol. 46, no. 1, pp. 59–65, 2013.
- [8] S. Q. Zhang, X. P. Zhai, and X. Dong, "Application of EMD and Duffing oscillator to fault line detection in un-effectively grounded system," *Proceedings of the CSEE*, vol. 33, no. 10, pp. 161–167, 2013.
- [9] H. C. Shu, W. Y. Zhao, and S. X. Peng, "Faulty line selection based on HHT detection for hybrid distribution network," *Electric Power Automation Equipment*, vol. 29, no. 5, pp. 4–10, 2009.
- [10] S. Wang and J. F. Zhu, "Faulty line selection of single-phase to ground fault in distribution network based on improved correlation analysis method," *Power System Protection and Control*, vol. 40, no. 15, pp. 76–81, 2012.
- [11] H. C. Shu, L. Xu, and S. X. Peng, "An approach for fault feeder line detection in resonant earthed system using correlation analysis," *Electric Power Automation Equipment*, vol. 28, no. 9, pp. 6–9, 2008.
- [12] H. C. Shu and S. X. Peng, "Distribution network fault line detection using the full waveband complex relative entropy of wavelet energy," *High Voltage Engineering*, vol. 35, no. 7, pp. 1559–1564, 2009 (Chinese).
- [13] F. Immovilli, M. Cocconcelli, A. Bellini, and R. Rubini, "Detection of generalized-roughness bearing fault by spectral-kurtosis energy of vibration or current signals," *IEEE Transactions on Industrial Electronics*, vol. 56, no. 11, pp. 4710–4717, 2009.

- [14] F. Millioz and N. Martin, "Circularity of the STFT and spectral kurtosis for time-frequency segmentation in Gaussian environment," *IEEE Transactions on Signal Processing*, vol. 59, no. 2, pp. 515–524, 2011.
- [15] H. R. Karimi, P. J. Maralani, B. Lohmann, and B. Moshiri, " H_∞ control of parameter-dependent state-delayed systems using polynomial parameter-dependent quadratic functions," *International Journal of Control*, vol. 78, no. 4, pp. 254–263, 2005.
- [16] G. Cheng, Z. G. Liu, and Q. G. Zhang, "Partial discharge signals extraction based on spectral kurtosis," *Electric Power Automation Equipment*, vol. 33, no. 8, pp. 95–110, 2013.
- [17] J. Li, T. Y. Jiang, S. Grzybowski, and C. K. Cheng, "Scale dependent wavelet selection for de-noising of partial discharge detection," *IEEE Transactions on Dielectrics and Electrical Insulation*, vol. 17, no. 6, pp. 1705–1714, 2010.
- [18] S. Yin, S. X. Ding, A. Haghani, H. Hao, and P. Zhang, "A comparison study of basic data-driven fault diagnosis and process monitoring methods on the benchmark Tennessee Eastman process," *Journal of Process Control*, vol. 22, no. 9, pp. 1567–1581, 2012.
- [19] H. R. Karimi, "A computational method for optimal control problem of time-varying state-delayed systems by Haar wavelets," *International Journal of Computer Mathematics*, vol. 83, no. 2, pp. 235–246, 2006.
- [20] J. Li, C. Cheng, T. Jiang, and S. Grzybowski, "Wavelet denoising of partial discharge signals based on genetic adaptive threshold estimation," *IEEE Transactions on Dielectrics and Electrical Insulation*, vol. 19, no. 2, pp. 543–549, 2012.
- [21] X. N. Kang, X. Liu, J. L. Suonan, C. Ma, C. Wang, and L. Yang, "New method for fault line selection in non-solidly grounded system based on matrix pencil method," *Automation of Electric Power Systems*, vol. 36, no. 12, pp. 88–93, 2012.
- [22] S. Yin, H. Luo, and X. Steven, "Real-time implementation of fault-tolerant control systems with performance optimization," *IEEE Transactions on Industrial Electronics*, vol. 61, no. 5, pp. 2402–2411, 2014.
- [23] Q. L. Wang and Z. X. Fu, "An adaptive method for single-phase-to-ground fault line selection based on energy entropy measure," *Automation of Electric Power Systems*, vol. 36, no. 5, pp. 103–107, 2012.
- [24] H. R. Karimi, "Robust H_∞ filter design for uncertain linear systems over network with network-induced delays and output quantization," *Modeling, Identification and Control*, vol. 30, no. 1, pp. 27–37, 2009.
- [25] Q. Q. Jia, L. L. Shi, and N. Wang, "A fusion method for ground fault line detection in compensated power networks based on evidence theory and information entropy," *Transactions of China Electrotechnical Society*, vol. 27, no. 6, pp. 192–197, 2012.
- [26] Y. N. Wang, B. I. Huo, and H. Wang, "New criterion for earth fault line selection based on wavelet packets in small current neutral grounding system," *Proceedings of the Chinese Society of Electrical Engineering*, vol. 24, no. 6, pp. 54–58, 2004.
- [27] H. R. Karimi, N. A. Duffie, and S. Dashkovskiy, "Local capacity control for production networks of autonomous work systems with time-varying delays," *IEEE Transactions on Automation Science and Engineering*, vol. 7, no. 4, pp. 849–857, 2010.
- [28] H. S. Zhang, Z. Y. He, and J. Zhang, "Frequency spectrum characteristic analysis of single-phase grounding fault in resonant grounded systems," *Automation of Electric Power Systems*, vol. 36, no. 6, pp. 79–83, 2012.
- [29] T. S. Sidhu and Z.-H. Xu, "Detection of incipient faults in distribution underground cables," *IEEE Transactions on Power Delivery*, vol. 25, no. 3, pp. 1363–1371, 2010.
- [30] M. Kizilcay and P. la Seta, "Digital simulation of fault arcs in medium-voltage distribution networks," in *Proceedings of the 15th Power Systems Computation Conference*, pp. 1–7, 2005.

Research Article

Estimating Target Heights Based on the Earth Curvature Model and Micromultipath Effect in Skywave OTH Radar

Hou Chengyu, Wang Yuxin, and Chen Jiawei

School of Electronics and Information Engineering, Harbin Institute of Technology, Harbin 150001, China

Correspondence should be addressed to Hou Chengyu; houcy@hit.edu.cn

Received 23 April 2014; Accepted 19 June 2014; Published 14 July 2014

Academic Editor: Weichao Sun

Copyright © 2014 Hou Chengyu et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Skywave over-the-horizon (OTH) radar systems have important long-range strategic warning values. They exploit skywave propagation reflection of high frequency signals from the ionosphere, which provides the ultra-long-range surveillance capabilities to detect and track maneuvering targets. Current OTH radar systems are capable of localizing targets in range and azimuth but are unable to achieve reliable instantaneous altitude estimation. Most existing height measurement methods of skywave OTH radar systems have taken advantage of the micromultipath effect and been considered in the flat earth model. However, the flat earth model is not proper since large error is inevitable, when the detection range is over one thousand kilometers. In order to avoid the error caused by the flat earth model, in this paper, an earth curvature model is introduced into OTH radar altimetry methods. The simulation results show that application of the earth curvature model can effectively reduce the estimation error.

1. Introduction

Skywave OTH radar works in the high frequency band (3–30 MHz), which uses the ionosphere to scatter the electromagnetic wave, observing the air and ground targets from top to bottom. Thus, skywave OTH radar has important long-range strategic warning value and a wide range of applications [1].

Skywave OTH radar systems are capable of localizing targets in range and azimuth but are unable to achieve reliable altitude estimation [2]. The altitude parameter is an important parameter for target recognition and tracking. Most existing altitude-finding theories for skywave OTH radar systems are based on the micromultipath effect. This effect refers to the fact that the radar signals have at least four kinds of paths through the relatively stable ionosphere and ground or sea surface reflection, when transmitted from the transmitting antenna and scattered by targets back to receiving antenna [3].

Currently, the methods that skywave OTH radar systems used to estimate the instantaneous altitude of the aircrafts can be divided into four kinds. (1) The superresolution method: this method uses modern spectral analysis method

to distinguish echo delay of each propagation path and then estimates the target altitude. By this method, the height estimation is divided into three types: low altitude, intermediate altitude, and high altitude. But this technique can only be applied to analog data [3]. (2) The state-space multipath fading model: after a series of delays, the micromultipath echoes interfered by each other will cause distance-Doppler peak attenuation. Then, using of the Fourier analysis to the amplitude attenuation echoes, we can obtain altitude information [4]. (3) The high-resolution ionospheric sounder: through the experiment data by the high-resolution ionosphere sounder, the relation between the micromultipath effect and the target altitude is studied [5]. (4) The matched-field processing [6, 7]: the echo signal is used to rematch delay-Doppler space. And then the maximum likelihood estimation (MLE) is adopted to estimate the target altitude [8].

The methods mentioned above mostly consider the radar altimetry problem in the flat earth model. However, for skywave OTH radar systems, the detection range is thousands of kilometers, it will cause large errors to estimate target altitude with the flat earth model. Therefore, this paper will focus on how to introduce the earth curvature model into the altimetry method.

2. Signal Model and Maneuvering Target Model

2.1. Establishing Echo Signal Model. In this paper, a monostatic multiple-input multiple-output (MIMO) radar system is considered, which consists of N_t closely spaced transmit antennas and N_r closely spaced receive antennas. Denote $S \in \mathbb{C}^{N_t \times T}$ as the orthogonal narrowband waveform matrix which is transmitted from the N_t transmit antennas over the pulse repetition period T . Assume that the orthogonal waveform is completed in the coherent integral time. s_i is the i th row of matrix S . When $i \neq m$, s_i is orthogonal to s_m ; then $SS^H = I_{N_t}$. So, for the k th pulse, the signal received at the m th receive antenna is expressed as [8, 9]

$$u_{m,k} = \sum_{i=1}^I \sum_{n=1}^{N_t} \rho_{i,k} s_n e^{j\phi_{i,m,k}^{[A]}} e^{j\phi_{i,n,k}^{[D]}} + w_{m,k}, \quad (1)$$

where $I = 4$ is the number of the multipath components and $\rho_{i,k}$ is the complex response of the i th multipath that considers the target reflection and propagation delays. The superscripts $[D]$ and $[A]$ denote the departure and the arrival modes, respectively. $d_m^{[A]}$ is the location of receive antennas. $\phi_{i,m,k}^{[A]}$ is determined by the m th receive antenna and is expressed as

$$\phi_{i,m,k}^{[A]} = \left(\frac{2\pi}{\lambda} \right) d_m^{[A]} \sin(\theta_{i,k}^{[A]}), \quad (2)$$

where $\theta_{i,k}^{[A]}$ is the elevation angle of arrival of the i th path. Similarly, $\phi_{i,n,k}^{[D]}$ is determined by the location of the n th transmit antenna $d_n^{[D]}$, and $\phi_{i,n,k}^{[D]}$ is expressed as

$$\phi_{i,n,k}^{[D]} = \left(\frac{2\pi}{\lambda} \right) d_n^{[D]} \sin(\theta_{i,k}^{[D]}), \quad (3)$$

where $\theta_{i,k}^{[D]}$ is the elevation angle of departure of the i th path. For simplicity, we assume that the clutter has been filtered. Thus $w_{m,k}$ is the additive noise, which is independent and identically distributed complex Gaussian random variable and independent of the target echo signal.

So the k th pulse echo, which is transmitted by the m th transmit antenna and received by the n th receive antenna, can be expressed as

$$r_{m,n,k} = \sum_{i=1}^I \rho_{i,k} e^{j\phi_{i,m,k}^{[A]}} e^{j\phi_{i,n,k}^{[D]}} + w_{m,n,k}. \quad (4)$$

Denote $\varphi_{i,k} = [e^{j\phi_{i,1,k}^{[A]} + j\phi_{i,1,k}^{[D]}}, \dots, e^{j\phi_{i,N_t,k}^{[A]} + j\phi_{i,N_t,k}^{[D]}}]^T \in \mathbb{C}^{N \times 1}$ as the signal steering vector of the virtual array corresponding to the i th path at the k th pulse, where $N = N_t N_r$. Denote $A_k = [\varphi_{1,k}, \dots, \varphi_{I,k}]$, $\rho_k = [\rho_{1,k}, \dots, \rho_{I,k}]^T$, and $W_k = [w_{1,1,k}, \dots, w_{N_t, N_r, k}]^T$, so the measurement radar data vector can be expressed as

$$r_k = [r_{1,1,k}, \dots, r_{N_t, N_t, k}]^T = A_k \rho_k + W_k. \quad (5)$$

In order to be consistent with the motion model, the k th pulse of the measurement data is expressed with the time t corresponding to the k th pulse.

TABLE 1: The parameters of the motion target.

Parameter	Notation	Value
Initial ground range	x_0	2000 km
Ionosphere height	H	350 km
Aircraft initial height	z_0	10 km
Maximum horizontal velocity	$v_x(0)$	500 km/h
Maximum descending velocity	$v_{z, \max}$	90 km/h
Carrier frequency	f_C	20 MHz
Impulse repetition frequency	f_S	50 Hz

2.2. Establish Maneuvering Target Model. In this paper, for generality, we consider a maneuvering aircraft which makes a 180° circular turn in $T = 30$ s [10]. The motion parameters are listed in Table 1. All the multipath signals are considered to fall within the same range cell. The ground distance $x(t)$ and flight altitude of the target $z(t)$, respectively, are expressed as

$$\begin{aligned} x(t) &= x(0) - \frac{v_x(0)T}{\pi} \sin\left(\frac{t\pi}{T}\right), \\ z(t) &= z(0) + \frac{v_z(0)T}{\pi} \left(1 - \cos\left(\frac{t\pi}{T}\right)\right). \end{aligned} \quad (6)$$

The target tracks in two dimensions are shown in Figures 1 and 2.

3. Analysis of Micromultipath Effect Based on the Flat Earth Model

3.1. Analysis of Micromultipath Effect. An OTH radar system considering the micromultipath effect based on the flat earth model is illustrated in Figure 3.

Since our purpose in this paper is to estimate the instantaneous height of the target, for simplicity and without loss of generality, we only consider the 2D position and velocity (range direction x and altitude direction z). Then the target state at the time t is expressed as

$$\bar{x}(t) = [x(t) \ z(t) \ v_x(t) \ v_z(t)]^T, \quad (7)$$

where $v_x(t)$ and $v_z(t)$ are the velocity of the range direction and the altitude direction, respectively.

As the effect of the cross-range array apertures on the height estimation is small, only the effect of array apertures which lies in the range direction will be considered. Correspondingly both the transmit arrays and receive arrays are considered to be linear and located in the x -axis.

As the micromultipath effect exists, the combination of the transmit path and receive path has four components: Path I ($l_1 : l_1$), Path II ($l_2 : l_2$), Path III ($l_1 : l_2$), and Path IV ($l_2 : l_1$). Path III and Path IV are with the same length [6]. The slant range $l_1(t)$ and $l_2(t)$ is expressed with the distance $x(t)$, the height of the ionosphere H , and the target height $z(t)$ as

$$\begin{aligned} l_1(t) &= \sqrt{x(t)^2 + (2H - z(t))^2}, \\ l_2(t) &= \sqrt{x(t)^2 + (2H + z(t))^2}. \end{aligned} \quad (8)$$

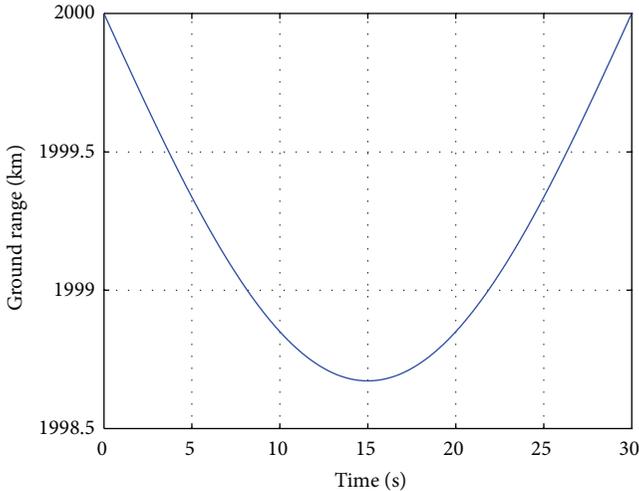


FIGURE 1: The target's track in the range direction.

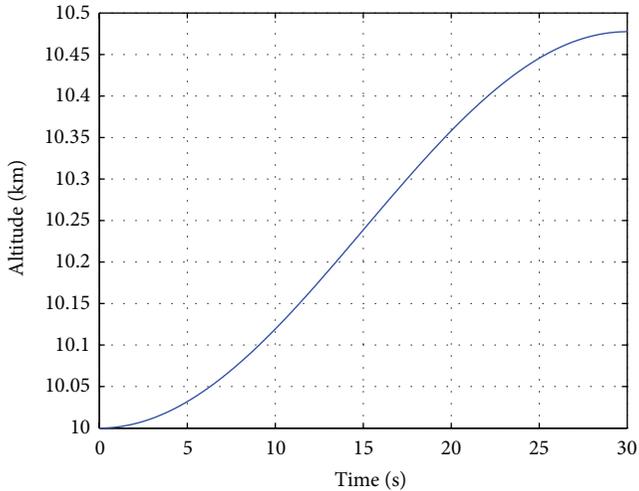


FIGURE 2: The target's track in the altitude direction.

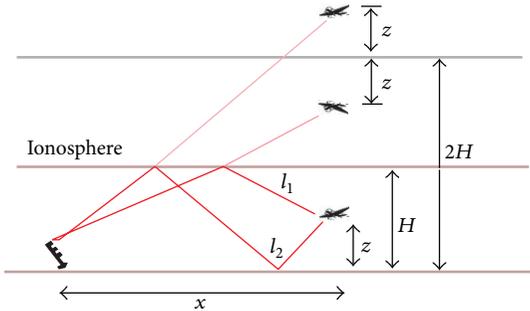


FIGURE 3: The micromultipath effect in the flat earth model.

As $x \gg H \gg z$, $l_1(t)$ and $l_2(t)$ are simplified as

$$\begin{aligned} l_1(t) &\approx x(t) + \frac{2H^2 - 2Hz(t)}{x(t)}, \\ l_2(t) &\approx x(t) + \frac{2H^2 + 2Hz(t)}{x(t)}. \end{aligned} \quad (9)$$

Then $L_1(t)$ and $L_2(t)$, the length of Path I and Path II, are expressed as

$$\begin{aligned} L_1(t) &= 2l_1(t) \approx 2x(t) + \frac{4H^2 - 4Hz(t)}{x(t)}, \\ L_2(t) &= 2l_2(t) \approx 2x(t) + \frac{4H^2 + 4Hz(t)}{x(t)}. \end{aligned} \quad (10)$$

$L_3(t)$ and $L_4(t)$, the length of Path III and Path IV, are written as

$$L_3(t) = L_4(t) = l_1(t) + l_2(t) \approx 2x(t) + \frac{4H^2}{x(t)}. \quad (11)$$

3.2. Instantaneous Target Altitude Estimation. In (4), $\rho_i(t)$ comprises the targets' Doppler information. Here we express $\rho_i(t)$ as

$$\rho_i(t) = \sigma_i(t) e^{-j2\pi L_i(t)/\lambda}, \quad (12)$$

where $i = 1, \dots, 4$. i is the index of the path. $\lambda = c/f_c$ is the wavelength corresponding to carrier frequency f_c , and $\sigma_i(t)$ represents the combined effect of transmit power, target reflection, which is a function of radar cross-section (RCS), and ionospheric effect. $L_i(t)$ is the two-way slant range, and the expressions of $L_i(t)$ are shown in the previous section.

In general, the target's track consists of the movement in the range and altitude direction. This section will discuss the Doppler frequency of the target flight in two directions and then deduce the instantaneous expression of the target height.

Suppose that both the range between the target and radar and the height of the target are the functions of time. As the height of the ionosphere H is changing slowly, we assume that H is a constant value in the coherent time. So we can get

$$\begin{aligned} \frac{dl_1(t)}{dt} &\approx J(t) v_x(t) - \frac{2H}{x(t)} v_z(t), \\ \frac{dl_2(t)}{dt} &\approx J(t) v_x(t) + \frac{2H}{x(t)} v_z(t), \end{aligned} \quad (13)$$

where $J(t) = (1 - 2H^2/x^2(t))$, $v_x(t) = dx(t)/dt$ is the target's velocity of the range direction, and $v_z(t) = dz(t)/dt$ is the target's velocity of the altitude direction. So the Doppler frequencies of the four paths are

$$f_1(t) = \frac{2f_c}{c} \frac{dl_1(t)}{dt} \approx \frac{2f_c}{c} J(t) v_x(t) - \frac{4f_c H}{x(t)c} v_z(t), \quad (14)$$

$$f_2(t) = \frac{2f_c}{c} \frac{dl_2(t)}{dt} \approx \frac{2f_c}{c} J(t) v_x(t) + \frac{4f_c H}{x(t)c} v_z(t), \quad (15)$$

$$f_3(t) = f_4(t) = \frac{f_c}{c} \frac{dl_1(t) + dl_2(t)}{dt} \approx \frac{2f_c}{c} J(t) v_x(t). \quad (16)$$

From (14)–(16), we find that the Doppler frequencies of the four paths have the same part $2f_c J(t)v_x(t)/c$, which is determined by $v_x(t)$. We call this the dominant Doppler component. c is the speed of light. f_1 and f_2 also have the part caused by $v_z(t)$. We call this the minor Doppler component, which contains the important information of the target movement in the altitude direction. Let $f_2(t) - f_1(t)$; then $v_z(t)$ is obtained by the expression

$$v_z(t) = \frac{[f_2(t) - f_1(t)] x(t) c}{8f_c H}. \quad (17)$$

To get the target height, it is necessary to know the target initial height. In practice, we can get $L_1(t)$, $L_2(t)$ and $L_3(t)$ by the time delay, and the transmit angle α and the ionospheric height H can also be known by the other measurement equipment. Then these parameters can be used to calculate the initial height z_0 . According to (8), we can get two estimated values z_0 by

$$z_0 = 2H - \sqrt{\left(\frac{L_1(0)}{2}\right)^2 - x_0^2}, \quad (18)$$

$$z_0 = \sqrt{\left(\frac{L_2(0)}{2}\right)^2 + x_0^2} - 2H.$$

However, z_0 gotten by (18) is quite different from the actual value. It is because z_0 is very small compared to $x(t)$. If $x(t)$ is with a small error, there will be a great impact on z_0 .

For elimination of the impact by the error of $x(t)$, we can use the following to calculate $z(t)$:

$$\begin{aligned} x(t)^2 + [2H - z(t)]^2 &= L_1^2, \\ x(t)^2 + [2H + z(t)]^2 &= L_2^2. \end{aligned} \quad (19)$$

$z(t)$ is obtained by (19) as

$$z(t) = \frac{L_2(t)^2 - L_1(t)^2}{8H}. \quad (20)$$

Let $t = 0$, z_0 be gotten. Then $z(t)$ can be written as

$$z(t) = \frac{L_2(0)^2 - L_1(0)^2}{8H} + \frac{[f_2(t) - f_1(t)] x(t) c}{8f_c H} \times t. \quad (21)$$

4. Analysis of Micromultipath Effect Based on Earth Curvature Model

4.1. Analysis of Micromultipath Effect. In this section, the micromultipath effect model, which is under the influence of the earth curvature, will be discussed. The micromultipath effect model in an OTH radar system of the earth curvature model is illustrated in Figure 4. $x(t)$ is expressed as the ground range between the target and the radar. In order to be consistent with the definition in the flat earth model, $z(t)$

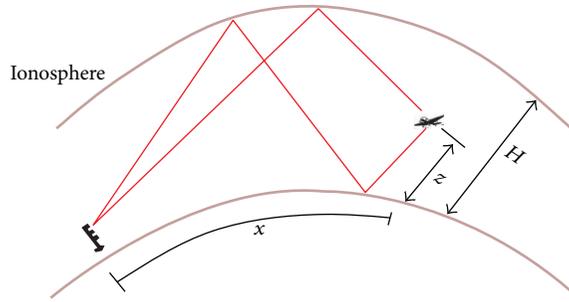


FIGURE 4: The micromultipath effect in the earth curvature model.

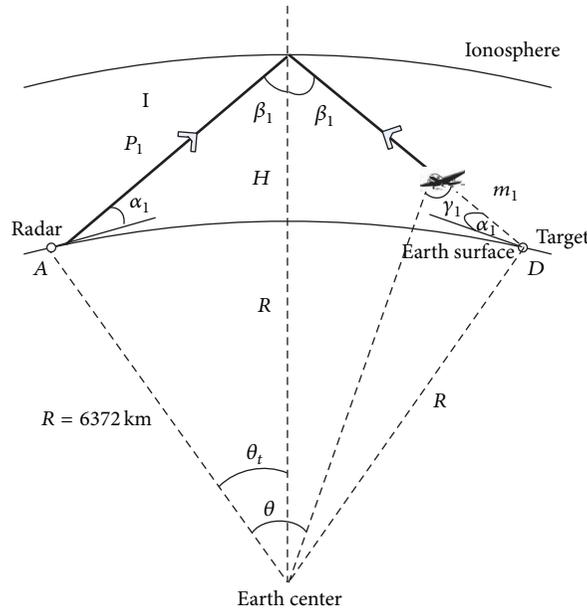


FIGURE 5: Path I in the earth curvature model.

indicates the vertical height, H is the height of the ionosphere, and R is the earth radius (it is set as a constant value in the paper; $R = 6372$ km). Same as the flat earth model, the radar signal propagation paths in the earth curvature model have four combinations: Path I ($l_1 : l_1$), Path II ($l_2 : l_2$), Path III ($l_1 : l_2$), and Path IV ($l_2 : l_1$).

Path I is shown in Figure 5, in which α_1 indicates the departure angle of radar transmitting antenna, P_1 is the distance of electromagnetic signal from radar transmitting antenna to ionosphere, m_1 is the extended line, and β_1, γ_1 indicate the auxiliary angles.

According to the geometrical relationship, α_1 meets the following condition:

$$\frac{x(t)}{R} + \alpha_1 + 2 \arcsin \left[\frac{R \cos(\alpha_1)}{R + H} \right] - \arcsin \left[\frac{R}{R + z} \cos(\alpha_1) \right] = \frac{\pi}{2}. \quad (22)$$

The accurate value of α_1 can be calculated by (22). The auxiliary angles β_1 and γ_1 can be expressed as the function of α_1 :

$$\beta_1 = \arcsin \left[\frac{R \cos(\alpha_1)}{R + H} \right], \quad \gamma_1 = \arcsin \left[\frac{R \cos(\alpha_1)}{R + z(t)} \right]. \quad (23)$$

P_1 and m_1 can be, respectively, expressed as

$$P_1 = \frac{R \cos(\alpha_1 + \beta_1)}{\sin \beta_1}, \quad (24)$$

$$m_1 = P_1 + \frac{H + R}{\sin \gamma} \cos \left[\frac{x(t)}{R} + \alpha_1 + \beta \right].$$

Then $l_1(t)$ is written as

$$l_1(t) = 2P_1 - m_1 = R \frac{\cos(\alpha_1 + \beta_1)}{\sin \beta_1} - \frac{H + R}{\sin \gamma_1} \cos \left[\frac{x(t)}{R} + \alpha_1 + \beta_1 \right]. \quad (25)$$

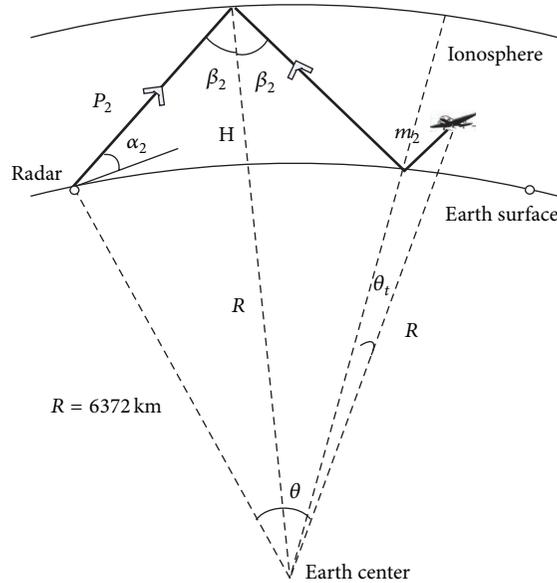


FIGURE 6: Path II in the earth curvature model.

Path II is illustrated in Figure 6. Similarly, α_2 indicates the departure angle of radar transmitting antenna, P_2 is the distance of electromagnetic signal from radar transmitting antenna to ionosphere, m_2 is the extended line, and β_2 and γ_2 indicate the auxiliary angles.

From Figure 6, α_2 should meet the following conditions:

$$\frac{x(t)}{R} + \arcsin\left(\frac{R}{R+z} \cos \alpha_2\right) + 2 \arcsin\left(\frac{R}{R+H} \cos \alpha_2\right) + 3\alpha_2 = \frac{3\pi}{2}. \quad (26)$$

Then α_2 can be calculated by (26). The auxiliary angles β_2 and γ_2 can be expressed as the function of α_2 :

$$\beta_2 = \arcsin\left(\frac{R \cos \alpha_2}{R+H}\right), \quad \gamma_2 = \arcsin\left[\frac{R \cos \alpha_2}{R+z(t)}\right]. \quad (27)$$

P_2 and m_2 can be, respectively, expressed as

$$P_2 = \frac{R \cos(\alpha_2 + \beta_2)}{\sin \beta_2}, \quad m_2 = \frac{\cos(\alpha_2 + \gamma_2)}{\cos \alpha_2} [R + z(t)]. \quad (28)$$

Then the expression of transmission oblique diameter $l_2(t)$ is as follows:

$$l_2(t) = 2P_2 + m_2 = 2R \frac{\cos(\alpha_2 + \beta_2)}{\sin \beta_2} + \frac{\cos(\alpha_2 + \gamma_2)}{\cos \alpha_2} [R + z(t)]. \quad (29)$$

So $L_1(t)$ and $L_2(t)$, the length of Path I and Path II, are

$$L_1(t) = 2l_1(t) = 2R \frac{\cos(\alpha_1 + \beta_1)}{\sin \beta_1} - 2 \frac{H+R}{\sin \gamma_1} \cos\left[\frac{x(t)}{R} + \alpha_1 + \beta_1\right], \quad (30)$$

$$L_2(t) = 2l_2(t) = 4R \frac{\cos(\alpha_2 + \beta_2)}{\sin \beta_2} + 2 \frac{\cos(\alpha_2 + \beta_2)}{\cos \alpha_2} [R + z(t)].$$

And $L_3(t)$ and $L_4(t)$, the length of Path III and Path IV, are shown as

$$L_3(t) = L_4(t) = l_1(t) + l_2(t) = R \frac{\cos(\alpha_1 + \beta_1)}{\sin \beta_1} - \frac{H+R}{\sin \gamma_1} \cos\left[\frac{x(t)}{R} + \alpha_1 + \beta_1\right] + 2R \frac{\cos(\alpha_2 + \beta_2)}{\sin \beta_2} + \frac{\cos(\alpha_2 + \gamma_2)}{\cos \alpha_2} [R + z(t)]. \quad (31)$$

4.2. Instantaneous Target Altitude Estimation. Consistent with the processing methods in the flat earth model, we

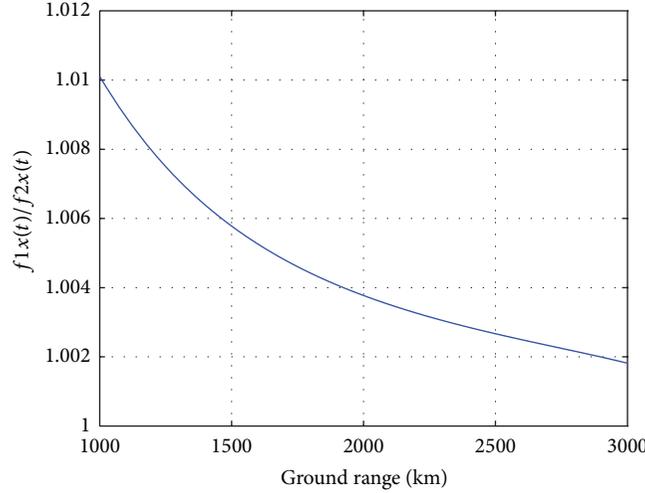


FIGURE 7: The ration of 11x to 12x.

suppose that H is constant in the coherent time. Then the time derivatives of the $l_1(t)$, $l_2(t)$, $l_3(t)$, and $l_4(t)$ are obtained by

$$\begin{aligned}
 \frac{dl_1(t)}{dt} &= \frac{\sin[\alpha_1 + \beta_1 + x(t)/R][R + z(t)](H + R)}{R^2 \cos \alpha_1} v_x(t) \\
 &\quad - \frac{\cos[\alpha_2/2 + \beta_1 + x(t)/R](H + R)}{R \cos \alpha_1} v_z(t) \\
 &= l_{1x} v_x(t) + l_{1z} v_z(t), \\
 \frac{dl_2(t)}{dt} &= \left(\left(\cos \left[\frac{\alpha_2}{2} - \frac{\pi}{4} + \frac{\gamma}{2} + \frac{x(t)}{2R} \right] (H + R) \right) \right. \\
 &\quad \times (R \cos \alpha_2)^{-1} \left. \right) v_x(t) \\
 &\quad + \left\{ \frac{\cos(\alpha_2 + \gamma_2)}{\cos \alpha_2} + \frac{R \sin(\alpha_2 + \gamma_2)}{[R + z(t)] \cos \gamma_2} \right. \\
 &\quad \left. - \left(R \cos \left[\frac{\alpha_2}{2} - \frac{\pi}{4} + \frac{\gamma}{2} + \frac{x(t)}{2R} \right] (H + R) \right) \right. \\
 &\quad \left. \times \left(\sqrt{1 - \left[\frac{R \cos \alpha_2}{R + z(t)} \right]^2} [R + z(t)]^2 \right)^{-1} \right\} v_z(t), \\
 &= l_{2x} v_x(t) + l_{2z} v_z(t) \\
 \frac{dl_3(t)}{dt} &= \frac{dl_4(t)}{dt} = \frac{1}{2} \left[\frac{dl_1(t)}{dt} + \frac{dl_2(t)}{dt} \right]. \tag{32}
 \end{aligned}$$

l_{1x} is the coefficient of the part of $dl_1(t)/dt$ caused by $v_x(t)$. l_{1z} is the coefficient of the part of $dl_1(t)/dt$ caused by $v_z(t)$. The expressions of l_{2x} and l_{2z} are similar to l_{1x} and l_{1z} , but replace

l_1 with l_2 . So the Doppler frequencies produced by the moving target of the four paths can be, respectively, written as

$$\begin{aligned}
 f_1(t) &= \frac{2f_c}{c} \frac{dl_1(t)}{dt}, \\
 f_2(t) &= \frac{2f_c}{c} \frac{dl_2(t)}{dt}, \\
 f_3(t) &= f_4(t) = \frac{1}{2} [f_1(t) + f_2(t)]. \tag{33}
 \end{aligned}$$

The four Doppler frequencies both consist of the components caused by $v_x(t)$ and $v_z(t)$. The ratio of l_{1x} to l_{2x} under the condition that $x(t)$ is made to change from 1000 km to 3000 km is shown in Figure 7. From the figure, it can be found that the values of Doppler frequency components caused by $v_x(t)$ in $f_1(t)$ and $f_2(t)$ are very close. And it indicates that the Doppler frequency component in each path caused by $v_x(t)$ is approximately equal. This conclusion is same as the one in the flat earth model. Therefore, $v_z(t)$ can be obtained by $f_2(t) - f_1(t)$ as

$$v_z(t) = \frac{c}{2f_c} \frac{f_2(t) - f_1(t)}{l_{2z} - l_{1z}}. \tag{34}$$

To get the instantaneous height of the moving target, the initial height of the target z_0 is needed. Different from the situation in the flat earth model, z_0 can be calculated accurately by using $L_1(t)$ or $L_2(t)$. In order to reduce random errors, $L_1(t)$ and $L_2(t)$ are used to calculate the initial heights z_{01} , z_{02} and then make the average of z_{01} , z_{02} as z_0 .

For Path I, z_{01} can be shown as follows:

$$z_{01} = \left[\frac{L_1(0)}{2} - \frac{\sin \theta_1}{\sin \beta_1} \right] \frac{\sin \beta_1}{\sin \theta_2} - R, \tag{35}$$

where

$$\beta_1 = \arcsin \left(\frac{R \cos \alpha_1}{H + R} \right), \tag{36}$$

$$\theta_2 = \frac{x(t)}{R} - \frac{\pi}{2} + \alpha_1 + \beta_1.$$

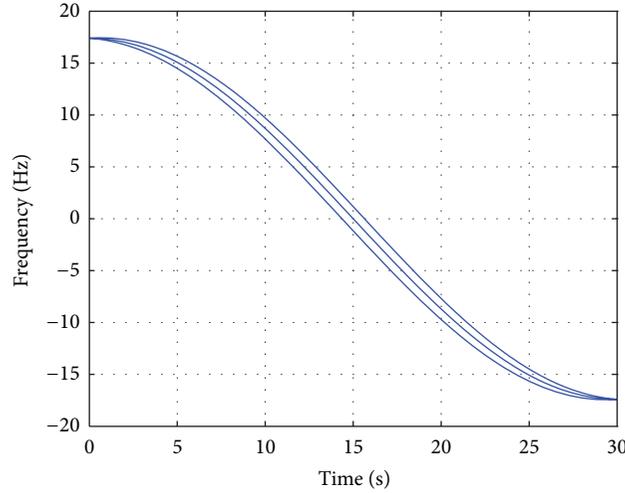


FIGURE 8: The Doppler frequency curves of the moving target in the flat earth model.

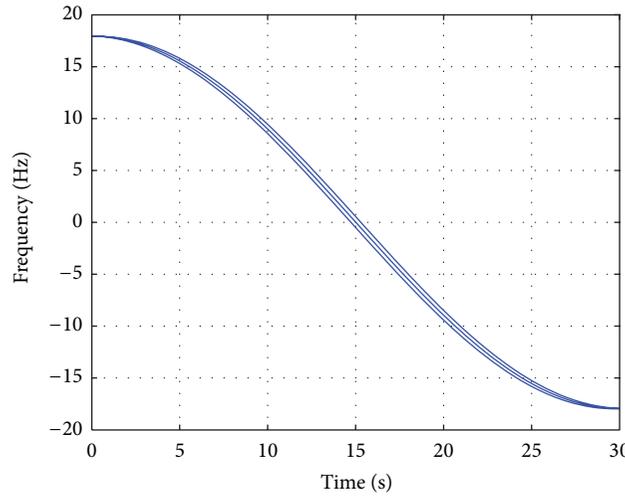


FIGURE 9: The Doppler frequency curves of the moving target in the earth curvature model.

For Path II, z_{02} can be written as

$$z_{02} = \frac{R \cos \alpha_2}{\sin \gamma_0} - R, \quad (37)$$

where

$$\begin{aligned} \gamma_2 &= \frac{3\pi}{2} - \frac{x(t)}{R} - 2\alpha_2 - 2\beta_2, \\ \beta_2 &= \arcsin\left(\frac{R}{R+H}\right) \times \cos \alpha_2. \end{aligned} \quad (38)$$

Then z_0 is the average value of z_{01} and z_{02} :

$$\begin{aligned} z_0 &= \left(\left[\frac{L_1(0)}{2} - \frac{\sin \theta_1}{\sin \beta_1} \right] \frac{\sin \beta_1}{\sin \theta_2} + \frac{R \cos \alpha_2}{\sin \gamma_0} - 2R \right) \\ &\quad \times (2)^{-1}. \end{aligned} \quad (39)$$

So the instantaneous height $z(t)$ can be expressed as

$$z(t) = z_0 + v_z(t) \times t. \quad (40)$$

5. Comparison of the Difference of Micromultipath Effects between Two Kinds of Models

The aim of this section is to analyze and compare the difference of micromultipath effects between the flat earth model and the earth curvature model. The moving target model established in Section 2.2 is used in this section.

5.1. The Difference in the Doppler Frequencies Produced by the Target's Movement. The curves of Doppler frequencies of the four paths produced by the target's movement in the flat earth model and the earth curvature model are shown in Figure 8 and Figure 9, respectively. As can be seen from the figures,

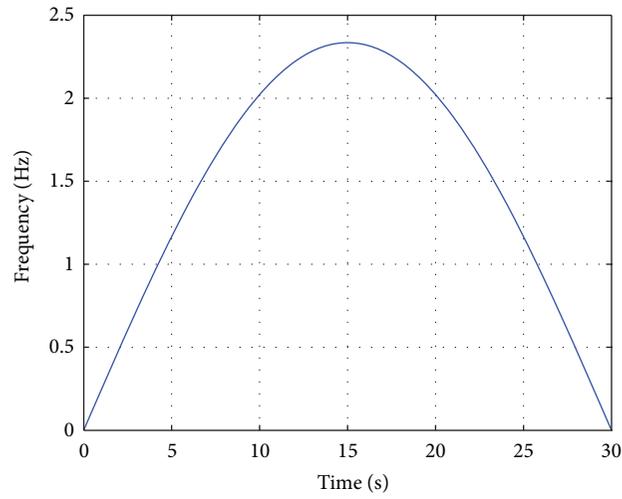


FIGURE 10: The Doppler frequency difference between Path I and Path II in the flat earth model.

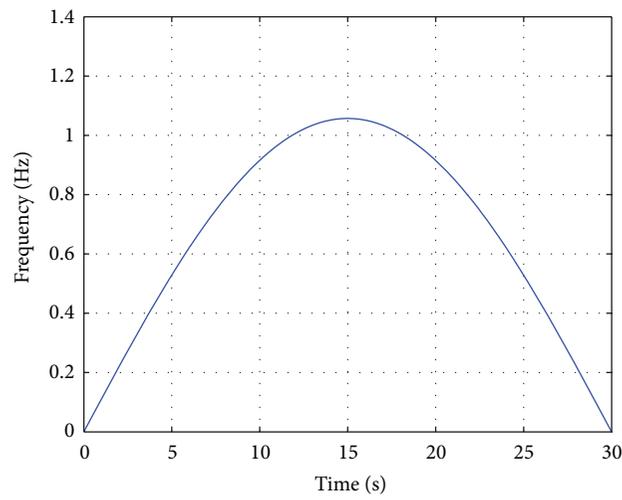


FIGURE 11: The Doppler frequency difference between Path I and Path II in the earth curvature model.

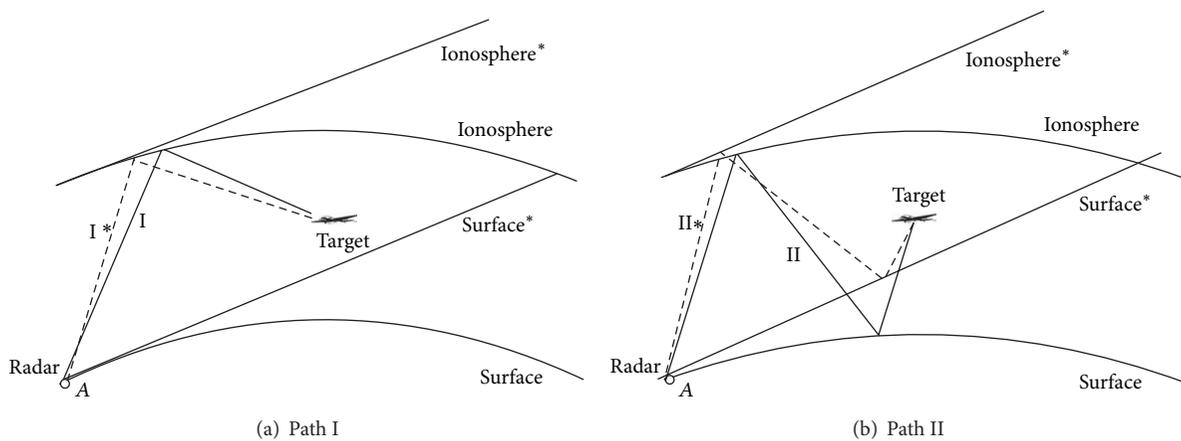


FIGURE 12: The Path I and Path II in two kinds of models.

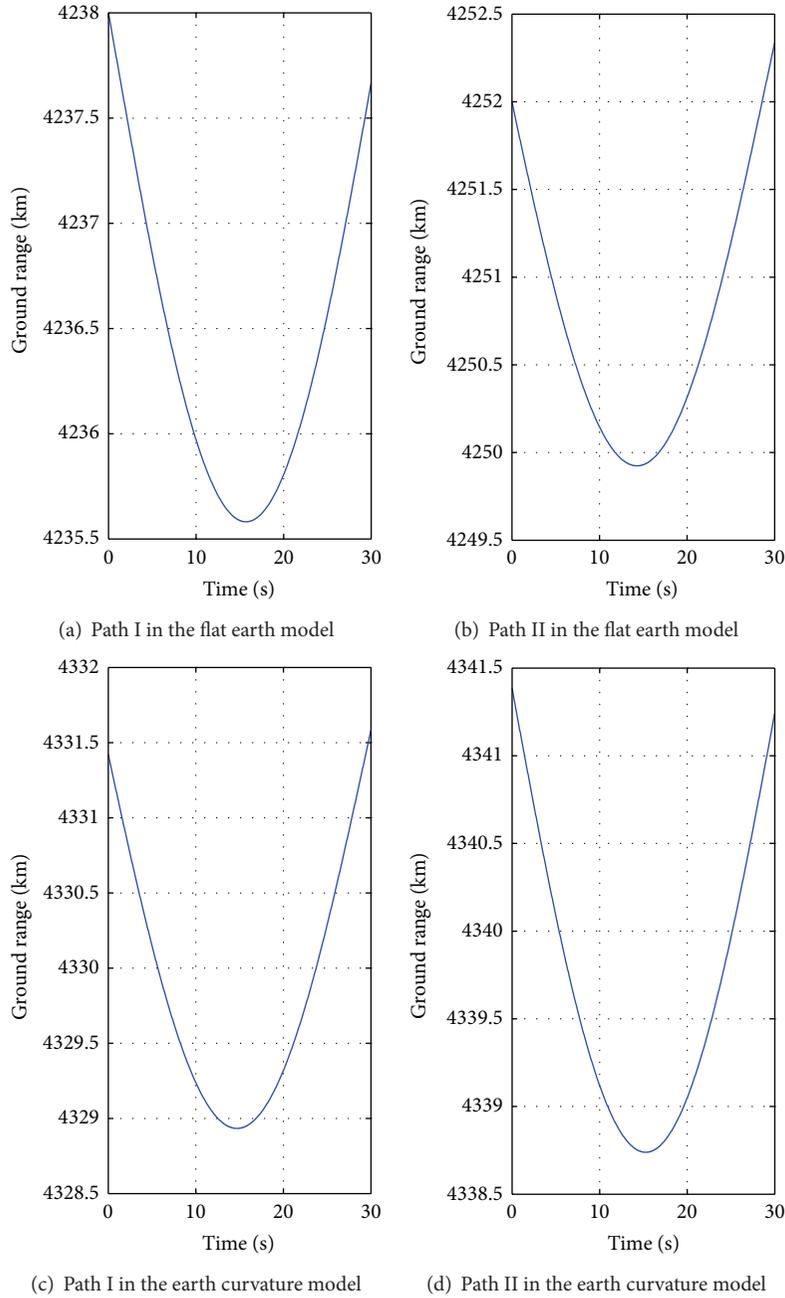


FIGURE 13: The simulation results of Path I and Path II in two kinds of models.

the target turns an angle of 180 degrees on the ground, so the value of Doppler frequency reduces from the maximum to the minimum. At the start and the end points of the movement period, for the altitude velocities are both zero, the Doppler frequency of each path is the same. For the absolute value of $v_x(t)$ is maximal, the absolute value of Doppler frequency comes to the maximum.

At the moment $T/2$, $v_x(t)$ comes to zero, and the Doppler frequencies of every path are close to zero, while $v_z(t)$ reaches its maximum. Therefore, the Doppler frequencies of every path come to its maximum difference value at this moment,

which is the best moment to obtain the Doppler frequency differences produced by $v_z(t)$.

By comparison of Figures 8 and 9, the Doppler frequency differences of the every path in the earth curvature model are smaller than the corresponding results in the flat earth model. Figures 10 and 11 illustrate, respectively, the values of the Doppler frequency differences between Path I and Path II in the flat earth model and the earth curvature model within one movement period. As seen, the maximum value of Doppler frequency difference in the flat earth model is 2.4 Hz and the one in the earth curvature model only is 1.3 Hz. It indicates

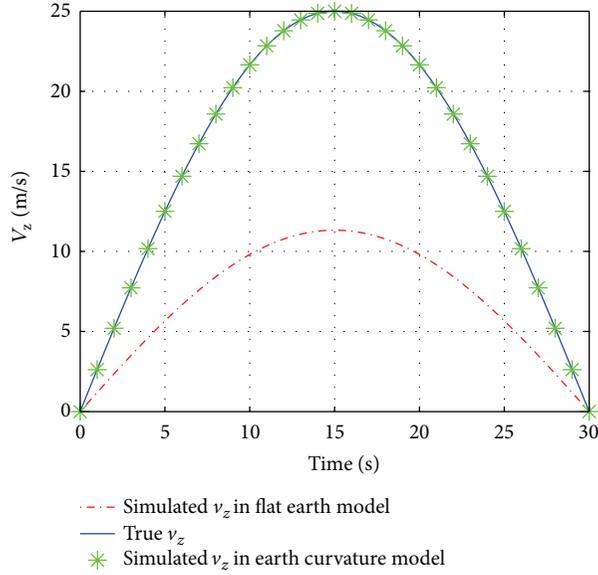


FIGURE 14: The comparison of v_z calculated by two kinds of earth models.

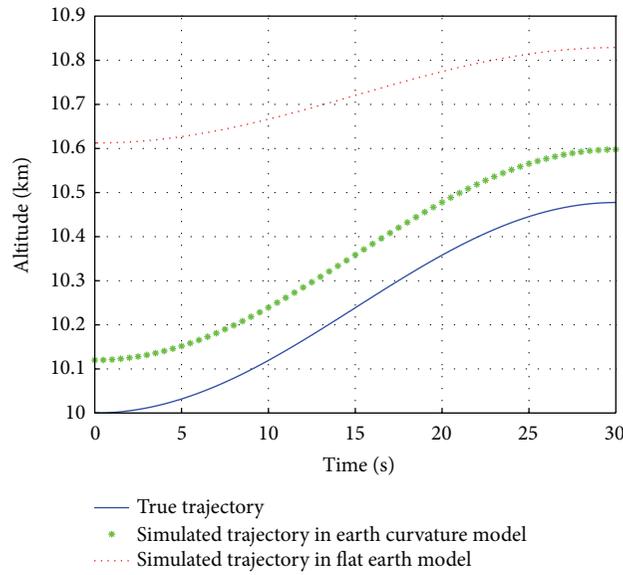


FIGURE 15: The comparison of the movement tracks calculated in two kinds of earth models.

that the error of the Doppler frequency of the flat earth model is close to 1.1 Hz, which will bring larger error in the altitude estimation.

5.2. *The Difference in the Target Track.* Figures 12(a) and 12(b) illustrate Path I and Path II in two kinds of models, in which the solid line represents the propagation path in the earth curvature model and the dotted line represents the propagation path in the flat earth model. The superscript * denotes the reference planes in the flat earth model. From these figures, we cannot determine directly in which model the propagation path is longer. So the tracks within one movement period are shown in Figure 13. In this figure, the

varied tendencies of the signal propagation path in two kinds of models are basically identical, the length of Path I in the earth curvature model is about 94 km longer than it is in the flat earth model, and the length of Path II in the earth curvature model is about 89 km longer than it is in the flat earth model.

The Doppler frequency difference f_{diff} , between Path I and Path II in the flat earth model, can be obtained by (14) and (15) as

$$f_{diff} = \frac{8f_c H}{x(t)c} v_z(t). \tag{41}$$

From (41), f_{diff} is inversely proportional to $x(t)$. For the signal propagation path in the earth curvature model is

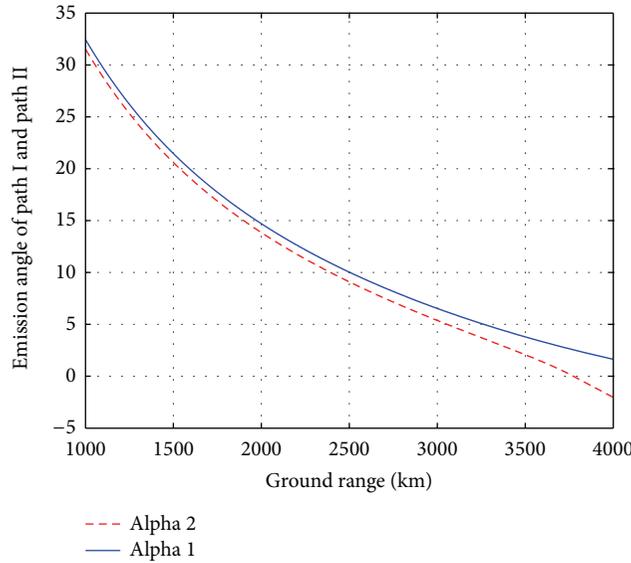


FIGURE 16: The curve of the emission angles α_1 and α_2 changing with the ground distance.

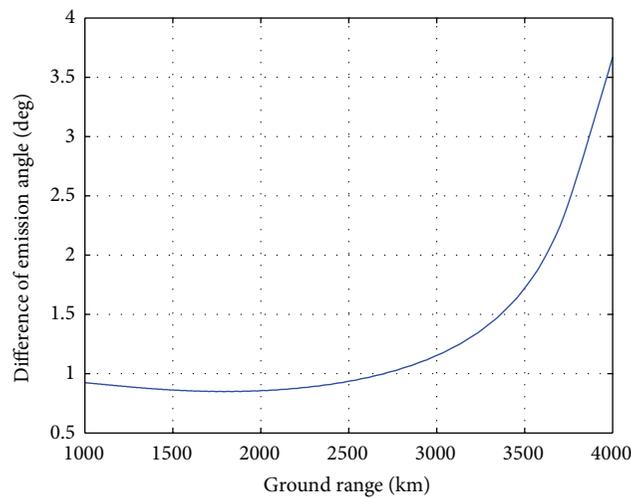


FIGURE 17: The curve of the difference α_{diff} changing with $x(t)$.

longer than the one in the flat earth model; $x(t)$ in the earth curvature model is larger than the one in the flat earth model. This is why the Doppler frequency differences of each path in the earth curvature model are smaller than the one in the flat earth model.

Figure 14 illustrates the comparison of the measured values $\hat{v}_z(t)$ of the height velocity in two kinds of earth models with its actual value $v_z(t)$. It can be seen that $\hat{v}_z(t)$ calculated in the earth curvature model is identical to the actual value, and $\hat{v}_z(t)$ obtained in the flat earth model has great differences from the actual one. The error of $\hat{v}_z(t)$ comes to 12 m/s when $v_z(t)$ is with the maximum value, so the altitude estimation would have great bias if the flat earth model is adopted.

Figure 15 shows the track in the altitude direction with two kinds of models. In the earth curvature model, $\hat{v}_z(t)$ has no error, so the varied tendency of the calculated track

(the dotted line) coincides with the actual one (the real line). The initial height z_0 calculated by (39) has a slight bias about 120 m. So the whole track has a fixed error with the true track. While $\hat{v}_z(t)$ and z_0 calculated in the flat earth model both have a larger error, the varied tendency of the whole movement track and the altitude value at every moment have great difference from the real value.

5.3. The Application Condition of the Earth Curvature Model

5.3.1. The Requirement of the Skywave Radar Altimetry to the Transmission Antenna Beam Width Ω . In order to use the effect of micromultipath, the transmission beam is requested to have a certain width to receive the echoes of four paths. Therefore, the transmission antenna beam width Ω must be larger than two times of the difference α_{diff} between the

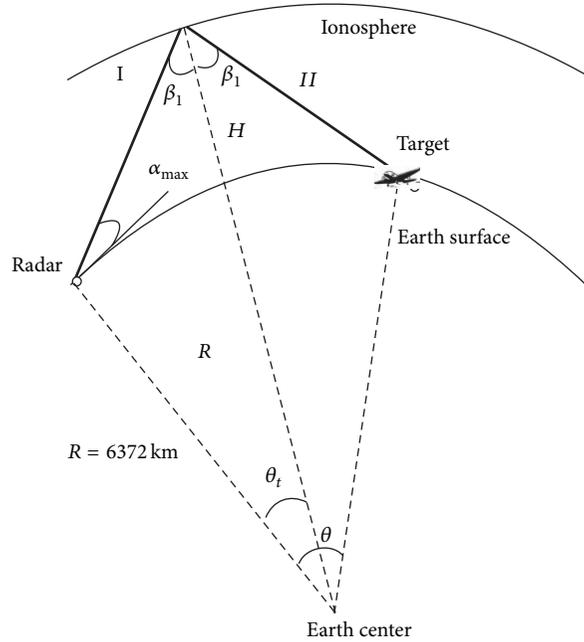


FIGURE 18: The illustration of the minimum detection range of the skywave radar.

emission angle α_1 and α_2 corresponding to Path I and Path II. In Figure 16, the changing curves of the emission angle α_1 (the real line) and α_2 (the dotted line) corresponding to Path I and Path II are shown when the ground distance $x(t)$ increases from 1000 km to 4000 km. Figure 17 illustrates the curve of the difference α_{diff} between the emission angles of two paths changing with $x(t)$. We can find that the difference α_{diff} between the emission angles corresponding to Path I and Path II changes little (its biggest value is about 1.2°) within the detection range from 1000 km to 3000 km. When the detection range exceeds 3000 km, α_{diff} rises sharply, and it reaches about 3.7° when the detection range is 4000 km. Accordingly, when the skywave radar use micromultipath effect to measure the height, its beam width needs to be larger than 2.4° if the detection range is smaller 3000 km and the beam width of the transmission antenna has to be larger than 7.4° when the detection range exceeds 3000 km.

5.3.2. The Detection Range of the Earth Curvature Model. The detection range of the skywave radar is limited in practice, for the emission angle α can only change in a certain range. The maximum value and the minimum value of α are written, respectively, as α_{max} and α_{min} .

When the target's altitude is close to the ground and the emission angle is α_{max} , the sky wave radar can realize the detection of the minimum range, as shown in Figure 18. When the emission angle is α_{min} , the maximum detection is obtained, as shown in Figure 19.

The emission angle α_{max} is generally taken as 30° . Then the minimum detection range x_{min} can be derived as follows:

$$x_{min} = \frac{2\pi R}{3} - 2R \arcsin\left(\frac{\sqrt{3}R}{2(R+H)}\right). \quad (42)$$

The emission angle α_{min} is always 0° ; then the maximum detection range x_{max} can be derived as

$$x_{max} = \pi R - 2R \arcsin\left(\frac{R}{R+H}\right). \quad (43)$$

Taking the earth radius as $R = 6372$ km and the ionosphere height as $H = 350$ km, then $x_{min} = 1072$ km and $x_{max} = 4131$ km. The results conform to the practical detection range of the skywave radar.

6. Conclusion

Skywave OTH radar systems can make use of the micromultipath effect to estimate the instantaneous altitude of maneuvering targets. The flat earth model is often used. However, ignoring the impact of the curvature of the earth, the analytical expressions of instantaneous target altitude are relatively simple, making it suitable for descriptive analysis. And the initial altitude has a great difference from actual altitude (600 meters) in the flat earth model. Moreover, the curvature of the earth should not be ignored for the thousands of kilometers' detection in the skywave radar.

This paper focuses on the theoretical derivations about how to introduce the earth curvature model into the estimation of the target's altitude. The expressions of the Doppler frequency in four paths are deduced. And the difference of two kinds of models is compared. Although the analytical expressions are quiet complicated, the estimation of the initial altitude is close to the actual one. Therefore, any altimetry methods that are based on real data can improve the estimation results by exploiting the altimetry correction factors derived from earth curvature model.

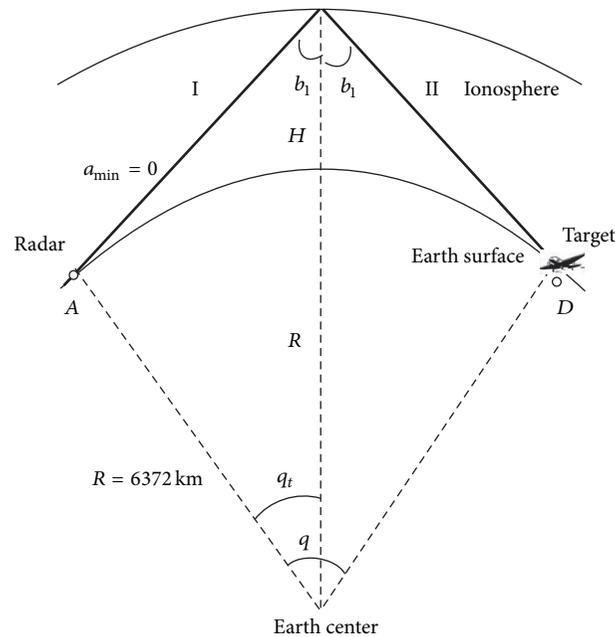


FIGURE 19: The illustration of the maximum detection range of the skywave radar.

Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

Acknowledgments

At the point of finishing this paper, the authors would like to express their sincere thanks to the National Natural Science Foundation (no. 61201303) and the Fundamental Research Funds for the Central Universities (HIT.NSRIF.2013027).

References

- [1] P. Jiang and J. Chen, "Research of OTH sky wave radar height finding based on matching domain processing," *Journal of Air Force Radar Academy*, pp. 157–165, 2012.
- [2] X. Guo, J. Ni, W. Su, and G. Liu, "Development of sky wave over-the-horizon radar," *Acta Aeronautica et Astronautica Sinica*, vol. 23, no. 6, pp. 495–500, 2002.
- [3] W. C. Anderson, D. S. Green, and S. P. Kingsley, "HF skywave radar: estimating aircraft heights using super-resolution in range," *IEEE Proceedings on Radar, Sonar and Navigation*, vol. 143, no. 4, pp. 281–285, 1996.
- [4] R. H. Anderson, S. Kraut, and J. L. Krolik, "Robust altitude estimation for over-the-horizon radar using a state-space multipath fading model," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 39, no. 1, pp. 192–201, 2003.
- [5] J. Prashifka, L. J. Durbridge, and J. Lane, "Investigation of target altitude estimation in skywave OTH radar using a high-resolution ionospheric sounder," in *Proceedings of the International Radar Conference on Surveillance for a Safer World (RADAR '09)*, pp. 1–6, IEEE, Bordeaux, France, October 2009.
- [6] M. Papazoglou and J. L. Krolik, "Electromagnetic matched-field processing for target height finding with over-the-horizon radar," in *Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP '97)*, pp. 559–562, April 1997.
- [7] M. Papazoglou and J. L. Krolik, "Matched-field estimation of aircraft altitude from multiple over-the-horizon radar revisits," *IEEE Transactions on Signal Processing*, vol. 47, no. 4, pp. 966–976, 1999.
- [8] M. Papazoglou and J. L. Krolik, "Electromagnetic matched-field processing for target height finding with over-the-horizon radar," in *Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP '97)*, pp. 559–562, April 1997.
- [9] Y. D. Zhang, M. G. Amin, and B. Himed, "Altitude estimation of maneuvering targets in MIMO Over-the-horizon-radar," in *Proceedings of the IEEE 7th Sensor Array and Multichannel Signal Processing Workshop*, pp. 257–230, 2012.
- [10] Y. Zhang, M. G. Amin, and G. J. Frazer, "High-resolution time-frequency distributions for manoeuvring target detection in over-the-horizon radars," *IEE Proceedings: Radar, Sonar and Navigation*, vol. 150, no. 4, pp. 299–304, 2003.

Research Article

Research on Torque Ratio Based on the Steering Wheel Torque Characteristic for Steer-by-Wire System

Yandong Han,¹ Lei He,¹ Xiang Wang,² and Changfu Zong¹

¹ State Key Laboratory of Automotive Simulation and Control, Jilin University, Changchun 130022, China

² Automobile Engineering Institute, Guangzhou Automobile Group Co., Ltd. Guangzhou 511434, China

Correspondence should be addressed to Lei He; jlu_helei@jlu.edu.cn

Received 25 April 2014; Accepted 22 June 2014; Published 6 July 2014

Academic Editor: Hamid R. Karimi

Copyright © 2014 Yandong Han et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Steer-by-wire system can improve the performance of vehicle handling stability. Removing the mechanical linkages between the front wheels and the steering wheel leads to a key technique of force feedback for steer-by-wire system. In view of the characteristic of variable torque transmission ratio for steer-by-wire system, this paper proposes a method for designing torque ratio based on the steering wheel torque characteristic for steer-by-wire system. It converts the torque ratio design into equivalent assist torque design by analyzing their relationship. It achieves the torque ratio design at different conditions based on the negative equivalent assist torque characteristic curve. Simulations and vehicle experiments are conducted by the proposed method, and the results show that the design goal has been achieved and the steering wheel torque characteristic obtained is very similar to that of the reference car.

1. Introduction

Steer-by-wire (SBW) system removes the mechanical linkages between the steering wheel (SW) and the front wheels, which makes the transmission ratio of SBW system variable. The torque transmission ratio α , namely, the ratio of the pinion torque and the SW torque, determines the SW torque characteristic. The simulation method of road feel is one of the key technologies for SBW system. A good simulation method of road feel can help to improve steering handiness at low speed and road feel at high speed and lead to a better handling stability and driving safety. For this reason, the study of road feel has drawn much attention. Im et al. utilized torque sensors to obtain the steering resistance between the front wheels and road and fed it back to the driver with a certain torque ratio [1]. A kind of road feel algorithm was proposed by utilizing the linear relationship between the electromagnetic torque and electric current of the steering motor to obtain the torque. Then the torque was fed back to the driver with a certain torque ratio [2–4]. A kind of virtual force feedback characteristic was also designed based on vehicle speed, steering angle, and other vehicle parameters in the conventional steering system [5, 6]. Based on the analysis

of the assist characteristic and the active damping control strategy of the EPS system, an integrated road feel algorithm was proposed [7]. Mehdizadeh et al. proposed the concept of virtual vehicle to provide the driver with a familiar force feedback [8]. And friction, stiffness, and damping compensation were also used to improve the performance of road feel control [9]. Sliding mode control was applied to the control of the steer-by-wire vehicles to obtain strong robustness [10].

The first two methods focus on the establishment of steering resistance observers and ignore the effects of the torque ratio, so they can only reflect the SW torque characteristic partly. The third method can only provide a virtual road feeling and not reflect the real adhesion state of the wheels and road. The fourth and fifth methods are to be further perfected. The last two methods just stress discussing the control algorithms. It is obvious that most current researches focus on how to obtain the steering resistance and ignore the influences of the torque ratio on handling stability because of the structure restriction of SBW system. A torque ratio design method based on the SW torque characteristic for SBW system is proposed and verified in this paper.

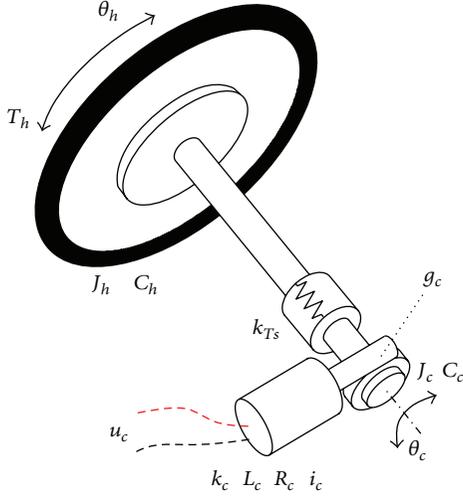


FIGURE 1: Structure of the SW block.

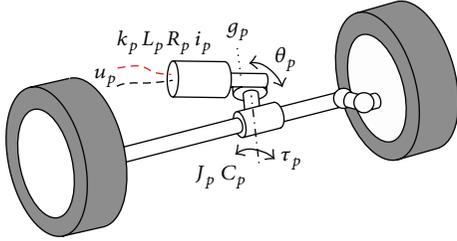


FIGURE 2: Structure of the steering executive block.

2. Materials and Methods

2.1. SBW System and Vehicle Model

2.1.1. SBW System Model. The SBW system consists of two blocks: the SW block and the steering executive block. Since the torsional stiffness of the torque sensor is low relatively, the SW block can be divided into two parts: the SW assembly and the steering column assembly. The SW assembly consists of the SW and part of the column that is the district between the SW and the torque sensor. The steering column assembly consists of the remaining part of the steering column, road feel motor, and its reducer. The structure of the SW block is shown in Figure 1.

Each of the two assemblies has one rotational degree, the angle of the SW, and the angle of the steering column. The dynamical formulas of the SW block are as follows:

$$\begin{aligned} T_h - T_s &= J_h \ddot{\theta}_h + C_h \dot{\theta}_h, \\ T_s &= k_{Ts} (\theta_h - \theta_c), \\ T_s - g_c k_c i_c &= J_c \ddot{\theta}_c + C_c \dot{\theta}_c, \\ u_c &= g_c k_c \dot{\theta}_c + L_c \dot{i}_c + R_c i_c, \end{aligned} \quad (1)$$

where θ_h and θ_c are the angle of the SW and the steering column, respectively. k_{Ts} is the stiffness of the torque sensor. T_h is the torque of the driver input and T_s is the measured

TABLE 1: Parameters of SW block.

Parameters	Values
Stiffness of torque sensor k_{Ts}	176 Nm/rad
Equivalent moment of inertia J_h	0.022 kgm ²
Equivalent moment of inertia J_c	0.116 kgm ²
Equivalent damping coefficient C_h	0.02 Nms/rad
Equivalent damping coefficient C_c	0.13 Nms/rad
Reduction ratio of worm and gear g_c	15

TABLE 2: Parameters of steering executive block.

Parameters	Values
Equivalent friction torque $T_{friction-p}$ (column hysteresis in Carsim)	2.1 Nm
Equivalent damping coefficient at rack (front steering damping in Carsim)	2.1 Ns/mm
Equivalent moment of inertia J_p (system inertia in Carsim)	0.12 kgm ²
Ratio of steering gear (C factors in Carsim)	49.93 mm/rev
Reduction ratio of worm and gear g_p	15

torque by torque sensor. J_h and J_c are the equivalent moment of inertia of the two parts at steering column, respectively. C_h and C_c are the equivalent damping coefficient of the two parts at steering column, respectively. g_c is the reduction ratio of worm and gear reducer. u_c , k_c , L_c , R_c , and i_c are the voltage, the back EMF constant (torque constant), the inductance, the resistance, and the current of the road feel motor, respectively.

The structure of the steering executive block is shown in Figure 2. The dynamical formulas of steering executive block are as follows:

$$g_p k_p i_p - \tau_p = J_p \ddot{\theta}_p + C_p \dot{\theta}_p, \quad (2)$$

$$u_p = g_p k_p \dot{\theta}_p + L_p \dot{i}_p + R_p i_p, \quad (3)$$

where θ_p is the angle of the pinion. τ_p is the equivalent resistance between the front wheels and the road at the pinion shaft. J_p and C_p are the equivalent moment of inertia and the equivalent damping coefficient of the block at the pinion shaft, respectively. g_p is the reduction ratio of worm and gear reducer. u_p , k_p , L_p , R_p , and i_p are the voltage, the back EMF constant (torque constant), the inductance, the resistance, and the current of the steering motor, respectively.

2.1.2. Vehicle Model. The SBW vehicle tests are conducted to verify assist torque characteristic in parking condition. The SBW test vehicle, as shown in Figure 3, is refitted from a domestic A-class car. The electrical parameter values of two blocks are obtained by the least-squares method, as discussed elsewhere [11]. The mechanical parameter values of two blocks are shown in Tables 1 and 2. The experiment approach of parameter identification is discussed by Wang in detail in [12].

A Carsim vehicle model is used to verify assist torque characteristic in the on-center, linear, and nonlinear region

TABLE 3: On-center evaluation indices of SW torque characteristic.

Performance		Evaluation indices
Hysteresis	Lateral acceleration-SW torque	Lateral acceleration at 0 Nm
	SW angle-SW torque	Steering hysteresis Steering fiction torque
Sensitivity	SW torque	Average steering stiffness
Road feel	Torque gradient	Steering torque gradient at 0 g Steering torque gradient at 0.1 g
	Torque	SW torque at 0.1 g



FIGURE 3: The SBW test vehicle.

and limit working conditions. The parameters, assembly, and size of the Carsim model are all very similar to those of the SBW test vehicle. SBW system model is built by MATLAB/Simulink. The vehicle dynamics simulations are achieved by the cosimulation of Carsim and Simulink. SBW system employs the bilateral control method of torque drive/angle feedback [13].

2.2. Design Basis of Torque Ratio

2.2.1. Objective Evaluation Indices of the SW Torque Characteristic. According to the research on objective evaluation system of handling stability, the SW torque characteristic is in the category of maneuverability and can be divided into three factors: the on-center torque characteristic, the steering effort characteristic for keeping a given control, and the SW torque level in parking condition [14]. On-center torque characteristic describes the variation of the SW torque, while the SW angle or lateral acceleration alters under the condition of high vehicle speed and small lateral acceleration. By analyzing and filtering the evaluation indices, the streamlined evaluation indices are shown in Table 3 [15]. The test method refers to [16].

Steering effort characteristic for keeping a given control describes the variation of the SW torque, while the speed and lateral acceleration alter under the condition of medium or large lateral acceleration. The SW torque should increase with

the increase of the speed and lateral acceleration. When the tires are close to the limit of adhesion, the SW torque should be decreased to help driver acquire the adhesion state of the tires. When the SW torque reaches the maximum value, the steering maneuverability and comfort should not be affected. The improper maximum SW torque leads to heavy steering and hinders the driver from turning the SW quickly. Over the entire range of lateral acceleration, the SW torque should be able to reflect the lateral acceleration and the adhesion state between tires and road truthfully. Relevant researches show that the SW torque desired by drivers is affected by gender, age, and ethnic factors [17–19]. So this paper designs the torque ratio based on the steering effort preference.

SW torque level in parking condition describes the SW torque, while steering in situ or at low speed and the SW closed to the limit position. The SW torque level is assessed by the indices of the steering effort at 70% of the entire steering angle range under the condition of steering in situ. It is reasonable that the torque is no more than 5 Nm for the power steering system in general, which applies to SBW system as well. The SW work level is assessed by the homonymic indices in steering handiness test. The test method refers to [20]. In addition, the SW torque delivered to the driver should not decrease with the increase of the SW angle when the SW is close to limit position. Otherwise it will generate a steering impact to the front wheels, which is not what we want.

2.2.2. Design Basis of Torque Ratio. If we can transplant the assist characteristic of power steering system to SBW system, the SW torque characteristic and yaw response of the SBW vehicle will be very similar to the vehicle equipped with power steering system. Considering the advantages of SBW system, the ratio designed should not only enable the performance SBW vehicle to match with the vehicle equipped with power steering system but also be adjustable to meet different groups of people and diverse working conditions. Summarizing relevant researches and test results, the relative ideal SW torque characteristic should meet the following requirements.

- (1) The SW torque should increase in proportion to the increase of speed and lateral acceleration in steady steering condition. The torque should increase significantly with the increase of lateral acceleration in linear region and not significantly in nonlinear region. When the tires are close to the limit of adhesion, the torque should be declined with the increase of lateral acceleration [21, 22].
- (2) The design of the on-center SW torque characteristic can refer to the parameters in Table 4. The parameters are the relevant on-center objective evaluation indices of one domestic car, which is called the reference car hereinafter.
- (3) The steering effort at 70% of the entire steering angle range should be less than 4.5 Nm in parking condition. When the SW is close to the limit position, the SW torque should increase smoothly, which can become a reminder to the driver.

TABLE 4: On-center evaluation indices of the reference car.

Evaluation	Value
Lateral acceleration at 0 Nm (g)	0.042
Steering hysteresis (deg)	5.32
Steering fiction torque (Nm)	1.64
Average steering stiffness (Nm/deg)	0.36
Steering torque gradient at 0 g (Nm/g)	32.0
Steering torque gradient at 0.1 g (Nm/g)	9.06
SW torque at 0.1 g (Nm)	2.75

2.3. Conversion of Torque Ratio into Equivalent Assist Torque.

The above objective evaluation indices provide a reasonable SW torque characteristic and we can design a relatively ideal torque ratio for SBW system by analyzing the relationship between the ratio and the indices. The damping, inertia, and friction of SBW system increase by about 200%–370% because of the two motors. That will cause great impacts on vehicle performance. So the SBW control system should compensate for these mechanical parameters. The compensation method should be adjusted and can improve vehicle performance. The specific compensation methods refer to [12]. Ignoring the effects of damping and inertia on the SW torque characteristic and considering the equivalent friction of steering executive block at the pinion, (2) can be rewritten as

$$T_p - \tau_p = J_p \ddot{\theta}_p + C_p \dot{\theta}_p + \text{sgn}(\dot{\theta}_p) T_{\text{friction}_p}, \quad (4)$$

where T_{friction_p} is the equivalent friction of steering block at the pinion shaft. τ_p is the resistance between the front wheels and road. T_p is the control torque of steering motor and the specific control method is similar to that of EPS system [23, 24]. T_p can be specifically expressed as

$$T_p = T_\alpha + T_{Jp} + T_{Cp} + T_{fp}, \quad (5)$$

where $T_\alpha = \alpha T_s$ is the direct drive torque. T_{Jp} , T_{Cp} , and T_{fp} are the compensation moment of inertia, damping, and fiction, respectively. The contents of the objective evaluation system are for the SW torque characteristic in steady state and quasi steady state. The damping and inertia of steering system, which greatly affects dynamic performance, need to be debugged and calibrated by the subjective evaluation. Therefore only considering the friction compensation and ignoring the impact of damping and inertia on the SW torque characteristic, (4) can be rewritten as

$$\alpha T_s + T_{fp} = \tau_p + T_{\text{friction}_p}(\dot{\theta}_p). \quad (6)$$

Road feel intensity E , as shown in (7), which is the ratio of torque increment of the SW dT_h and the output torque increment of steering gear dF_p , is used to measure the intensity of road feeling [25]:

$$E = \frac{dT_h}{dF_p}. \quad (7)$$

Ignoring the friction and the compensation torque of the steering executive block, T_h is approximately equal to T_s

TABLE 5: On-center evaluation indices of negative equivalent assist characteristic curve.

Evaluation indices	Value
Lateral acceleration at 0 Nm (g)	0.019
Steering hysteresis (deg)	4.69
Steering fiction torque (Nm)	1.52
Average steering stiffness (Nm/deg)	0.32
Steering torque gradient at 0 g (Nm/g)	33.5
Steering torque gradient at 0.1 g (Nm/g)	9.31
SW torque at 0.1 g (Nm)	2.44

under the condition of small SW angular speed and acceleration. So E can be expressed as

$$E = \frac{dT_h}{dF_p} = r_p \frac{dT_h}{d\tau_p} = r_p \frac{dT_s}{d\tau_p} = r_p \frac{1}{\alpha(T_s)}, \quad (8)$$

where r_p is the radius of the pinion. Because r_p is a constant gain, $1/\alpha$ can be used to measure the intensity of road feeling. Referring to the design method of EPS system [23, 24], the relationship between equivalent assist torque T_{assis} of EPS system and torque ratio α of SBW system can be expressed as

$$\alpha = \frac{T_{\text{assis}} + T_s}{T_s}. \quad (9)$$

Obviously, “ αT_s ” can be replaced by “ $T_{\text{assis}} + T_s$ ” and the torque ratio can be designed based on (9). So the study of torque ratio is replaced by the study of equivalent assist torque hereinafter.

3. Results and Discussion

Equivalent assist torque characteristic is not only related to the SW torque, but also varies with vehicle speed. To facilitate the calibration of SBW equivalent assist torque characteristic, a series of characteristic speed should be determined firstly. Then a series of characteristic curves, which show the relationship between equivalent assist torque and the SW torque at characteristic speed, should be designed. The curves at other speeds are designed by linear interpolation. The partition method of characteristic speed is not discussed here. The design method of equivalent assist torque characteristic is illustrated by designing the characteristic at 100 km/h. The curve in Figure 4 is negative equivalent to assist torque characteristic curve, which is proposed in [12] by Wang. It is obvious that the characteristic curve includes four factors: the on-center slope, the SW torque range corresponding to large curvature region, the SW torque in linear and nonlinear region, and the maximum equivalent assist torque. The on-center evaluation indices of the characteristic curve in Figure 4 are shown in Table 5. The torque ratio design method proposed is based on this curve. We can achieve the design goal by adjusting the characteristic curve.

3.1. Equivalent Assist Characteristic in Linear, Nonlinear, and Limit Region. The steering effort preference is summarized

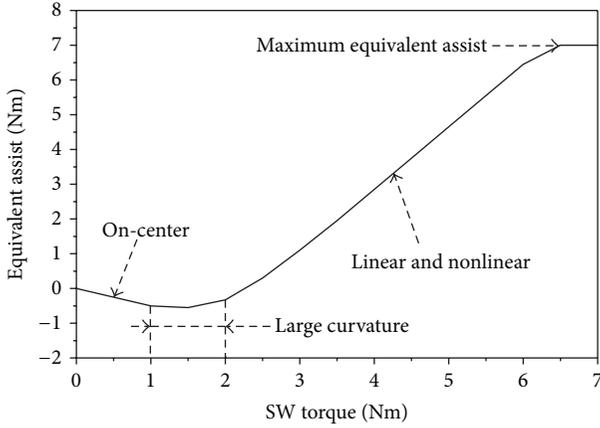


FIGURE 4: Negative equivalent assist characteristic curve.

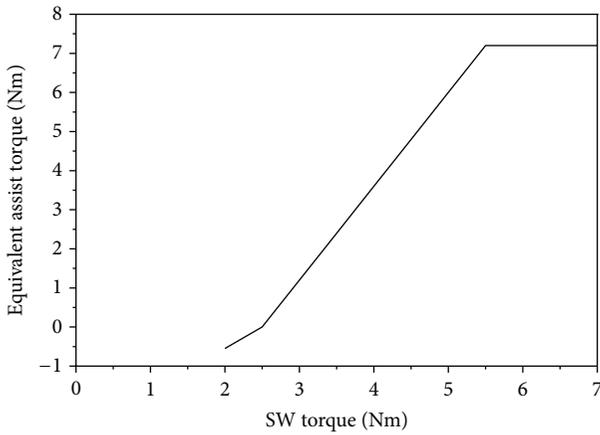


FIGURE 5: Equivalent assist characteristic in linear region at 100 km/h.

with VTI driving simulator in [17]. The steering effort preference of Chinese drivers by using ADSL driving simulator is summarized in [22]. Referring to the above research results and Table 3, the steering effort characteristic for keeping a given control in linear region is designed. The target SW torque at lateral acceleration of 0.1g is 2.6 Nm and at 0.3g is 2.6 Nm. In the test the driver turns the SW at rotation speed of 360 deg/min until the lateral acceleration reaches 0.75 g, while the vehicle speed is 100 km/h.

The characteristic curve in Figure 5, which results from adjusting the SW torque region from 2Nm to 4Nm of the curve in Figure 3, helps to achieve the design goals of the SW torque at lateral acceleration of 0.1g and 0.3g, as shown in Figure 6. The test results show that when the SW torque is more than 2.5 Nm, there is a linear relationship between feedback torque and steering resistance at the lateral acceleration region from 0.1g to 0.42 g and a nonlinear relation from 0.42 g to 0.65 g. When the lateral acceleration is over 0.65 g, feedback torque decreases a little. So the SW torque characteristic can reflect the steering resistance faithfully in linear, nonlinear, and limit region. The driver can feel the adhesion state and the lateral acceleration precisely.

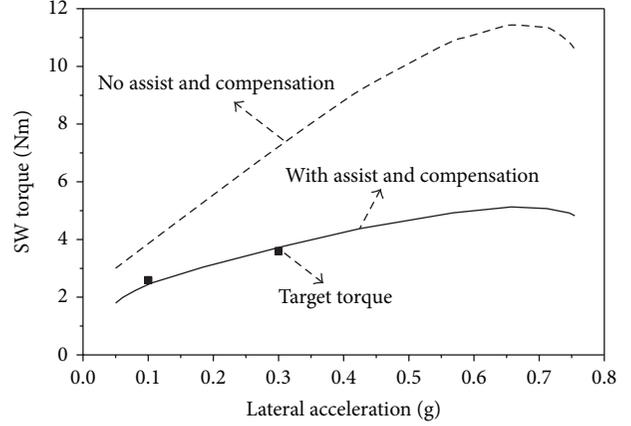


FIGURE 6: Steering effort characteristic for keeping a given control at 100 km/h.

3.2. *On-Center Equivalent Assist Characteristic.* Comparing the on-center evaluation indices in Tables 4 and 5, we can find the following.

- (1) The hysteresis of lateral acceleration and the SW angle relative to the SW torque of the model are significantly lower than those of the reference car. On one hand it results from negative equivalent assist characteristic and friction compensation. On the other hand it results from the differences of hysteresis characteristic between a model and a car.
- (2) Average steering stiffness and the SW torque gradient at lateral acceleration of 0 g of the model are close to those of the reference car. So these two indices will not be adjusted. It is found in the test that increasing the on-center slope of the equivalent assist characteristic curve can improve the characteristics.
- (3) The SW torque and torque gradient of the model at lateral acceleration of 0.1g has determined by the steering effort characteristic for keeping a given control shown in Figure 6.
- (4) The SW torque level corresponding to the large curvature region of the equivalent power characteristic curve of the model should be consistent with that of the reference car. The torque specifically refers to the SW torque corresponding to the lateral acceleration region from 0 to 0.1g. The torque range of the reference car is from 1.01 Nm (half of the steering friction torque) to 2.75 Nm (the SW torque at lateral acceleration of 0.1 g).

The last equivalent assist characteristic curve on-center and large curvature region is shown in Figure 7. The corresponding indices of the model and the reference car are shown in Table 6. By a targeted calibration, SBW vehicle model has a similar on-center road feel and the SW torque characteristic to the reference car. And the model has a smaller delay. In other words, the aligning performance of SBW vehicle model is better.

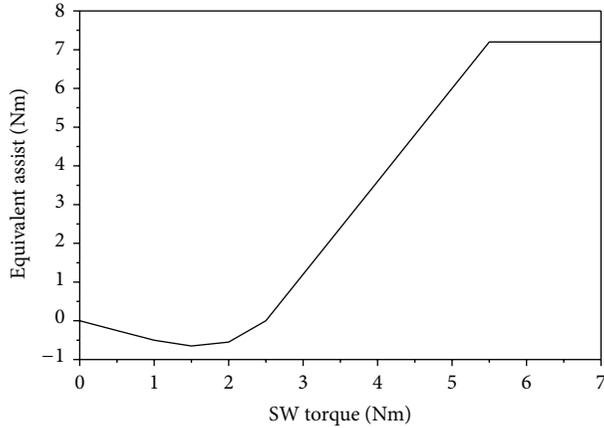


FIGURE 7: On-center equivalent assist characteristic at 100 km/h.

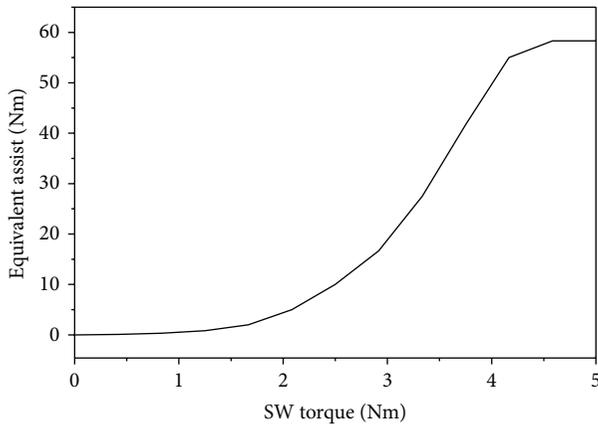


FIGURE 8: Equivalent assist characteristic at 0 km/h.

TABLE 6: On-center evaluation indices at 100 km/h.

Evaluation indices	Reference car	model
Lateral acceleration at 0 Nm (g)	0.042	0.021
Steering hysteresis (deg)	5.32	4.98
Steering fiction torque (Nm)	1.64	1.62
Average steering stiffness (Nm/deg)	0.36	0.29
Steering torque gradient at 0 g (Nm/g)	32.0	33.7
Steering torque gradient at 0.1 g (Nm/g)	9.06	8.73
SW torque at 0.1 g (Nm)	2.75	2.61

3.3. *Equivalent Assist Characteristic in Parking Condition.* The test of steering in situ is conducted to verify the characteristic in parking condition. In the test the driver turns the SW at rotation speed of 10~15 r/min to the right limit position, then to the left limit position, and at last go to the right limit position, that is, one cycle. One test has two cycles and repeats the test three times [20]. The equivalent assist characteristic in parking condition is shown in Figure 8. The characteristic curve of the SW torque versus the SW angle is shown in Figure 9. The SW torque at 70 percent of the entire steering angle range is 3.73 Nm, which represents a good handiness.

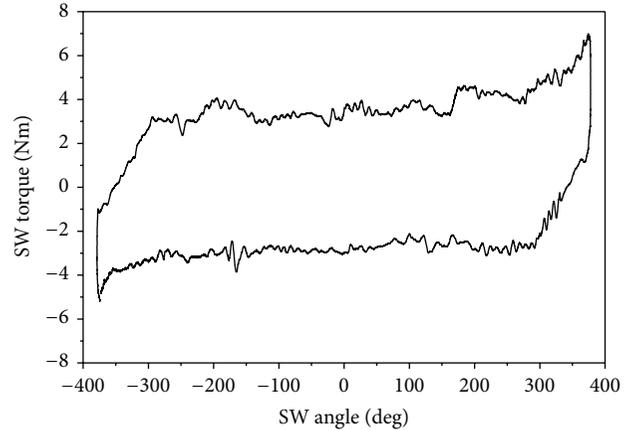


FIGURE 9: SW torque versus SW angle at 0 km/h.

4. Conclusions

A torque ratio design method based on the SW torque characteristic for SBW system is proposed. According to the classification of the handling evaluation indices, the SW torque characteristic can be divided into three factors: the on-center torque characteristic, the steering effort characteristic for keeping a given control, and the SW torque level in parking condition. By satisfying the general requirements of these three aspects, the torque ratio for SBW system, which is based on negative equivalent assist torque characteristic, improves road feel and the on-center aligning performance achieves the target goals and provides a better handling and stability in different working conditions. Converting the study of the torque ratio into the study of the equivalent assist torque characteristic helps to make better use of current research results of the equivalent assist torque characteristic and promote the research of the torque ratio of SBW system. Simulation and vehicle test results show that the SW torque characteristic obtained is very similar to that of the reference vehicle.

Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

Acknowledgments

This research was supported by Foundation of State Key Laboratory of Automotive Simulation and Control (no. 20120111) and the National High Technology Research and Development Program of China (863 Program) (no. 2012AA110904).

References

- [1] J.-S. Im, F. Ozaki, N. Matsunaga, and S. Kawaji, "Bilateral control for steer-by-wire vehicles," in *Proceedings of the SICE-ICASE International Joint Conference*, pp. 528–533, Busan, Republic of Korea, October 2006.

- [2] S. Asai, H. Kuroyanagi, S. Takeuchi, T. Takahashi, and S. Ogawa, "Development of a steer-by-wire system with force feedback using a disturbance observer," SAE Technical Paper 2004-01-1100, 2004.
- [3] Y. Yokokohji and T. Yoshikawa, "Bilateral control of master-slave manipulators for ideal kinesthetic coupling - formulation and experiment," *IEEE Transactions on Robotics and Automation*, vol. 10, no. 5, pp. 605–619, 1994.
- [4] C. Jeffrey, R. Cutajar, A. Richardson, S. Prosser, M. Lickess, and S. Riches, "The integration of on-line monitoring and reconfiguration functions into a safety critical automotive electronic control unit," *Journal of Electronic Testing: Theory and Applications*, vol. 21, no. 4, pp. 405–416, 2005.
- [5] S. Oh, S. Yun, H. Chae et al., "The development of an advanced control method for the steer-by-wire system to improve the vehicle maneuverability and stability," SAE Technical Paper, 2003.
- [6] M. Kamel Salaani, G. J. Heydinger, and P. A. Grygier, "Experimental steering feel performance measures," SAE Technical Paper 2004-01-1074, 2004.
- [7] H. Zheng, C. Zong, and L. Yu, "Road feel feedback design for vehicle steer-by-wire via electric power steering," SAE Technical Paper 2013-01-2898, 2013.
- [8] E. Mehdizadeh, M. Kabganian, and R. Kazemi, "A new force feedback for Steer-by-Wire vehicles via virtual vehicle concept," in *Proceedings of the 50th IEEE Conference on Decision and Control and European Control Conference*, pp. 2281–2286, December 2011.
- [9] U. B. Mandhata, M. J. Jensen, J. R. Wagner, F. S. Switzer, D. M. Dawson, and J. D. Summers, "Evaluation of a customizable haptic feedback system for ground vehicle steer-by-wire interfaces," in *Proceedings of the American Control Conference (ACC '12)*, pp. 2781–2787, Montreal, Canada, June 2012.
- [10] M. T. Do, Z. Man, C. Zhang, H. Wang, and F. S. Tay, "Robust sliding mode-based learning control for steer-by-wire systems in modern vehicles," *IEEE Transactions on Vehicular Technology*, vol. 63, no. 2, pp. 580–590, 2014.
- [11] S. S. Saab and R. A. Kaed-Bey, "Parameter identification of a DC motor: an experimental approach," in *Proceedings of the 8th IEEE International Conference on Electronics, Circuits and Systems (ICECS '01)*, vol. 2, pp. 981–984, September 2001.
- [12] X. Wang, *Research on bilateral control and variable ratio characteristics for steer-by-wire automobile [Ph.D. thesis]*, Jilin University, Changchun, China, 2013.
- [13] X. Wang, C. Zong, H. Xing, R. Hu, and X. Xie, "Bilateral control method of torque drive/angle feedback used for steer-by-wire system," SAE Technical Paper 2012-01-0792, 2012.
- [14] X. Na, *Study on passenger cars handling objective benchmark targets setting [M.S. thesis]*, Jilin University, Changchun, China, 2009.
- [15] C. Wang, *Study on objective evaluation index system of on-center handling for passenger car [M.S. thesis]*, Jilin University, Changchun, China, 2012.
- [16] ISO 13674-1:2003, *Road Vehicles-Test Method for The Quantification of On-Center Handling-Part1: Weave Test*, ISO, 2003.
- [17] G. P. Bertolini and R. M. Hogan, "Applying driving simulator to quantify steering effort preference as a function of vehicle speed," SAE Technical Paper 1999-01-0394, 1999.
- [18] C. Zong, X. Na, and D. Hu, "Study on steering effort preference of drivers based on driving simulator," SAE Technical Paper 2007-01-3575, 2007.
- [19] Y. W. Chai, "A study of effect of steering gain and steering torque on driver's feeling for SBW vehicle," in *Proceedings of the FISITA World Automotive Congress*, Barcelona, Spain, 2004.
- [20] JASO C 705-72, *Static Steering Efforts Test Procedure*, JASO, Japan, 2004.
- [21] B. Heissing and J. H. Brandl, *Subjektive Beurteilung des Fahrverhaltens*, Vogel Business Media, 2002.
- [22] C. Zong, L. Mai, D. Wang, and Y. Li, "Study on steering effort preference of drivers based on driving simulator," *Chinese Journal of Mechanical*, vol. 18, no. 8, pp. 1001–1005, 2007.
- [23] T. Kifuku and S. Wada, *An Electric Power-Steering System*, Mitsubishi Electric Advance, 1997.
- [24] S. Rongwei, L. Yi, T. Xiaohong, and S. Guobiao, "Research on modeling and compensation control strategy of electric power steering system," *Transactions of the Chinese Society of Agricultural Machinery*, vol. 38, no. 7, pp. 5–9, 2007.
- [25] W. Wang, *Automobile Design*, China Machine Press, Beijing, China, 2004.

Research Article

Study on the Detectability of the Sky-Surface Wave Hybrid Radar

Hou Chengyu,¹ Ke Guo,¹ Shi Tiange,² and Wang Yuxin³

¹ School of Electronics and Information Engineering, Harbin Institute of Technology, Harbin, Hei Longjiang 150001, China

² School of Astronautics, Harbin Institute of Technology, Harbin, Hei Longjiang 150001, China

³ Institute of Electronic and Information Engineering, Harbin Institute of Technology, Harbin, Hei Longjiang 150001, China

Correspondence should be addressed to Hou Chengyu; houcy@hit.edu.cn

Received 23 April 2014; Accepted 10 June 2014; Published 6 July 2014

Academic Editor: Hamid R. Karimi

Copyright © 2014 Hou Chengyu et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Working in the HF (high-frequency) band and the transmitter and receiver locating separately, the sky-surface wave hybrid radar both has the capabilities of the OTHR (over-the-horizon radar) and the advantage of the bistatic radar. As the electromagnetic wave will be disturbed by the ionosphere, interfered by the sea clutter and attenuated by the sea surface, the detectability of this radar system is more complex. So, in this paper, we will discuss the problem detailedly. First of all, the radar equation is deduced based on the propagation of the electromagnetic wave. Then, how to calculate the effect of the ionosphere and the propagation loss is discussed. And an example based on the radar equation is given. At last, the ambiguity function is used to analyze the range and velocity resolution. From the result, we find that the resolution has relation with the location of the target and the height of reflection point of the ionosphere. But compared with the location, the effect of the ionospheric height can be ignored.

1. Introduction

Sky-wave and surface-wave OTH radar take advantage of characteristics of HF electromagnetic wave refraction by the ionosphere and diffraction on the sea surface respectively to achieve the targets detection. The sky-surface wave hybrid radar is a new type of radar systems, which has strong application prospects on detecting the stealth target, countering antiradiation missiles, enhancing the radar survivability, and other aspects. However, the detectability of single sky-wave or ground-wave radar is difficult to improve. So applying the MIMO (multi input multi output) radar theory into the sky-wave and ground-wave radar system to improve the performance of the OTH radar is one of new research fields in recent years [1].

The sky-surface wave hybrid radar is a MIMO radar system. This system is based on the technology of the existing sky-wave and ground-wave OTH radar system and works as a bistatic radar. So the detectability and target recognition performance can be further improved. In this system,

the transmitter of the sky-wave radar is located inland far away from the coast, while the receiver of the ground-wave radar is on the coast, as the coastal part of the system only receives the echo, which prevents the receiver from the anti-radiation missile. Meanwhile, the multireceivers receive the target echoes from different directions, which can effectively solve the problem of RCS (Radar Cross Section) flickering and improve detectability of the low-altitude target.

Nowadays, this hybrid radar is at the experimental stage. Papers [2–4] describe the research status of this system in Canada, USA, and Australia. From these papers, we find that the test systems are built with the same working principle. These systems transmit the wave with the sky wave mode and the receivers are placed close to the target. Since the receivers are close to the target, the target echo is relatively stronger which can improve the detection probability. Researchers in Wuhan University try to use the broadcast signal DRM to realize the target detection in the hybrid mode [5]. As a new radar system, only the sea clutter characteristics, the ionospheric influence on the resolution, and the time

availability and the asynchronous filtering are mentioned in papers [6–10]. Few works on the analysis of the radar detectability have been reported. So we will discuss this topic thoroughly in this paper.

This paper mainly studies two aspects of the radar detectability. On one hand, the hybrid sky-surface wave radar equation is deduced. Then the attenuation effect of the ionosphere and sea surface on the radar detectability is analyzed. And an example is given to illustrate how to use the radar equation to calculate the detection range. On the other hand, the ambiguous function of the hybrid radar with FMCW (frequency modulated continuous wave) is analyzed and used to study the range and velocity resolutions of the hybrid sky-surface wave radar. And the results are compared with the ones of monostatic OTH radar.

2. The Radar Equation and the Propagation Loss

In this section, we will describe the working principle of the sky-surface wave hybrid radar and deduce the radar equation. In the equation, the most important parameter is the propagation loss. So we will discuss how to calculate it. Then, an example is given.

2.1. The Geometry of Sky-Surface Wave Hybrid Radar. The work pattern of the sky-surface wave hybrid radar is different from the sky-wave and the surface-wave OTH radar. The transmitter located far from the coast emits the electromagnetic waves and the receiver located at the coast receives the echo signal. So the radar equation should be based on the equations of sky-wave and surface-wave OTH radar equation, and the effect of the ionosphere and sea surface propagation should also be taken into consideration. The geometrical model of this radar is shown in Figure 1. For simplification, the curvature of the earth and the curve propagation in the ionosphere are not considered in this paper.

In Figure 1, L is the straight line distance from the transmitter to the receiver; h is the vertical distance from ionospheric reflection point to the ground; R_1 is the distance from the transmitter to reflection point; R_2 is the distance from the target to reflection point; R_{wg} is the straight line distance from the transmitter to the target; R_r is the distance between receiver and the target; α is the emission elevation of the radar beam; β is the bistatic angle of transmitter, receiver, and target; θ_r is the angle between the baseline of the transmitter to the receiver and the line of receiver to the target; ν is the target velocity; Φ is the angle between the target velocity direction and the bisectors of β . From the geometrical relationship, it can be found that $R_1 = R_2$ and the propagation distance of the wave is $R = R_1 + R_2 + R_r$.

2.2. The Radar Equation of the Hybrid Radar. Without considering the ionospheric effect, the power density of the wave emitted by the transmitter at the target location is

$$S_t = \frac{P_t \sigma_t}{4\pi R_s^2} \quad (\text{W/m}^2), \quad (1)$$

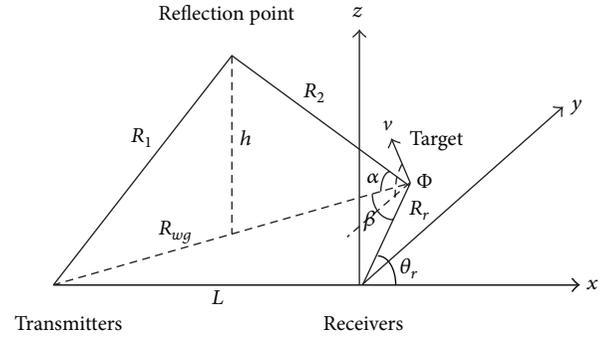


FIGURE 1: The geometrical model of the HF sky-surface hybrid radar.

where P_t denotes the transmitting power of the sky-wave radar (W) and R_s denotes the propagation distance from the transmitter to the target. The scattering power at the target is

$$P_{t2} = S_t \times \sigma_t \times 10^{-3} = \frac{P_t \sigma_t}{4\pi R_s^2} \times 10^{-3} \quad (\text{kW}), \quad (2)$$

where σ_t is the RCS. The field intensity received by the receiver is

$$E_r(R_r) = E(R_r) \sqrt{\frac{P_{t2}}{3}} \quad (\mu\text{V/m}). \quad (3)$$

The effective cross-sectional area of the receiving antenna is $A_r = D_r \lambda^2 / 4\pi$, where λ is the working wavelength of the radar; then the power received by the receiving antenna is

$$P_r = \frac{E_r^2(R_r)}{120\pi} \times 10^{-12} \times A_r \quad (\text{kW}). \quad (4)$$

$E_r(R_r)$ is the radiation field intensity ($\mu\text{V/m}$) where the distance to the receivers is R_r and the transmitter is located at the coast with the transmitting power 1 kW. Then P_r can be written as

$$P_r = \frac{5.6 \times 10^{-21} E^2(R_r) P_t D_t D_r \lambda^2 \sigma_t}{R_s^2} \quad (\text{W}). \quad (5)$$

A total loss factor L is used to express the effects of different factors on the radar detection performance. And L should include the following factors: ① the ionospheric propagation loss L_p , which includes the absorption loss, polarization loss generated by the Faraday effect, multipath effects, and so on; ② the radar system loss L_s , which contains the loss produced by the weighting processing, the beam forming, the antenna efficiency, and so on; ③ sea wave additional loss L_w .

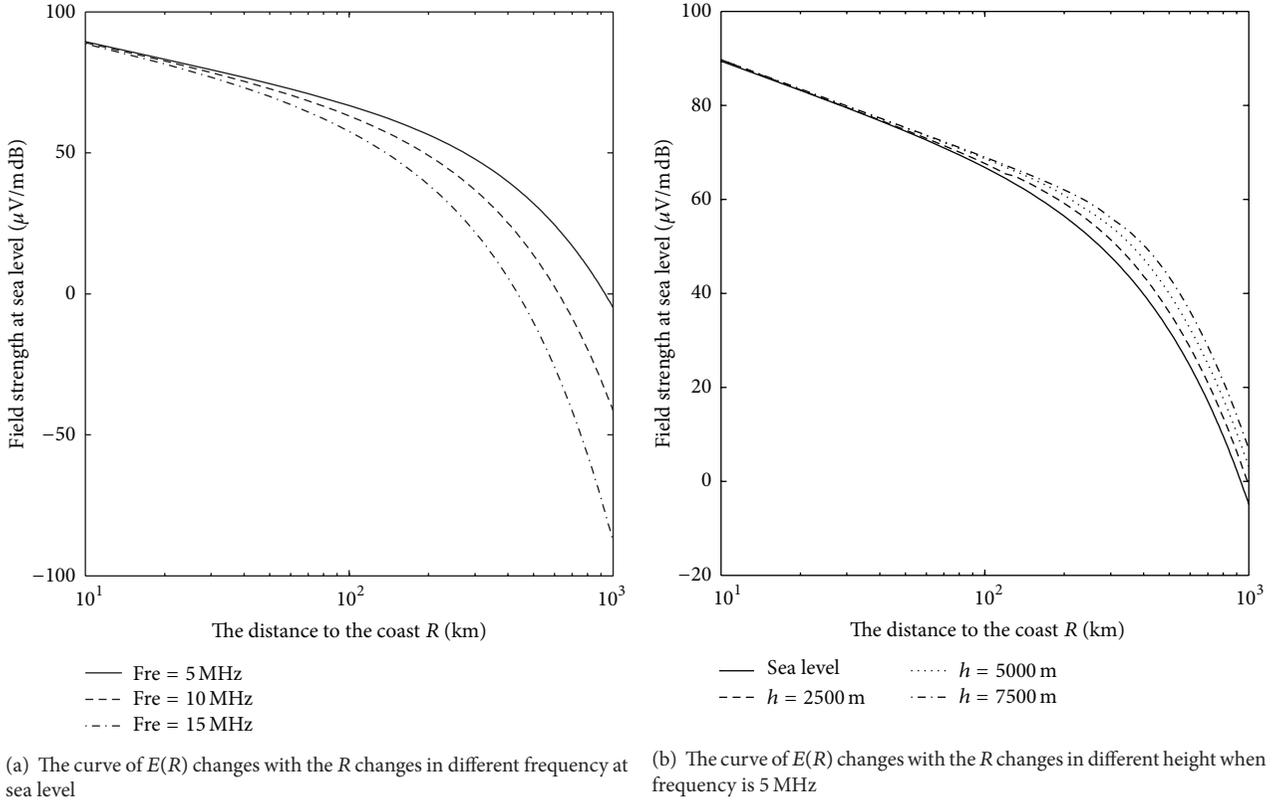


FIGURE 2: The characters of $E(R)$.

If the coherent time is T_c , the received effective energy E_{ra} can be written as

$$E_{ra} = \frac{P_r \times T_c}{L} = 5.6 \times 10^{-21} E^2 (R_r) P_t D_t D_r \lambda^2 \sigma_t \times \frac{T_c}{R_s^2 L_p L_s L_w} \text{ (W)}. \quad (6)$$

For the target with a low radial velocity, the signal to clutter energy ratio (SCR) should be taken into consideration in the radar equation, while for the target with a high one, the signal to noise ratio (SNR) should be considered. The power of the noise $P_n(f) = KT_o F_a$. Then the radar equation is

$$\frac{S}{N} = \frac{E_{ra}}{P_n(f)} = 5.6 \times 10^{-21} E^2 (R_r) P_t D_t D_r \lambda^2 \sigma_t \times \frac{T_c}{R_s^2 L_p L_s L_w K T_o F_a}, \quad (7)$$

where P_t is the mean transmission power, K is Boltzmann's constant, T_o is the absolute temperature, and F_a is the external noise factor.

To calculate (7), $E(R)$ and the system loss should be known. For $E(R)$, we can get the value in the ITU-Radio wave Propagation (Vol.2000) report, as shown in Figure 2. For the latter, the sea wave additional loss and ionospheric propagation loss are mainly considered and how to calculate them is discussed in Sections 2.3 and 2.4.

2.3. The Sea Wave Additional Loss Factor. The value of $E(R)$ in Figure 2 is gotten when the sea surface is smooth. However, the sea wave, sea wind, and the weather will cause an additional loss L_w . And L_w is related to the radar working frequency, sea-state, and the propagation distance.

L_w can be obtained in paper [11], as shown in Figure 3. The sea surface additional loss increases sharply as the frequency, sea wind, and propagation distance increase. To reduce the effect of L_w , the working frequency of the this hybrid radar should be selected in the low band.

Figure 3 shows L_w varies with different sea states when the working frequency is 3 MHz, 5 MHz, 10 MHz, and 20 MHz. Using the interpolation method, we can approximately calculate L_w when the frequency is 3~20 MHz and the sea state is from state 1 to state 6.

2.4. The Ionospheric Propagation Loss. The ionospheric propagation loss includes the polarization loss and the ionospheric

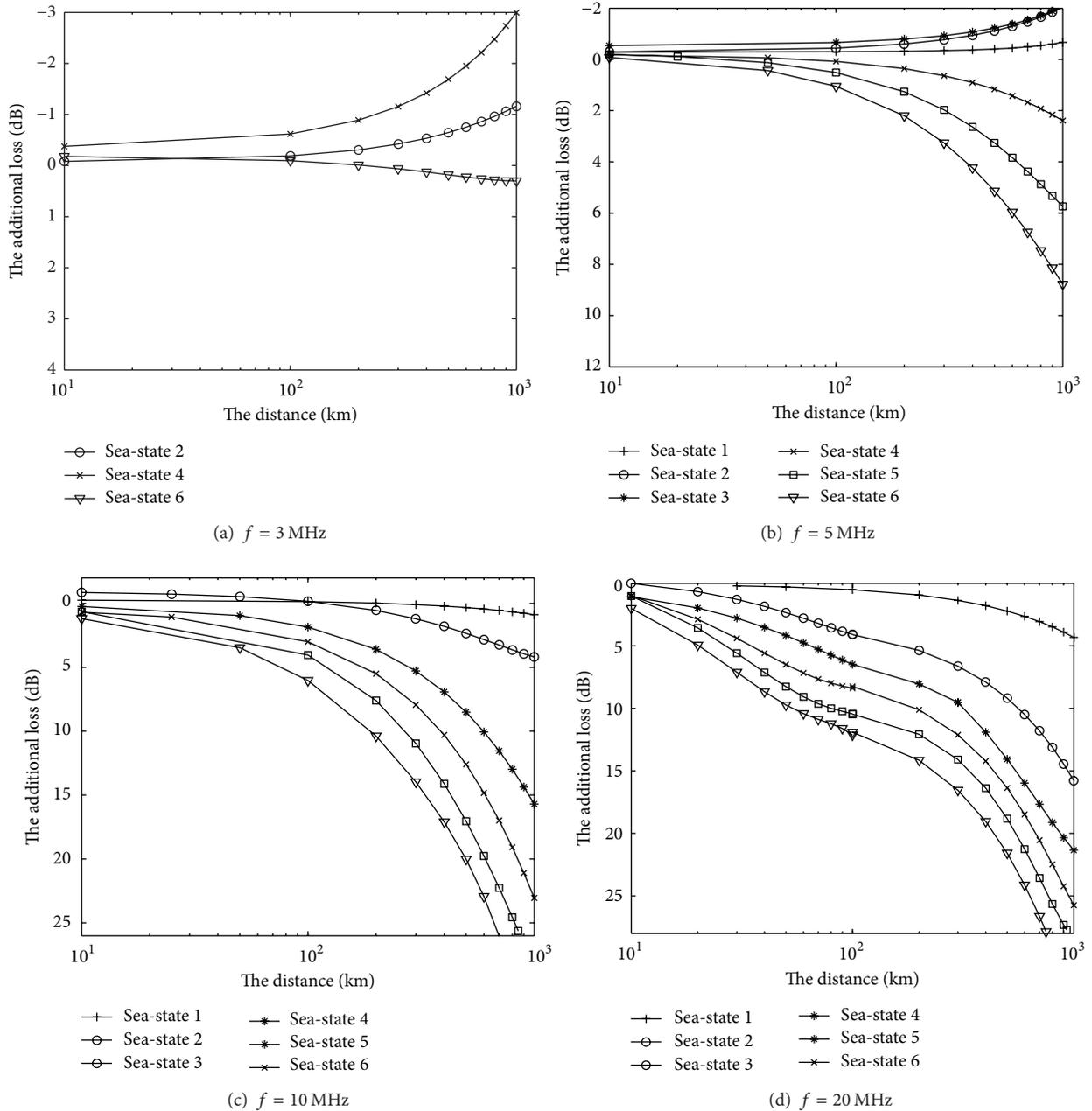


FIGURE 3: The curve of L_w with different sea states (S-T is sea state; W-V is wind velocity).

absorption loss L . The polarization loss is usually 3 dB. And the ionospheric absorption loss usually is obtained with

$$L = \frac{677.2 \sec \theta_0}{(f + f_H)^{1.98} + 10.2} I. \quad (8)$$

The unit of L is dB. θ_0 is the incident angle at 100 km high, and $\theta_0 = \arcsin(0.985 \cdot \cos \Delta)$. Δ is the launch elevation, f is the working frequency (MHz), and f_H is the magnetic gyro frequency (MHz) at 100 km high. I is an absorption index [12, 13].

2.5. An Example of HF Sky-Surface Wave Hybrid Radar. Without considering the curvature of the earth, the geometrical model of this hybrid radar is shown in Figure 1. Assuming the working frequency $f_o = 10$ MHz, the straight distance between the transmitting station and the receiving station $L = 1000$ km and the height from ionospheric reflection point to the ground $h = 200$ km. As the detection range of the surface-wave radar is usually less than 200 km, let $R_r = 230$ km.

Let the transmitting gain $D_t = 100$, let the receiving gain $D_r = 14$, let the aircraft $\sigma_t = 100 \text{ m}^2$, let the coherent integration time $T_c = 60$ s, let the radar system loss $L_s = 10$ dB,

$T_o = 290\text{K}$, let $K = 1.38 \times 10^{-23}$, and let the noise factor $F_a = 40\text{ dB}$.

When $R_r = 230\text{ km}$ and $f_o = 10\text{ MHz}$, we can get $E(R) = 45.3\text{ dB}\mu\text{V/m}$ in Figure 2.

When the sea-state is 5, the additional wave loss $L_w = 8.6\text{ dB}$ is calculated by the interpolation based on the data as shown in Figure 3. But the value of L_w is the back and forth attenuation in Figure 3. As L_w is the attenuation of the back propagation in (7), L_w should be half of 8.6 dB. $L_w = 4.3\text{ dB}$.

By (8), the ionospheric attenuation $L = 7.1\text{ dB}$. And the attenuation loss of the Faraday effect is 3 dB. Taking these values into (7), we can get the transmitting power $P_t = 32.7\text{ kW}$, which is similar to an actual value of the OTH radar. So (8) is correct theoretically.

3. Analysis of the Ambiguous Function of the Hybrid Radar

We often use the ambiguity function to analyze the range-velocity resolution of the radar system and the effect factors of the resolution. So in this section, we will discuss the ambiguous function of the hybrid radar [14, 15].

3.1. Calculation of the Geometrical Parameters. From Figure 1, the distance between the transmitting station to the target is

$$\begin{aligned} R_{wg}^2 &= R_r^2 + L^2 - 2R_r L \cos(\pi - \theta_r) \\ &= R_r^2 + L^2 + 2R_r L \cos \theta_r. \end{aligned} \quad (9)$$

Then,

$$\begin{aligned} R_1^2 &= h^2 + \left(\frac{1}{2}R_{wg}\right)^2 \\ &= h^2 + \frac{1}{4}(R_r^2 + L^2 + 2R_r L \cos \theta_r). \end{aligned} \quad (10)$$

The total propagation distance is

$$R = R_r + 2\sqrt{h^2 + \frac{1}{4}(R_r^2 + L^2 + 2R_r L \cos \theta_r)}. \quad (11)$$

Assuming the transmitting station and the receiving station are stationary, the range rate of the target relative to the transmitting station is the velocity component along the direction of target to the ionospheric reflection point, which can be expressed as

$$\frac{d}{dt}R_t = v \cos\left(\phi - \frac{\beta}{2}\right) \cos \alpha. \quad (12)$$

The range rate relative to the receiving station is

$$\frac{d}{dt}R_r = v \cos\left(\phi + \frac{\beta}{2}\right). \quad (13)$$

Then the change rate of the total propagation distance is

$$\begin{aligned} \frac{d}{dt}R &= \frac{d}{dt}(R_t + R_r) \\ &= v \left[\cos \phi \cos\left(\frac{\beta}{2}\right) (1 + \cos \alpha) \right. \\ &\quad \left. + \sin \phi \sin\left(\frac{\beta}{2}\right) (\cos \alpha - 1) \right], \end{aligned} \quad (14)$$

where

$$\cos \beta = \frac{R_{wg}^2 + R_r^2 - L^2}{2R_{wg}R_r} = \frac{R_r + L \cos \theta_r}{\sqrt{R_r^2 + L^2 + 2R_r L \cos \theta_r}}. \quad (15)$$

The relationship between the launch elevation α and θ_r is

$$\alpha = \arctan\left(\frac{2h}{R_{wg}}\right) = \arctan\left(\frac{2h}{\sqrt{R_r^2 + L^2 + 2R_r L \cos \theta_r}}\right). \quad (16)$$

Then the total propagation distance can be obtained by the initial value R_o and the range rate. The expression is

$$\begin{aligned} R(t) &= R_o + vt = R \\ &= [R_r + 2\sqrt{D}] \\ &\quad + v \left[A_1 \sqrt{\frac{1}{2} + \frac{B}{2\sqrt{C}}} + A_2 \sqrt{\frac{1}{2} - \frac{B}{2\sqrt{C}}} \right] t, \end{aligned} \quad (17)$$

where

$$\begin{aligned} A_1 &= \cos \phi (1 + \cos \alpha), & A_2 &= \sin \phi (-1 + \cos \alpha), \\ B &= R_r + L \cos \theta_r, & C &= R_r^2 + L^2 + 2R_r L \cos \theta_r, \end{aligned} \quad (18)$$

$$D = h^2 + \frac{1}{4}C.$$

3.2. The Ambiguity Function of the Hybrid Radar. The FMCW signal is often used as the transmitting signal in the OTH radar system, which is expressed as below:

$$S_T(t) = \cos \left\{ 2\pi \left[f_o t + \frac{\mu t^2}{2} \right] + \phi_o \right\}, \quad (19)$$

where T is the sweeping period, $\mu = B_w/T$ is the sweeping slope, and B_w is the sweeping bandwidth.

For simplification, assume there are only two targets in the area that can be detected. The positions of the two targets in the radar plane are (R_r, θ_r) and $(R_r + \delta, \theta_r)$, respectively. Only the velocity component of the target along direction of the bistatic angle bisector can be detected, That is, $\Phi = 0$. Let

the velocity component along the bistatic angle of two targets be $V, V + \zeta$. Then, the echo time delay of Target 1 is

$$\begin{aligned}\tau_1(t) &= \frac{R_1(t)}{c} = [R_r + 2\sqrt{D}] \frac{1}{c} \\ &\quad + \frac{V}{c} A_1 \sqrt{\frac{1}{2} + \frac{B}{2\sqrt{C}}} t \\ &= E_1 + F_1 t.\end{aligned}\quad (20)$$

The echo time delay of Target 2 is

$$\tau_2(t) = \frac{R_2(t)}{c} = \left[R_r + \delta + 2\sqrt{D + \frac{1}{2}B\delta + \frac{1}{4}\delta^2} \right] \frac{1}{c}$$

$$\begin{aligned}\xi^2 &= \int_{-T/2}^{T/2} |S_1(t) - S_2(t)|^2 dt \\ &= 2E - 2 \int_{-T/2}^{T/2} S_1(t) S_2^*(t) dt \\ &= 2E - 2E \operatorname{Re} \left\{ \int_{-T/2}^{T/2} e^{j2\pi\{[f_o(E_2-E_1)+(\mu/2)(E_1^2-E_2^2)]+[f_o(F_2-F_1)+\mu(E_2-E_1)+(E_1F_1-E_2F_2)]t+\mu[(F_2-F_1)+(1/2)(F_1^2-F_2^2)]t^2\}} dt \right\},\end{aligned}\quad (23)$$

where $E = \int_{-T/2}^{T/2} |S_1(t)|^2 dt = \int_{-T/2}^{T/2} |S_2(t)|^2 dt = \text{contant}$.

$$\chi(\delta, \zeta) = \int_{-T/2}^{T/2} e^{j2\pi\{f_o(F_2-F_1)+\mu[(E_2-E_1)+(E_1F_1-E_2F_2)]t+\mu[(F_2-F_1)+(1/2)(F_1^2-F_2^2)]t^2\}} dt \approx \int_{-T/2}^{T/2} e^{j2\pi\{f_o(F_2-F_1)+\mu(E_2-E_1)t\}} dt.\quad (24)$$

Equation (24) can be simplified as

$$\begin{aligned}\chi(\delta, \zeta) &= \sin c \left(\pi T \frac{\zeta}{\lambda} \left[A_1 \sqrt{\frac{1}{2} + \frac{B + \delta}{2\sqrt{C + 2B\delta + \delta^2}}} \right. \right. \\ &\quad \left. \left. + \frac{\pi B_w}{c} \left[\delta + 2\sqrt{D + \frac{1}{2}B\delta + \frac{1}{4}\delta^2} - 2\sqrt{D} \right] \right) \right).\end{aligned}\quad (25)$$

From (25), we find that $\chi(\delta, \zeta)$ is the function of R_r and θ_r , which indicates that the ambiguity function is related to the location of the target.

3.3. Range Resolution. When the range resolution is considered, the velocity components of the two targets along the direction of the bistatic bisector are the same, which means $\zeta = 0$. So the ambiguity function of the range resolution can be expressed as

$$\chi(\delta, 0) = \sin c \left[\frac{\pi B_w}{c} \left(\delta + 2\sqrt{D + \frac{1}{2}B\delta + \frac{1}{4}\delta^2} - 2\sqrt{D} \right) \right].\quad (26)$$

$$\begin{aligned}&+ \frac{V + \zeta}{c} \left(A_1 \sqrt{\frac{1}{2} + \frac{B + \delta}{2\sqrt{C + 2B\delta + \delta^2}}} \right) t \\ &= E_2 + F_2 t.\end{aligned}\quad (21)$$

In the period $[0, T]$, the echo can be expressed as

$$\begin{aligned}S_1(t) &= S_T [t - \tau_1(t)], \\ S_2(t) &= S_T [t - \tau_2(t)].\end{aligned}\quad (22)$$

The resolution depends on the difference between the two targets in the time period T , so

So ξ is related to the ambiguity function $\chi(\delta, \zeta)$:

The condition that two targets can be distinguished is $\chi(\delta, 0) \geq -3$ dB, so

$$\frac{B_w}{c} \left(|\delta| + 2\sqrt{D + \frac{1}{2}B|\delta| + \frac{1}{4}\delta^2} - 2\sqrt{D} \right) \leq a,\quad (27)$$

where $a = 0.443$. Then (27) can be written as

$$\begin{aligned}2\delta^2 + 2B|\delta| + 4\sqrt{D + \frac{1}{2}B|\delta| + \frac{1}{4}\delta^2} \\ \leq \frac{4ac\sqrt{D}}{B_w} + \frac{a^2c^2}{B_w^2}.\end{aligned}\quad (28)$$

As $2\delta^2 \ll 2B|\delta| + 4\sqrt{D + (1/2)B|\delta| + (1/4)\delta^2}$, $(1/2)B|\delta| + (1/4)\delta^2 \ll D$, $a^2c^2/B_w^2 \ll 4ac\sqrt{D}/B_w$, and $2|\delta| \leq \delta_R$, so the radar range resolution δ_R is

$$\begin{aligned}\delta_R &= \frac{4ac}{B_w(B/\sqrt{D} + 2)} \\ &= 4ac\end{aligned}$$

$$\times \left(B_w \left((R_r + L \cos \theta_r) \div \sqrt{h^2 + \frac{1}{4} (R_r^2 + L^2 + 2R_r L \cos \theta_r) + 2} \right) \right)^{-1} \quad (29)$$

From (29), it can be known that the range resolution of the hybrid radar is not only inversely proportional to the transmitting bandwidth, but also related to the location of the target and the height of reflection points in the ionosphere. Figure 4 shows the simulation results of δ_R with the working frequency is 15 MHz and the bandwidth is 20 kHz.

As shown in Figure 4(b), the range resolution increases when θ_r increases. And when $\theta_r \geq 30^\circ$, the increasing rate increases significantly. Figures 4(c) and 4(d) illustrate that the range resolution increases with R_r decreases and h increases respectively. We also find that the changes of h and R_r have little effect on the range resolution when θ_r is small. Only when θ_r is large, the changes have an obvious effect.

$$\begin{aligned} \zeta_v &\approx \frac{2a\lambda}{T} \frac{1}{A_1 \sqrt{1/2 + B/2\sqrt{C}}} \\ &= \frac{2a\lambda}{T} \left(1 \times \left(1 + \arctan \left(\frac{2h}{\sqrt{R_r^2 + L^2 + 2R_r L \cos \theta_r}} \right) \sqrt{\frac{1}{2} + \frac{R_r + L \cos \theta_r}{2\sqrt{R_r^2 + L^2 + 2R_r L \cos \theta_r}}} \right)^{-1} \right). \end{aligned} \quad (32)$$

By (32), it can be known that the velocity resolution of the hybrid radar is not only inversely proportional to the sweep period and proportional to the wavelength, but also related to the location of the target and the height of reflection points in the ionosphere. Figure 5 shows the simulation results about ζ_v with $T = 1$ s and $f = 15$ MHz.

As shown in Figure 5, we can get the similar conclusion of the velocity resolution with the one about the range resolution.

For the velocity resolution of the sky-wave or surface-wave radar is $\Delta V = \lambda/2T$, ΔV is only related to the sweeping period. When $T = 1$ s, $\Delta V = 10$ m/s. In comparison with Figure 5(b), when $\theta_r < 60.7^\circ$, the hybrid radar has a better velocity resolution than the monostatic OHT radar.

3.5. The Coupling Analysis of Range and Velocity. From the ambiguity theory [16], when the velocity and range of the two targets are not same, the time difference and frequency difference should be different. But under some conditions, $\chi(\delta, \zeta) = 0$ may occur, which means that two targets cannot

The range resolution of sky-wave or surface-wave radar is $\Delta R = c/2B$, which means the range resolution is only related to the bandwidth. When $B = 20$ kHz, $\Delta R = 7.5$ km. So, from Figure 4(b), when $\theta_r < 44^\circ$, the hybrid radar has a better range resolution than monostatic OHT radar.

3.4. The Velocity Resolution. When the velocity resolution is considered, the location of the two targets are same, which means $\delta = 0$. So the ambiguity function of the velocity resolution can be expressed as

$$\chi(0, \zeta) = \sin c \left[\pi T \frac{\zeta}{\lambda} \left(A_1 \sqrt{\frac{1}{2} + \frac{B}{2\sqrt{C}}} \right) \right]. \quad (30)$$

As the condition that two targets can be distinguished is $\chi(0, \zeta) \geq -3$ dB, so

$$T \frac{|\zeta|}{\lambda} \left[A_1 \sqrt{\frac{1}{2} + \frac{B}{2\sqrt{C}}} \right] \leq a. \quad (31)$$

As $2|\zeta| \leq \zeta_v$, the velocity resolution ζ_v is

be distinguished. By (25), let $\chi(\delta, \zeta) = 0$; it can be obtained that

$$\begin{aligned} T\zeta \left[A_1 \sqrt{\frac{1}{2} + \frac{B + \delta}{2\sqrt{C + 2B\delta + \delta^2}}} \right] \\ + \frac{B_w}{c} \left[\delta + 2\sqrt{D + \frac{1}{2}B\delta + \frac{1}{4}\delta^2} - 2\sqrt{D} \right] = 0. \end{aligned} \quad (33)$$

The equation above is rewritten as

$$\begin{aligned} \left(2B + 4\sqrt{D + \frac{1}{2}B\delta + \frac{1}{4}\delta^2} \right) \delta + \frac{5}{4}\delta^2 \\ = -\frac{4Tc\zeta}{\lambda B_w} A_1 \sqrt{\frac{1}{2} + \frac{B + \delta}{2\sqrt{C + 2B\delta + \delta^2}}} \\ + \left(\frac{Tc\zeta}{\lambda B_w} A_1 \right)^2 \left(\frac{1}{2} + \frac{B + \delta}{2\sqrt{C + 2B\delta + \delta^2}} \right). \end{aligned} \quad (34)$$

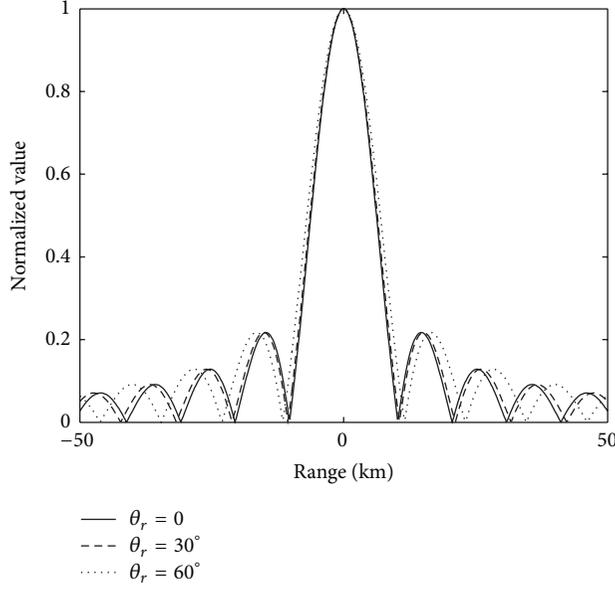
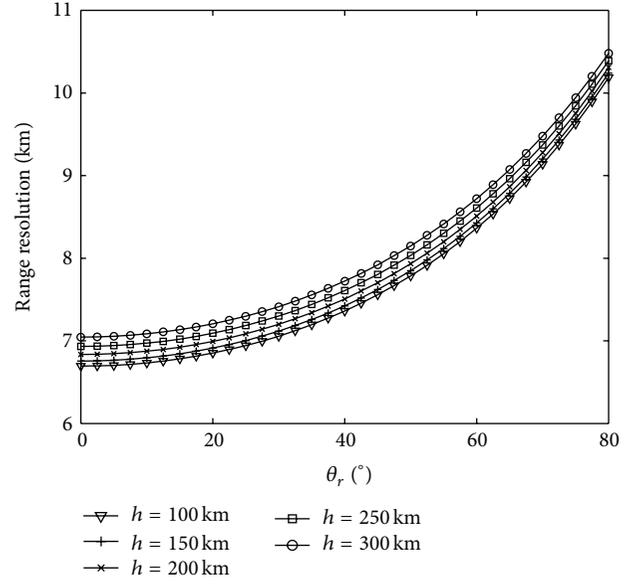
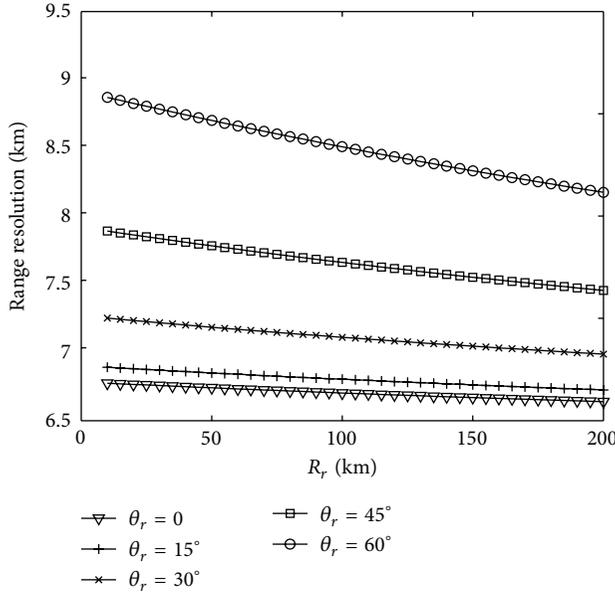
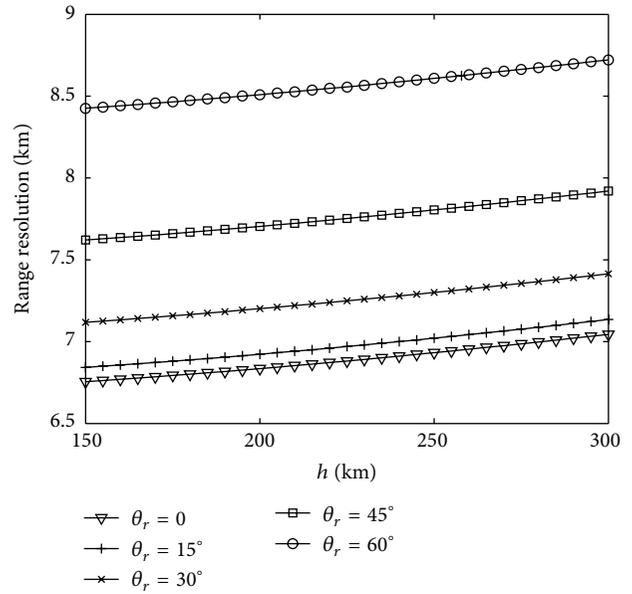
(a) The range ambiguous function when θ_r is 0, 30°, or 60°(b) When $R_r = 150$ km, the range resolution changes with θ_r and h (c) When $h = 200$ km, the range resolution changes with h and θ_r (d) When $R_r = 150$ km, the range resolution changes with h and θ_r

FIGURE 4: The ambiguous function and the range resolution of the hybrid radar.

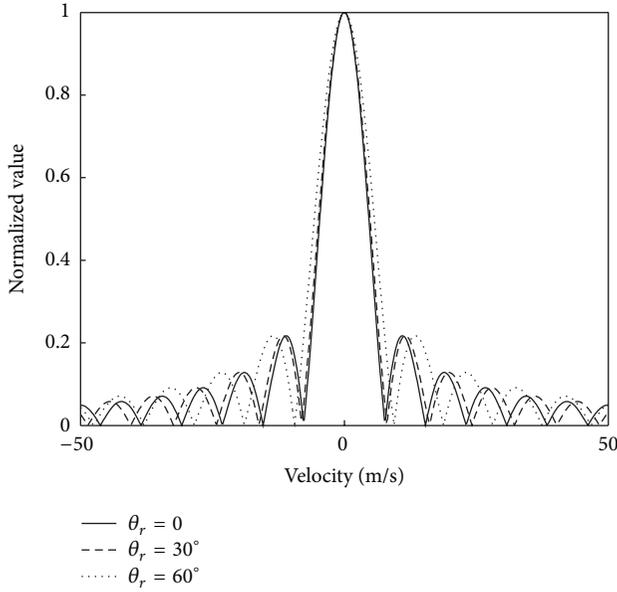
The second term of the left side of (34) can be ignored compared with the first term, and $\sqrt{D + (1/2)B\delta + (1/4)\delta^2} \approx \sqrt{D}$, $\sqrt{1/2 + (B + \delta)/2\sqrt{C + 2B\delta + \delta^2}} \approx \cos(\beta/2)$, so

$$\frac{\delta}{\zeta} \approx -\frac{4Tc\zeta}{\lambda B_w} \frac{1}{2B/\sqrt{D}} \cos\left(\frac{\beta}{2}\right). \quad (35)$$

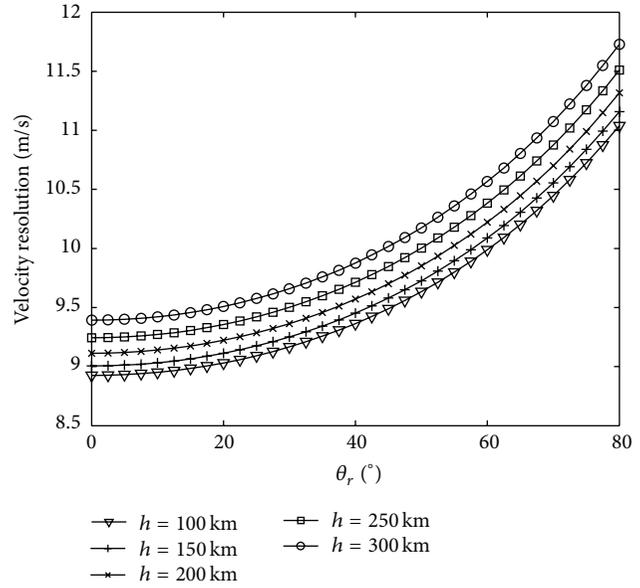
When the difference of the range-velocity between the two targets satisfies (35), the two targets cannot be distinguished.

4. Conclusion

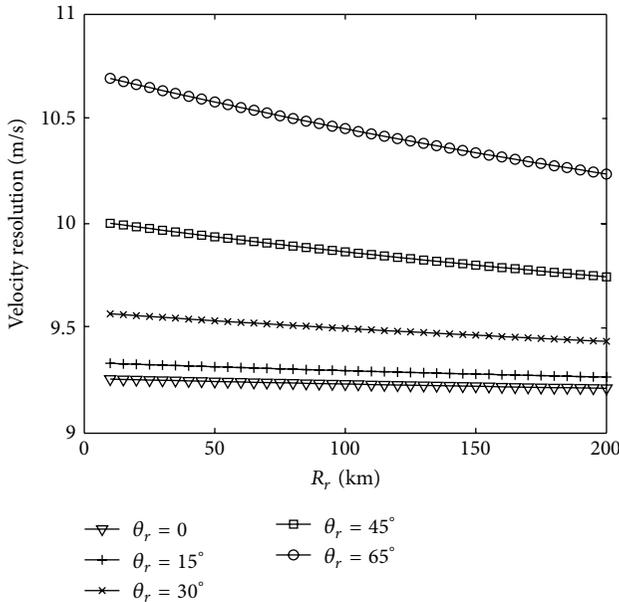
As a new type of bistatic radar, the hybrid sky-surface wave OTH radar is different with the sky-wave and the surface-wave OTH radar. So the radar equation and the range-velocity resolution, the two aspects of the detection performance of this hybrid radar, are studied in this paper. According to the ones of the sky-wave and the surface-wave radar, the radar equation of this hybrid radar which concludes the ionospheric loss and the wave additional loss is derived. Then, an example is given to testify the validity. From the



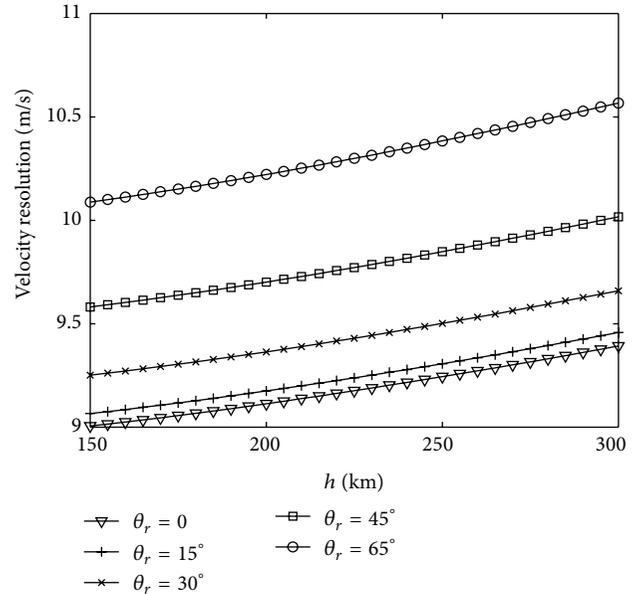
(a) The velocity ambiguous function when θ_r is 0, 30°, or 60°



(b) When $R_r = 150$ km, the velocity resolution changes with θ_r and h



(c) When $h = 200$ km, the velocity resolution changes with h and θ_r



(d) When $R_r = 150$ km, the velocity resolution changes with h and θ_r

FIGURE 5: The ambiguous function and the velocity resolution of the hybrid radar.

analysis of the ambiguity function, we find that the range-velocity resolution of this radar is related to the location of the target and the reflection height in the ionosphere. In a certain launch elevation, the range-velocity resolution is better with respect to the one of the monostatic OHT radar, and the effect of the target location and the ionospheric reflection height can be ignored to some extent. It can be predicted that with more MIMO radar theory is applied to this radar system, the performance of the OHT radar will be further enhanced.

In this paper, the detectability of the radar system with one sky-wave radar transmitter and one surface-wave radar receiver is studied. How to evaluate the performance of

the netted radar system with multiple surface-wave radar receivers and networked control [17] will be the future work.

Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

Acknowledgment

At the point of finishing this paper, the author would like to express his sincere thanks to the National Natural Science

Foundation (no. 61201303) and the Fundamental Research Funds for the Central Universities (HIT.NSRIF.2013027).

References

- [1] E. Fishler, A. Haimovich, R. Blum, D. Chizhik, L. Cimini, and R. Valenzuela, "MIMO radar: an idea whose time has come," in *Proceedings of the IEEE Radar Conference*, pp. 71–78, IEEE Press, Philadelphia, Pa, USA, April 2004.
- [2] R. J. Riddolls, "Limits on the detection of low-Doppler targets by a high frequency hybrid sky-surface wave radar system," in *Proceedings of the IEEE Radar Conference (RADAR '08)*, pp. 1–4, Rome, Italy, May 2008.
- [3] P. Vouras and B. Freburger, "Application of adaptive beamforming techniques to HF radar," in *Proceedings of the IEEE Radar Conference (RADAR '08)*, pp. 795–800, May 2008.
- [4] G. J. Frazer, "Forward-based receiver augmentation for OTHR," in *Proceedings of the IEEE Radar Conference*, pp. 373–378, Boston, Mass, USA, April 2007.
- [5] Y. Li and Y. Wei, "Analysis of first-order sea clutter spectrum characteristics for HF sky-surface wave radar," in *Proceedings of the International Conference on Radar (Radar '13)*, pp. 368–373, Adelaide, Australia, September 2013.
- [6] Y. Li, Y. Wei, R. Xu, and Z. Wang, "Influence of ionosphere on resolution cell of HF hybrid sky-surface wave radar," in *Proceedings of the IEEE Antennas and Propagation Society International Symposium (APSURSI '13)*, pp. 1028–1029, 2013.
- [7] L. Zhang and H. Gao, "Asynchronously switched control of switched linear systems with average dwell time," *Automatica*, vol. 46, no. 5, pp. 953–958, 2010.
- [8] Z. Zhixin, W. Xianrong, Z. Delei, and C. Feng, "An experimental study of HF passive bistatic radar via hybrid sky-surface wave mode," *Transactions on Antennas and Propagation*, vol. 61, no. 1, pp. 415–424, 2013.
- [9] Y. Longquan, F. Junmei, J. Peinan, and C. Hongyan, "Time availability of hybrid sky-surface wave radar," *Chinese Journal of Radio Science*, vol. 28, no. 5, pp. 829–834, 840, 2013.
- [10] L. Zhang, N. Cui, M. Liu, and Y. Zhao, "Asynchronous filtering of discrete-time switched linear systems with average dwell time," *IEEE Transactions on Circuits and Systems I: Regular Papers*, vol. 58, no. 5, pp. 1109–1118, 2011.
- [11] Z. Wenyu and J. Peinan, *OTH Radar Technology*, Publishing House of Electronics Industry, 2004.
- [12] L. Jiaying and P. Jiao, "Nomograms for tactical skywave communication in tropical forest regions," *Chinese Journal of Radar Science*, vol. 9, no. 1, pp. 63–72, 1994.
- [13] V. Jodalen and E. V. Thrane, "A study of the relation between ionospheric absorption and predicted HF propagation parameters at high latitudes," in *Proceedings of the 6th International Conference on HF Radio Systems and Techniques*, pp. 64–68, New York, NY, USA, July 1994.
- [14] M. C. Jackson, "The geometry of bistatic systems," *IEE Proceedings*, vol. 133, no. 7, pp. 604–612, 1986.
- [15] D. Trizna and J. Gordon, "Results of a bistatic HF radar surface wave sea scatter experiment," in *Proceedings of the IEEE International Geoscience and Remote Sensing Symposium (IGARSS '02)*, vol. 3, pp. 1902–1904, Piscataway, NJ, USA, June 2002.
- [16] T. Tsao, M. Slamani, P. Varshney, D. Weiner, H. Schwarzlander, and S. Borek, "Ambiguity function for a bistatic radar," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 33, no. 3, pp. 1041–1051, 1997.
- [17] L. Zhang, H. Gao, and O. Kaynak, "Network-induced constraints in networked control systems—a survey," *IEEE Transactions on Industrial Informatics*, vol. 9, no. 1, pp. 403–416, 2013.

Research Article

Urban and Indoor Weak Signal Tracking Using an Array Tracker with MVA and Nonlinear Filtering

Jicheng Ding,¹ Guoli Zhang,² and Lin Zhao¹

¹ College of Automation, Harbin Engineering University, Harbin 150001, China

² The Unit 92941, PLA, Huludao 125000, China

Correspondence should be addressed to Jicheng Ding; aaron.heu@163.com

Received 4 May 2014; Accepted 17 June 2014; Published 2 July 2014

Academic Editor: Yuxin Zhao

Copyright © 2014 Jicheng Ding et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

We focus on the need for weak GPS signal tracking technique at a receiver powered on in urban or indoor environment; the tracking loop is unlocked and data bit edge position is unknown. A modified Viterbi algorithm (MVA) based on dynamic programming is developed and it is applied to GPS bit synchronization to improve bit edge position detection probability. Meanwhile, two combination carrier tracking schemes based on central difference Kalman filter (CDKF) and MVA module are designed for tracking very weak GPS signal. The testing results indicate that the methods can successfully detect bit edge position with high detection probability whether or not the tracking loop is locked. The tested combination tracking scheme is still able to work well when the signal quality deteriorates to 20 dB-Hz without additional large store space.

1. Introduction

Urban and indoor positioning techniques have become a major focus for research over the past decade for applications like location based services (LBS). Such requirements are the main thrust in the technology development for a pedestrian navigation system (PNS). Global positioning system (GPS) is undoubtedly a viable option for use in such systems due to its high accuracy positioning capability. A pedestrian typically moves in dense urban areas (along sidewalks), inside buildings, in tunnels, and under foliage, where GPS signals often fail to reach the user. That is, the receiver technologies to overcome weak signal environment have been a topic of radiobiological interest.

In the present day practice, a receiver may get “assistance” in the form of additional acquisition aiding messages from a server or base station or Internet based [1]. However, the E-911 mandate requires GPS receivers to provide a position solution in any kind of environment [2]. So, providing this type of assistance requires additional infrastructure and may not be available in all places. Also, the receiver requires additional hardware to receive the aiding messages. Therefore developing a GPS receiver which is able to work well without

assistant information under weak or indoor signal conditions is very important. Further, there is a need as in the case of E-911.

Conventional GPS receivers process signals with a C/N_0 above 33–35 dB-Hz successfully. With advances in GPS receiver technology, high sensitivity GPS (HSGPS) technology provides the ability to process signal at lower power levels using longer predetection integration times (PIT) [3]. At present, several studies indicate that coherent integration, noncoherent integration, and differential coherent integration have led to growing acceptance of these weak signal acquisition approaches [4]. Followed by the signal acquisition strategy is weak signal tracking with long time integration model. Thus, it raises an interesting question: the data bit timing presents a navigation message bit offset ambiguity; the purpose of data bit edge detection is to avoid integration across a data bit edge which might cause loss of bit energy and errors in the navigation message extraction. Conventional wisdom states that histogram method (HM) and Kokkonen and Pietila method (K-P method) are able to combat the bit offset ambiguity. However, it is difficult to get ideal results using these methods when receiving signal degrading [5]. Lately, considerable progress has been made in this question,

such as Viterbi algorithm (VA) [5–7]. This idea using dynamic programming technology has led to a better result than the previous two methods [7]; that is, the VA with a cost function in the form of bit energy sum is able to determine data bit boundary with a higher probability detection (PD) in the presence of weak signal. The form of cost function would lead to error when detection occurs in those close to true bit edge positions. Furthermore, these approaches would be subjected to carrier phase and frequency error [8]. This means that they will be performed after the loop is locked or has wiped off carrier phase and frequency error. This will raise a very realistic problem: if a receiver is powered on under very weak signal conditions, its loop will not be locked forever and the data bit edge detection methods as mentioned would be failure. Received signal phase changes fierce in the presence of much noise. As a traditional PLL (phase locked loop) phase discriminator with nonlinear characteristic and the loop is easy to lose lock [9, 10]. It is necessary to develop an effective tracking method when a receiver is powered on in weak signal environments.

The structure of the paper is as follows. In Section 2, a modified Viterbi algorithm (MVA) based on hidden Markov model (HMM) with alternative cost function is developed for improving GPS bit synchronization. When detection occurs especially in those close-to-true bit edge positions, the new cost function will make the bit energy in different position appear of great differences and turn out to be more robust when working in weak signal environments. In Section 3, a CDKF weak signal tracking model and filtering process are designed by establishing the carrier tracking error model to avoid nonlinear affection coming from the traditional PLL discriminator. In Section 4, a weak signal tracking scheme which combined CDKF with MVA is put forward; the proposed array tracker scheme contains two realization methods. Performance analysis and simulation results are given in Section 5. Then conclusions are provided in the last section.

2. Modified Viterbi Algorithm to Bit Synchronization

Assuming code loop has been synchronized and $c_L(t_k)$ is local code, an approximate Doppler shift \hat{f}_d , Doppler rate \hat{f}_a , and phase $\hat{\theta}_0$ are obtained. They are used to construct the following model for the in-phase and quad-phase local signals [11]:

$$\begin{aligned} I_L(t_k) &= c_L(t_k) \cdot \cos(2\pi(f_{IF} + \hat{f}_d)t_k + \pi\hat{f}_a t_k^2 + \hat{\theta}_0), \\ Q_L(t_k) &= c_L(t_k) \cdot \sin(2\pi(f_{IF} + \hat{f}_d)t_k + \pi\hat{f}_a t_k^2 + \hat{\theta}_0). \end{aligned} \quad (1)$$

The I and Q components from correlators are

$$\begin{aligned} I_k &= A_k \cdot d_k \cdot \cos(\theta_k) + n_{I,k}, \\ Q_k &= A_k \cdot d_k \cdot \sin(\theta_k) + n_{Q,k}, \end{aligned} \quad (2)$$

where, d_k is the data bit associated with sample k , θ_k is carrier phase error, $n_{I,k}$ and $n_{Q,k}$ are the noise for the I and

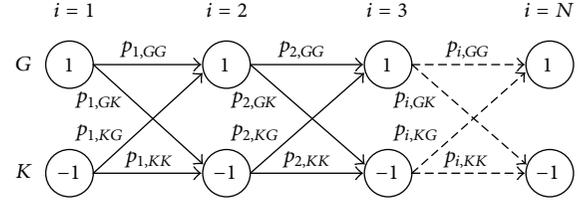


FIGURE 1: N steps state transition.

Q samples, which is assumed to be additive white Gaussian noise (AWGN), and A_k is the signal level.

Suppose there is no phase and frequency error between the local signal and received signal; loop is locked. Then, the θ_k is equal to zero in (2) and correlator output is y_k , $k = 1, \dots, 20(N + 1)$. When the sampling period is 1 millisecond (ms), $N + 1$ is in units of 20 milliseconds. y_k can be expressed as

$$y_k = A d_k + n_k. \quad (3)$$

If the 20 consecutive samples belong to the same message data bit, the sign of d_k will not change. Define the corresponding message as d_j ($j = 1, \dots, N + 1$) during the sampling period. For 20 milliseconds data length, suppose missing points is e . There would be N complete navigation message at least in $(N + 1)$ group 20 milliseconds sampling sequences due to the period of a GPS navigation message which is 20 milliseconds. Because the period of sampling is 1 millisecond, there is 20 possible data bit edges. We can get corresponding 20 continuous observation sequences and the length of each observation sequence $s_{i,\delta}$ is N . If δ indicates the bit edge position, then the corresponding observation sequence is [7]

$$\begin{aligned} s_{i,\delta} &= \frac{1}{20} \sum_{k=20(i-1)+\delta}^{20i+\delta-1} y_k \\ &= \frac{1}{20} \left(\sum_{k=20(i-1)+\delta}^{20i-e} A_k d_j \right. \\ &\quad \left. + \sum_{k=20i-e+1}^{20i+\delta-1} A_k d_{j+1} + \sum_{k=20(i-1)+\delta}^{20i+\delta-1} n_k \right), \end{aligned} \quad (4)$$

where $i = 1, \dots, N$; it represents the current integration interval and goes from 1 to N and N is the number of received data bits; y_k is the correlator samples.

Equation (4) can be written in the form

$$s_{i,\delta} = \bar{A}_{i,\delta} d_i + \bar{n}_{i,\delta}, \quad (5)$$

where the $\bar{A}_{i,\delta}$ is average level and $\bar{n}_{i,\delta}$ is noise.

Obviously, If δ is the expected data edge position, $|s_{i,\delta}|$ will reach the maximum.

For each bit edge candidate, a two-state trellis diagram has been set in Figure 1 [7]. Every diagram is constructed setting full connections between consecutive states; it means that every state will have two paths entering it and two paths leaving it. Since N is the total number of observations in

a set, each diagram presents exactly N steps. The algorithm processes them identifying the minimum weighting paths; the weights for each transition are expressed by (2)

$$p_i = s_{i,\delta} - Ad_i. \quad (6)$$

Here, A is the true signal level and d_i is the bit value associated with the state considered as destination in the current transition.

For the total $(N + 1) \cdot 20$ ms data, it contains at least N full data bits. The maximum of (5) can be represented as $\max_{\{d_i\}_{i=1}^N} \sum_{i=1}^N |s_{i,\delta}|$ [7]. However, when δ is close to the edge of the correct data bits, even on normal signal conditions, $\sum_{i=1}^N |s_{i,\delta}|$ and $\sum_{i=1}^N |s_{i,\delta \pm 1}|$ certainly are very close too. A typical example as we know is that 6 plus 7 is equal to 13 and 6 plus 8 is equal to 14. It is very close between 13 and 14. If we consider the effect of noise, the result will be worse. But 6 multiples 7 is equal to 42 and 6 multiples 8 is equal to 48; the result of "multiples" is better than "plus." Therefore, we use $\max_{\{d_i\}_{i=1}^N} \prod_{i=1}^N |s_{i,\delta}|$ instead of $\max_{\{d_i\}_{i=1}^N} \sum_{i=1}^N |s_{i,\delta}|$ or $\min_{\{d_i\}_{i=1}^N} \sum_{i=1}^N |s_{i,\delta} - Ad_i|$ as the cost function. In this way, it will have superior resistance to noise. Define

$$\Gamma_N = \prod_{i=1}^N |s_{i,\delta}|. \quad (7)$$

Obviously, the δ which makes Γ_N maximum is expected data bit edge.

When tracking loop is in bit synchronization tracking state, the sampling period is 1 millisecond. Data bit edge index (δ) range is from 1 to 20. A grid map that is shown in Figure 1 will always be achieved for each δ and a surviving path corresponding to $\max_{\{d_i\}_{i=1}^N} \prod_{i=1}^N |s_{i,\delta}|$ be generated. When the 20 δ 's candidates have been processed, we can get 20 surviving paths. In all 20 surviving paths, the path of maximum weight is the optimal path, that is, the maximum likelihood path. And the optimal path suggests a symbol sequences which is expected navigation message.

3. Nonlinear Filter Technique Improving PLL's Parameter Estimation Performance

The Kalman filter is essentially a recursive algorithm that implements a predictor corrector type estimator. The predictor is based on a system model and the corrector is based on the measurement model. The Kalman filter is optimal in the sense that it minimizes the estimated error covariance. Ping has established adaptive Kalman filtering model to track high dynamic GPS signal. It has a linear measurement model based on oscillator output. Psiaki realized weak GPS signal tracking using EKF [12–14]. But the EKF algorithm has some potential drawbacks such as Jacobi matrix calculation and linear error to solve nonlinear equation [15]. The sigma-point Kalman filters (SPKF), a family of filters based on derivativeness statistical linearization achieves higher performance than EKF for many problems and is applicable to areas where EKFs cannot be used, such as unscented Kalman filter (UKF) and

center difference Kalman filter (CDKF) [16–18]. As UKF had to select three parameters, α , β , and κ , it is very difficult in practice and UKF covariance of positive semidefinite cannot be fully guaranteed [16]. For CDKF, only one parameter h will be selected to complete filtering operation, and it can fully guarantee the positive semidefinite covariance when $h^2 \geq 1$ [17]. Therefore, CDKF is more suitable for practical application.

In this paper, the initial attempt is made for CDKF method tracking weak signal due to its precision which is equal to EKF's 2nd Taylor series expansion. An error tracking scheme is designed. The phase discriminator is substituted by CDKF in order to get better parameters estimation.

The estimated parameters are phase difference, Doppler frequency error, Doppler rate error, and amplitude. That is, $\mathbf{x} = [\theta_e, f_e, \alpha_e, A]$. An error tracking model is established as follows:

$$\begin{bmatrix} \theta_e \\ f_e \\ \alpha_e \\ A \end{bmatrix}_k = \begin{bmatrix} 1 & \Delta t & \frac{1}{2}\Delta t^2 & 0 \\ 0 & 1 & \Delta t & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \theta_e \\ f_e \\ \alpha_e \\ A \end{bmatrix}_{k-1} + \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} w_\theta \\ w_f \\ w_\alpha \\ w_A \end{bmatrix}. \quad (8)$$

Here, Δt is CDKF update time. Disturbances term is assumed to be independent; their variances can be defined. The established measurement equation with measurement noise R_k is

$$z_k = \begin{bmatrix} I_{p,k} \\ Q_{p,k} \end{bmatrix} + \begin{bmatrix} n_{I,p,k} \\ n_{Q,p,k} \end{bmatrix} = \begin{bmatrix} \bar{A}_k d_k \hat{d}_k \cos(\bar{\theta}_{e,k}) \\ \bar{A}_k d_k \hat{d}_k \sin(\bar{\theta}_{e,k}) \end{bmatrix} + R_k, \quad (9)$$

where \bar{A}_k is signal mean level during loop update and \hat{d}_k is message bit estimation; if $I_{p,k} > 0$, $\hat{d}_k = 1$, else $\hat{d}_k = -1$. Outputs of the CDKF are used to calculate the $\bar{\theta}_{e,k}$ and it will be used to control local NCO and make the phase constituted. The CDKF carrier tracking processing is as follows.

(1) *With the Initial Values.* Consider

$$\begin{aligned} \hat{\mathbf{x}}_0 &= E(\mathbf{x}_0), \\ \mathbf{P}_0 &= \text{Cov}(\mathbf{x}_0) = E(\mathbf{x}_0 - \hat{\mathbf{x}}_0)(\mathbf{x}_0 - \hat{\mathbf{x}}_0)^T. \end{aligned} \quad (10)$$

(2) *Time Updating.* According to the sigma point symmetric sampling theory, Sigma points $\xi_{i,k}$, $i = 0, 1, \dots, 2n$ are calculated by $\hat{\mathbf{x}}_k$ and \mathbf{P}_k :

$$\begin{aligned} \xi_{0,k} &= \hat{\mathbf{x}}_k, \\ \xi_{i,k} &= \hat{\mathbf{x}}_k + \left(h \sqrt{\mathbf{P}_k} \right)_i, \\ \xi_{i+n,k} &= \hat{\mathbf{x}}_k - \left(h \sqrt{\mathbf{P}_k} \right)_i, \\ & i = 1, 2, \dots, n. \end{aligned} \quad (11)$$

$\xi_{i,k}$ propagates to $\gamma_{i,k+1|k}$ through the nonlinear function $\mathbf{f}_k(\cdot)$; one can get one-step status prediction $\widehat{\mathbf{x}}_{k+1|k}$ and the error covariance matrix $\mathbf{P}_{k+1|k}$ from $\gamma_{i,k+1|k}$:

$$\begin{aligned} \gamma_{i,k+1|k} &= \mathbf{f}_k(\xi_{i,k}), \quad i = 0, 1, \dots, 2n, \\ \widehat{\mathbf{x}}_{k+1|k} &= \sum_{i=0}^{2n} W_i^m \gamma_{i,k+1|k} + \mathbf{q}_k = \sum_{i=0}^{2n} W_i^m \mathbf{f}_k(\xi_{i,k}) + \mathbf{q}_k. \\ \mathbf{P}_{k+1|k} &= \sum_{i=1}^n \left[W_i^{c1} (\gamma_{i,k+1|k} - \gamma_{i+n,k+1|k})^2 \right. \\ &\quad \left. + W_i^{c2} (\gamma_{i,k+1|k} + \gamma_{i+n,k+1|k} - 2\gamma_{0,k+1|k})^2 \right] + \mathbf{Q}_k. \end{aligned} \quad (12)$$

(3) *Measurement Updating*. Similarly, sigma points $\xi_{i,k+1|k}$ and $i = 0, 1, \dots, 2n$ are achieved using $\widehat{\mathbf{x}}_{k+1|k}$ and $\mathbf{P}_{k+1|k}$ according to the sigma point symmetric sampling theory; that is, $\xi_{i,k+1|k}$ can be written by

$$\begin{aligned} \xi_{0,k+1|k} &= \widehat{\mathbf{x}}_{k+1|k}, \\ \xi_{i,k+1|k} &= \widehat{\mathbf{x}}_{k+1|k} + \left(h \sqrt{\mathbf{P}_{k+1|k}} \right)_i, \\ \xi_{i+n,k+1|k} &= \widehat{\mathbf{x}}_{k+1|k} - \left(h \sqrt{\mathbf{P}_{k+1|k}} \right)_i, \\ &\quad i = 1, 2, \dots, n. \end{aligned} \quad (13)$$

$\xi_{i,k+1|k}$ propagates to $\chi_{i,k+1|k}$ through nonlinear function $\mathbf{h}_{k+1}(\cdot)$. The outputs prediction $\widehat{\mathbf{z}}_{k+1|k}$, autocovariance matrix $\mathbf{P}_{\widehat{\mathbf{z}}_{k+1|k}}$, and cross-covariance matrix $\mathbf{P}_{\widehat{\mathbf{z}}_{k+1|k} \widehat{\mathbf{x}}_{k+1|k}}$ are achieved using $\chi_{i,k+1|k}$:

$$\begin{aligned} \chi_{i,k+1|k} &= \mathbf{h}_{k+1}(\xi_{i,k+1|k}), \quad i = 0, 1, \dots, 2n, \\ \widehat{\mathbf{z}}_{k+1|k} &= \sum_{i=0}^{2n} W_i^m \chi_{i,k+1|k} + \mathbf{r}_{k+1} \\ &= \sum_{i=0}^{2n} W_i^m \mathbf{h}_{k+1}(\xi_{i,k+1|k}) + \mathbf{r}_{k+1}, \\ \mathbf{P}_{\widehat{\mathbf{z}}_{k+1|k}} &= \sum_{i=1}^n \left[W_i^{c1} (\chi_{i,k+1|k} - \chi_{i+n,k+1|k})^2 \right. \\ &\quad \left. + W_i^{c2} (\chi_{i,k+1|k} + \chi_{i+n,k+1|k} - 2\chi_{0,k+1|k})^2 \right] + \mathbf{R}_{k+1}, \\ \mathbf{P}_{\widehat{\mathbf{z}}_{k+1|k} \widehat{\mathbf{x}}_{k+1|k}} &= \sum_{i=1}^n \sqrt{W_i^{c1}} (\xi_{i,k+1|k} - \widehat{\mathbf{x}}_{k+1|k}) (\chi_{i,k+1|k} - \chi_{i+n,k+1|k})^T. \end{aligned} \quad (14)$$

After getting new measurement \mathbf{z}_{k+1} , the filter is updated:

$$\begin{aligned} \widehat{\mathbf{x}}_{k+1} &= \widehat{\mathbf{x}}_{k+1|k} + \mathbf{K}_{k+1} (\mathbf{z}_{k+1} - \widehat{\mathbf{z}}_{k+1|k}), \\ \mathbf{K}_{k+1} &= \mathbf{P}_{\widehat{\mathbf{z}}_{k+1|k} \widehat{\mathbf{x}}_{k+1|k}} \mathbf{P}_{\widehat{\mathbf{z}}_{k+1|k}}^{-1}, \\ \mathbf{P}_{k+1} &= \mathbf{P}_{k+1|k} - \mathbf{K}_{k+1} \mathbf{P}_{\widehat{\mathbf{z}}_{k+1|k}} \mathbf{K}_{k+1}^T. \end{aligned} \quad (15)$$

Here, \mathbf{K}_{k+1} is a gain matrix.

4. Array Nonlinear Filtering Tracker

The bit synchronization model starts working after the PLL has been locked by pulling frequency and phase for a typical receiver. If a receiver is powered on under weak signal environment, it should extend the loop integral time right. However, unknown navigation data bit edge makes it impossible to execute this strategy. The reason is that input signal is just very weak before the loop is locked. The typical 1 millisecond integration is not suitable for tracking loop since its processing gain is not enough for very weak signal. A practical program which can accurately identify the bit edge position before the PLL achieves lock state is needed.

In above circumstances, in order to make the receiver still work well, we combine designed bit synchronization module with CDKF signal tracking module, so that a bit synchronization tracking is realized for very weak signals. In this scheme, the CDKF module removes the phase and frequency errors on the state transition path which is in the bit synchronization module. The bit synchronization tracking parallel processing scheme (scheme I) is shown in Figure 2. If the first group of data delay 0 ms, the finally group of data will delay 19 milliseconds considering that the number of bit edge candidate is 20.

The bit synchronization in the red box refers to Section 2. As input signal is very weak, when the receiver is in tracking state, the sample period of input signal is 1 millisecond and CDKF loop update time is 20 ms. Every observation sequence which is used by bit synchronization comes from independent CDKF tracking loop outputs. Updated CDKF removes the phase error of observations before performance bit synchronization algorithm. Because the number of bit edge candidate is 20, 20 groups of CDKF and bit synchronization modules are designed to quickly complete bit synchronization tracking. After the bit synchronization, the integration time no longer crosses the data bit edge. The CDKF loop which has the maximum weight continues to work with 20 milliseconds integrated time. Other modules are to stop working. Tracking loop converts to single CDKF weak signal tracking status. Be noted that the above process has no complex process control.

Sometimes, for saving hardware resources, we can also use a serial manner (scheme II) that is realized by using a group of bit synchronization tracking module in Figure 2. That is, only one red box in Figure 2 will be realized for scheme II. The designed process of scheme II is as follows.

- (i) Deal with current 1 millisecond data; one can get an optimal weight and the path recording this information.
- (ii) Extract next 1 millisecond data and go into (i).
- (iii) Repeat (i) and (ii) 19 times; 20 optimal weights will be achieved.
- (iv) Find the maximum from the 20 optimal weights and extract the delay time (T milliseconds) corresponding to the maximum.

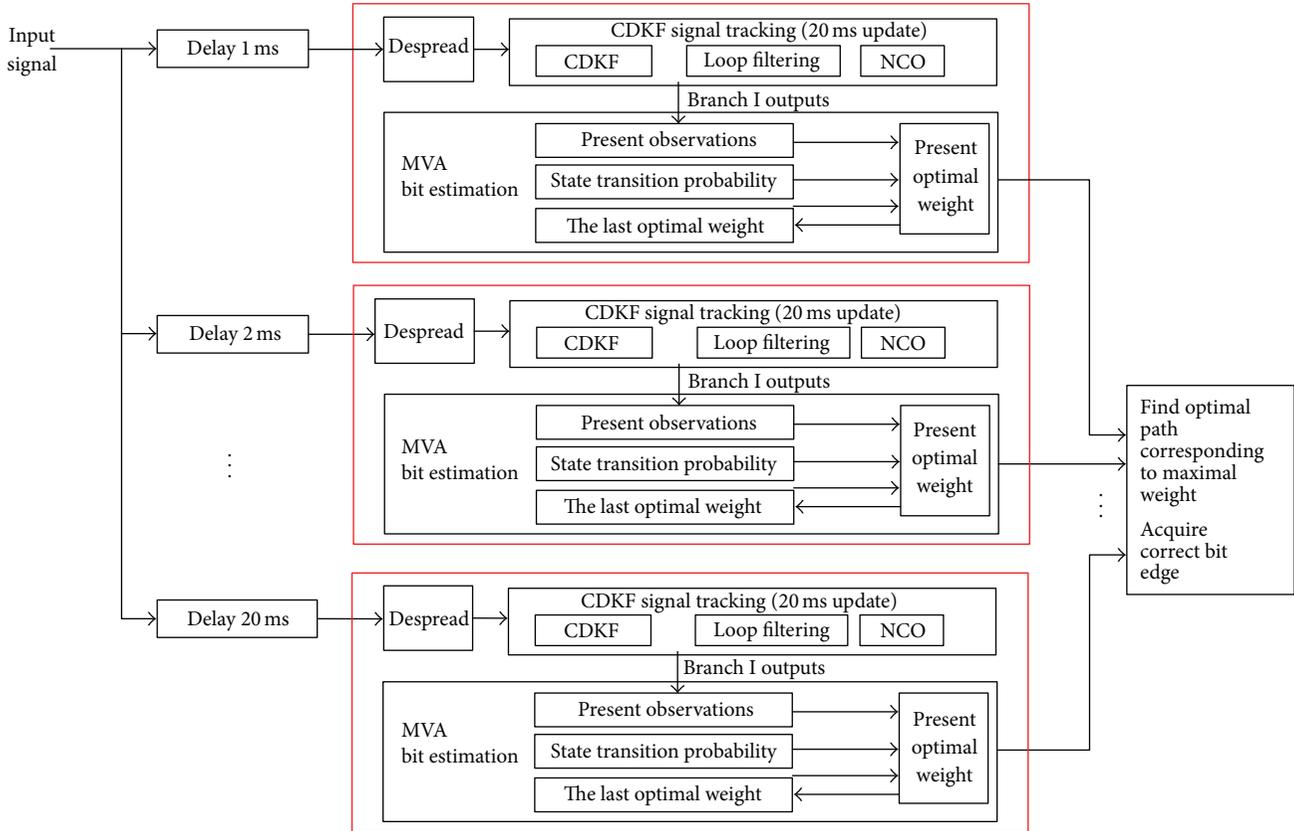


FIGURE 2: Parallel bit synchronized tracking scheme (scheme I).

(v) Delay new input data T milliseconds; it is the right bit edge position. Bit synchronization tracking is completed. The tracking loop uses single CDKF to work.

Obviously, scheme II (serial mode) is less than scheme I (parallel mode) in terms of hard resource consumption. However scheme II has more complex control and its bit synchronization tracking total time is 20 times more than the parallel mode.

Results of scheme II are shown in Figure 6. It is easier to implement using software than hardware. In terms of complexity, CDKF and optimal path dynamic programming are implemented by recursive algorithm and the quantity of calculation is very small. Further, once the current optimal weights are determined, it is not necessary that the used observations are saved on a large memory. It is very important to practice project.

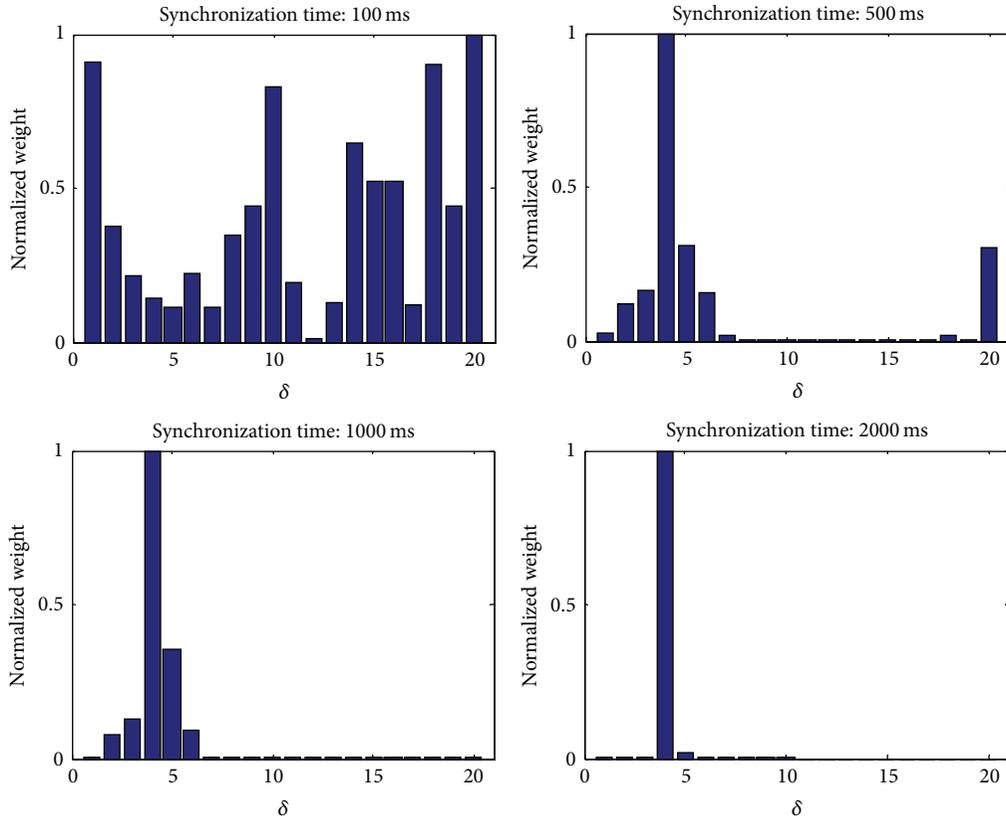
5. Test Results and Analysis

For comprehensive assessment of the proposed weak signal tracking scheme, the following three experiments were set up to determine bit synchronization time, bit synchronization performance assessment, phase and frequency estimation error, and demodulation error rate; the implementation complexity was analyzed at last. Taken together, these results

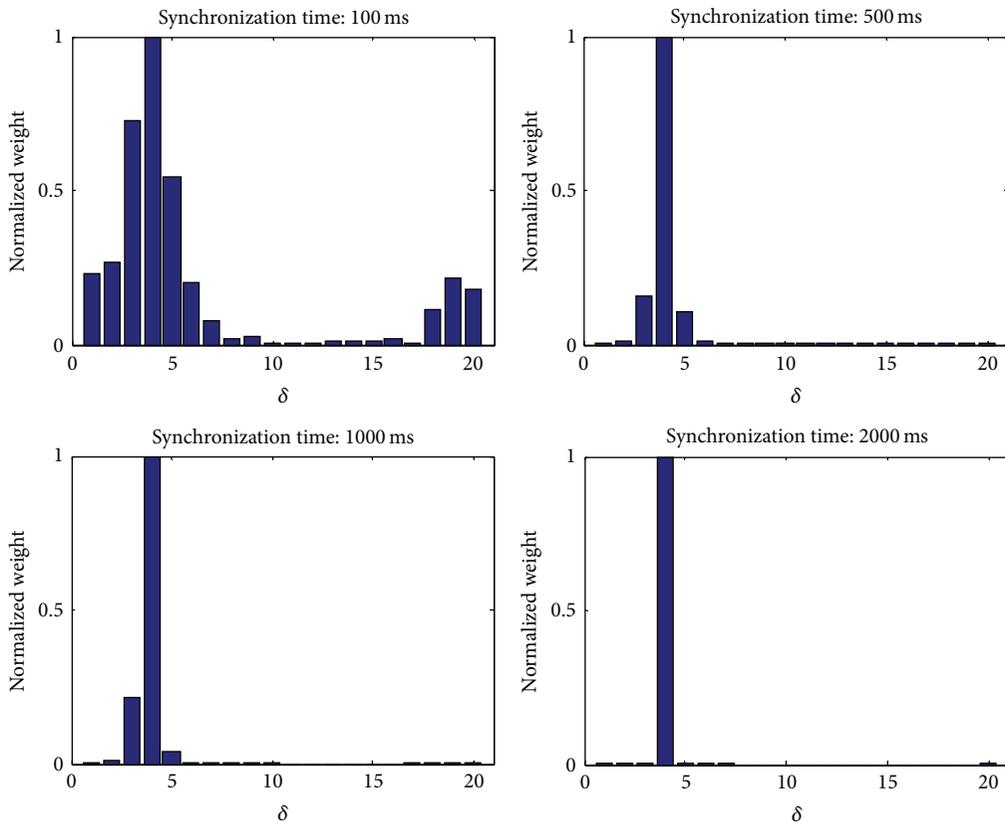
suggest that proposed method is very suitable for weak signal tracking. Particularly the bit edge is unknown when a receiver is powered on under very weak signal environment.

5.1. Determining Bit Synchronization Time. Firstly, assess suitable bit synchronization time. CNR (carrier-to-noise ratio) or C/N_0 of tested signal is set to 15 dB-Hz, 20 dB-Hz, 25 dB-Hz, and 30 dB-Hz. It is difficult to demodulate when using 1 millisecond integration time for very weak tested signal. A reconstructed output signal was generated and its phase and frequency error has been removed. Tested signal center frequency is 4.092 MHz, sampling frequency is 12.276 MHz, and data bit state obeys binomial distribution; that is, the probability of each state transition is 0.5 and bit synchronization time is set to 100 milliseconds, 500 milliseconds, 1000 milliseconds, and 2000 milliseconds, respectively. The tested results are shown in Figure 3.

As a result of Figure 3, the longer the synchronized time or the more the bits that contain the full number (N), the more accurate the bit synchronization, the shorter the bit synchronization time, and the more closely the optimal path weight values which are between other candidate bit edges and the expected bit edge. Obviously, the results are credible. When bit synchronization time arises to 2 seconds, it can also detect the right bit edge even for 15 dB-Hz signal and the weights may be quite different between the expected

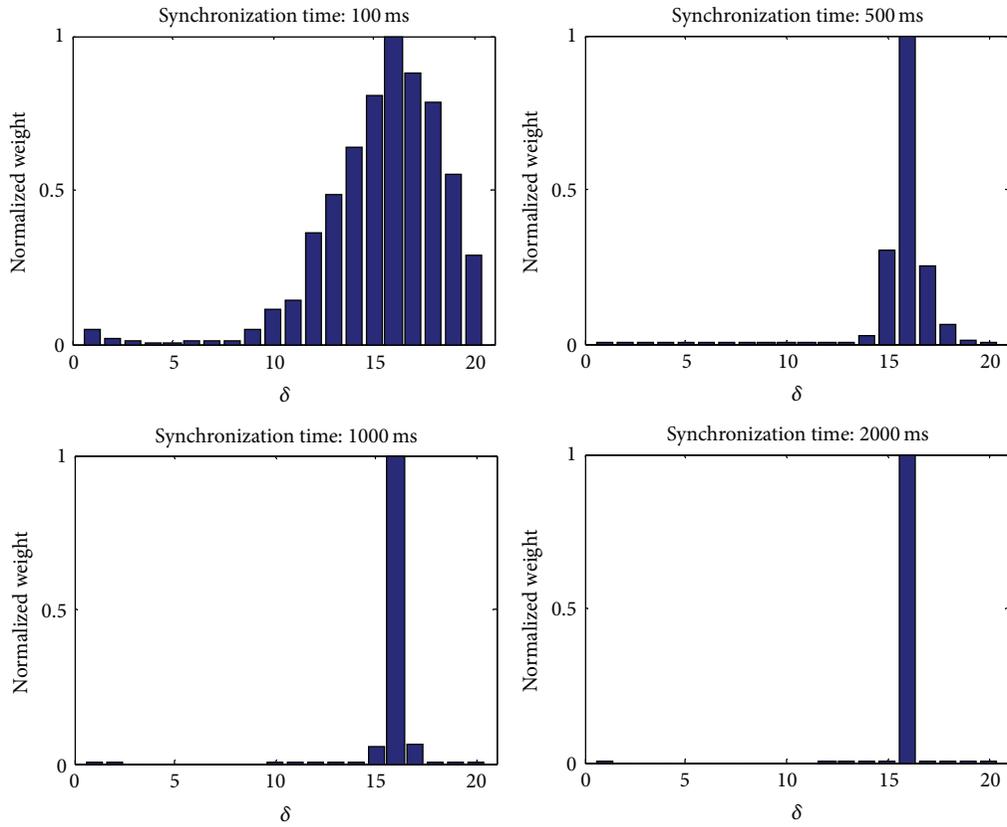


(a) $C/N_0 = 15$ dB-Hz, true bit edge position: $\delta = 4$

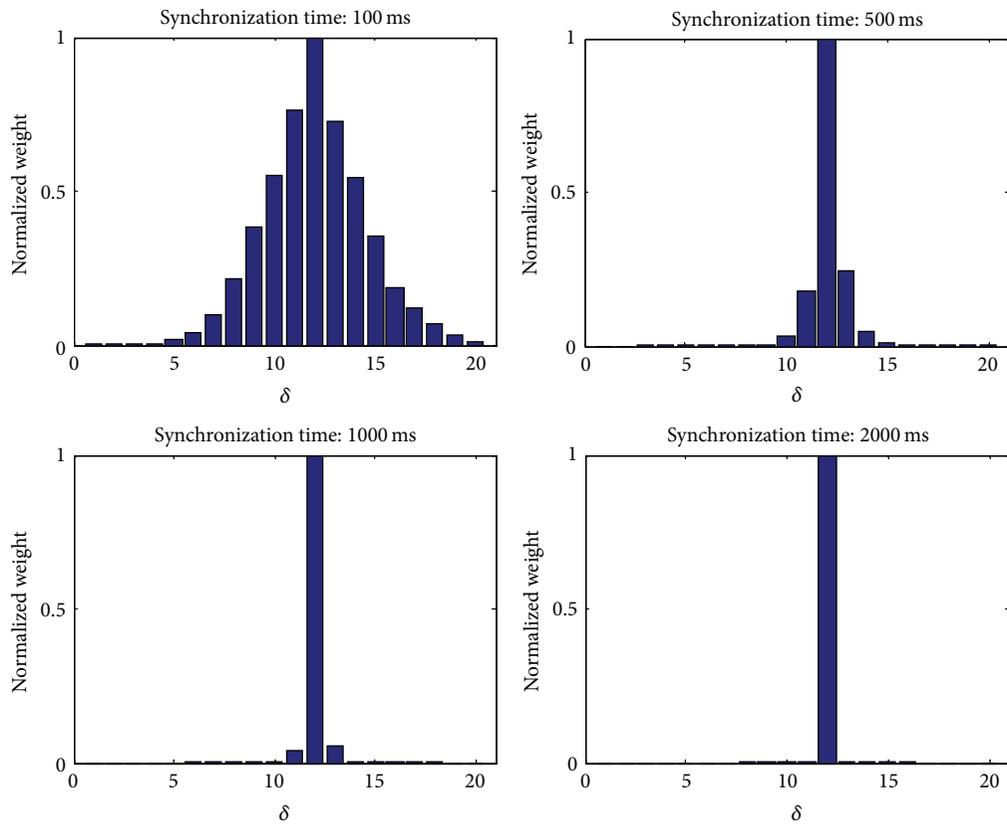


(b) $C/N_0 = 20$ dB-Hz, true bit edge position: $\delta = 4$

FIGURE 3: Continued.



(c) $C/N_0 = 25$ dB-Hz, true bit edge position: $\delta = 16$



(d) $C/N_0 = 30$ dB-Hz, true bit edge position: $\delta = 12$

FIGURE 3: Bit synchronization time test.

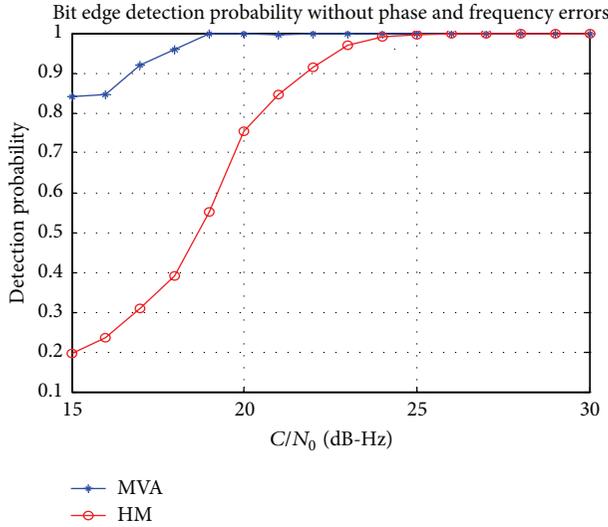


FIGURE 4: Without phase and frequency errors.

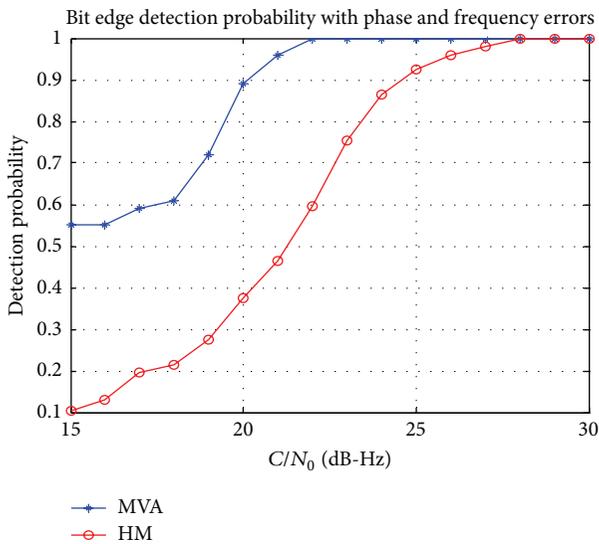


FIGURE 5: With phase and frequency errors.

edge position and the adjacent position. However, from the point of view of the weights distribution, the weak signal synchronization model has no guarantee of a higher bit EDR in very short time, for example, 100 milliseconds. Certainly, 2 seconds have been also short enough. More importantly, we should pay attention to the bit energy difference between the true bit edge position and its near position. The difference has been very significant using more than 500 milliseconds of bit synchronization time. 2 seconds of bit synchronization time will be adopted in the following assessment according to these test results.

5.2. Bit Synchronization Performance Evaluation. In order to compare detection performance between the MVA and

the traditional histogram method, we test the two methods and get different data bit edge detection probability to different CNR. The tested CNR range is from 15 to 30 dB-Hz, bit synchronization time is 2 seconds according to the previous results, and other settings in MVA are the same with the previous experiment; $NBS_1 = 50$ and $NBS_2 = 45$ in histogram method. Repeat 200 times for every CNR for the two methods. The results are shown in Figure 4, which has removed the phase and frequency errors. The test results suggest that the power of the input signal gradually increases with CNR; performance of both two methods can be continuously improved. As far as detection probability is concerned, the MVA is superior to the histogram method. Even for 20 dB-Hz, the MVA can also achieve 100% of the detection probability, but histogram method has to increase bit synchronization time to achieve better detection probability.

Meanwhile, the bit edge detection results are shown in Figure 5 in the presence of phase and frequency error, where the phase error and frequency error are normal distribution; the range of phase error is from 0 to 15 degrees and range of frequency error is from 0 to 5 Hz. The results suggest that the bit edge detection performance has been declined compared with no phase and frequency errors. This decline would be relieved by increasing bit synchronization time.

5.3. Signal Tracking Performance Evaluation. Usually, carrier is in low dynamic movement on weak signal environment, such as downtown and indoor. The phase and frequency errors estimation history and the first 75 demodulated data bits using CDKF loop with bit synchronization module for 20 dB-Hz and 28 dB-Hz weak GPS signal are demonstrated in Figure 6. As can be seen from Figure 6, when the loop has not been locked, the conventional method, without bit synchronization aided, cannot determine the right edge of the data bit edge position. The new tracking scheme using a bit synchronization module can work well when the C/N_0 is 20 dB-Hz; the frequency and phase errors are allowed within a certain range and navigation message has been correctly demodulated. In fact, the signal quality drops to 20 dB-Hz and the tracking loop is not always kept locked. The undesired result is due to the signal power which is too low.

The loop locked probability to different CNR using 100 times Monte Carlo tests is shown in Figure 7. The data suggest that the new tracking method retains lower loss locked probability during weak signal tracking.

5.4. Complexity Analysis. The computational complexity of CDKF is very low due to its recursion. A bit synchronization tracking module is able to achieve an optimal path when the state transition occurs every time. The current optimal weights can be obtained using four status transition probabilities and the optimal path weights which are the previous results; only four optimal path weights state transition probabilities and the previous weights are needed to be saved; that is, there are only five numbers of data that need to be stored; after obtaining the current optimal weights, the current observations will be discarded to relieve storage

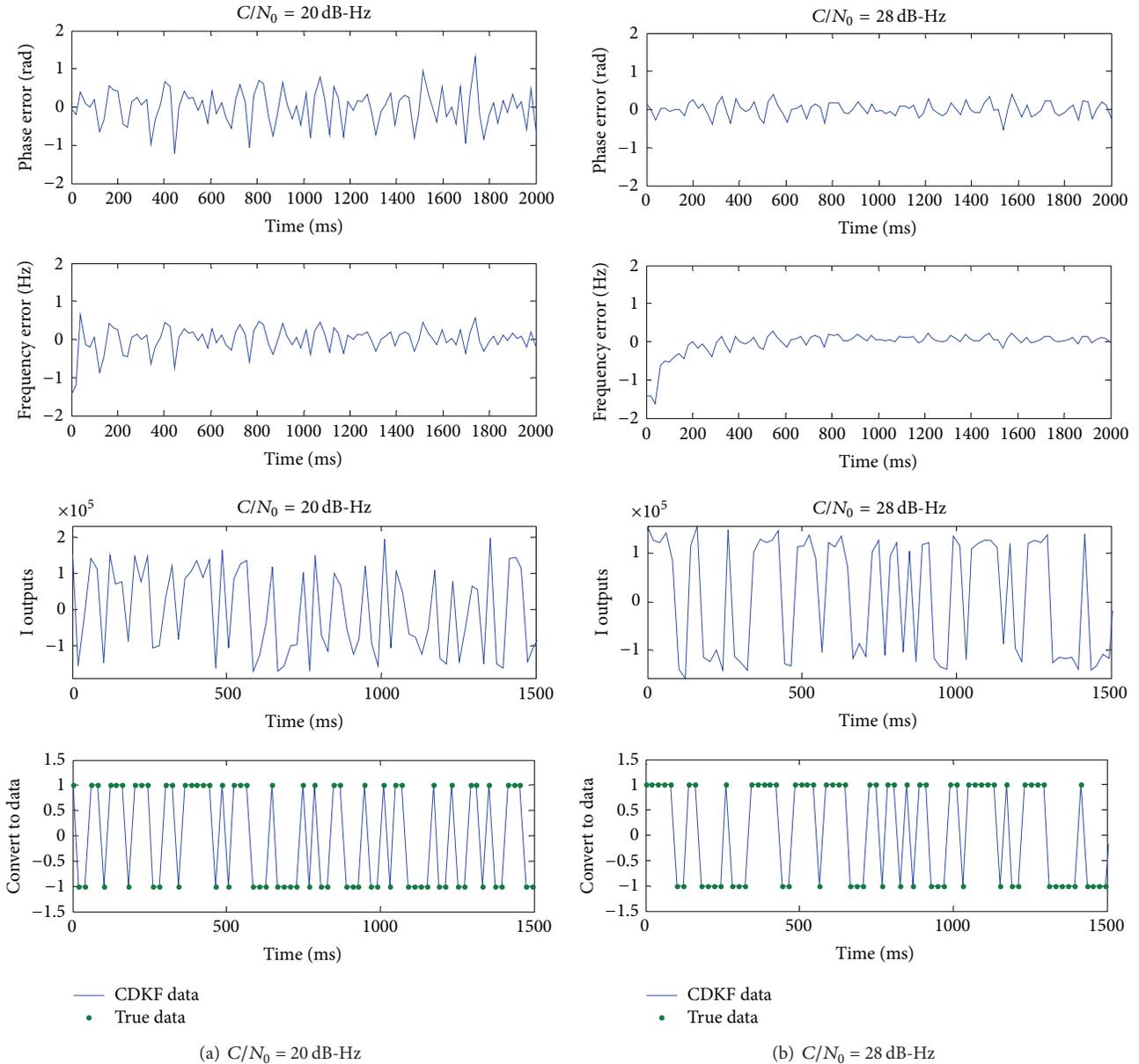


FIGURE 6: Estimated phase, frequency error history, and the first 75 data bits.

space. If the start time of a receiver needs to be improved, save the survivor path information corresponding to the optimal weights for navigation solution. Certainly, an additional store space is needed to temporarily store $21 \times 20 = 420$ number of data.

6. Summary

The paper endeavors to solute the problem of signal tracking when a receiver is powered on under weak signal environments. The loop will stay in the unlocked state and cannot find the data edge to implement weak signal tracking using

traditional 1 millisecond update loop approach. A tracking scheme based on CDKF has been designed to avoid the complexity of traditional EKF algorithm in computing non-linear equation. The MVA based on optimal path dynamic programming which comes from HMM has been developed and some better results have been obtained. Finally, serial and parallel weak signal tracking scheme combining CDKF estimation technology with MVA are designed and implemented. A suitable scheme would be adopted due to practical requirements. Test results and analysis confirmed the rationality and resource consumption advantages; furthermore, some acceptable tracking results proved 20 dB-Hz

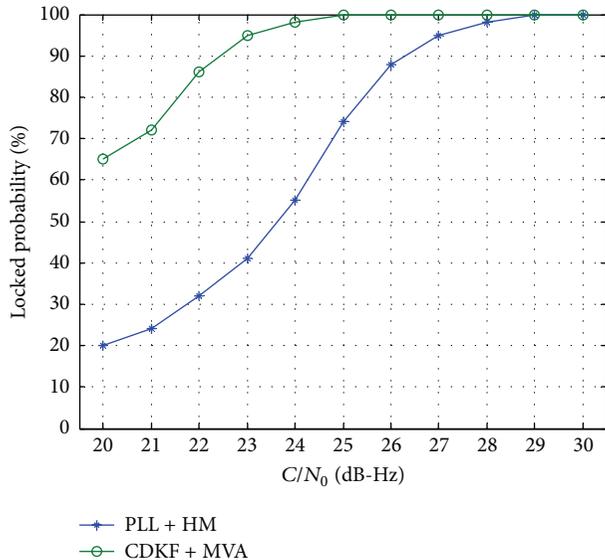


FIGURE 7: The loop stability to different CNR.

weak GPS signal can be processed. The phase and frequency estimation accuracy for weak GPS signal tracking will be evaluated in following research work.

Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

Acknowledgments

The financial support from the National Science Foundation of China under Grant (nos. 61304234 and 61273081) and the Fundamental Research Funds for the Central Universities (HEUCFX041403 and HEUCFR1114) is gratefully acknowledged.

References

- [1] L. Wang, D. P. Groves, and M. K. Ziebart, "GNSS shadow matching: improving urban positioning accuracy using a 3D city model with optimized visibility prediction scoring," in *Proceedings of the 25th International Technical Meeting of the Satellite Division of the Institute of Navigation (ION GNSS '12)*, pp. 423–437, Nashville, Tenn, USA, 2012.
- [2] S. Singh, "Comparison of assisted GPS (AGPS) performance using simulator and field tests," in *Proceedings of the ION GNSS Conference*, pp. 1–12, Fort Worth, Tex, USA, 2006.
- [3] S. Schön and O. Bielenberg, "On the capability of high sensitivity GPS for precise indoor positioning," in *Proceedings of the 5th Workshop on Positioning, Navigation and Communication (WPNC '08)*, pp. 121–127, March 2008.
- [4] B. Y. T. James, *Fundamentals of Global Positioning System Receivers: A Software Approach*, John Wiley & Sons, New York, NY, USA, 2nd edition, 2004.
- [5] B. Parkinson and J. Spilker, *Global Positioning System: Theory and Applications*, vol. 1, American Institute of Aeronautics and Astronautics, 1996.
- [6] M. Kokkonen and S. Pietilä, "A new bit synchronization method for a GPS receiver," in *Proceedings of the IEEE Position Location and Navigation Symposium*, pp. 85–90, April 2002.
- [7] N. I. Ziedan and J. L. Garrison, *Bit Synchronization and Doppler Frequency Removal at Very Low Carrier to Noise Ratio Using a Combination of the Viterbi Algorithm with an Extended Kalman Filter*, ION GPS/GNSS, Portland, Ore, USA, 2003.
- [8] S. Jeon, C. Kim, G. Kim et al., "Optimal signal tracking algorithm for GNSS signal using moving set-point LQG system," *International Journal of Control, Automation and Systems*, vol. 11, no. 6, pp. 1214–1222, 2013.
- [9] M. Anghiler, T. Pany, J. Won et al., "An algorithm for bit synchronization and signal tracking in software GNSS receivers," in *Proceedings of the ION GNSS 19th International Technical Meeting of the Satellite Division*, pp. 1836–1848, 2006.
- [10] O. Macchi and L. L. Scharf, "Dynamic programming algorithm for phase estimation and data decoding on random phase channels," *IEEE Transactions on Information Theory*, vol. 27, no. 5, pp. 581–595, 1981.
- [11] G. J. Gao, *INS-assisted high sensitivity GPS receivers for degraded signal navigation [Ph.D. thesis]*, Department of Geomatics Engineering, University of Calgary, Calgary, Canada, 2007.
- [12] M. L. Psiaki and H. Jung, "Extended Kalman filter methods for tracking weak GPS signal," in *Proceedings of the ION GPS 15th International Technical Meeting of the Satellite Division of The Institute of Navigation*, pp. 2539–2553, Portland, Ore, USA, 2002.
- [13] N. I. Ziedan and J. L. Garrison, "Extended Kalman filter-based tracking of weak GPS signals under high dynamic conditions," in *Proceedings of the 17th International Technical Meeting of the Satellite Division of the Institute of Navigation (ION GNSS '04)*, pp. 20–31, Long Beach, Calif, USA, September 2004.
- [14] A. Razavi, D. Gebre-Egziabher, and D. M. Akos, "Carrier loop architectures for tracking weak GPS signals," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 44, no. 2, pp. 697–710, 2008.
- [15] J. H. Won, T. Pany, and B. Eissfeller, "Characteristics of kalman filters for GNSS signal tracking loop," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 48, no. 4, pp. 3671–3681, 2012.
- [16] S. Julier, J. Uhlmann, and H. F. Durrant-Whyte, "A new method for the nonlinear transformation of means and covariances in filters and estimators," *IEEE Transactions on Automatic Control*, vol. 45, no. 3, pp. 477–482, 2000.
- [17] M. Nørgaard, N. K. Poulsen, and O. Ravn, "New developments in state estimation for nonlinear systems," *Automatica*, vol. 36, no. 11, pp. 1627–1638, 2000.
- [18] R. van der Merwe and E. A. Wan, "The square-root unscented Kalman filter for state and parameter-estimation," in *Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing*, pp. 3461–3464, May 2001.

Research Article

A Novel Method for Target Navigation and Mapping Based on Laser Ranging and MEMS/GPS Navigation

Jianhua Cheng,^{1,2} Rene Landry Jr.,² Daidai Chen,¹ and Dongxue Guan¹

¹ Marine Navigation Research Institute, College of Automation, Harbin Engineering University, Harbin 150001, China

² LASSENA Laboratory, Ecole de Technologie Supérieure, 1100 Notre-Dame Street West, Montreal, QC, Canada H3C 1K3

Correspondence should be addressed to Jianhua Cheng; ins.cheng@163.com

Received 16 March 2014; Accepted 9 June 2014; Published 29 June 2014

Academic Editor: Zheping Yan

Copyright © 2014 Jianhua Cheng et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Making the sensor rigidly mounted in the target is the common characteristic of conventional navigation system. However, it is difficult or impossible to realize that for special applications such as the positioning of hostile aircraft. A novel new algorithm for target navigation and mapping is designed based on the position, attitude, and ranging information provided by laser distance detector (electronic distance measuring, LDS) and MEMS/GPS navigation, which can solve problem of the target navigation and mapping without any sensor in the target. The detailed error analysis shows that attitude error of MEMS/GPS is the main error source which dominated the accuracy of the algorithm. Based on the error analysis, a calibration algorithm is designed so as to improve the accuracy to a large extent. The result shows that, by using this new algorithm, the performance of target positioning can be efficiently improved, and the positioning error is less than 2 meters for the target within 1 kilometer range.

1. Introduction

Motion state of the target can be described by its accelerometer, velocity, position, and attitude. Navigation system is the most convenient device to provide these information by means of optical navigation, electronics navigation, dynamics navigation, acoustic navigation, and so on [1]. For the mentioned navigation system, there is a common characteristic that the navigation sensor must be mounted in the target. For example, there must be optical observation device in the target which can measure the attitude vector from the target to another known point in terrestrial navigation and celestial navigation [2]. MEMS inertial navigation system must rigidly mount the gyros and accelerometers to the target so as to measure its linear and angular movement. In communication domain, the target position method based on the cell site is also realized by communication between the cell phone and the base station [3].

Sometimes, conventional navigation system cannot act the right function in special application. For the positioning of those hard to or unable to reach places, it is difficult or

expensive to mount the sensor in the target. For the positing of the enemy target in military, it is impossible [4, 5].

Optical measure device in mapping is a good choice for the target mapping because it is unnecessary to mount the sensor in the target. On the basis of the known self-position, the target position can be measured accurately by means of angle measurement [6]. For example, the total station can output precise slope distance from the instrument to a particular point by a theodolite integrated with an electronic distance meter (EDM). In a total station, the theodolite is used to measure angles in the horizontal and vertical planes, the distance between the instrument and the point is measured by the EDM. But for the mapping and navigation of the moving targets, optical measure device is unsuitable because it can only work in a relative static state [7].

MEMS/GPS integrated navigation system has the advantages of small size, light weight, and low cost; it can be applied in many navigation fields such as unmanned aircrafts, land vehicles, and robots [8]. If the MEMS/GPS navigation system is allowed to substitute the optical measure device, the framework of the MEMS gyros constituted can be used to describe

of $\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3$, the vector of \mathbf{oT} can also be described in $ox_t y_t z_t$. That is,

$$\mathbf{oT} = x_{oT}^t \mathbf{e}_1 + y_{oT}^t \mathbf{e}_2 + z_{oT}^t \mathbf{e}_3, \quad (4)$$

where $x_{oT}^t, y_{oT}^t, z_{oT}^t$ is the coordinate difference between T and o in $ox_t y_t z_t$.

In order to calculate the vector of \mathbf{oT} , assume a new vector of \mathbf{oC} in $ox_t y_t z_t$:

$$\mathbf{oC} = 0\mathbf{e}_1 + d\mathbf{e}_2 + 0\mathbf{e}_3, \quad (5)$$

where d is the distance between o and T .

From (5), the vector of \mathbf{oC} has, in fact, the same direction as oy_t axis. Its mold is equivalent to the ranging between o and T .

For the transform between \mathbf{oC} and \mathbf{oT} , two rotations of H and ϕ can accomplish the mission, because the final rotation of θ will not change the direction of \mathbf{oT} . In Figure 2, it can be seen that the oy_t axis has the same direction as the oy_b axis.

The rotation of H and ϕ can be described by the following matrix transform. The first rotation transform matrix of H angle is

$$\mathbf{C}_t^1 = \begin{bmatrix} \cos H & \sin H & 0 \\ -\sin H & \cos H & 0 \\ 0 & 0 & 1 \end{bmatrix}. \quad (6)$$

The second rotation transform matrix of ϕ angle is

$$\mathbf{C}_1^2 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos \phi & \sin \phi \\ 0 & -\sin \phi & \cos \phi \end{bmatrix}. \quad (7)$$

According to the definition of \mathbf{oT} and \mathbf{oC} , the vector of \mathbf{oT} can be expressed based on the former transform matrix:

$$\mathbf{oT} = \mathbf{C}_1^2 \mathbf{C}_t^1 \mathbf{oC}. \quad (8)$$

Involving (6) and (7) into (8), the final calculation of \mathbf{oT} can be got as follows:

$$\begin{aligned} \begin{bmatrix} x_{oT}^t \\ y_{oT}^t \\ z_{oT}^t \end{bmatrix} &= \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos \phi & \sin \phi \\ 0 & -\sin \phi & \cos \phi \end{bmatrix} \cdot \begin{bmatrix} \cos H & \sin H & 0 \\ -\sin H & \cos H & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 0 \\ d \\ 0 \end{bmatrix} \\ &= \begin{bmatrix} d \sin H \\ d \cos \phi \cos H \\ -d \sin \phi \cos H \end{bmatrix}. \end{aligned} \quad (9)$$

For the vector of \mathbf{oT} , its description in (2) and (4) is the same vector. On the other hand, the unit vector of $\mathbf{i}_1, \mathbf{i}_2, \mathbf{i}_3$ and the unit vector of $\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3$ have the following relationship:

$$\begin{bmatrix} \mathbf{i}_1 \\ \mathbf{i}_2 \\ \mathbf{i}_3 \end{bmatrix} = \mathbf{C}_e^i \begin{bmatrix} \mathbf{e}_1 \\ \mathbf{e}_2 \\ \mathbf{e}_3 \end{bmatrix}. \quad (10)$$

According to the relationship of (10), the coordinate difference between T and o in $ox_e y_e z_e$ can be got:

$$\begin{bmatrix} x_{oT}^e \\ y_{oT}^e \\ z_{oT}^e \end{bmatrix} = \mathbf{C}_t^e \begin{bmatrix} x_{oT}^t \\ y_{oT}^t \\ z_{oT}^t \end{bmatrix}. \quad (11)$$

Thus, the position of the target (x_T^e, y_T^e, z_T^e) can be got based on (3) and (11):

$$\begin{bmatrix} x_T^e \\ y_T^e \\ z_T^e \end{bmatrix} = \begin{bmatrix} x_A^e \\ y_A^e \\ z_A^e \end{bmatrix} + \mathbf{C}_t^e \begin{bmatrix} d \sin H \\ d \cos \phi \cos H \\ -d \sin \phi \cos H \end{bmatrix}. \quad (12)$$

All the variables in (12) are ideal. In order to calculate the target position, the position of o , the distance between o and T , and the orientation of \mathbf{oT} must be known by suitable measurement devices.

If the MEMS/GPS can be rigidly mounted in the observer along with the body coordinate system, the position of o can be calculated from the position output of MEMS/GPS and the orientation of \mathbf{oT} can be described by the attitude of MEMS/GPS. On the other hand, LDS can measure the distance if it is mounted along with the oy_b axis of the body coordinate system accurately.

Thus, define the new variables of the measurement:

- (1) latitude, longitude, and height measurement output of MEMS/GPS: $\tilde{\varphi}_o, \tilde{\lambda}_o, \tilde{h}_o$;
- (2) heading, pitch, and roll measurement output of MEMS/GPS: $\tilde{H}, \tilde{\phi}, \tilde{\theta}$;
- (3) ranging measurement output of LDS: \tilde{d} .

After getting the measurement information of MEMS/GPS and LDS, the Cartesian coordinates of observer o can be calculated according to the following equation [10]:

$$\begin{aligned} \tilde{x}_A^e &= (R_N + \tilde{h}_A) \cos \tilde{\varphi}_A \cos \tilde{\lambda}_A \\ \tilde{y}_A^e &= (R_N + \tilde{h}_A) \cos \tilde{\varphi}_A \sin \tilde{\lambda}_A \\ \tilde{z}_A^e &= [R_N (1 - e^2) + \tilde{h}_A] \sin \tilde{\varphi}_A, \end{aligned} \quad (13)$$

where $R_N = a / \sqrt{1 - e^2 \sin^2 \tilde{\varphi}_o}$, a is the long axle of the Earth, and e is the eccentricity of the Earth.

By substituting the ideal variables with the measurement result, the final of Cartesian coordinates of the target can be got:

$$\begin{bmatrix} \tilde{x}_T^e \\ \tilde{y}_T^e \\ \tilde{z}_T^e \end{bmatrix} = \begin{bmatrix} \tilde{x}_A^e \\ \tilde{y}_A^e \\ \tilde{z}_A^e \end{bmatrix} + \begin{bmatrix} \cos \tilde{\lambda}_A & -\sin \tilde{\lambda}_A \cos \tilde{\varphi}_A & \sin \tilde{\lambda}_A \cos \tilde{\varphi}_A \\ \sin \tilde{\lambda}_A & \cos \tilde{\lambda}_A \cos \tilde{\varphi}_A & -\cos \tilde{\lambda}_A \sin \tilde{\varphi}_A \\ 0 & \sin \tilde{\varphi}_A & \cos \tilde{\varphi}_A \end{bmatrix} \cdot \begin{bmatrix} \tilde{d} \sin \tilde{H} \\ \tilde{d} \cos \tilde{\phi} \cos \tilde{H} \\ -\tilde{d} \sin \tilde{\phi} \cos \tilde{H} \end{bmatrix}. \quad (14)$$

2.4. *Functions for the Target Mapping and Navigation.* Because MEMS/GPS and LDS can realize real-time measurement, the new algorithm can be used for both the static and the moving target. Besides the target position, the new algorithm can also measure many other pieces of information including longitude, latitude, height of the target, the slant range from one point to point, and the height difference and horizontal ranging between two targets. These pieces of information can all be calculated based on the position information of the target.

For the slant range (\tilde{D}) from one point to another point, \tilde{D} can be calculated by the following equation:

$$\tilde{D} = \sqrt{(\tilde{x}_{T_2}^e - \tilde{x}_{T_1}^e)^2 + (\tilde{y}_{T_2}^e - \tilde{y}_{T_1}^e)^2 + (\tilde{z}_{T_2}^e - \tilde{z}_{T_1}^e)^2}, \quad (15)$$

where $(\tilde{x}_{T_1}^e, \tilde{y}_{T_1}^e, \tilde{z}_{T_1}^e)$ and $(\tilde{x}_{T_2}^e, \tilde{y}_{T_2}^e, \tilde{z}_{T_2}^e)$ are, respectively, the measurement result of the Cartesian coordinates for targets T_1 and T_2 .

For the height difference ($\Delta \tilde{H}$) between the two targets,

$$\Delta \tilde{H} = \left| \frac{\sqrt{(\tilde{x}_{T_2}^e)^2 + (\tilde{y}_{T_2}^e)^2}}{\cos \tilde{\varphi}_{T_2}} - \frac{\sqrt{(\tilde{x}_{T_1}^e)^2 + (\tilde{y}_{T_1}^e)^2}}{\cos \tilde{\varphi}_{T_1}} \right|. \quad (16)$$

3. Error Analysis for the Algorithm

As mentioned in the former introduction, MEMS/GPS has a relatively low accuracy of attitude. For example, NAV440 GPS-aided MEMS inertial system has a 0.4° (RMS) accuracy of pitch and roll and a 1.0° (RMS) accuracy of heading for all motion [11]. ADIS16480 MEMS/GPS has an accuracy of 0.1° (pitch, roll) and 0.3° (heading) for static state [12]. Besides the attitude error, the positioning error of MEMS/GPS probably has a severe impact on the position calculation of the target. For example, Omni STAR HP of Trimble has accuracy of minor to 10 centimeters [13]. On the other hand, LDS also has measurement error. For example, the accuracy of AccuRange AR500 Laser Sensor is generally specified with a linearity of about $\pm 0.15\%$ of the range [14].

3.1. *Positioning Error Caused by the Positioning Error of MEMS/GPS.* The positioning error of MEMS/GPS will cause two errors which are tightly concerned with the calculation of the target. Those are the position error of point o and the cosine transform matrix of \tilde{C}_e^t . Based on (12) and (14), the target positioning error caused by the positioning error

MEMS/GPS can be calculated with the neglect of other error resources:

$$\begin{aligned} \Delta_1 &= \Delta_{11} + \Delta_{12} = \begin{bmatrix} \Delta x_o^e \\ \Delta y_o^e \\ \Delta z_o^e \end{bmatrix} + \Delta C_e^t \begin{bmatrix} d \sin H \\ d \cos \phi \cos H \\ -d \sin \phi \cos H \end{bmatrix} \\ &= \begin{bmatrix} \Delta x_o^e \\ \Delta y_o^e \\ \Delta z_o^e \end{bmatrix} + \begin{bmatrix} \Delta c_{11} & \Delta c_{12} & \Delta c_{13} \\ \Delta c_{21} & \Delta c_{22} & \Delta c_{23} \\ \Delta c_{31} & \Delta c_{32} & \Delta c_{33} \end{bmatrix} \begin{bmatrix} d \sin H \\ d \cos \phi \cos H \\ -d \sin \phi \cos H \end{bmatrix}, \end{aligned} \quad (17)$$

where $\Delta x_o^e = \tilde{x}_o^e - x_o^e$, $\Delta y_o^e = \tilde{y}_o^e - y_o^e$, $\Delta z_o^e = \tilde{z}_o^e - z_o^e$, $\Delta C_e^t = \tilde{C}_e^t - C_e^t$.

Comparing with the item of Δ_{11} , another item of Δ_{12} can be neglected because the error of the transform matrix is very small. For example, the first line of ΔC_e^t has the following characteristic:

$$\begin{aligned} \Delta c_{11} &= -\sin \lambda_o \Delta \lambda_o \leq |\Delta \lambda_o| \\ \Delta c_{12} &= \sin \lambda_o \sin \varphi_o \Delta \varphi_A - \cos \lambda_o \cos \varphi_o \Delta \lambda_o \\ &\leq |\Delta \varphi_o| + |\Delta \lambda_o| \\ \Delta c_{13} &= \sin \lambda_o \cos \varphi_o \Delta \varphi_o + \cos \lambda_o \sin \varphi_o \Delta \lambda_o \\ &\leq |\Delta \varphi_o| + |\Delta \lambda_o|. \end{aligned} \quad (18)$$

Then,

$$\begin{aligned} \Delta c_{11} d \sin H + \Delta c_{12} d \cos \phi \cos H + \Delta c_{13} (-d \sin \phi \cos H) \\ \leq 3d |\Delta \lambda_o| + 2d |\Delta \varphi_o|. \end{aligned} \quad (19)$$

The following result can be got by using the same method:

$$\begin{aligned} \Delta c_{21} d \sin H + \Delta c_{22} d \cos \phi \cos H + \Delta c_{23} (-d \sin \phi \cos H) \\ \leq 3d |\Delta \lambda_o| + 2d |\Delta \varphi_o| \\ \Delta c_{31} d \sin H + \Delta c_{32} d \cos \phi \cos H + \Delta c_{33} (-d \sin \phi \cos H) \\ \leq 2d |\Delta \varphi_o|. \end{aligned} \quad (20)$$

For the distance which is minor to 1 kilometer, the relationship between d and radius of the Earth will satisfy

$$d \ll R, \quad (21)$$

where R is the radius of the Earth.

Then,

$$\begin{aligned} d^2 [(\Delta\varphi_o)^2 + (\Delta\lambda_o)^2] \\ \ll R^2 [(\Delta\varphi_o)^2 + (\Delta\lambda_o)^2] \approx [(\Delta x_o^e)^2 + (\Delta y_o^e)^2 + (\Delta z_o^e)^2]. \end{aligned} \quad (22)$$

That is to say,

$$|\Delta_{12}| \ll |\Delta_{11}|. \quad (23)$$

Thus,

$$\Delta P_1 = |\Delta_1| \approx \sqrt{(\Delta x_o^e)^2 + (\Delta y_o^e)^2 + (\Delta z_o^e)^2}, \quad (24)$$

where ΔP_1 is the positioning error caused by the positioning error of MEMS/GPS.

3.2. Positioning Error Caused by the Error of LDS. Based on (12) and (14), the target positioning error caused by the error of LDS can be calculated with the neglect of other error resources:

$$\Delta_2 = \mathbf{C}_t^e \begin{bmatrix} (\tilde{d} - d) \sin H \\ (\tilde{d} - d) \cos \phi \cos H \\ -(\tilde{d} - d) \sin \phi \cos H \end{bmatrix} = \mathbf{C}_t^e \begin{bmatrix} \Delta d \sin H \\ \Delta d \cos \phi \cos H \\ -\Delta d \sin \phi \cos H \end{bmatrix}. \quad (25)$$

Thus,

$$\Delta P_2 = |\Delta_2| = |\Delta d|. \quad (26)$$

3.3. Positioning Error Caused by the Attitude Error of MEMS/GPS. The target positioning error caused by the attitude error of MEMS/GPS can also be calculated with the neglect of other error resources:

$$\begin{aligned} \Delta_3 &= \mathbf{C}_t^e \begin{bmatrix} d(\sin \tilde{H} - \sin H) \\ d(\cos \tilde{\phi} \cos \tilde{H} - \cos \phi \cos H) \\ -d(\sin \tilde{\phi} \cos \tilde{H} - \sin \phi \cos H) \end{bmatrix} \\ &= \mathbf{C}_t^e \begin{bmatrix} d[\sin(H + \Delta H) - \sin H] \\ d[\cos(\phi + \Delta\phi) \cos(H + \Delta H) - \cos \phi \cos H] \\ -d[\sin(\phi + \Delta\phi) \cos(H + \Delta H) - \sin \phi \cos H] \end{bmatrix}. \end{aligned} \quad (27)$$

Because $\Delta\phi$ and ΔH are all small error angles, then

$$\begin{aligned} \sin \Delta\phi &\approx \Delta\phi & \cos \Delta H &\approx 1 & \Delta\phi \cdot \Delta H &\ll \Delta\phi \\ \cos \Delta\phi &\approx 1 & \sin \Delta H &\approx \Delta H & \Delta\phi \cdot \Delta H &\ll \Delta H. \end{aligned} \quad (28)$$

Based on the condition of (28), the error can be got as follows:

$$\Delta_3 = \mathbf{C}_t^e \begin{bmatrix} d\Delta H \cos H \\ -d\Delta H \cos \phi \cos H - d\Delta\phi \sin \phi \cos H \\ d\Delta H \sin \phi \sin H - d\Delta\phi \cos \phi \cos H \end{bmatrix} \quad (29)$$

$$\begin{aligned} \Delta P_3 &= |\Delta_3| \\ &= d \left((\Delta H)^2 (\cos^2 H + \cos^2 \phi - \cos 2\phi \sin^2 H) \right. \\ &\quad \left. + (\Delta\phi)^2 \cos^2 H + \Delta H \Delta\phi \sin 2\phi \cos H \right. \\ &\quad \left. \times (\cos H - \sin H) \right)^{1/2}. \end{aligned} \quad (30)$$

3.4. Simulation for the Positioning Error of the Target. Assume that the initial position of the observer is east longitude 126° and north latitude 45°. The longitude and latitude error of MEMS/GPS is 0.2 m and the heading and pitch error of MEMS/GPS are, respectively, 0.2° and 0.1° [12]. LDS has a linearity error of 0.15% of the range [14]. Corresponding to three conditions, three simulation results are given in Figure 3. By the way, the heading and pitch are in swaying state in order to follow a moving target: $H = 30^\circ + 7^\circ \sin(2\pi * t/7)$ and $\phi = 2^\circ + 1^\circ \sin(2\pi * t/7)$.

Condition 1. Only the position error of MEMS/GPS is involved.

Condition 2. The position error of MEMS/GPS and the linearity error of LDS are involved.

Condition 3. All errors of MEMS/GPS and LDS are involved.

As shown in Figure 3(a), the positioning error of target T which is caused by the positioning error of MEMS/GPS is mainly concerned with Δ_{11} , because the calculation of $\sqrt{0.2^2 + 0.2^2}$ is 0.2828, and the error is almost independent of the parameter d . The positioning error of the target in Figure 3(b) will reach 1.7 m when the parameter d is about 1 kilometer. The main reason is the linear error of LDS besides the error of 0.28 m stimulated by condition 1. It also can be seen that the positioning error of the target added quickly after the attitude error is involved in Figure 3(c). In the meantime, the error caused by ΔH and $\Delta\phi$ is linear to the parameter d .

4. Calibration for the Algorithm

For navigation and mapping with higher accuracy demand, the positioning algorithm probably cannot meet the demands because of the measure error. Former simulation has explained the reason which is because of errors of MEMS/GPS and LDS. Among all errors, attitude error is the most important because it dominates the accuracy of the target position. That is to say, the positioning error of point o and the ranging error of d can be neglected when the attitude error of MEMS/GPS is bigger than 0.1°, especially when the distance d is minor to 500 m.

Besides the attitude error mentioned, the attitude accuracy of MEMS/GPS will decline because the gyro scale factor

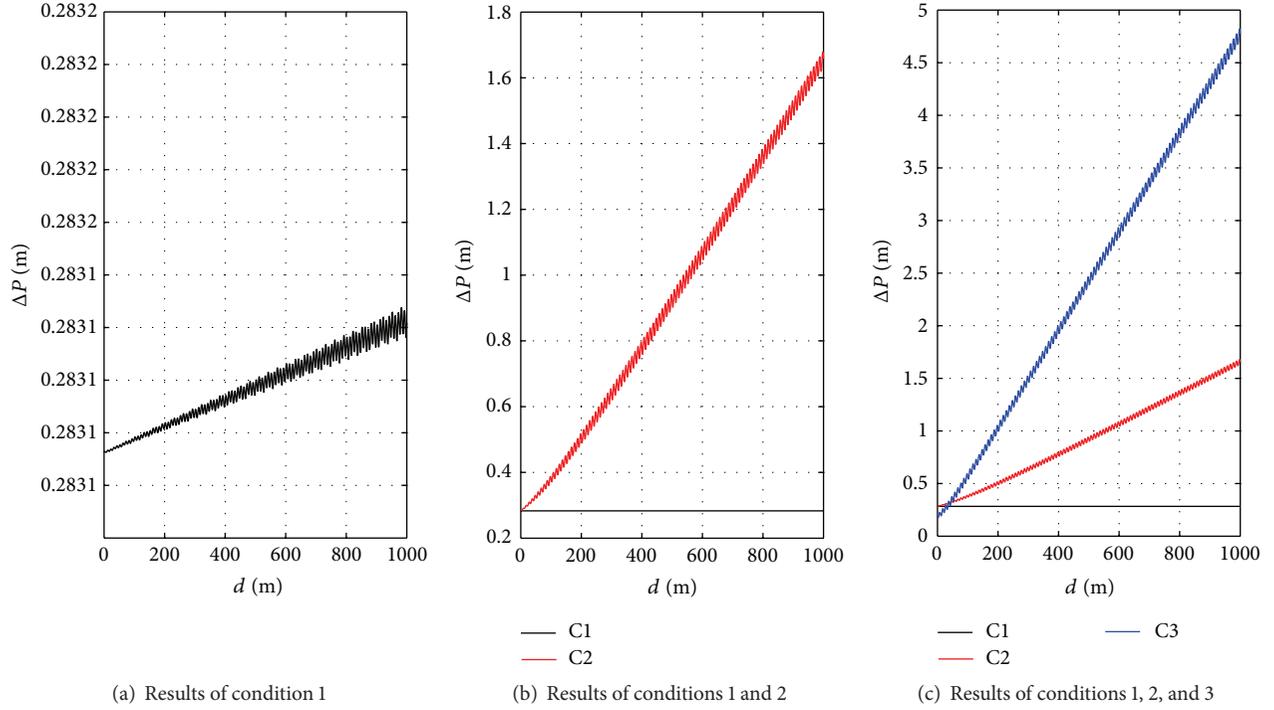


FIGURE 3: Simulation of error analysis caused by MEMS/GPS and LDS.

and the misalignment parameters will change with time [15]. These parameters can only be calibrated in laboratory. Thus, a new calibration method for the system should be designed to improve and maintain the performance of the system.

4.1. Description for the Known Reference Point. With the condition of another point (R) nearby the observer, its position can be got based on the former algorithm. That is,

$$\begin{bmatrix} \tilde{x}_R^e \\ \tilde{y}_R^e \\ \tilde{z}_R^e \end{bmatrix} = \begin{bmatrix} \tilde{x}_o^e \\ \tilde{y}_o^e \\ \tilde{z}_o^e \end{bmatrix} + \tilde{\mathbf{C}}_t^e \begin{bmatrix} \tilde{d}_{oR} \sin \tilde{H} \\ \tilde{d}_{oR} \cos \tilde{\phi} \cos \tilde{H} \\ -\tilde{d}_{oR} \sin \tilde{\phi} \cos \tilde{H} \end{bmatrix}. \quad (31)$$

Because error of Δx_o^e , Δy_o^e , Δz_o^e , and Δd can be neglected when the parameter d is minor to 500 m, The following equation can be got:

$$\begin{bmatrix} \tilde{x}_R^e \\ \tilde{y}_R^e \\ \tilde{z}_R^e \end{bmatrix} = \begin{bmatrix} x_o^e \\ y_o^e \\ z_o^e \end{bmatrix} + \mathbf{C}_t^e \begin{bmatrix} d_{oR} \sin \tilde{H} \\ d_{oR} \cos \tilde{\phi} \cos \tilde{H} \\ -d_{oR} \sin \tilde{\phi} \cos \tilde{H} \end{bmatrix}. \quad (32)$$

On the other hand, the ideal position of R can be described based on (12):

$$\begin{bmatrix} x_R^e \\ y_R^e \\ z_R^e \end{bmatrix} = \begin{bmatrix} x_o^e \\ y_o^e \\ z_o^e \end{bmatrix} + \mathbf{C}_t^e \begin{bmatrix} d_{oR} \sin H \\ d_{oR} \cos \phi \cos H \\ -d_{oR} \sin \phi \cos H \end{bmatrix}. \quad (33)$$

4.2. Attitude Calculation Based on the Known Reference Point.

If the position of the reference can be got in advance by means of MEMS/GPS, the subtraction of (32) and (33) can construct a new measurement variable which can be used to calculate the parameter H and ϕ . That is,

$$\begin{aligned} \begin{bmatrix} \Delta x_R^e \\ \Delta y_R^e \\ \Delta z_R^e \end{bmatrix} &= \begin{bmatrix} \tilde{x}_R^e - x_R^e \\ \tilde{y}_R^e - y_R^e \\ \tilde{z}_R^e - z_R^e \end{bmatrix} \\ &= \mathbf{C}_t^e \begin{bmatrix} d_{oR} (\sin \tilde{H} - \sin H) \\ d_{oR} (\cos \tilde{\phi} \cos \tilde{H} - \cos \phi \cos H) \\ -d_{oR} (\sin \tilde{\phi} \cos \tilde{H} - \sin \phi \cos H) \end{bmatrix}. \end{aligned} \quad (34)$$

In (34), most variables can be measured by MEMS/GPS and LDS except the ideal attitude angle of ϕ and H . That is to say, angle of ϕ and H can be calculated so as to replace $\tilde{\phi}$ and \tilde{H} . But (34) is the type of transcendental equation which is very difficult to solve. In order to get the ideal attitude, (34) should be converted into linear equation which can make the solution easy.

The parameters $\tilde{\phi}$ and \tilde{H} can be substituted as

$$\begin{aligned} \phi &= \tilde{\phi} - \Delta\phi \\ H &= \tilde{H} - \Delta H. \end{aligned} \quad (35)$$

Involving (35) into (34) and using the same simplification as (28), the transcendental equation can be converted into linear equation:

$$\begin{bmatrix} \Delta x_R^e \\ \Delta y_R^e \\ \Delta z_R^e \end{bmatrix} = \mathbf{C}_t^e \begin{bmatrix} d_{oR} \cos \tilde{H} \Delta H \\ -d_{oR} \sin \tilde{\phi} \cos \tilde{H} \Delta \phi - d_{oR} \cos \tilde{\phi} \sin \tilde{H} \Delta H \\ -d_{oR} \cos \tilde{\phi} \cos \tilde{H} \Delta \phi + d_{oR} \sin \tilde{\phi} \sin \tilde{H} \Delta H \end{bmatrix}. \quad (36)$$

Equation (36) can also be described as

$$\begin{bmatrix} \Delta x_R^e \\ \Delta y_R^e \\ \Delta z_R^e \end{bmatrix} = \mathbf{C}_t^e \begin{bmatrix} 0 & d_{oR} \cos \tilde{H} \\ -d_{oR} \sin \tilde{\phi} \cos \tilde{H} & -d_{oR} \cos \tilde{\phi} \sin \tilde{H} \\ -d_{oR} \cos \tilde{\phi} \cos \tilde{H} & d_{oR} \sin \tilde{\phi} \sin \tilde{H} \end{bmatrix} \begin{bmatrix} \Delta \phi \\ \Delta H \end{bmatrix} \\ = \mathbf{C}_t^e \mathbf{C}_1 \begin{bmatrix} \Delta \phi \\ \Delta H \end{bmatrix}. \quad (37)$$

Three equations are redundant for the solution of two unknown variables. Assuming $\mathbf{C} = \mathbf{C}_t^e \mathbf{C}_1$ and involving \mathbf{C}^T , $\Delta \phi$ and ΔH can be estimated by the least-square method [16]:

$$\begin{bmatrix} \Delta \hat{\phi} \\ \Delta \hat{H} \end{bmatrix} = (\mathbf{C}^T \cdot \mathbf{C})^{-1} \mathbf{C}^T \begin{bmatrix} \Delta x_R^e \\ \Delta y_R^e \\ \Delta z_R^e \end{bmatrix}. \quad (38)$$

Then, $\hat{\phi} = \tilde{\phi} - \Delta \hat{\phi}$ and $\hat{H} = \tilde{H} - \Delta \hat{H}$ can substitute $\tilde{\phi}$ and \tilde{H} . The calibration can improve the attitude of the MEMS/GPS, so as to get the target position with higher accuracy.

5. Simulation for the Calibration and the Target Positioning

5.1. Simulation for the Calibration. Performance of the calibration is very important for the positioning accuracy of the target. In order to verify the performance of the calibration, simulation results are provided by giving various conditions. During the simulation, all errors of MEMS/GPS and LDS are the same as the former. For the reference point, there is a random error of 0.05 m which is influenced by measurement. Three conditions are given during the simulation.

Condition 1. Only the attitude error of MEMS/GPS and random error of R are involved: $\Delta H = 0.2^\circ$, $\Delta \phi = 0.1^\circ$, and $d_{oR} = 200$ m.

Condition 2. All errors of MEMS/GPS, LDS, and R are involved: $\Delta H = 0.2^\circ$, $\Delta \phi = 0.1^\circ$, and $d_{oR} = 200$ m.

Condition 3. All errors of MEMS/GPS, LDS, and R are involved: $\Delta H = 0.2^\circ$, $\Delta \phi = 0.1^\circ$, and $d_{oR} = 400$ m.

Condition 4. Only the attitude error of MEMS/GPS and random error of R are involved, $\Delta H = 0.2^\circ \sin(2\pi t/50)$, $\Delta \phi = 0.1^\circ \sin(2\pi t/50)$, and $d_{oR} = 200$ m.

Condition 5. All errors of MEMS/GPS, LDS, and R are involved: $\Delta H = 0.2^\circ \sin(2\pi t/50)$, $\Delta \phi = 0.1^\circ \sin(2\pi t/50)$, and $d_{oR} = 200$ m.

Condition 6. All errors of MEMS/GPS, LDS, and R are involved: $\Delta H = 0.2^\circ \sin(2\pi t/50)$, $\Delta \phi = 0.1^\circ \sin(2\pi t/50)$, and $d_{oR} = 400$ m.

As shown in Figure 4(a), the calibration algorithm can estimate the angle of ΔH and $\Delta \phi$, the residual errors of ΔH and $\Delta \phi$ are all minor to 0.02° ; this is a very high accuracy of attitude for the target navigation and mapping. After involving errors of MEMS/GPS and LDS, the accuracy of the calibration decreased because of the error sources. The residual error of ΔH is about 0.07° , and the residual error of $\Delta \phi$ is about 0.01° in Figure 4(b). But with the increasing of d_{oR} , the residual error of ΔH will decrease to 0.04° and that of $\Delta \phi$ to 0.005° . That is to say, the performance of the calibration will increase if there is a relatively long distance of d_{oR} . If the type of the angle error is changed, the calibration has the same performance. The comparison of ideal and estimated heading error in Figures 4(d), 4(e), and 4(f) can demonstrate it.

5.2. Simulation for the Target Positioning after Calibration. Former simulations demonstrate that the calibration algorithm can estimate most attitude errors of MEMS/GPS. This will be very helpful for the accuracy improvement of the target position. The following simulation gave the final error analysis for the target position after the attitude error of MEMS/GPS is compensated. In order to compare the positioning performance to the uncalibrated algorithm, all errors of MEMS/GPS are the same as the former during the simulation except the error source demanded in the condition.

Condition 1. Only the attitude error of MEMS/GPS and random error of R are involved: $\Delta H = 0.2^\circ$ and $\Delta \phi = 0.1^\circ$, $d_{oR} = 400$ m.

Condition 2. All errors of MEMS/GPS, LDS, and R are involved: $\Delta H = 0.2^\circ$, $\Delta \phi = 0.1^\circ$, and $d_{oR} = 400$ m.

Condition 3. All errors of MEMS/GPS, LDS, and R are involved: $\Delta H = 0.2^\circ$, $\Delta \phi = 0.1^\circ$, and $d_{oR} = 200$ m.

As shown in Figures 5(a) and 5(b), the positioning error has been decreased to a large extent after the attitude error of MEMS/GPS is compensated (curve of Cal). For more information in Figures 5(a) and 5(b), $\Delta \hat{\phi}$ and $\Delta \hat{H}$ are in fact the equivalent estimation of attitude error which will be influenced by the position error of o and ranging error of d . Because the positioning error of the target is also decreased when the distance is zero, this means that the position error of MEMS/GPS is partly compensated by the calibration algorithm. On the other hand, error of condition 2 is bigger than error of condition 3 because there is more residual error when

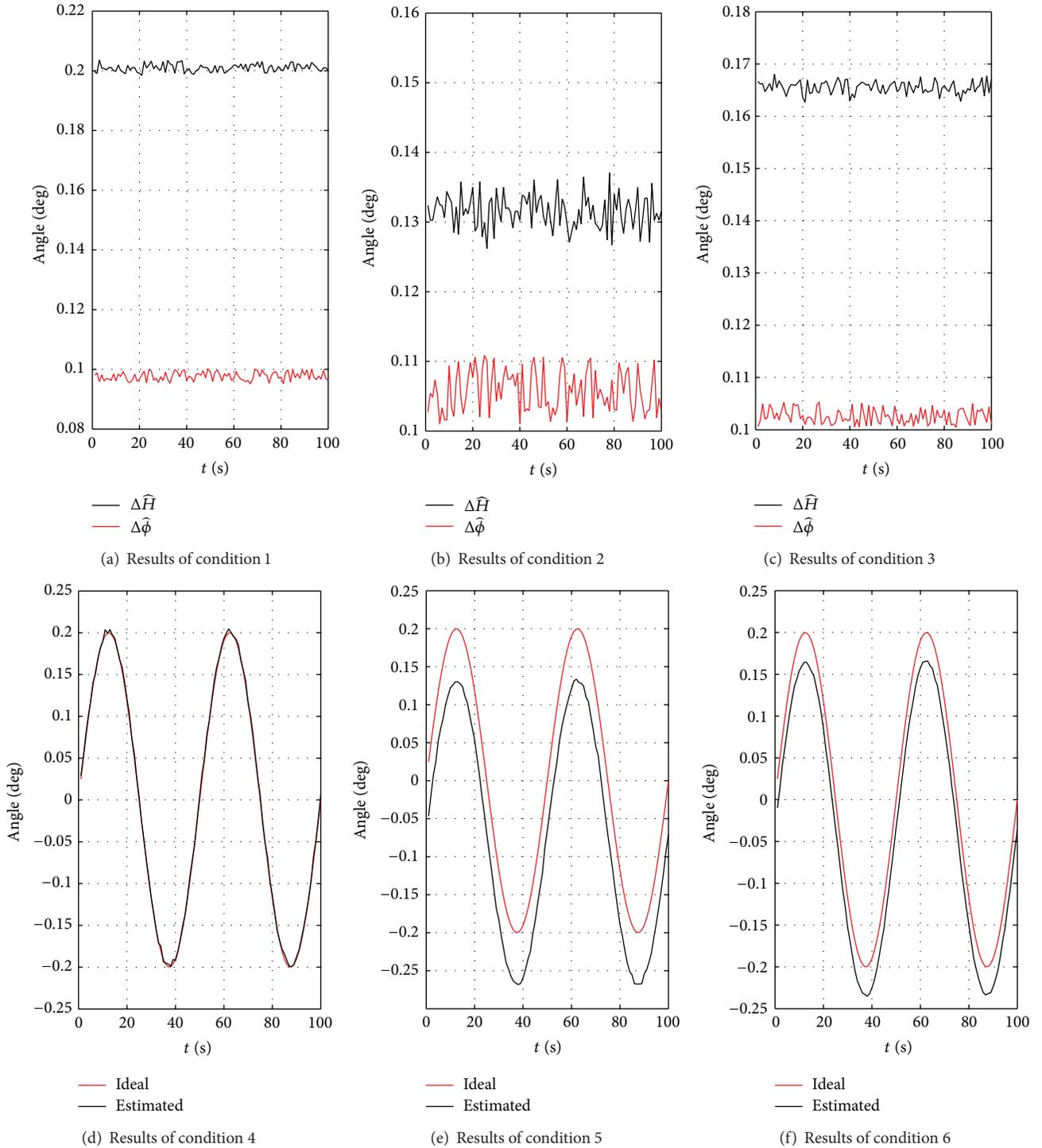


FIGURE 4: Performance verification for the calibration algorithm.

$d_{OR} = 200$ m. This is perfectly matched with the simulation result of Figure 4.

6. Conclusions

This study has developed a new method for target navigation and mapping. The results indicate that the proposed method

can measure position, height, height difference, slant range, and horizontal range for static and moving target. After the calibration of MEMS/GPS, the algorithm can realize precise positioning with the error minor to 2 meters for the static and moving target within 1 kilometer range. The stable and superior performance make the new algorithm very suitable for the target mapping and navigation on land.

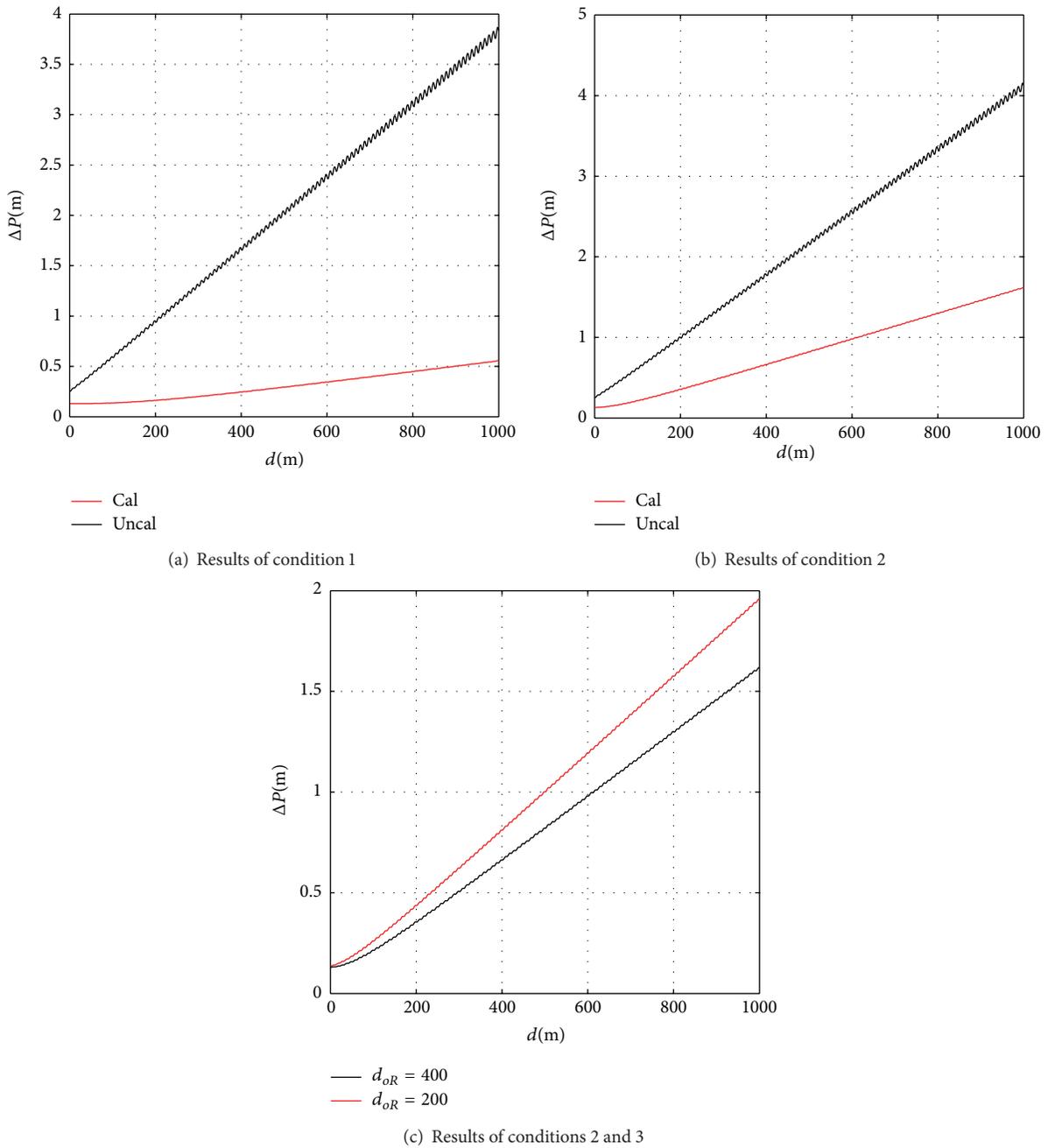


FIGURE 5: Performance verification for the calibration algorithm.

Conflict of Interests

The authors declare that they have no conflict of interests regarding the publication of this paper.

Acknowledgments

Funding for this work was provided by the National Nature Science Foundation of China under Grant no. 61374007 and no. 61104036 and the Fundamental Research Funds for the Central Universities under the Grant of HEUCFX41309. The

authors would like to thank all the editors and anonymous reviewers for improving this paper.

References

- [1] M. Liu and H. Wang, "Research on agriculture survey and evaluation UAV navigation system," *INMATEH-Agricultural Engineering*, vol. 41, no. 3, pp. 31–38, 2013.
- [2] A. R. Underwood and M. L. Thein, "A comparative study of celestial navigation-based controller designs for extraterrestrial surface navigation," in *Proceedings of the AIAA Guidance, Navigation, and Control Conference (GNC '13)*, pp. 1–21, Boston, Mass, USA, August 2013.

- [3] M. Wang, S. D. Schrock, N. Vander Broek, and T. Mulinazzi, "Estimating dynamic origin-destination data and travel demand using cell phone network data," *International Journal of Intelligent Transportation Systems Research*, vol. 11, no. 2, pp. 76–86, 2013.
- [4] M. P. Naylor, E. M. Atkins, and S. Roderick, "Visual target recognition and tracking for autonomous manipulation tasks," in *Proceedings of the AIAA Guidance, Navigation, and Control Conference and Exhibit*, pp. 194–209, Hilton Head, SC, USA, August 2007.
- [5] Q. Q. Wen, Y. Y. Cui, Q. L. Xia, and D. W. Liu, "Firing table modifying based on acquisition ground zone of seeker for laser guidance bomb," *Infrared and Laser Engineering*, vol. 41, no. 8, pp. 2053–2057, 2012.
- [6] Z. Lušić, "The use of horizontal and vertical angles in terrestrial navigation," *Transactions on Maritime Science*, vol. 2, no. 1, pp. 5–14, 2013.
- [7] F. Keller and H. Sternberg, "Multi-sensor platform for indoor mobile mapping: system calibration and using a total station for indoor applications," *Remote Sensing*, vol. 5, no. 11, pp. 5805–5824, 2013.
- [8] R. Jaffe, B. Scott, and M. Asad, "MEMS GPS/INS for micro air vehicle application," in *Proceedings of the 17th International Technical Meeting of the Satellite Division of the Institute of Navigation (ION GNSS '04)*, pp. 819–824, September 2004.
- [9] X. J. Niu, N. El-Sheimy, and S. Nassar, "An accurate land-vehicle MEMS IMU/GPS navigation system using 3D auxiliary velocity updates," *Journal of Navigation*, vol. 54, no. 3, pp. 177–188, 2007.
- [10] L. Zhao, J. H. Cheng, and Y. X. Zhao, *Marine Navigation System*, Press of National Defense, Washington, Wash , USA, 2011.
- [11] http://pdf.directindustry.com/pdf/crossbow/nav440-gps-aided-mems-inertial-system/19299-72230-_2.html.
- [12] http://www.analog.com/static/imported-files/data_sheets/ADIS16480.pdf.
- [13] http://trl.trimble.com/docushare/dsweb/Get/Document-572430/022503-864_Correction_Services_Bro_0611_LR_pbp.pdf.
- [14] http://www.disensors.com/downloads/products/AR500%20Laser%20Manual_1350.pdf.
- [15] G. A. Aydemir and A. Saranlı, "Characterization and calibration of MEMS inertial sensors for state and parameter estimation applications," *Measurement: Journal of the International Measurement Confederation*, vol. 45, no. 5, pp. 1210–1225, 2012.
- [16] Q. Fu, Y. Qin, J. Zhang, and S. Li, "Rapid recursive least-square fine alignment method for SINS," *Journal of Chinese Inertial Technology*, vol. 20, no. 3, pp. 278–282, 2012.

Research Article

Synchronous Routing for Personal Rapid Transit Pods

Mehdi Mrad,¹ Olfa Chebbi,² Mohamed Labidi,¹ and Mohamed Ali Louly¹

¹ Department of Industrial Engineering, College of Engineering, King Saud University, P.O. Box 800, Riyadh 11421, Saudi Arabia

² High Institute of Management, University of Tunis, 41 Rue de la Liberte, 2000 Le Bardo, Tunisia

Correspondence should be addressed to Mohamed Labidi; mlabidi@ksu.edu.sa

Received 7 February 2014; Revised 3 June 2014; Accepted 6 June 2014; Published 26 June 2014

Academic Editor: M. Chadli

Copyright © 2014 Mehdi Mrad et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Personal rapid transit (PRT) is a public and automated transport system in which a fleet of small driverless vehicles operate in order to transport passengers between a set of stations through a network of guided ways. Each customer is carried from one station to another directly with no stop in intermediate stations. This mode of transport can result in a high level of unused capacity due to the empty moves of the vehicles. In this paper, we model the problem of minimizing the energy consumed by the PRT system while assuming predetermined list of orders; then we solve it using some constructive heuristics. Experiments are run on 1320 randomly generated test problems with various sizes. Our algorithms are shown to give good results over large trip instances.

1. Introduction

Using public transportation means, such as train, metro, or bus, can lead to lengthy and suffering journey. The consecutive stops at the intermediate stations, the long waiting times at the stations, and the multiple line changing make the trip less comfortable. Using private cars seems to be a solution. However, the high cost related to the fuel consumption, pollution, and congestion problems presents serious drawbacks to private car users (see Donatos and Kioulafas [1] and Banister [2]).

In order to create a new transportation technology, an increasing effort is provided by researchers worldwide in different research areas including economy, energy, urban planning, safety, sustainable development, and transport environment (see Townsend and Zacharias [3] and Chadli et al. [4]).

In this context, the personal rapid transit (PRT) is considered as an interesting emerging technology. The basic idea of PRT can be traced back to 1953. The first academic article that treated the PRT system was written by Fichter [5]. And the first real implementation of a PRT system was made in the 70s by the University of West Virginia in Morgan town to connect the three-town campus as well as the down town area. In fact, PRT is an innovative system which combines the convenience of private travel and the advantage of public transport. PRT is

composed generally of an electrical driverless fleet of vehicles called pods (Figure 1) that can take passengers directly from their departure station to their destination without any stop. This is possible due to the bypasses that exist in every station (Figure 2) which allow PRT pods to bypass the intermediate stations until arriving to their destination without stopping (see Won et al. [6] and Schweizer and Mantecchini [7]). A central automatic system guides the pods through a network of dedicated guides ways [8].

There was no real interest in the PRT system until the 90s when technology progresses related to the PRT system such as battery vehicle, stations structures, and automated control systems were made (see Dahmani et al. [9] and Dahmani et al. [10]). During the 90s, different PRT projects were prospected and studied. In fact, the taxi 2000 system [11] was selected for testing and construction in O'hare airport (Rosemont, USA). At present, there are a number of PRT systems that are operational as the Heathrow airport PRT system (London, UK) which allows people to travel from the car parking to Terminal 5 in the airport and the Masdar city (Abu Dhabi, UAE) which is planned to be the first zero carbon emission city where PRT is the only powered intercity transportation system.

Two routing approaches for the PRT pods exist. The first one is the synchronous (or clear path) approach where a list of scheduled trips is planned before pods leave the depot. In this



FIGURE 1: PRT vehicle.



FIGURE 2: PRT station.

case, the central automatic system that guides the pods should prepare the route of each pod before starting the execution of the different trips [12]. The second one is the asynchronous approach where pods move from one station to another basically upon the request of customers. This approach seems more comfortable for customers but harder to control for the system supervisor. In this latter case, it becomes difficult to predict the congestion problems in the PRT network. The decision of the automatic system that guides the pods will be restricted to the assignment of the nearest available pod to the customer (see Kornhauser and McEvaddy [13] and Irving [14]).

A comparative study of the two approaches was made by Xithalis [15]. It is shown that despite the fact that the clear path approach offers less flexibility to customers, it is still much more interesting than the asynchronous approach because of its high capacity, congestion-free operation, and less topology limitations and also because it offers communications that are not real-time critical [8].

Unfortunately, both of routing approaches can result in a high level of unused capacity and wasted energy. However due to its deterministic aspect, the clear path approach offers the opportunity to optimize the wasted energy by designing optimal routing of the PRT pods based on the list of scheduled trips and the characteristics of the PRT network. In this context, many exact methods based on linear programming are presented for the vehicle routing problem with time window (see Baldacci et al. [16], Baños et al. [17], and Çetinkaya et al. [18]). These exact methods are time consuming procedures; therefore, some heuristic

algorithms are presented; among them we quote briefly (Pang [19], Ursani et al. [20], and Belhaiza et al. [21]). The aim of this paper is to minimize the total consumed energy in the case of synchronous routing approach of PRT pods while ensuring all the scheduled trips.

This paper is organized as follows: Section 2 presents the problem definition and the mathematical formulation; Section 3 includes a description of the constructive heuristics proposed to solve the problem. We present the results of the proposed methods in Section 4. The conclusion follows.

2. Problem Definition and Mathematical Formulation

In this section, we present the problem definition as presented in the work of Mrad and Hidri [22]. Consider a connected network of personal rapid transit (PRT) with a finite number of stations (M) and one depot. We assume that we have a deterministic list of trips and that the vehicles have a limited battery capacity (B). Thus, the road from the depot to itself while visiting some stations must never use more than B electricity charge. We suppose also that we have an unlimited number of vehicles to serve all the passengers demands.

The set of trips has the cardinality $|T| = n$. Each trip i ($i = 1, \dots, n$) has a departure time Dt_i , a departure station Ds_i , an arrival time At_i , and an arrival station As_i . It is worth noting that the arrival time At_i is the departure time Dt_i plus the duration of the shortest path from the departure to the arrival stations. Furthermore, $Sp_{(i,j)}$ denotes the lasted duration taking the shortest path from station i to station j .

In the sequel, we present an integer programming formulations, based on the following underlying network representation. Consider the graph $G = (V, E)$, with the set of nodes $V = T \cup \{s, t\}$, where s and t are two dummy nodes and E is the set of arcs. Denote $V^* = V \setminus \{s, t\}$ and $E^* = E \setminus \{(i, j); i = s \text{ or } j = t\}$, and sort the trips in the increasing order of their departure time. The set of arcs E is defined as follows.

- (i) If $i, j \in V^*$ with $j > i$ and $At_i + Sp_{(As_i, Ds_j)} \leq Dt_j$, then we add an arc (i, j) with cost c_{ij} representing the energy consumed from the arrival station of trip i (As_i) to the arrival station of trip j (As_j).

```

(1) integer: GlobalCost  $\leftarrow$  0,  $i \leftarrow$  0
(2) Procedure RelaxedProb(integer TotalCost[], integer PathStructure[][]))
(3) for ( $i < \text{Size}(\text{TotalCost})$ ) do
(4)   if ( $\text{TotalCost}[i] > B$ ) then
(5)     GlobalCost  $\leftarrow$  GlobalCost + GreedyCost(PathStructure[i])
(6)   else
(7)     GlobalCost  $\leftarrow$  GlobalCost + TotalCost[i]
(8)   end if
(9) end for
(10) Return (GlobalCost)

```

ALGORITHM 1: Greedy correction.

```

(1) integer GlobalCost  $\leftarrow$  0
(2) Procedure RelaxedProb(integer TotalCost[], integer PathStructure[][]))
(3) Function integer InitialProb(integer trip[], integer Cost[][]))
(4) for ( $j < \text{Size}(\text{TotalCost})$ ) do
(5)   if ( $\text{TotalCost}[j] > B$ ) then
(6)     integer trip[]
(7)     for ( $i < \text{Size}(\text{PathStructure}[j])$ ) do
(8)       trip[i]  $\leftarrow$  PathStructure[j][i]
(9)     end for
(10)    GlobalCost  $\leftarrow$  ProbInitial (trip, Cost)
(11)   else
(12)    GlobalCost  $\leftarrow$  GlobalCost + TotalCost(j)
(13)   end if
(14) end for
(15) Return (GlobalCost)

```

ALGORITHM 2: Optimal correction.

- (ii) For each node i , we add an arc (s, i) (the cost of this arc is c_{si} and it represents the energy used to reach the arrival station of trip i , from the depot).
- (iii) For each node i , we add an arc (i, t) (the cost of this arc is c_{it} and it represents the energy used from the arrival station of trip i to the depot).

2.1. *An Assignment-Based Formulation.* In this section, an assignment-based formulation is presented. To that aim, the following decision variables and definitions are introduced.

- (i) $x_{ij} = 1$, if the station j is visited immediately after station i and 0, otherwise.
- (ii) z_i is the amount of charge used to reach the node $i \in V^*$ from the depot.
- (iii) Consider $a_i = c_{si}$ for $i \in V^*$.
- (iv) Consider $b_i = B - c_{it}$ for $i \in V^*$.

In addition, we define the following:

- (i) $\delta^+(i)$ is the set of nodes j such that an arc (i, j) exists;
- (ii) $\delta^-(i)$ is the set of nodes j such that an arc (j, i) exists.

Hence, the minimum charge assuring the trips is equal to the optimal value of the following programming model:

$$\text{PRT (1): Minimize } \sum_{(i,j) \in E} c_{ij} x_{ij} \quad (1)$$

$$\sum_{j \in \delta^+(i)} x_{ij} = 1 \quad \forall i \in V^* \quad (2)$$

$$\sum_{j \in \delta^-(i)} x_{ji} = 1 \quad \forall i \in V^* \quad (3)$$

$$z_i + c_{ij} \leq z_j + (b_i - a_j + c_{ij})(1 - x_{ij}) \quad \forall (i, j) \in E^* \quad (4)$$

$$a_i \leq z_i \leq b_i \quad \forall i \in V^* \quad (5)$$

$$x_{ij} \in \{0, 1\} \quad \forall (i, j) \in E \quad (6)$$

$$z_i \geq 0 \quad \forall i \in V^*. \quad (7)$$

Objective (1) is to minimize the total used charge. Constraints (2) and (3) require that each node $i \in V^*$ must be visited and left only one time, respectively. Constraints (4) ensure the following conditions.

- (i) If $x_{ij} = 1$, then $z_i + c_{ij} \leq z_j$ with $a_i \leq z_i \leq b_i \quad \forall i, j \in V^*$.

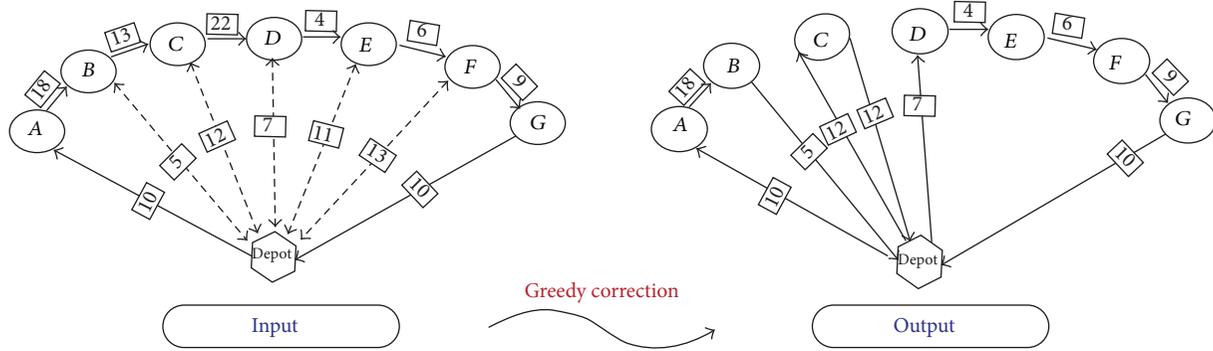


FIGURE 3: Illustration of greedy correction heuristic.

Constraints (5) present a trivial bound limitation on the charge needed to perform the trip i . Finally, constraints (6) and (7) indicate that the decision variables x_{ij} are binary-valued and z_i are positive real variables.

3. Constructive Methods

In this section, we present three constructive heuristics for the problem considered above.

We first solve the following relaxed linear program:

$$\begin{aligned}
 \text{RLPRT: } \text{Minimiser } & \sum_{(i,j) \in E} c_{ij} x_{ij} \\
 & \sum_{j \in \delta^+(i)} x_{ij} = 1 \quad \forall i \in V^* \\
 & \sum_{j \in \delta^-(i)} x_{ji} = 1 \quad \forall i \in V^* \\
 & x_{ij} \in \{0, 1\} \quad \forall (i, j) \in E.
 \end{aligned} \tag{8}$$

However, by solving this linear program, we can get infeasible tour which consumes more energy than what the battery allows. To fix this dilemma, we propose three different algorithms which are presented hereafter.

3.1. Greedy Correction Algorithm

3.1.1. Approach. Each infeasible path can be divided into a set of feasible paths using the following method. Let S_j be the sequence of trips that should be visited by a PRT car in the j th infeasible path. We start by the first trip in the sequence S_j and we continue to add trips while the cumulative energy required by these trips is less or equal to the battery capacity. Once the battery capacity is exceeded, a new path should be considered starting from the first noncovered node in the sequence S_j . We continue to build feasible paths till all trips of S_j are covered. This method will be applied on all infeasible paths.

3.1.2. Illustration. It is supposed that after the resolution of the relaxed linear program, which does not consider the

battery's capacity, we will find a set of infeasible paths. Each infeasible path begins and returns to the depot in order to make trips A, B, \dots, G Figure 3 includes an example illustrating the correction of infeasible routes based on the heuristic described above. The capacity of the battery in this example is equal to 40.

3.1.3. Algorithm. Algorithm 1 describes in detail the different steps of the heuristic. We suppose that we have the procedure `RelaxedProb` (integer `TotalCost []` and integer `PathStructure [][]`) which takes two parameters: an empty array `TotalCost` which is filled by the total cost of each path provided using this method. The second parameter is the matrix `PathStructure` which will be completed by all the different paths of itinerary. We suppose also that the function `size` (array) takes an array as an argument and returns its size. The function `GreedyCost` should take the structure of an infeasible path and try to correct it in order to respect the battery's capacity. It will split the infeasible route to a set of feasible routes.

3.2. Split Correction Algorithm

3.2.1. Approach. The route-first cluster second approach was introduced by Beasley [23]. The main idea of this approach consists of building an auxiliary graph, with the vertices of an infeasible path including the depot. An arc (i, j) such that $i < j$ in the graph represents a feasible tour from i to j which visits the vertices $i + 1, i + 2, \dots, j$. The final extremity of an arc represents the end of the tour, and the cost of the arc is the cost of the tour. Only arcs having cost less than the capacity of the battery are considered in the graph. An optimal partition is obtained by solving the shortest path problem on the obtained auxiliary graph. Each arc (i, j) considered in the solution of the shortest path problem represents a feasible tour including trips $(i, i + 1, \dots, j)$ that should be considered in the solution of the PRT problem.

3.2.2. Illustration. Figure 4 represents an application of the Split procedure to get the optimal splitting of a tour for the VRP problem. Figure 4(a) shows a sequence $S = (a, b, c, d, e)$ and the costs are written on the arcs. The middle graph

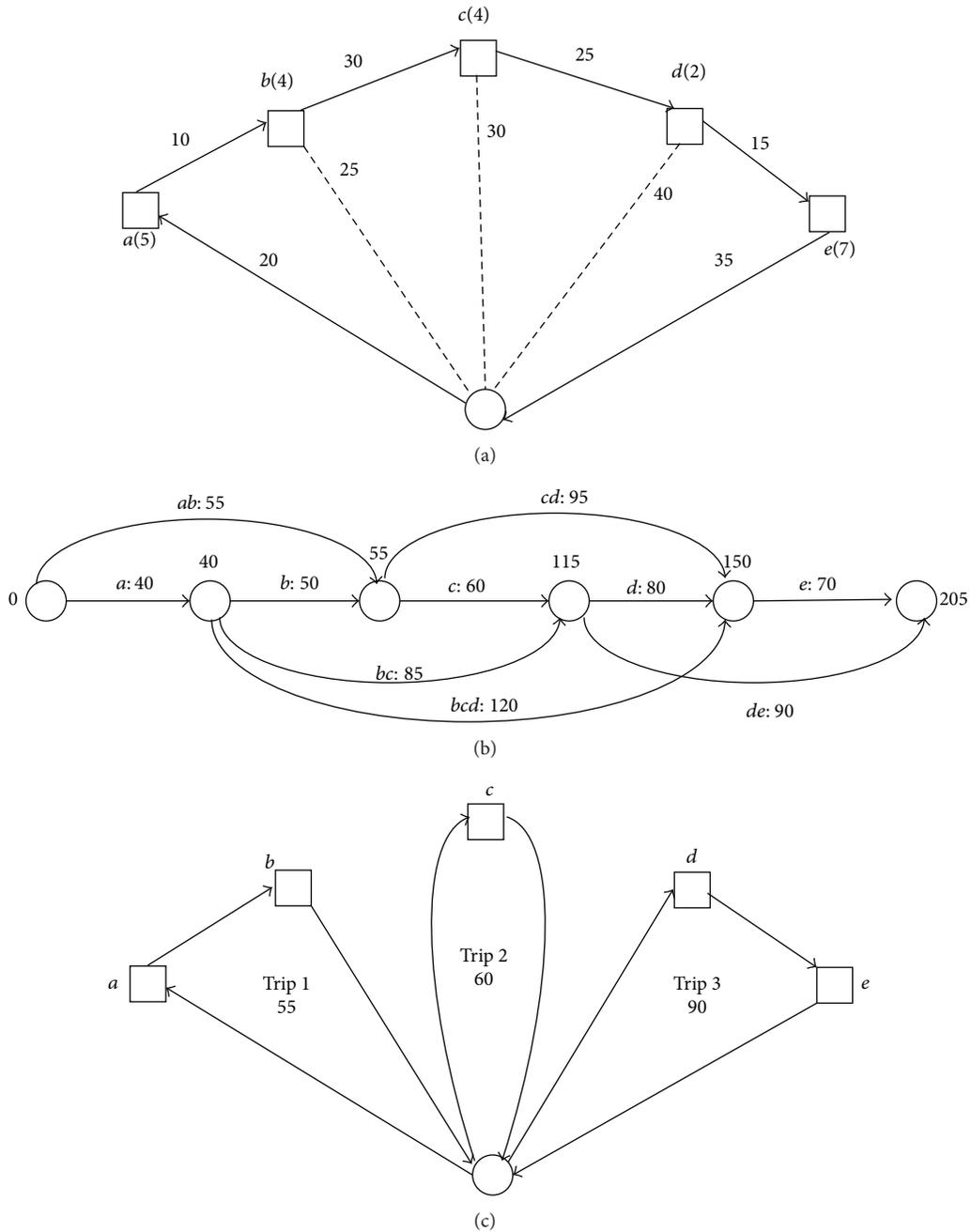


FIGURE 4: Example of Split procedure [18].

contains such an arc ab with a cost of 55 for the trip $(0, a, b, 0)$. The bold path has three arcs, and its cost is 255. The lower part gives the resulting solution with three trips.

3.3. Optimal Correction Algorithm

3.3.1. Approach. Solving the relaxed PRT problem (RLPRT) will lead to a set of paths. Unfortunately, it is possible that some infeasible paths may exist in this set. In this heuristic, we will correct the infeasible resulting tours while considering

that the set of trips included in an infeasible path is in itself a new subproblem that can be solved optimally using the formulation PRT(1). In fact, assuming that the number of trips that belong to only one path is reduced, then the resulting subproblem can be solved to optimality in a short time using integer linear programming.

3.3.2. Illustration. Figure 5 illustrates how a feasible solution can be found through two steps. In the first step, the resolution of the relaxed integer linear program (RLPRT)

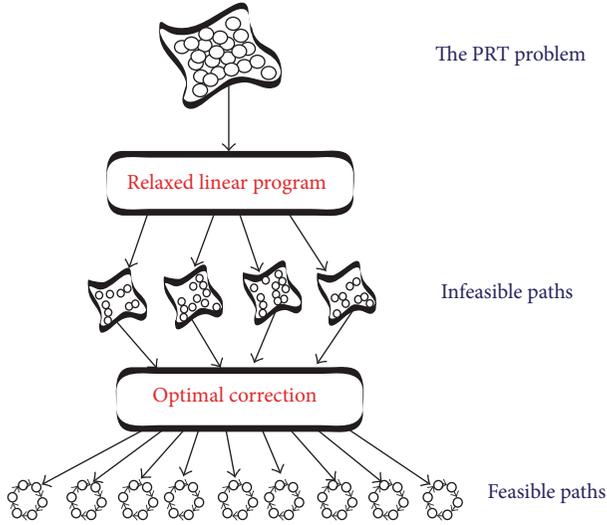


FIGURE 5: Illustration of optimal correction heuristic.

TABLE 1: Instance assumptions.

Number of stations	12
Cost of arcs	Generated randomly between 1 and 15
Departure station	Generated randomly between 1 and 12
Arrival station	Generated randomly between 1 and 12
Departure time	Generated randomly between 1 and 3600
Arrival time	Sum of the departure time and the trip duration
Battery capacity	40 minutes
Max-wait	1000 seconds

will provide a set of paths potentially infeasible because no constraint related to the battery capacity is considered. In the second step, each set of trips belonging to an infeasible path will be considered as the input of the integer linear program PRT(1) which ensures the feasibility of any solution regarding the requirement of the PRT system as described in this paper.

3.3.3. Algorithm. Algorithm 2 describes the process of solving the problem using two main functions. The first Relaxed-Prob (integer TotalCost [] and integer PathStructure [[]]) solves the relaxed integer linear program (RLPRT). The second function InitialProb (integer trip [] and integer Cost [[]]) solves the original linear program PRT(1) which takes into account the constraints of the battery. This function returns the total cost of the different tours built from the set of trips of the corrected infeasible path.

4. Computational Results

All the algorithms were coded in C++ and compiled with Visual Studio 2010. The LP instances have been solved using Cplex 12.2. All the computational experiments were carried out on a Dual Core 2.70 GHz Personal Computer with 1 GB RAM under Windows XP.

TABLE 2: Results of the greedy correction heuristic.

Size	Average gap (%)	Average time (sec)
10	3.491	0.044
15	3.880	0.033
20	3.777	0.023
25	6.104	0.028
30	7.408	0.027
35	7.222	0.037
40	6.354	0.040
45	5.768	0.042
50	7.400	0.053
55	7.605	0.174
60	7.568	0.082
65	7.046	0.170
70	7.671	0.085
75	8.243	0.150
80	9.669	0.178
85	8.109	0.294
90	7.901	0.183
95	8.148	0.121
100	8.148	0.120
110	8.361	0.137
120	8.478	0.169
130	8.758	0.181
140	9.973	0.193
150	11.083	0.301
160	8.347	0.224
170	9.410	0.237
180	9.424	0.231
190	10.613	0.287
200	9.498	0.304
250	8.617	0.613
300	12.005	0.879
350	11.373	1.182
400	13.603	1.685
Average	8.214	0.258

4.1. Data Sets. We generated 40 instances for 33 values of n ranging in [10, 400]. So, in general we did test the proposed heuristics on 1320 different instances. Table 1 summarizes the generating principle of each type of instances.

In order to evaluate the results, we computed the average relative gap (GAP) defined by

$$\text{GAP} = \left(\frac{S_{\text{Heuristic}} - \text{LB}}{\text{LB}} \right) * 100. \quad (9)$$

(i) $S_{\text{Heuristic}}$ is the solution of the heuristic.

(ii) LB is the lower bound obtained by the linear relaxation of PRT(1) (see Section 2).

4.2. Results. The performance of all the developed heuristics are shown in Tables 2, 3, and 4. We could note that the best

TABLE 3: Results of the split correction heuristic.

Size	Average gap (%)	Average time (sec)
10	2.529	0.022
15	2.701	0.018
20	2.476	0.018
25	4.055	0.021
30	5.536	0.017
35	5.544	0.020
40	4.596	0.027
45	4.199	0.025
50	5.464	0.030
55	5.774	0.035
60	5.642	0.243
65	4.782	0.195
70	5.727	0.107
75	5.591	0.187
80	7.158	0.114
85	5.911	0.089
90	5.561	0.212
95	5.717	0.097
100	6.036	0.125
110	6.229	0.127
120	6.056	0.153
130	6.348	0.211
140	7.198	0.609
150	8.499	0.216
160	6.330	0.165
170	6.869	0.173
180	6.799	0.266
190	7.829	0.267
200	6.692	0.297
250	6.188	0.450
300	8.882	0.808
350	8.465	1.065
400	10.605	1.479
Average	6.000	0.239

TABLE 4: Results of the optimal correction heuristic.

Size	Average gap (%)	Average time (sec)
10	2.529	0.051
15	2.603	0.034
20	2.436	0.047
25	4.005	0.055
30	5.174	0.068
35	5.475	0.095
40	4.566	0.089
45	4.169	0.106
50	5.318	0.155
55	5.663	0.154
60	5.406	0.201
65	4.757	0.152
70	5.634	0.168
75	5.470	0.200
80	7.008	0.277
85	5.828	0.242
90	5.447	0.262
95	5.594	0.254
100	5.936	0.282
110	6.162	0.284
120	5.956	0.402
130	6.252	0.300
140	7.056	0.420
150	8.359	0.532
160	6.260	0.567
170	6.809	0.619
180	6.718	0.658
190	7.718	0.691
200	6.637	0.791
250	6.112	1.113
300	8.802	1.610
350	8.383	2.285
400	10.503	2.987
Average	5.901	0.489

heuristic in terms of average gap is the optimal correction heuristic with a value of 5.901% and in terms of computational time the best one is the split correction heuristic which has an average time of 0.239 seconds.

5. Conclusion

Given the ecological, economic, and social interests, the personal rapid transit (PRT) is a research area that attracts more and more researchers from different domains. Optimizing the energy used by a PRT system is an important and challenging problem that has a direct impact on the cost of this new transportation mean. Since the exact approaches are time consuming and are not able to solve large instances (see Mrad and Hidri [22]), we propose in this paper to solve heuristically larger instances. In this work, three different constructive heuristics were presented. Our methods can

find a high quality solution for almost all the different randomly generated instances. As an extension to this work, metaheuristics can be developed and other versions of the PRT problem can be considered by changing the charging strategy of the batteries or by limiting the fleet size of PRT vehicles.

Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

Acknowledgment

The authors would like to extend their sincere appreciation to the Deanship of Scientific Research at King Saud University

for its funding of this research through the Research Group Project no. RGP-VPP-296.

References

- [1] G. S. Donatos and K. E. Kioulafas, "A quantitative analysis of new car sales and advertising in Greece," *European Journal of Operational Research*, vol. 48, no. 3, pp. 311–317, 1990.
- [2] D. Banister, "Cities, mobility and climate change," *Journal of Transport Geography*, vol. 19, no. 6, pp. 1538–1546, 2011.
- [3] C. Townsend and J. Zacharias, "Built environment and pedestrian behavior at rail rapid transit stations in Bangkok," *Transportation*, vol. 37, no. 2, pp. 317–330, 2010.
- [4] M. Chadli, S. Aouaouda, H. R. Karimi, and P. Shi, "Robust fault tolerant tracking controller design for a VTOL aircraft," *Journal of the Franklin Institute*, vol. 350, no. 9, pp. 2627–2645, 2013.
- [5] D. Fichter, *Individualized Automatic Transit and the City*, 1964.
- [6] J. Won, H. Choe, and F. Karray, "Optimal design of personal rapid transit," in *Proceedings of the IEEE Intelligent Transportation Systems Conference (ITSC '06)*, pp. 1489–1494, Toronto, Canada, September 2006.
- [7] J. Schweizer and L. Mantecchini, *Performance Analysis of Large Scaleprt Networks: Theoretical Capacity and Microsimulations*, Bologna University, DISTART Trasporti.
- [8] J. E. Anderson, *Transit Systems Theory*, Lexington Books, Lexington, Ky, USA, 1978.
- [9] H. Dahmani, M. Chadli, A. Rabhi, and A. El Hajjaji, "Road curvature estimation for vehicle lane departure detection using a robust Takagi-Sugeno fuzzy observer," *Vehicle System Dynamics*, vol. 51, no. 5, pp. 581–599, 2013.
- [10] H. Dahmani, M. Chadli, A. Rabhi, and A. El Hajjaji, "Vehicle dynamic estimation with road bank angle consideration for rollover detection: theoretical and experimental studies," *Vehicle System Dynamics*, vol. 51, no. 12, pp. 1853–1871, 2013.
- [11] J. E. Anderson, "The taxi 2000 personal rapid transit system," *Journal of Advanced Transportation*, vol. 22, no. 1, pp. 1–15, 1988.
- [12] J. E. Anderson, "Control of personal rapid transit systems," *Journal of Advanced Transportation*, vol. 32, no. 1, pp. 57–74, 1998.
- [13] A. L. Kornhauser and P. McEvaddy, "A quantitative analysis of synchronous vs. quasi-synchronous network operations of automated transit systems," *Transportation Research*, vol. 9, no. 4, pp. 241–248, 1975.
- [14] J. H. Irving, *Fundamentals of Personal Rapid Transit: Based on a Program of Research, 1968–1976 at the Aerospace Corporation*, Lexington Books, El Segundo, Calif, USA, 1978.
- [15] C. Xithalis, "Synchronous control method for personal rapid transit systems," 2008.
- [16] R. Baldacci, A. Mingozzi, and R. Roberti, "Recent exact algorithms for solving the vehicle routing problem under capacity and time window constraints," *European Journal of Operational Research*, vol. 218, no. 1, pp. 1–6, 2012.
- [17] R. Baños, J. Ortega, C. Gil, A. L. Márquez, and F. de Toro, "A hybrid meta-heuristic for multi-objective vehicle routing problems with time windows," *Computers & Industrial Engineering*, vol. 65, no. 2, pp. 286–296, 2013.
- [18] C. Çetinkaya, I. Karaoglan, and H. Gökçen, "Two-stage vehicle routing problem with arc time windows: a mixed integer programming formulation and a heuristic approach," *European Journal of Operational Research*, vol. 230, no. 3, pp. 539–550, 2013.
- [19] K. W. Pang, "An adaptive parallel route construction heuristic for the vehicle routing problem with time windows constraints," *Expert Systems with Applications*, vol. 38, no. 9, pp. 11939–11946, 2011.
- [20] Z. Ursani, D. Essam, D. Cornforth, and R. Stocker, "Localized genetic algorithm for vehicle routing problem with time windows," *Applied Soft Computing Journal*, vol. 11, no. 8, pp. 5375–5390, 2011.
- [21] S. Belhaiza, P. Hansen, and G. Laporte, "A hybrid variable neighborhood tabu search heuristic for the vehicle routing problem with multiple time windows," *Computers and Operations Research*, 2013.
- [22] M. Mrad and L. Hidri, "Optimal consumed electric energy while sequencing vehicle trips in a personal rapid transit transportation system," *Computers and Industrial Engineering*. In press.
- [23] J. Beasley, "Route first-cluster second methods for vehicle routing," *Omega*, vol. 11, no. 4, pp. 403–408, 1983.

Research Article

Nonlinear Dynamical Analysis of Hydraulic Turbine Governing Systems with Nonelastic Water Hammer Effect

Junyi Li and Qijuan Chen

Power and Mechanical Engineering, Wuhan University, Wuhan, Hubei 430072, China

Correspondence should be addressed to Junyi Li; whujyli@whu.edu.cn

Received 27 March 2014; Revised 5 June 2014; Accepted 5 June 2014; Published 25 June 2014

Academic Editor: Hamid Reza Karimi

Copyright © 2014 J. Li and Q. Chen. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

A nonlinear mathematical model for hydroturbine governing system (HTGS) has been proposed. All essential components of HTGS, that is, conduit system, turbine, generator, and hydraulic servo system, are considered in the model. Using the proposed model, the existence and stability of Hopf bifurcation of an example HTGS are investigated. In addition, chaotic characteristics of the system with different system parameters are studied extensively and presented in the form of bifurcation diagrams, time waveforms, phase space trajectories, Lyapunov exponent, chaotic attractors, and Poincare maps. Good correlation can be found between the model predictions and theoretical analysis. The simulation results provide a reasonable explanation for the sustained oscillation phenomenon commonly seen in operation of hydroelectric generating set.

1. Introduction

Many hydropower plants have been built [1–3] worldwide to harness the energy of falling or running water for electricity purpose. An important part of hydropower plant is the hydraulic turbine governing system (HTGS), which serves to maintain safe, stable, and economical operation of hydropower generating unit [4]. The HTGS is in nature a complex nonlinear, multivariable, time-varying, and non-minimum phase system, which involves the interactions between hydraulic system, mechanical system, and electrical system [5]. The complex dynamic behaviors of the HTGS significantly influence the operation conditions of hydroelectric generating set. For instance, oscillatory problems in hydroelectric generating units were reported to be closely related to the possible Hopf bifurcation and chaotic oscillatory behaviors in HTGS [6–9]. In the absence of a model to conveniently predict the dynamic behaviors of the system, means to address the practical operational issues of hydroelectric generating units, that is, sustained oscillation phenomenon [6, 10], have to be limited.

The literature review reveals that considerable research efforts have been devoted to the modeling of each individual part of the HTGS [11]. For instance, elastic model

and nonelastic model [12, 13] for the conduit system have been adopted in long pipeline and short pipeline systems, respectively. In addition, various hydroturbine models were also available in [10, 14–26]. While Sanathanan [14] claimed that the output of hydroturbine was proportional to hydraulic head and volume flow, Hannett et al. [15] approximated this term by its first-order Taylor formula and Kishor et al. [16, 17] proposed six detailed expressions about these six transfer coefficients with turbine speed and head, which were shown to provide reasonably accurate predictions for HTGS with small disturbance. X. Liu and C. Liu [10] studied small disturbance stability of hydropower plant with complex conduit with a linear turbine model, but they fell short of including elastic water-hammer effects. Bakka et al. made significant contributions to modeling, simulation [18–22], and control [23–26] of wind turbine system, which stimulates modeling of hydroturbine in this paper.

In regards to the power generator, models with different orders could also be found in [27–29]. While higher order power generator models typically represent better accuracy, the overall computational efficiency would be greatly sacrificed especially in case of the modeling of a fully coupled governing system. To compromise with the computational efficiency, the power generator model has to be selected

carefully such that it provides accurate approximation with acceptable computational effort.

Aside from the abovementioned models for individual part of the HTGS, investigations on bifurcation and chaos for the HTGS at the system level were reported only in a few studies [6, 30, 31]. For instance, Konidaris and Tegopoulos [31] investigated the oscillatory problems in hydraulic generating units; Mansoor et al. [6] successfully reproduced an oscillatory phenomenon whose causes were difficult to be identified with limited recorded data. He also proposed a methodology to improve the stability of the control system. While these studies were instrumental in identifying and documenting the oscillations in hydraulic generating units, neither of them noticed the effects of Hopf bifurcation on the oscillatory behaviors. Ling and Tao [30] first reported the influence of bifurcation phenomenon on the sustained oscillations in his study of Hopf bifurcation behavior of HTGS with saturation. However, his model was oversimplified by employing first-order model for the power generator and the PI governing system. Chen et al. [32] developed a novel nonlinear dynamical model for hydroturbine governing system with a surge tank and studied exhaustively the influences of different parameters for the first time. However, they failed to include adequate theoretical analysis and description on bifurcation. Determination of the existence and detailed calculation of Hopf bifurcation were also not considered in their work.

The above literature review shows that numerous well-developed models for each individual part of HTGS are available, but investigations of bifurcation and chaotic oscillations in HTGS with a fully coupled model of the system are rarely seen. As such, this paper aims to develop a fully coupled nonlinear dynamical model for HTGS and investigate the bifurcation and chaotic oscillatory behaviors of the HTGS. In addition, a theorem for existence determination of Hopf bifurcation for four-dimensional nonlinear system has been proposed for convenient prediction of the bifurcation of critical points, which would otherwise be impractical especially for high-dimensional systems due to tremendous computational demand by conventional analysis methods, that is, Lyapunov-Schmidt (L-S) method, center manifold method [33, 34], or normal form theory [35].

There are three main contributions of this paper compared with prior works. First, a new four-dimensional fully coupled nonlinear mathematical model of HTGS was presented and the parameters were from a practical power station, which made the work more consistent with actual project compared with [10, 30] work. Second, the theorem for stability, Hopf bifurcation, and dynamic quality analysis of four-dimensional system that can avoid excessive and tedious calculations was firstly introduced in the paper, providing a new approach for HTGS analysis and computation. Third, nonlinear dynamical behaviors of the above system with different parameters were studied in detail and necessary numerical simulation results were presented.

The paper is outlined as follows. First, the formulations of a fully coupled nonlinear dynamical model will be presented. Next, the investigations of the Hopf bifurcation and chaotic behaviors of the system will be described. Finally, a brief conclusion will be given.

2. Nonlinear Mathematical Model of HTGS

HTGS consists of five parts, that is, conduit system, hydro-turbine, governor, electrohydraulic servo system, and power generator. Model for each individual part has been well developed. Water from reservoir enters tunnel first and then flows through penstock before reaching turbine gate. Next, it flows into scroll casing to promote the hydroturbine to rotate. The power generator and hydroturbine are connected by a shaft coupling. Water that flows into the hydroturbine can be regulated by wicket gates, which are controlled by the governor system. The governor system operates accordingly given the deviation between electric demand and developed torque [35].

2.1. Conduit System Model. A no-elastic model [28] is employed for the conduit system in this study. The unsteady flow partial differential equations in pressure pipes can be described as

$$\begin{aligned} \text{Momentum equation: } \quad & \frac{\partial H}{\partial x} + \frac{1}{gA} \frac{\partial Q}{\partial t} + \frac{fQ^2}{2gDA^2} = 0, \\ \text{Continuity equation: } \quad & \frac{\partial Q}{\partial x} + \frac{gA}{a^2} \frac{\partial H}{\partial t} = 0, \end{aligned} \quad (1)$$

where D , f , A are parameters of the penstock. They denote the diameter, head loss, and area of the pipeline, respectively. H , Q are hydraulic head and turbine flow in penstock in operating condition, a is pressure wave velocity, and x is the length from upstream.

The head and flow equation between two sections of penstock can be deduced from (1). It can be described as [28]

$$\begin{bmatrix} H_A(s) \\ Q_A(s) \end{bmatrix} = \begin{bmatrix} ch(r\Delta x) & -Z_c sh(r\Delta x) \\ -\frac{sh(r\Delta x)}{Z_c} & ch(r\Delta x) \end{bmatrix} \begin{bmatrix} H_B(s) \\ Q_B(s) \end{bmatrix}. \quad (2)$$

Subscripts B , A are symbols of upstream and downstream section of pipeline, respectively. r and Z_c are the composite equations of parameters of the penstock, $r = \sqrt{LCs^2 + RCs}$, $Z_c = r/Cs$.

L , R , C can be written as

$$L = \frac{Q_0}{gAH_0}, \quad C = \frac{gAH_0}{(a^2Q_0)}, \quad R = \frac{(fQ_0^2)}{(gDA^2H_0)}. \quad (3)$$

Providing that the head loss is negligible, r and Z_c can be rewritten as follows:

$$r = \frac{1}{a}s, \quad Z_c = 2h_w, \quad h_w = \frac{aQ_0}{(2gAH_0)}. \quad (4)$$

With the hydraulic friction losses being trivial and $H_B(s) = 0$ (tunnel connects with reservoir directly), the head and flow function is simplified as follows:

$$\frac{H_A(s)}{Q_A(s)} = -Z_c \frac{sh(r\Delta x)}{ch(r\Delta x)} = -Z_c th(r\Delta x). \quad (5)$$

It can be seen from (5) that the water hammer transfer function is a nonlinear hyperbolic tangent function, which was inconvenient to use and should be expanded by series. Substituting (4) into (5), the equation for the water strike transfer function $G_h(s)$ at point A is obtained as

$$\begin{aligned} G_h(s) &= -2h_w th(0.5T_r s) \\ &= -2h_w \frac{\sum_{i=0}^n ((0.5T_r s)^{2i+1} / (2i+1)!) }{\sum_{i=0}^n ((0.5T_r s)^{2i} / (2i)!)}. \end{aligned} \quad (6)$$

Typically, there are the two models, $i = 0, 1$, as follows:

$$G_h(s) = -2h_w th(0.5T_r s) = -2h_w \frac{(1/48)T_r^3 s^3 + (1/2)T_r s}{(1/8)T_r^2 s^2 + 1}, \quad (7)$$

$$G_h(s) = -2h_w th(0.5T_r s) = -T_w s, \quad (8)$$

where $T_r = 2\Delta x/a$, $T_w = LQ_0/AgH_0$. Equations (6) to (8) represent three kinds of water hammer models: the first two models are called elastic water hammer model and the last one is rigid water hammer model that is employed in the paper. In the above equations, the pipeline is assumed to have constant cross-sectional area over the full length, which is impossible in actual engineering. Relevant parameters in the field are obtained by the following formulas:

$$\begin{aligned} L &= \sum_{i=1}^n L_i, & A &= \frac{\sum_{i=1}^n L_i}{\sum_{i=1}^n (L_i/A_i)}, & T_r &= \sum_{i=1}^n \frac{2L_i}{a_i}, \\ T_w &= \frac{\sum_{i=1}^n (L_i/A_i) Q_0}{gH_0} \end{aligned} \quad (9)$$

where L_i , A_i , a_i denote the length, cross-sectional area, and the velocity of pipeline i , respectively, and there are n pipes in total.

2.2. Turbine Model of Rigid Water Hammer. For small perturbation around the rated operating point, the equation of the turbine can be represented as below:

$$\begin{aligned} m_t &= e_x x + e_y y + e_h h, \\ q &= e_{qx} x + e_{qy} y + e_{qh} h. \end{aligned} \quad (10)$$

The six constants of hydroturbine e_x , e_y , e_h , e_{qx} , e_{qy} , e_{qh} are the partial derivatives of the torque and flow with respect to turbine speed, guide vane, and head, respectively. These constants may vary as the operating point changes.

In the dynamic models of the turbine and conduit system [12, 14] where the relative deviation is used to represent the state variables, the relationship between turbine torque and its output power is

$$P_m = m_t + \Delta\omega. \quad (11)$$

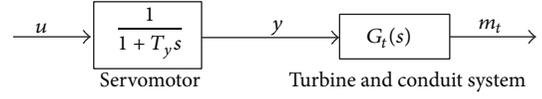


FIGURE 1: Dynamic model of turbine and conduit system.

As the unit speed changes little, the speed deviation $\Delta\omega = 0$, leading to $P_m = m_t$. Then the transfer function of the turbine and conduit system is

$$G_t(s) = e_y \frac{1 + eG_h(s)}{1 - e_{qh}G_h(s)}, \quad (12)$$

where $G_h(s)$ is transfer function of conduit system defined in (6).

In case of rigid water hammer, the coefficient n in (6) is 0 and the dynamic equation for the turbine output torque, as shown in Figure 1, becomes

$$\dot{m}_t = \frac{1}{e_{qh}T_w} \left[-m_t + e_y y - \frac{ee_y T_w}{T_y} (u - y) \right]. \quad (13)$$

2.3. Generator Model. A synchronous generator [36], which connected to an infinite bus through a transmission line, is considered as the target system. The second-order nonlinear dynamical model, after making standard considerations, can be written as

$$\begin{aligned} \dot{\delta} &= \omega_0 \omega, \\ \dot{\omega} &= \frac{1}{T_{ab}} (m_t - m_e - K\omega), \end{aligned} \quad (14)$$

where δ , ω , K , and T_{ab} denote the rotor angle, relative speed deviation, damping coefficient, and mechanical starting time, respectively. The electromagnetic torque of the generator m_e is equal to its electromagnetic power P_e :

$$m_e = P_e. \quad (15)$$

The electromagnetic power can be calculated with the following formula:

$$P_e = \frac{E'_q V_s}{x'_{dx} \Sigma} \sin \delta + \frac{V_s^2}{2} \frac{x'_{dx} \Sigma - x_{qx} \Sigma}{x'_{dx} \Sigma x_{qx} \Sigma} \sin 2\delta, \quad (16)$$

where the effects of speed deviations, damping coefficient, and torque variations are all included in the analysis of the generator dynamic characteristics

$$\begin{aligned} x'_{d\Sigma} &= x'_d + x_T + \frac{1}{2} x_L, \\ x_{q\Sigma} &= x_q + x_T + \frac{1}{2} x_L. \end{aligned} \quad (17)$$

Equations (14) to (17) are the simplified second-order nonlinear generator model based on the turbine model of

rigid water hammer, which has been widely applied in nonlinear controller design and stability analysis of HTGS. It is often used for the stability characteristics and dynamic quality analysis of HTGS from the perspective of power system. Higher order nonlinear generator model could be employed according to research needs.

2.4. Hydraulic Servo System Model. The servomotor, which acts as the actuator, is used to amplify the control signals and provide power to operate the guide vane. Its transfer function can be written as

$$G_s(s) = \frac{1}{1 + T_y s}, \quad (18)$$

where T_y is the engager relay time constant.

2.5. Governor Model. At present, a parallel PID controller [18, 20, 22–26] is widely used in hydraulic turbine governors [27, 30] in filed. Its transfer function is given as

$$G_2(s) = \left(k_p + \frac{k_i}{s} + k_d s \right). \quad (19)$$

Substituting (19) into (18), the results can be written in state-space form. At last, the mixed function can be obtained as follows:

$$\frac{dy}{dt} = \frac{1}{T_y} \left(-k_p \omega - k_i \int \Delta\omega - k_d \dot{\omega} - y \right). \quad (20)$$

Based on the discussions above, the differential equations that coupled each individual part of the turbine nonlinear control system can be written as

$$\begin{aligned} \dot{\delta} &= \omega_0 \omega, \\ \dot{\omega} &= \frac{1}{T_{ab}} \left(m_t - D\omega - \frac{E'_q V_s}{x'_{dx\Sigma}} \sin \delta \right. \\ &\quad \left. - \frac{V_s^2 x'_{dx\Sigma} - x_{qx\Sigma}}{2 x'_{dx\Sigma} x_{qx\Sigma}} \sin 2\delta \right), \\ \dot{m}_t &= \frac{1}{e_{qh} T_w} \left(-m_t + e_y y \right. \\ &\quad \left. - \frac{e e_y T_w}{T_y} \left(-k_p \omega - \frac{k_i}{\omega_0} \delta - k_d \dot{\omega} - y \right) \right), \\ \dot{y} &= \frac{1}{T_y} \left(-k_p \omega - \frac{k_i}{\omega_0} \delta - k_d \dot{\omega} - y \right). \end{aligned} \quad (21)$$

Equation (21) is the four-dimensional water-electromechanical coupled model of HTGS that integrates the turbine model of rigid water hammer, the water pipes linear model, and the nonlinear dynamic generator model. Compared to the linear model, it could reflect the complex nonlinear nature problem within the system much better. Equation (21) could

be applied to analyze and simulate the dynamic characteristics of the HTGS.

At the equilibrium point $(0, 0, 0, 0)$, the following condition has to be satisfied:

$$\begin{aligned} \omega &= 0, \\ m_t &= e_y y, \\ y &= -k_i \frac{\delta}{\omega_0}, \end{aligned} \quad (22)$$

$$e_y k_i \frac{\delta}{\omega_0} + \frac{E'_q V_s}{x'_{dx\Sigma}} \sin \delta + \frac{V_s^2 x'_{dx\Sigma} - x_{qx\Sigma}}{2 x'_{dx\Sigma} x_{qx\Sigma}} \sin 2\delta = 0.$$

Equations (21) and (22) can be numerically solved to investigate the nonlinear behaviors of HTGS such as Hopf bifurcation points, bifurcation surface of PID adjustment coefficients, time domain response waveforms of state variables, and Lyapunov exponent with MATLAB by using a variable-step continuous solver based on the four-order Runge-Kutta formula. Time step is 0.01 in the study.

3. The Existence of Dynamic Hopf Bifurcation

3.1. Existence Determination of Hopf Bifurcation. For a four-dimensional nonlinear system, the criteria for the existence of Hopf bifurcation are given in the following theorem. Meanwhile, the bifurcation value collection can also be determined when Hopf bifurcation occurs.

Theorem 1. For a nonlinear system $y = F(x, \mu)$, $x \in R^4$, $\mu \in R^1$ is the bifurcation parameter, $x = 0$ is the equilibrium point, and the Jacobi matrix characteristic polynomial at the equilibrium point is

$$f(\lambda, \mu) = \lambda^4 + p_3(\lambda) \lambda^3 + p_2(\lambda) \lambda^2 + p_1(\lambda) \lambda + p_0(\lambda). \quad (23)$$

If $\mu = 0$, the following conditions hold.

(1) The coefficient $p_i > 0$ ($i = 0, 1, 2, 3$),

$$p_3 p_2 p_1 = p_3^2 p_0 + p_1^2. \quad (24)$$

(2)

$$p'_0 \neq \left(\frac{p_3 p_0}{p_1} - \frac{p_1}{p_3} \right) \frac{p_3 p'_1 - p'_3 p_1}{p_3^2} + \frac{p'_2 p_1}{p_3}, \quad (25)$$

where $^{*'} = d^*/d\mu$, when $|\mu|$ is sufficiently small, Hopf bifurcation exists in one side of $\mu = 0$, and the period of limit cycle can be described as follows:

$$T = 2\pi \sqrt{\frac{p_3}{p_1}}. \quad (26)$$

This theorem is the direct algebraic criterion to determine the existence of Hopf bifurcation for a four-dimensional system, which can avoid excessive and tedious calculations. However,

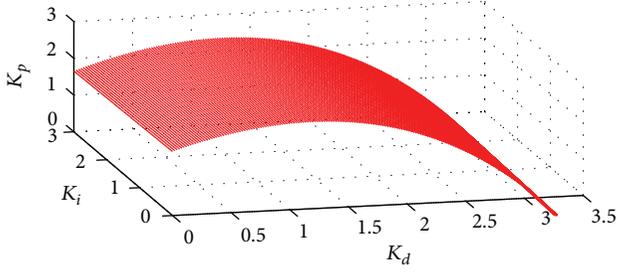


FIGURE 2: Bifurcation surface of K_p, K_i, K_d .

the theorem does not give bifurcation direction. Each point on the surface that is determined by (24) is bifurcation critical point when Hopf bifurcation occurs (the whole bifurcation critical points constitute the entire bifurcation collection). For example, periodic oscillation phenomenon, namely, Hopf bifurcation, will occur, if the system parameters and the PID parameters satisfy conditions 1 and 2 of the above theorem in a certain operating condition. The surface determines the scope of the system's stability region also. In fact, from the perspective of nonlinear dynamical systems structure stability analysis, when the PID parameters can meet with some certain conditions, the structure stability of the system will change dramatically and then comes to the system structure instability, which leads to complex nonlinear oscillations of the system.

3.2. Analysis and Simulation for Hopf Bifurcation. Taking a practical power plant, for example, setting parameters $\omega_0 = 314, T_w = 0.8 \text{ s}, T_{ab} = 9.0 \text{ s}, T_y = 0.1 \text{ s}, D = 2, E'_q = 1.35, V_s = 1, x'_{dx\Sigma} = 1.15, x'_{qx\Sigma} = 1.474, e_{qh} = 0.5, e_y = 1.0,$ and $e = 0.7,$ respectively, analyzes theoretical calculation and simulation in the paper based on the data. The followings are the system Jacobi matrix characteristic polynomial coefficients at equilibrium point $(0, 0, 0, 0)$:

$$\begin{aligned} p_3 &= 12.72 - 1.56 * K_d; \\ p_2 &= 2.78 * K_d - 1.56 * K_p + 62.07; \\ p_1 &= 2.78 * K_p - 1.56 * K_i + 434.15; \\ p_0 &= 2.78 * K_i + 857.2. \end{aligned} \tag{27}$$

It requires all the characteristic polynomial coefficients $p_i > 0$ ($i = 0, 1, 2, 3$) by condition 1 when using Theorem 1 to determine the existence of Hopf bifurcation. For example, $p_3 > 0$; then $K_d < 8.17857$, which means that the Hopf bifurcation area is limited to the range of $K_d < 8.17857$. Similarly, the ranges of K_i and K_p can be obtained. After substituting p_i into (24), a three-parameter implicit polynomial is obtained, and a space bifurcation surface can be made according to the polynomial, as shown in Figure 2.

All the points on the surface are Hopf bifurcation critical points that would lead to the amplitude oscillation of the unit; stability region lies beneath the surface; PID parameters in stability region ensure the stable and safe operation of units; the instability region is located above surface in certain area.

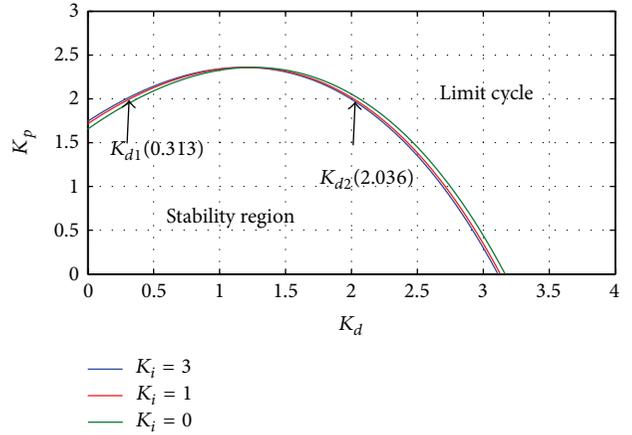


FIGURE 3: Curves of bifurcation points with $K_i = 0, 1, 3$.

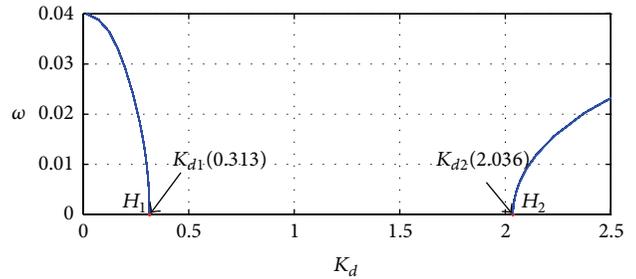


FIGURE 4: Bifurcation diagram of rotor speed ω .

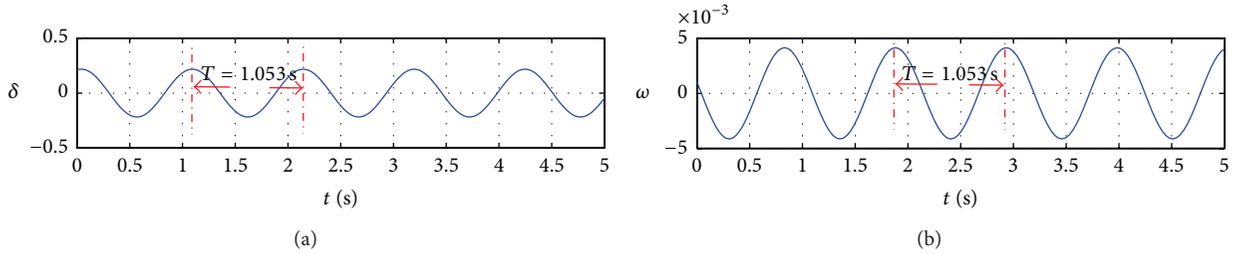
Values of the PID parameters that are above the surface or on the surface should be avoided when a PID or a gain-scheduling PID scheme is used in the system. Otherwise, the Hopf bifurcation or undesirable oscillations may occur in the governing system. Figure 2 can be used to study the relationship between the stability of system and PID parameters, and stable PID parameters could be chosen from it directly. When taking $\mu_d = K_d - K_{d^*}$ or $\mu_p = K_p - K_{p^*}$ as the bifurcation parameter, it can be calculated that the left hand side of (25) is equal to zero, while the right hand side is not; thus condition 2 of the theorem holds.

According to Theorem 1, when $|\mu|$ is sufficiently small, Hopf bifurcation phenomenon of the system will take place. For example, when $K_i^* = 1$ and $K_p^* = 2$, the bifurcation parameters $K_{d1^*} = 0.313$ and $K_{d2^*} = 2.036$ can be calculated from (24). K_{d1^*} and K_{d2^*} can also be got by the intersection values between the curve ($K_i = 1$) and the line ($K_p = 2$), as shown in Figure 3. The curve ($K_i = 0, 1, 3$) is the boundary between stability region and limit cycle; point that located in limit cycle, for example, $K_i = 1, K_d = 2.5,$ and $K_p = 2,$ would lead to the undesirable oscillation. Figure 3 is more intuitive to tell the stale and unstable PID parameters compared with Figure 2.

Figure 4 is the bifurcation diagram of unit speed ω . When $K_p = 2, K_i = 1,$ and K_d is the bifurcation parameter, there are two bifurcation points in the figure; that is, $K_{d1} = 0.313$

TABLE 1: Changes of system properties with the bifurcation parameter K_d .

Bifurcation parameter K_d	$K_d < K_{d1^*}$	$K_d = K_{d1^*} = 0.313$	$K_{d1^*} < K_d < K_{d2^*}$	$K_d = K_{d2^*} = 2.036$	$K_d > K_{d2^*}$
Jacobi matrix eigenvalues	Two complex conjugate eigenvalues, negative real part one and positive real part one	-2.46 -9.77 $-4.5e-5 \pm 5.98i$	Two complex conjugate eigenvalues with negative real parts	-2.76 -6.80 $-2.7e-6 \pm 6.77i$	Two complex conjugate eigenvalues, negative real part one and positive real part one
Property	Stable limit cycle	center of bifurcation in a supercritical state	Stable focus, stable domain	Center of bifurcation in a supercritical state	Stable limit cycle

FIGURE 5: Time waveforms of δ and ω with $K_i = 1$, $K_p = 2$, and $K_d = 0.309$.

and $K_{d2} = 2.036$ as presented calculated by the Matcont software. It is observed that the system is stable and system state variable ω will converge to the equilibrium point 0 when $K_{d1} < K_d < K_{d2}$; it will converge to a stable limit cycle when $K_d < K_{d1}$ or $K_d > K_{d2}$, and the amplitude of the limit cycle oscillation corresponds to the ordinate value of limit cycle curve. The bifurcation thresholds K_{d1} and K_{d2} in Figure 4 are consistent with the intersection values in Figure 3 and the results calculated from (24) as shown in Figure 2. Table 1 shows the properties of the system and the changes of Jacobi matrix eigenvalues with different K_d . When the system is in stable domain ($K_{d1^*} < K_d < K_{d2^*}$), all eigenvalues of Jacobi matrix have negative real parts and the system tends to be stable. As the parameter changes to the critical bifurcation point (K_{d1^*}, K_{d2^*}), the eigenvalues of Jacobi matrix have complex conjugate eigenvalues with zero ($-4.5e-5, -2.7e-6$) real part; that is, the system is in a supercritical state. When $K_d < K_{d1^*}$ or $K_d > K_{d2^*}$, the matrix has two complex conjugate eigenvalues; one has a negative real part and the other has a positive real part, system state variables cannot converge to the equilibrium point at this time, and the system is in stable amplitude oscillation (limit cycle) in this region.

Table 1 provides theoretical explanation for the phenomenon that occurs in Figure 4. Dynamic behaviors of nonlinear systems in critical bifurcation point on both sides can be seen from Figure 4 and Table 1 clearly. Furthermore, it can be obtained from simulation and theoretical analysis that the farther from the two bifurcation points ($K_{d1^*} = 0.313$ and $K_{d2^*} = 2.036$) and the nearer to middle place the K_d is, the quicker the convergence rate is and the more stable the system is. In practical applications, value of K_d should be between two critical places and the farther the better, which

will be verified by the simulation results in Figures 5, 6, 7, and 8.

Time domain response waveforms of δ and ω after stabilization can be seen in Figure 5 where the PID parameters are $K_i = 1$, $K_p = 2$, and $K_d = 0.309$, which are in limit cycle area in Figure 3, indicating that system state variables δ and ω will converge to a stable limit cycle. The motion of the system at this value of adjustment coefficients is sustained periodic oscillation and the oscillations amplitude of ω is the limit cycle curve ordinate value at point $K_d = 0.309$, as can be seen from Figure 4. In addition, the oscillation cycle of δ and ω at bifurcation point is about 1.053 s. Meanwhile, it can be got that $p_3 = 12.24$, $p_1 = 438.15$ easily from (27), so the period of limit cycle $T = 1.050$ s can be calculated from (26) at the Hopf bifurcation point, which is consistent with the simulation result. These results indicate that the governing system makes a periodic vibration and further reveal that the system is not able to be stable. Similarly, the waveforms of δ and ω for the value of $K_d = 0.6, 0.8, 1.15$, which are beneath the limit cycle curve ($K_i = 1$), namely, the stable parameters, are presented in Figures 6 to 8. After a period of time, each state (δ and ω) will converge to equilibrium point 0 and the system keeps steady at these values of adjustment coefficients; the biggest δ and ω value are 0.5, 0.01, respectively. It is observed that when value of K_d is farther from the two critical points ($K_{d1} = 0.313$ and $K_{d2} = 2.036$) and nearer to middle place (1.175), the convergence rate is quicker and the system is more stable.

Figures 9 and 10 show the phase space orbits of the system variables. It can be observed that the closed orbits limit cycles for these parameters formed in a certain region in the phase space orbit when $\mu_d = 0.002$; thus $K_i = 1$, $K_p = 2$, and $K_d = 2.038$. The system shows a diffused and nonlinear growth

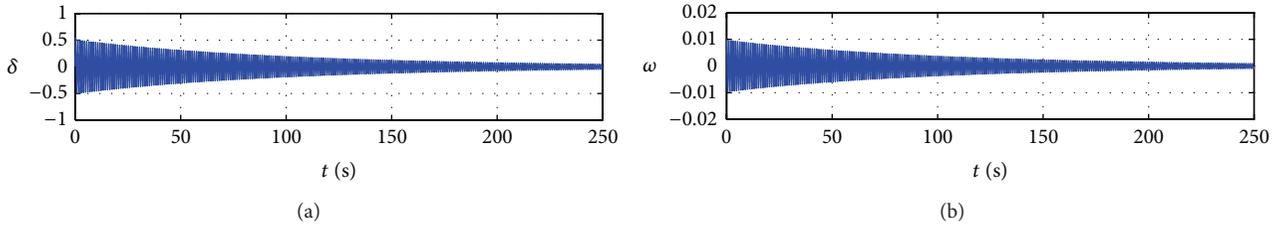


FIGURE 6: Time waveforms of δ and ω with $K_i = 1, K_p = 2, K_d = 0.6$.

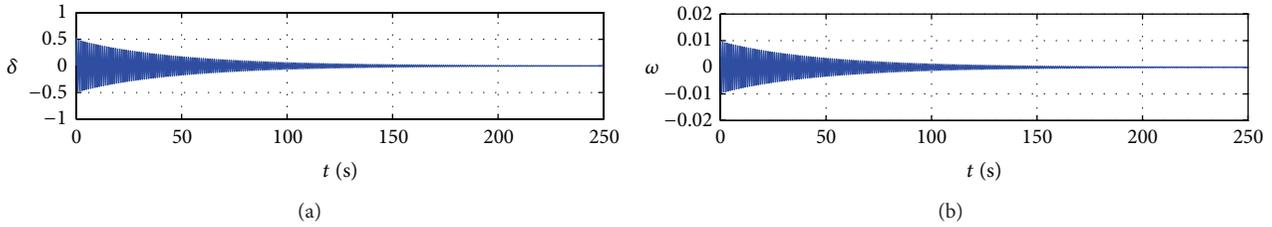


FIGURE 7: Time waveforms of δ and ω with $K_i = 1, K_p = 2,$ and $K_d = 0.8$.

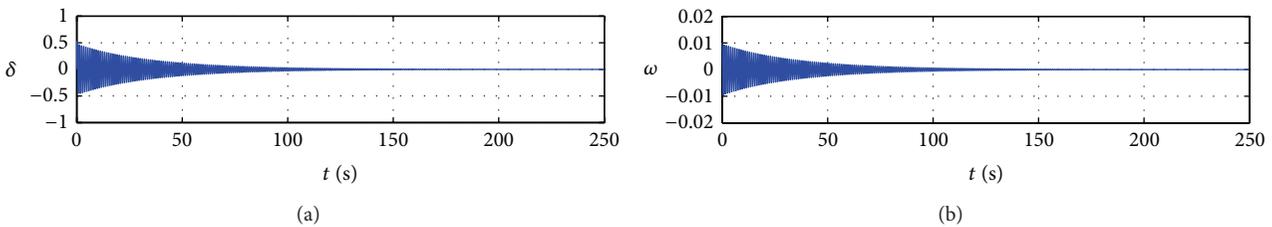


FIGURE 8: Time waveforms of δ and ω with $K_i = 1, K_p = 2,$ and $K_d = 1.15$.

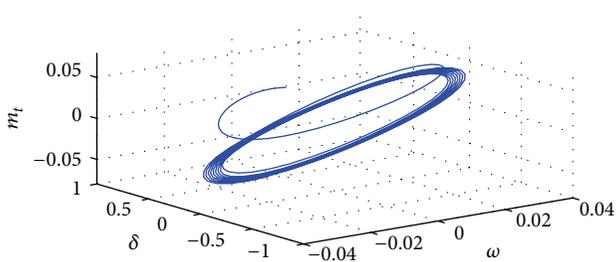


FIGURE 9: Phase orbit of $\delta, \omega,$ and m_t .

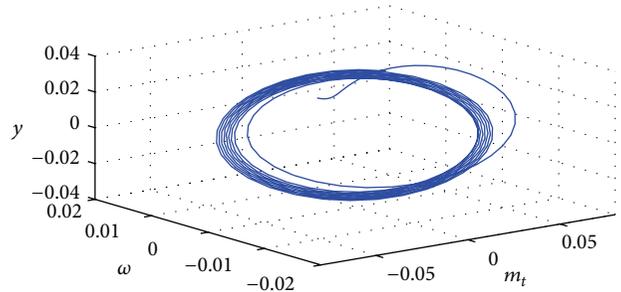


FIGURE 10: Phase orbit of $m_t, \omega,$ and y .

oscillations curve. Hopf bifurcation occurs and the system tends to be a stable oscillation state where $-0.06 < m_t < 0.06, -0.04 < \omega < 0.04, -1 < \delta < 1, -0.04 < y < 0.04$, indicating that the system is in sustained periodic oscillation state; this value of the differential adjustment coefficient cannot be used in practical applications. Therefore, when the turbine regulating system uses PID (or gain scheduling PID) strategy, the values of adjustment coefficients should be beneath the limit cycle curve. Otherwise, there may be a sustained and

stable oscillations, instead of tending to be stable, and the governing system would lose stability finally, leading to an unstable control.

4. Analysis and Simulation for Chaos

When $K_d > 2.5$, with the continually increasing of K_d , unit speed ω on bifurcation diagram remains as a single rising curve until $K_d = 5.72$. The bifurcation diagram gets

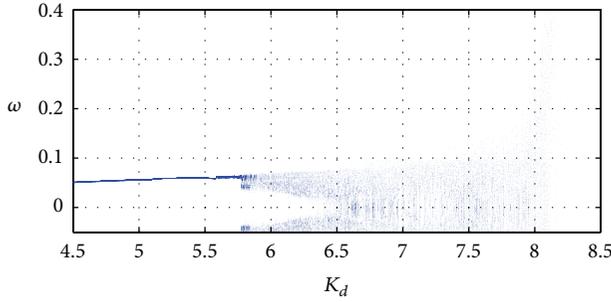


FIGURE 11: Bifurcation diagram of ω with $K_p = 2, K_i = 1$.

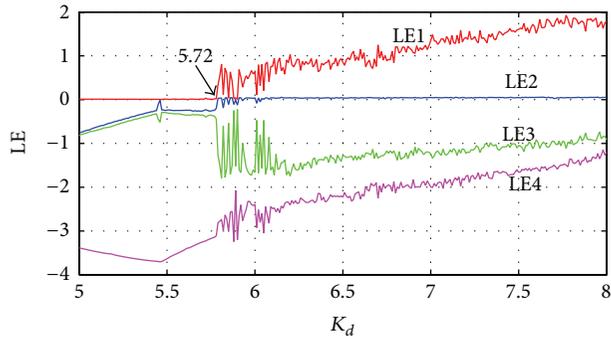


FIGURE 12: Lyapunov exponent with K_d .

an intricate pattern when $K_d > 5.72$, the system state parameters converge neither to the equilibrium point nor to a stable limit cycle, and it undergoes a random motion without any rules, namely, the chaotic motion. There are rich and complex nonlinear dynamical behaviors in this area, as shown in Figure 11.

Lyapunov exponent is a quantitative indicator to measure the system dynamic behavior, which indicates the average rate of convergence or divergence of the system in phase space among different adjacent tracks. The existence of chaotic dynamics for the system can be judged intuitively by the largest Lyapunov exponent depending on whether it is greater than 0 or not. The system chaotic motion can be obtained by Lyapunov exponent with different K_d , as presented in Figure 12. When $K_d \leq 5.72$, the largest Lyapunov exponent LE1 is zero, indicating that the system is in periodic motion state (limit cycle). In the vicinity of $5.72 \leq K_d \leq 6.0$, LE1 changes between 0 and 1, and thus, the system movement state alternates between periodic motion and chaotic motion; LE1 and LE2 are both greater than zero till $K_d = 8.0$ and means chaotic motion that leads to the unstable state of the control system exists.

Figure 13 shows the system stability domain, limit cycles, and chaotic region of the system when $K_i = 1.0$. There are three different areas in the Fig, namely, the stable region, limit cycle, and chaos region, respectively; boundary between periodic motion and chaotic zone is determined by the three PID parameters especially K_d when the maximum Lyapunov exponent LE1 is greater than zero.

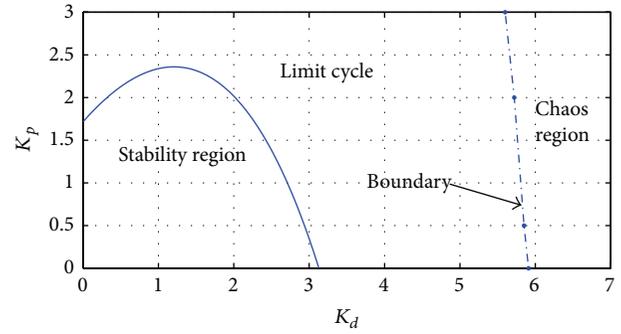


FIGURE 13: Stability domain, limit cycles, and chaotic region of K_p, K_d with $K_i = 1$.

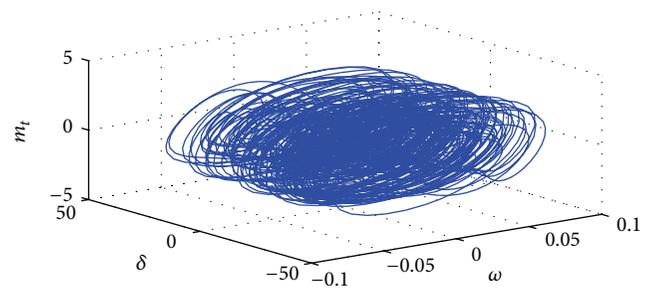


FIGURE 14: The chaotic attractor with $K_i = 1, K_p = 2$, and $K_d = 6.5$.

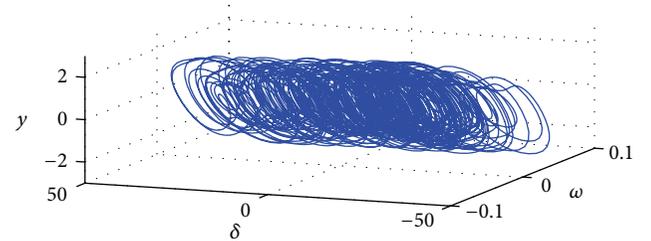


FIGURE 15: The chaotic attractor with $K_i = 1, K_p = 2$, and $K_d = 6.5$.

Chaotic motion of nonlinear systems can generate attractors in the phase space with unique nature. The original stable periodic motion turns to instability after entering into the chaotic region. Chaotic attractor has a complex motion internal with countless unstable periodic orbits studded in it. Overall, the system trajectories are always stretching and varying within a certain range and quite disordered in some part, leading to the chaotic strange attractor. When $K_p = 2.0, K_i = 1.0$, and $K_d = 6.5$, chaotic attractors that are similar to the one with scroll structure in Chua's circuit [37] generate in the δ, ω, m_t and δ, ω, y phase space, as shown in Figures 14 and 15. It is observed that they have the characteristic of being globally bounded, but being local unstable, which would do great harm to the HTGS in operation, further reveals that the system would lose stability finally. Figure 16 is the time domain response waveforms of δ and ω , result shows that the system makes a disordered and aperiodic

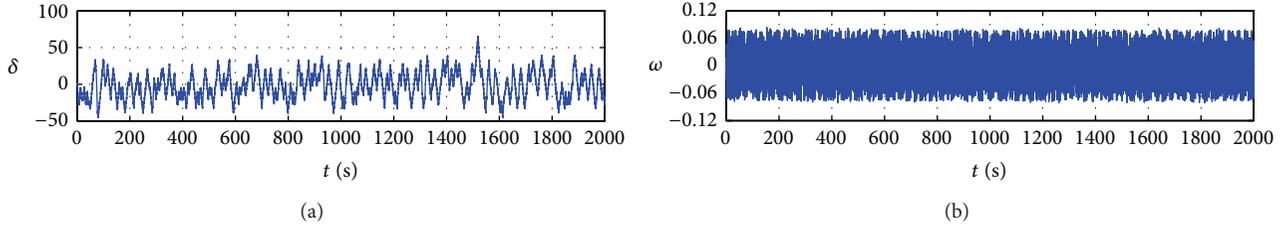


FIGURE 16: Time waveforms of δ and ω with $K_i = 1$, $K_p = 2$, and $K_d = 6.5$.

oscillation, and behaviors of the system are hardly to be predicted. The rotor angle δ and turbine speed ω have left their original position largely, oscillating between -50 and 50 -0.1 to 0.1 , respectively, with a random period, and the system is in a chaos, which means that this value of the differential adjustment coefficient should be avoided.

Poincare maps are shown in Figures 17 and 18. It can be seen that numerous intersections of the continuous trajectory turn up in the Poincare maps on section of $\omega = 0$ and $m_t = 0$. However, they neither fill with the entire phase space nor distribute discretely in the δ, y, m_t and δ, ω, y phase space but occur in a certain region frequently, constituting a piece of dense point just like cloud. These results indicate that the system is in chaotic state and the system cannot tend to be stable. Figures 11, 12, 13, 14, 15, 16, 17, and 18 verify the chaotic and aperiodic movement of the system from mathematical perspective.

5. Conclusion

In this paper, a fully coupled nonlinear mathematical model for HTGS has been developed. Extensive investigations of the nonlinear behaviors of an example HTGS were conducted using the proposed model. The existence and stability of Hopf bifurcation were studied first. Both the simulation results and theoretical analysis predicted that Hopf bifurcation would occur when the governor's parameters satisfy certain criteria. It was also found that the convergence rate and stability of the system were closely related to the bifurcation parameters, indicating that the PID adjustment coefficient parameters have to be carefully selected to guarantee the safe and stable operation of hydroelectric generating unit. Relevant stability ranges of PID adjustment coefficients for the example HTGS were presented in the paper. In addition to the study of Hopf bifurcation, chaotic behaviors of the system were also investigated extensively using different methods. It was found that in certain region of PID parameter space, chaos phenomenon would occur and lead to the chaotic oscillatory and unstable control of the system, which has been identified as a cause of decreased availability and reduced stability of the system. The above investigations on Hopf bifurcation and chaos phenomenon of the example HTGS clearly illustrated the complex nonlinear nature of the system. The simulation results based on the proposed model were shown to have good correlation with the theoretical analysis predictions. The model presented in this study for the analysis

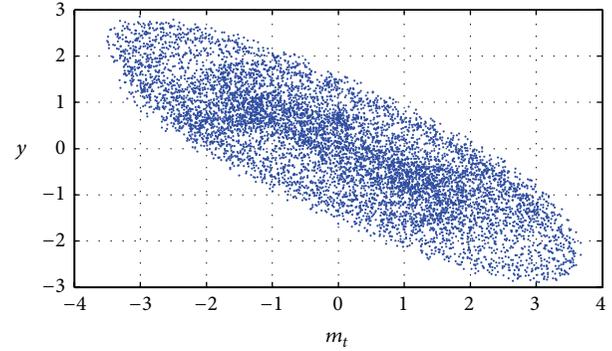


FIGURE 17: Poincare map with $\omega = 0$.

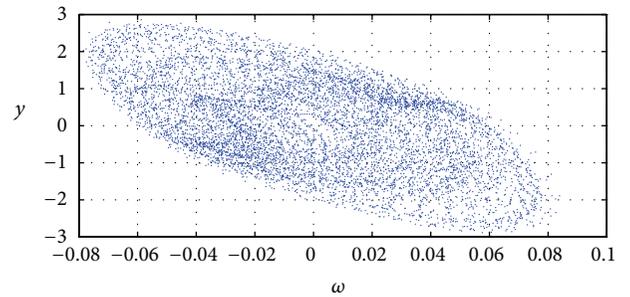


FIGURE 18: Poincare map with $m_t = 0$.

of the dynamic behaviors of HTGS can be further extended to a more sophisticated system.

Some related topics, such as data-driven framework [38], chaos control, and Hopf Bifurcation in HTGS with nonlinear saturation, would be considered in the future work to achieve more practical oriented results.

Nomenclature

- $H(s)$: Laplace transform of h , p.u.
- $Q(s)$: Laplace transform of q , p.u.
- Q_0 : Initial turbine flow of turbine, m^3/s
- H_0 : Initial hydraulic head of turbine, m
- g : Gravitational acceleration
- l : Penstock length, m
- A : Penstock area, m^2
- D : Penstock diameter, m

f :	Head loss coefficient, p.u.
a :	Pressure wave velocity, m/s
T_r :	Elastic time constant, s
T_w :	Water starting time, s
y :	Incremental deviation of guide vane/wicket gate position, p.u.
e_x, e_y, e_h :	Partial derivatives of turbine torque with respect to head, guide vane, and speed, p.u.
e_{qx}, e_{qy}, e_{qh} :	Partial derivatives of the flow with respect to head, guide vane, and turbine speed, p.u.
δ :	Rotor angle, rad
ω :	Turbine/rotor speed, rad/s
ω_0 :	Base angular speed, rad/s
$\Delta\omega$:	Speed deviation
E'_q :	Transient electric potential of q -axis, p.u.
V_s :	Infinite bus voltage, p.u.
$x'_{dx\Sigma}$:	Transient reactance of d -axis, p.u.
$x'_{qx\Sigma}$:	Synchronous reactance of q -axis, p.u.
x'_d :	Direct axis transient reactance, p.u.
x'_q :	Quadrature axis reactance, p.u.
x_T :	Transformer short circuit reactance, p.u.
x_L :	Transmission line reactance, p.u.
T_{ab} :	Mechanical starting time, s
m_t :	Mechanical torque of turbine, N.m
m_e :	Electrical torque, N.m
K :	Damping factor, p.u.
P_e :	Terminal active power
P_m :	Output power
T_y :	Engager relay time constant
K_p :	Proportional adjustment coefficient
K_i :	Integral adjustment coefficient
K_d :	Differential adjustment coefficient.

Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

References

- [1] J. K. Kaldellis, "Critical evaluation of the hydropower applications in Greece," *Renewable and Sustainable Energy Reviews*, vol. 12, no. 1, pp. 218–234, 2008.
- [2] B. Lehner, G. Czisch, and S. Vassolo, "The impact of global change on the hydropower potential of Europe: a model-based analysis," *Energy Policy*, vol. 33, no. 7, pp. 839–855, 2005.
- [3] S. R. Shakya and R. M. Shrestha, "Transport sector electrification in a hydropower resource rich developing country: energy security, environmental and climate change co-benefits," *Energy for Sustainable Development*, vol. 15, no. 2, pp. 147–159, 2011.
- [4] C. S. Li and J. Z. Zhou, "Parameters identification of hydraulic turbine governing system using improved gravitational search algorithm," *Energy Conversion and Management*, vol. 52, no. 1, pp. 374–381, 2011.
- [5] C. Jiang, Y. Ma, and C. Wang, "PID controller parameters optimization of hydro-turbine governing systems using deterministic-chaotic-mutation evolutionary programming (DCMEP)," *Energy Conversion and Management*, vol. 47, no. 9-10, pp. 1222–1230, 2006.
- [6] S. P. Mansoor, D. I. Jones, D. A. Bradley, F. C. Aris, and G. R. Jones, "Reproducing oscillatory behaviour of a hydroelectric power station by computer simulation," *Control Engineering Practice*, vol. 8, no. 11, pp. 1261–1272, 2000.
- [7] D. M. Dobrijevic and M. V. Jankovic, "An improved method of damping of generator oscillations," *IEEE Transactions on Energy Conversion*, vol. 14, no. 4, pp. 1624–1629, 1999.
- [8] A. A. P. Lerm and A. S. Silva, "Avoiding Hopf bifurcations in power systems via set-points tuning," *IEEE Transactions on Power Systems*, vol. 19, no. 2, pp. 1076–1084, 2004.
- [9] Z. Jing, D. Xu, Y. Chang, and L. Chen, "Bifurcations, chaos, and system collapse in a three node power system," *International Journal of Electrical Power & Energy Systems*, vol. 25, no. 6, pp. 443–461, 2003.
- [10] X. Liu and C. Liu, "Eigenanalysis of oscillatory instability of a hydropower plant including water conduit dynamics," *IEEE Transactions on Power Systems*, vol. 22, no. 2, pp. 675–681, 2007.
- [11] E. De Jaeger, N. Janssens, B. Malfliet, and F. Van De Meulebroeke, "Hydro turbine model for system dynamic studies," *IEEE Transactions on Power Systems*, vol. 9, no. 4, pp. 1709–1715, 1994.
- [12] P. Pennacchi, S. Chatterton, and A. Vania, "Modeling of the dynamic response of a Francis turbine," *Mechanical Systems and Signal Processing*, vol. 29, pp. 107–119, 2012.
- [13] A. W. Guo and J. D. Yang, "Self-tuning PID control of hydro-turbine governor based on genetic neural networks," in *Advances in Computation and Intelligence*, vol. 4683 of *Lecture Notes in Computer Science*, pp. 520–528, 2007.
- [14] C. K. Sanathanan, "Accurate low order model for hydraulic turbine-penstock," *IEEE Transactions on Energy Conversion*, vol. 2, no. 2, pp. 196–200, 1966.
- [15] L. N. Hannett, J. W. Feltes, and B. Fardanesh, "Field tests to validate hydro turbine-governor model structure and parameters," *IEEE Transactions on Power Systems*, vol. 9, no. 4, pp. 1744–1751, 1994.
- [16] N. Kishor, R. P. Saini, and S. P. Singh, "A review on hydropower plant models and control," *Renewable and Sustainable Energy Reviews*, vol. 11, no. 5, pp. 776–796, 2007.
- [17] N. Kishor, S. P. Singh, and A. S. Raghuvanshi, "Dynamic simulations of hydro turbine and its state estimation based LQ control," *Energy Conversion and Management*, vol. 47, no. 18-19, pp. 3119–3137, 2006.
- [18] T. Bakka and H. R. Karimi, "H $_{\infty}$ static output-feedback control design with constrained information for offshore and turbine system," *Journal of the Franklin Institute*, vol. 350, no. 8, pp. 2244–2260, 2013.
- [19] Y. Si, H. R. Karimi, and H. Gao, "Modeling and parameter analysis of the OC3-hywind floating wind turbine with a tuned mass damper in nacelle," *Journal of Applied Mathematics*, vol. 2013, Article ID 679071, 10 pages, 2013.
- [20] T. Bakka and H. R. Karimi, "Linear parameter-varying modeling and control of an offshore wind turbine with constrained information," *IET Control Theory & Applications*, vol. 8, no. 1, pp. 22–29, 2014.
- [21] T. Bakka and H. R. Karimi, "Bond graph modeling and simulation of wind turbine systems," *Journal of Mechanical Science and Technology*, vol. 27, no. 6, pp. 1843–1852, 2013.

- [22] T. Bakka and H. R. Karimi, "Robust \mathcal{H}_∞ dynamic output feedback control synthesis with pole placement constraints for offshore wind turbine systems," *Mathematical Problems in Engineering*, vol. 2012, Article ID 616507, 18 pages, 2012.
- [23] H. R. Karimi, P. Jabedar Maralani, B. Moshiri, and B. Lohmann, "Numerically efficient approximations to the optimal control of linear singularly perturbed systems based on Haar wavelets," *International Journal of Computer Mathematics*, vol. 82, no. 4, pp. 495–507, 2005.
- [24] H. R. Karimi, "Robust synchronization and fault detection of uncertain master-slave systems with mixed time-varying delays and nonlinear perturbations," *International Journal of Control, Automation and Systems*, vol. 9, no. 4, pp. 671–680, 2011.
- [25] H. R. Karimi, "Robust H_∞ filter design for uncertain linear systems over network with network-induced delays and output quantization," *Modeling, Identification and Control*, vol. 30, no. 1, pp. 27–37, 2009.
- [26] H. R. Karimi, M. J. Yazdanpanah, R. V. Patel, and K. Khorasani, "Modeling and control of linear two-time scale systems: applied to single-link flexible manipulator," *Journal of Intelligent & Robotic Systems*, vol. 45, no. 3, pp. 235–265, 2006.
- [27] H. Fang, L. Chen, N. Dlakavu, and Z. Shen, "Basic modeling and simulation tool for analysis of hydraulic transients in hydroelectric power plants," *IEEE Transactions on Energy Conversion*, vol. 23, no. 3, pp. 834–841, 2008.
- [28] Z. Y. Shen, *Hydraulic Turbine Regulation*, China Waterpower Press, Beijing, China, 1998.
- [29] J. I. Inayat-Hussain, "Nonlinear dynamics of a statically misaligned flexible rotor in active magnetic bearings," *Communications in Nonlinear Science and Numerical Simulation*, vol. 15, no. 3, pp. 764–777, 2010.
- [30] D. J. Ling and Y. Tao, "An analysis of the Hopf bifurcation in a hydroturbine governing system with saturation," *IEEE Transactions on Energy Conversion*, vol. 21, no. 2, pp. 512–515, 2006.
- [31] D. N. Konidaris and J. A. Tegopoulos, "Investigation of oscillatory problems of hydraulic generating units equipped with Francis turbines," *IEEE Transactions on Energy Conversion*, vol. 12, no. 4, pp. 419–425, 1997.
- [32] D. Chen, C. Ding, X. Ma, P. Yuan, and D. Ba, "Nonlinear dynamical analysis of hydro-turbine governing system with a surge tank," *Applied Mathematical Modelling*, vol. 37, no. 14-15, pp. 7611–7623, 2013.
- [33] B. D. Hassard, N. D. Kazarinoff, and Y.-H. Wan, *Theory and Applications of Hopf Bifurcation*, Cambridge University Press, 1981.
- [34] J. Caff, *Applications of Center Manifold Theory*, Springer, New York, NY, USA, 1981.
- [35] E. Ott, *Chaos in Dynamical Systems*, Cambridge University Press, Cambridge, UK, 2nd edition, 2002.
- [36] M. Dehghani and S. K. Y. Nikraves, "Nonlinear state space model identification of synchronous generators," *Electric Power Systems Research*, vol. 78, no. 5, pp. 926–940, 2008.
- [37] S. Kahan and A. C. Sicardi-Schifino, "Homoclinic bifurcations in Chua's circuit," *Physica A: Statistical Mechanics and Its Applications*, vol. 262, no. 1-2, pp. 144–152, 1999.
- [38] S. Yin, S. X. Ding, A. Haghani, H. Hao, and P. Zhang, "A comparison study of basic data-driven fault diagnosis and process monitoring methods on the benchmark Tennessee Eastman process," *Journal of Process Control*, vol. 22, no. 9, pp. 1567–1581, 2012.

Research Article

Performance of a Nonlinear Real-Time Optimal Control System for HEVs/PHEVs during Car Following

Kaijiang Yu and Junqi Yang

School of Electrical Engineering and Automation, Henan Polytechnic University, Jiaozuo 454000, China

Correspondence should be addressed to Kaijiang Yu; yu.kaijiang@163.com

Received 20 March 2014; Accepted 23 May 2014; Published 19 June 2014

Academic Editor: Yuxin Zhao

Copyright © 2014 K. Yu and J. Yang. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

This paper presents a real-time optimal control approach for the energy management problem of hybrid electric vehicles (HEVs) and plug-in hybrid electric vehicles (PHEVs) with slope information during car following. The new features of this study are as follows. First, the proposed method can optimize the engine operating points and the driving profile simultaneously. Second, the proposed method gives the freedom of vehicle spacing between the preceding vehicle and the host vehicle. Third, using the HEV/PHEV property, the desired battery state of charge is designed according to the road slopes for better recuperation of free braking energy. Fourth, all of the vehicle operating modes engine charge, electric vehicle, motor assist and electric continuously variable transmission, and regenerative braking, can be realized using the proposed real-time optimal control approach. Computer simulation results are shown among the nonlinear real-time optimal control approach and the ADVISOR rule-based approach. The conclusion is that the nonlinear real-time optimal control approach is effective for the energy management problem of the HEV/PHEV system during car following.

1. Introduction

In recent years, hybrid electric vehicles (HEVs) and plug-in hybrid electric vehicles (PHEVs) have become a research hotspot due to the rising price of fossil fuels and environmental problems. HEVs and PHEVs (referred as XEVs) use a battery to add an extra degree of freedom to the power sources. It can downsize the internal combustion engine, optimize the engine operating point, use the battery electricity, and regenerate dissipation kinematic energy during deceleration, which help to improve fuel economy and reduce emissions [1, 2].

A lot of works have been published on the energy management problem of HEV systems. These approaches are typical in a family of optimal control techniques. They can be subdivided into four categories: numerical optimization, analytical optimal control theories, instantaneous optimization, and heuristic control techniques [3]. The most representative of numerical optimization is dynamic programming (DP) [3, 4]. However DP is based on fixed speed patterns which are impossible to get in reality. A kind of analytical optimal control techniques is Pontryagin's minimum principle [5].

It gives necessary conditions that the optimal solution must satisfy. It also needs to know the entire driving cycle in advance. The instantaneous optimization includes the equivalent consumption minimization strategy (ECMS) [3, 6]. It is based on instantaneous optimization and is easy to implement in real-time. However, it cannot guarantee the optimality over the whole driving cycle. Heuristic control techniques like rule-based control strategies are robust, but they are impossible to guarantee the optimality. In [7], dynamic programming, quadratic programming, and model predictive control (MPC) solutions of HEV energy management problems were presented. A model predictive control approach was used to investigate the energy management problem of a power-split HEV over standard driving cycles in [8]. A new charge/discharge control system for hybrid electric vehicles based on the use of car navigation information was proposed in [9].

The literature related to PHEV energy management problems provides a lot of approaches using the ideas of modelling and controlling the powertrain components for better fuel economy. A stochastic optimal control approach for power management in plug-in hybrid electric vehicles was

proposed by [2]. Energy-optimal control of plug-in hybrid electric vehicles for real-world driving cycles was proposed by [10]. The Gipps car following model was applied to the local road trip modeling of plug-in hybrid electric vehicle power management using historical traffic data on flat roads in [11]. The authors of [12] proposed a new approach to optimal power management of plug-in hybrid electric vehicles in the charge-depletion mode with driving cycle modeling based on the historic traffic information. Both the dynamic programming and the equivalent consumption minimization strategy (see [13–15]) were utilized to optimize the battery state of charge (SOC) profile with terrain, vehicle speed, and trip distance information for a PHEV power management problem in [16].

Although model predictive control is also in the numerical optimization class, its advantage is its predictive nature. The method can use road traffic information in the near future [9] and be applicable to the unfixed speed pattern [17]. Based on a simple and accurate model of the system, MPC can provide real-time control for the system. This paper examines energy management problems of both HEVs and PHEVs. The fuel economy optimization results of XEVs rely strongly on the future road load. The battery SOC can be scheduled optimally using the future road load.

When the slope information in advance is predicted, the battery can be depleted slowly in the PHEV case or be charged up in the HEV case, before the upslope. So the XEV can make best use of the battery charged power to assist the vehicle driving. Then the battery SOC is reduced to be prepared for the upcoming downhill battery recharging. At last the battery is charged up by the free regenerative braking energy. These make better use of the high efficiency points of the engine and the regenerative braking energy. The preceding vehicle is assumed to be equipped with an ecological driver-assistance system developed by the authors of [17] and is controlled by an HEV energy management nonlinear model predictive controller proposed by the authors of [18, 19]. Since ordinary drivers have some intelligence, it is more reasonable to assume an eco-driving preceding vehicle (PV) than a conventional proportional-integral controlled preceding vehicle. For example, a driver will accelerate the vehicle before the up slope and decelerate the vehicle before the down slope to make good use of the vehicle inertia kinetic energy. This intelligent driver behavior was realized by [18]. When this preceding vehicle eco-driving behavior is predicted, the following vehicle can schedule the speed and the vehicle spacing optimally using a nonlinear real-time optimal control approach. Especially for MPC, the future road load can be incorporated in the predictive model to better optimize the future speed profile and energy use. In other words, a decentralized nonlinear real-time optimal control system can be developed to model more real vehicle driving situations to get better fuel economy.

Recently, the vehicle GPS-based navigation technology, digital map databases, and laser sensors have been developed quickly. Prediction of future vehicle road loads like road slopes and preceding vehicle position and speed is becoming realistic, which was impossible in three decades ago. Research on look-ahead control using the GPS road

slope information for the fuel optimization of a conventional powertrain heavy truck was accomplished in [20]. Both the dynamic programming and the equivalent consumption minimization strategy were utilized to optimize the battery state of charge profile with terrain information for an HEV power management problem in [6]. For adaptive cruise control (ACC), the emphasis is on safely increasing driving comfort rather than increasing road capacity. Therefore, normally a constant headway or other safe following policies are used to determine the following distance [21]. In [22], two different longitudinal control policies for automatically controlled vehicles were investigated. One was based on maintaining a constant spacing between the vehicles while the other was based upon maintaining a constant headway (or time) between successive vehicles.

In our work [18, 19], a nonlinear model predictive control algorithm using a simplified model for a power-split HEV was proposed to optimize the fuel economy. We assumed that the engine always worked along its optimal operating line which was an industrial traditional energy management strategy for commercially available HEVs in [23]. In this paper, we release the above general rule to search whole areas of the engine fuel consumption map for better fuel economy. This work differs from our previous work in that it presents a real-time implementable algorithm using slope and traffic information under a model predictive control structure.

The rest of this paper is organized as follows. In Section 2, the nonlinear plant model is derived. Section 3 formulates the nonlinear real-time optimal control algorithm. Section 4 presents comparative simulation results among the nonlinear real-time optimal control approach, the dynamic programming approach, the Pontryagin's minimum principle approach, and the ADVISOR [24] rule-based approach (see [25]). Section 5 provides conclusions.

2. Nonlinear Plant Model

The configuration of the power-split HEV/PHEV system is shown in Figure 1. The power-split device is the key component of the power-split HEV/PHEV system and has both functionality of a speed coupler and CVT. There are five dynamic components: the engine, the battery, two motor/generators (M/Gs), and the wheels in this power-split HEV/PHEV system. The power-split device property which reveals the torque and speed relationships among the engine, M/Gs, and the road load can be expressed as follows [4, 26]:

$$\begin{aligned}
 I_{M/G1}\dot{\omega}_{M/G1} &= \tau_{M/G1} + fS, \\
 \left(I_{M/G2} + \frac{I_w}{g_f^2} + m\frac{r_w^2}{g_f^2} \right) \dot{\omega}_{M/G2} \\
 &= \tau_{M/G2} - \frac{\tau_{resist} + \tau_{brake}}{g_f} + fR, \\
 I_{eng}\dot{\omega}_{eng} &= \tau_{eng} - f(R + S),
 \end{aligned} \tag{1}$$

$$\tau_{resist} = r_w mg (\mu \cos(\theta) + \sin(\theta)) + \frac{1}{2} \rho C_D A r_w v^2, \tag{2}$$

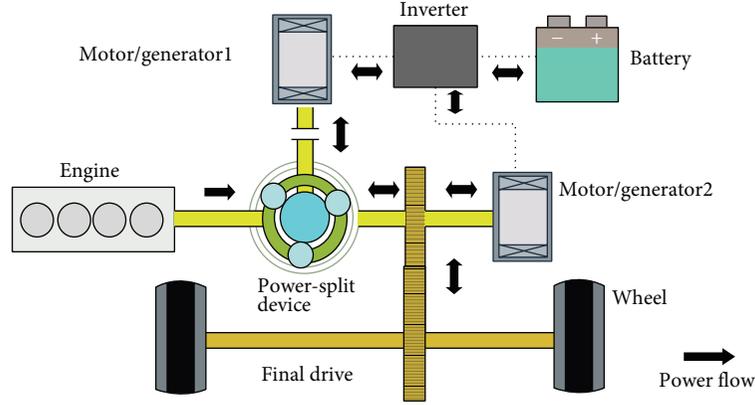


FIGURE 1: Model of the HEV/PHEV. Diagram adapted from [4].

where S and R are the number of the sun gear and the ring gear teeth, respectively; $\tau_{M/G1}$, $\tau_{M/G2}$, τ_{resist} , τ_{brake} , and τ_{eng} are the torque of M/G1, M/G2, the vehicle resistance, the friction brake, and the engine, respectively; $\omega_{M/G1}$, $\omega_{M/G2}$, and ω_{eng} are the angular velocities of M/G1, M/G2, and the engine, respectively; g_f is the final drive gear ratio; v is the vehicle speed; $I_{M/G1}$, $I_{M/G2}$, I_w , and I_{eng} are the inertia of M/G1, M/G2, the wheels, and the engine, respectively; r_w is the wheel radius; f is the internal force of the power-split device on the pinion gears; ρ , C_D , A , m , g , μ , and θ are the air density, the air drag coefficient, the frontal area of the vehicle, the vehicle mass, the gravity acceleration, the rolling resistance coefficient, and the road grade, respectively.

The slope information at location p is approximated by the sigmoid functions piecewise linearly as follows:

$$\theta(p) = \frac{s_{10}}{1 + e^{(s_{30}(p-s_{20}))}} + \frac{s_{11}}{1 + e^{(s_{31}(p-s_{21}))}} + \dots, \quad (3)$$

where s_{10} and s_{11} decide the gradient angle. Parameters s_{20} and s_{21} decide the position when the gradient angle is changed. Parameters s_{30} and s_{31} are slope shaped parameters.

The relationships among the speed of the powertrain components are given as

$$\begin{aligned} S\omega_{M/G1} + R\omega_{M/G2} - (S + R)\omega_{eng} &= 0, \\ \omega_{M/G2} &= \frac{g_f}{r_w}v. \end{aligned} \quad (4)$$

The power-balancing constraint needs to be considered. Since the power is a multiplication of the torque and the speed, the torque balance is presented in (1); the power-balancing constraint is addressed identically. The power-balancing can be implicitly considered.

Using (1) and (4) and eliminating the interaction force f , we can obtain the dynamics of the engine and M/G2:

$$\begin{bmatrix} I_{eng} + \left(\frac{S+R}{S}\right)^2 I_{M/G1} & -\frac{R(S+R)}{S^2} I_{M/G1} \\ -\frac{R(S+R)}{S^2} I_{M/G1} & I_{M/G2} + \frac{I_w}{g_f^2} + \left(\frac{R}{S}\right)^2 I_{M/G1} + m\frac{r_w^2}{g_f^2} \end{bmatrix} \times \begin{bmatrix} \dot{\omega}_{eng} \\ \dot{\omega}_{M/G2} \end{bmatrix} = \begin{bmatrix} \tau_{eng} + \frac{S+R}{S}\tau_{M/G1} \\ \tau_{M/G2} - \frac{R}{S}\tau_{M/G1} - \frac{\tau_{resist} + \tau_{brake}}{g_f} \end{bmatrix}. \quad (5)$$

We evaluate the fuel consumption rate using the Willan's line method [27]. It uses the function of the engine speed and torque to approximate the engine fuel consumption rate map directly, which leads to more accurate results than those of polynomial approximations. The fuel consumption rate can be expressed as follows [28]:

$$\dot{m}_f = \frac{a\tau_{eng}\omega_{eng} + b\omega_{eng} + c\omega_{eng}^3}{h + k\omega_{eng} + l\omega_{eng}^2}, \quad (6)$$

where \dot{m}_f is the fuel consumption rate; a , b , c , h , k , and l are constant parameters.

Based on the previous analysis, the system dynamics is reduced to the battery dynamics, the engine dynamics, and the vehicle dynamics. The nonlinear system model is then represented by

$$\dot{x} = f(x, u),$$

$$x = [\omega_{eng} \quad p \quad \omega_{M/G2} \quad x_{SOC}]^T,$$

$$u = [\tau_{eng} \quad \tau_{M/G2} \quad \tau_{M/G1} \quad \tau_{brake}]^T,$$

$$f(x, u) = \begin{bmatrix} M\tau_{\text{eng}} + N\tau_{\text{M/G2}} + \left(M\frac{S+R}{S} - N\frac{R}{S} \right) \tau_{\text{M/G1}} - N\frac{\tau_{\text{resist}} + \tau_{\text{brake}}}{g_f} \\ \frac{r_w}{g_f} \omega_{\text{M/G2}} \\ N\tau_{\text{eng}} + P\tau_{\text{M/G2}} + \left(N\frac{S+R}{S} - P\frac{R}{S} \right) \tau_{\text{M/G1}} - P\frac{\tau_{\text{resist}} + \tau_{\text{brake}}}{g_f} \\ -\frac{V_{\text{OC}} - \sqrt{V_{\text{OC}}^2 - 4P_{\text{batt}}R_{\text{batt}}}}{2R_{\text{batt}}Q_{\text{batt}}} \end{bmatrix},$$

$$\begin{bmatrix} M & N \\ N & P \end{bmatrix}^{-1} = \begin{bmatrix} I_{\text{eng}} + \left(\frac{S+R}{S} \right)^2 I_{\text{M/G1}} & -\frac{R(S+R)}{S^2} I_{\text{M/G1}} \\ -\frac{R(S+R)}{S^2} I_{\text{M/G1}} & I_{\text{M/G2}} + \frac{I_w}{g_f^2} + \left(\frac{R}{S} \right)^2 I_{\text{M/G1}} + m\frac{r_w^2}{g_f^2} \end{bmatrix}^{-1}, \quad (7)$$

where p is the vehicle position; x_{SOC} is the battery SOC; V_{OC} , R_{batt} , and Q_{batt} are the battery open circuit voltage, the battery internal resistance, and the battery capacity, respectively; and M , N , and $P \in \mathbb{R}$. We use the vehicle position and the vehicle speed to represent the vehicle dynamics. The battery power P_{batt} is governed by

$$P_{\text{batt}} = \tau_{\text{M/G1}} \omega_{\text{M/G1}} + \tau_{\text{M/G2}} \omega_{\text{M/G2}}. \quad (8)$$

The motor/generator energy loss is ignored due to the two motor/generators high efficiency.

Since an approximate continuous and differentiable mathematical engine model is needed for the model predictive control algorithm, the nonlinear engine fuel model is used as above. However, as for the fuel economy evaluation in the ultimate simulation, the high fidelity engine model which is a lookup table that provides the engine fuel rate and efficiency as a function of instantaneous engine speed and engine torque is used. The configuration of the HEV/PHEV in this work can stop idling of the engine using the electric CVT. The efficiency of other components of the HEV/PHEV system like the CVT and the power electronics is assumed to be 1 due to their high efficiency. The approximate models of the engine and the battery for optimal control are used to integrate with the high fidelity engine map model in the simulation.

The nonlinear model of the power-split HEV/PHEV with slope information includes the vehicle speed dynamics. The vehicle speed profile can be optimized with this nonlinear model.

3. Nonlinear Real-Time Optimal Control

The optimal control problem is defined as

$$\begin{aligned} \min_u \quad & J = \int_t^{t+T} L(x(\tau|t), u(\tau|t)) d\tau \\ \text{subject to} \quad & \tau_{\text{M/G2 min}} \leq \tau_{\text{M/G2}}(\tau|t) \leq \tau_{\text{M/G2 max}} \\ & \tau_{\text{M/G1 min}} \leq \tau_{\text{M/G1}}(\tau|t) \leq \tau_{\text{M/G1 max}} \\ & 0 \leq \tau_{\text{brake}}(\tau|t) \leq \tau_{\text{brakemax}}, \end{aligned} \quad (9)$$

where T is the prediction horizon; $\tau_{\text{M/G2 max}}$, $\tau_{\text{M/G2 min}}$, $\tau_{\text{M/G1 max}}$, $\tau_{\text{M/G1 min}}$, and τ_{brakemax} denote the bounds of the control inputs.

The following objectives are considered in this optimal control problem.

L_x : the fuel consumption is minimized.

L_y : the vehicle deceleration or acceleration is moderated.

L_z : the vehicle speed is kept near to its desired value.

L_d : the battery SOC is kept near to its desired value. This is one of the cores of the proposed approach. We adapt the battery energy to the vehicle future energy requirements by setting the desired battery SOC as a function of road slopes which represent the main part of the future road load.

L_e : the battery SOC constraint is kept satisfied.

L_f : the engine speed constraint is kept satisfied.

L_g : the M/G2 speed constraint is kept satisfied.

L_h : the mechanical brake use is minimized.

L_i : the M/G1 speed constraint is kept satisfied.

L_j : the battery power constraint is kept satisfied.

L_k : the engine torque constraint is kept satisfied.

L_l : the following distance constraint is kept satisfied. This is also one of the cores of the proposed approach. The following distance constraint is kept in a predictive controller structure. The host vehicle maneuvers are independent of the preceding vehicle. The proposed approach does not require intervehicle communication. The following distance is varied above the minimum following distance, which improves the freedom of eco-driving car following control to optimize the driving profile for better fuel economy.

L_m : the battery power use is moderated.

The cost function L is defined as follows:

$$\begin{aligned} L = & w_x L_x + w_y L_y + w_z L_z + w_d L_d + w_e L_e + w_f L_f \\ & + w_g L_g + w_h L_h + w_i L_i + w_j L_j + w_k L_k \\ & + w_l L_l + w_m L_m, \end{aligned}$$

$$L_x = \dot{m}_f,$$

$$L_y = \frac{1}{2} \left(\frac{r_w}{g_f} \dot{\omega}_{\text{M/G2}} + g \sin(\theta) \right)^2,$$

$$\begin{aligned}
L_z &= \frac{1}{2}(v - v_d)^2, \\
L_e &= -\ln(x_{\text{SOC}} - \text{SOC}_{\min}) - \ln(\text{SOC}_{\max} - x_{\text{SOC}}), \\
L_f &= -\ln(\omega_{\text{eng}}) - \ln(\omega_{\text{eng max}} - \omega_{\text{eng}}), \\
L_g &= -\ln(\omega_{\text{M/G2}}) - \ln(\omega_{\text{M/G2 max}} - \omega_{\text{M/G2}}), \\
L_h &= \frac{1}{2}(\tau_{\text{brake}})^2, \\
L_i &= -\ln(\omega_{\text{M/G1}} - \omega_{\text{M/G1 min}}) - \ln(\omega_{\text{M/G1 max}} - \omega_{\text{M/G1}}), \\
L_j &= -\ln(P_{\text{batt}} - P_{\text{batt min}}) - \ln(P_{\text{batt max}} - P_{\text{batt}}), \\
L_k &= -\ln(\tau_{\text{eng}}) - \ln(\tau_{\text{eng max}} - \tau_{\text{eng}}), \\
L_m &= P_{\text{batt}}.
\end{aligned} \tag{10}$$

The parameter v_d is the desired vehicle speed. Fuel economy is affected by the steady cruise speed. The best fuel economy occurs at a steady speed around 60 km/h. This speed is chosen as the desired vehicle speed. The parameters $w_x, w_y, w_z, w_d, w_e, w_f, w_g, w_h, w_i, w_j, w_k,$ and w_l are the weights. The parameters $\text{SOC}_{\min}, \text{SOC}_{\max}, \omega_{\text{eng max}}, \omega_{\text{M/G2 max}}, \omega_{\text{M/G1 min}}, \omega_{\text{M/G1 max}}, P_{\text{batt min}}, P_{\text{batt max}},$ and $\tau_{\text{eng max}}$ denote the bounds of the parameters. The parameter $\tau_{\text{eng max}}$ is a state variant control input. It is a function of the engine speed.

The last term concerning the battery power use is inspired by the equivalent consumption minimization strategy. The equivalent consumption minimization strategy assumes that the current battery energy use will cost the same amount of fuel energy in the future as it does in the current driving conditions [1]. So this term can control the speed of battery charge depleting.

The battery energy is made better use of by the following term:

$$L_d = \frac{1}{2}(x_{\text{SOC}} - \text{SOC}_d(p))^2, \tag{11}$$

where $\text{SOC}_d(p)$ is the desired battery SOC value. It is designed using the road slope information beforehand to make better use of the battery energy.

The vehicle spacing optimization term is as follows:

$$L_l = -\ln(p_p(t) + v_p(t)(\tau - t) - p(t) - l_p - d_{\min}), \tag{12}$$

where the parameters $p_p, v_p, l_p,$ and d_{\min} denote the preceding vehicle position, the preceding vehicle speed, the preceding vehicle length, and the minimum vehicle spacing. The preceding vehicle speed in the prediction horizon is assumed to be constant, and its value is the same as the observed preceding vehicle speed at the beginning of the nonlinear real-time optimal control algorithm. In this way, the vehicle spacing can be kept above the minimum vehicle spacing in the prediction horizon.

Both the engine operating point and the driving profile can be optimized using the above performance index. The

optimal driving profile can be decided in real-time, which is different from the predetermined profile in other papers.

In [29–31], the authors stated that the conventional engine OOL idea is valid only if the power transmission loss is negligible or if it shows only a mild change throughout the operation condition. However, HEVs have far more complicated and irregular power transmission mechanisms and characteristics than conventional vehicles do. This complexity is mainly due to the electrical power transmission paths which involve nonlinear power conversion losses in M/Gs. In other words, HEVs have an energy buffer like batteries whose efficiency is highly nonlinear to the input road loads. The battery can utilize the free regenerative braking energy to improve fuel economy significantly. The engine optimal operation for HEVs corresponding to the system optimality needs to be reconsidered. For real-time implementation, the fuel model needs to be continuous and differentiable. The fuel model needs simplifications and also to be accurate enough. It is different from the engine map model which cannot be predictable and implemented in real-time. As for the physical constraint, it is guaranteed by the second term of (10). The discontinuous jump of the engine speed will cause the discontinuous fly of the vehicle speed because of the planetary set. This will lead a very large punitive value of the vehicle acceleration which is included in the second term of (9).

The log barrier functions are introduced as penalizing terms for violations of the state constraints and the state variant control input constraint. The value of the performance index becomes very large when the constraints are being violated. By doing so, the state constraints and the state variant control input constraint of the system are satisfied. The general rule that the engine always works along its optimal operating line does not promise optimal fuel economy. Due to lack of future road load information, the engine may work in the low efficiency parts of the engine optimal operating line. We do not follow the industrial tradition which assumes that the engine always works along its optimal operating line in the commercially available HEV/PHEV energy management strategy. The fuel economy is optimized using the only term concerning the fuel consumption rate in the cost function. It will search the whole areas of the engine fuel consumption map for better fuel economy. In this way we want to develop the full strength of HEVs/PHEVs.

The inequality constraint in the optimal control problem is converted to an equality constraint by introducing a dummy input u_d for computation simplicity as follows:

$$C(x(t), u(t)) = u^2(t) + u_d^2(t) - u_{\max}^2 = 0, \tag{13}$$

where u_{\max} denotes the bound of the control input.

To solve this optimal control problem with the calculus of variation method [32], the Hamiltonian function is defined by

$$H(x, u, \lambda, \psi) = L(x, u) + \lambda^T f(x, u) + \psi^T C(x, u), \tag{14}$$

where λ denotes the costate and ψ denotes the Lagrange multiplier associated with the equality constraint.

The first-order necessary conditions for the optimal control input u , the multiplier ψ , and the costate λ are obtained using the calculus of variation as

$$\begin{aligned} \dot{x} &= f(x, u), & x(t_0) &= x_0, \\ \dot{\lambda} &= -\frac{\partial H}{\partial x}, & \lambda(t+T) &= 0, \\ \frac{\partial H}{\partial u} &= 0, \\ C(x, u) &= 0, \end{aligned} \quad (15)$$

where t_0 is the initial time and x_0 is the initial state.

The derivative of the costate λ concerning the slope information and the battery SOC is obtained as

$$\begin{aligned} \dot{\lambda}_2 &= -\lambda_1 \frac{\partial \dot{w}_{\text{eng}}}{\partial p} - \lambda_3 \frac{\partial \dot{w}_{M/G2}}{\partial p} - w_y \frac{\partial L_y}{\partial p} - w_l \frac{\partial L_l}{\partial p} \\ &\quad - w_d \frac{\partial L_d}{\partial p}, \\ \dot{\lambda}_4 &= -w_d (x_{\text{SOC}} - \text{SOC}_d) \\ &\quad - w_e \left(\frac{1}{\text{SOC}_{\text{max}} - x_{\text{SOC}}} - \frac{1}{x_{\text{SOC}} - \text{SOC}_{\text{min}}} \right). \end{aligned} \quad (16)$$

It reveals that the costate of the vehicle position is related to the two power devices, the terms concerning the vehicle acceleration or deceleration, the desired battery SOC, and the vehicle spacing in the cost function. And the battery SOC costate is affected by the battery desired SOC and the bounds of the battery SOC. A large costate will lead to the small variation of the battery SOC. A small costate will lead to the large variation of the battery SOC. A well-tuned performance index and a set of well-tuned weights can lead to a better system.

The structure of the nonlinear model predictive control system is shown in Figure 2. The system inputs contain the control inputs. The system outputs consist of the vehicle states. The predictive controller uses terrain information from a digital map to calculate $\text{SOC}_d(p)$ and $\theta(p)$. The energy management problem can be viewed as an optimal control problem which is addressed here using a nonlinear real-time optimal control approach.

At each time t , the optimal control input is computed by solving the above optimal control problem during the prediction horizon T . Only the first element of the optimal control sequence is applied. At the next time step, the prediction horizon moves forward and the process is repeated [33].

4. Computer Simulations

4.1. Comparison Controllers. There are four simulations in this work. They are the vehicle tracking nonlinear real-time optimal control (ROC) approach for the HEV, the vehicle tracking ADVISOR rule-based (TRB) approach for the HEV, the vehicle tracking nonlinear real-time optimal control

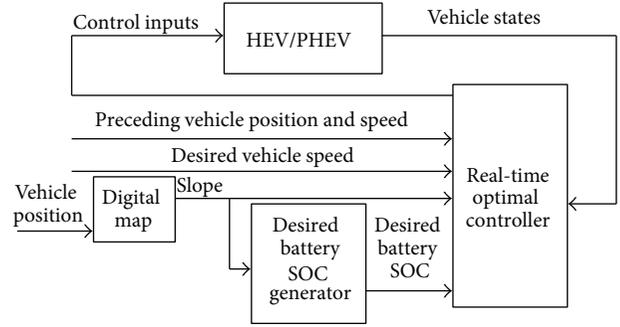


FIGURE 2: Structure of the nonlinear real-time optimal control system.

approach for the PHEV, and the vehicle tracking ADVISOR rule-based approach for the PHEV. The two vehicle tracking nonlinear real-time optimal control approaches for HEVs/PHEVs are the proposed approaches. The rule-based control for HEVs introduces a set of rules to decide the power split between the engine and the battery after the vehicle states are observed. The ADVISOR rule-based control approach is utilized as a comparison for the proposed vehicle tracking nonlinear real-time optimal control approach for the HEV. The PHEV all-electric charge depletion followed by charge sustenance control strategy used in the ADVISOR simulation software is utilized as a comparison for the proposed vehicle tracking nonlinear real-time optimal control approach for the PHEV. The driving pattern used in the vehicle tracking ADVISOR approach is obtained from an adaptive cruise control (ACC) method [22]. The control input of the tracking vehicle using the ACC method is calculated as follows [22]:

$$\begin{aligned} u_{\text{ACC}}(t) &= \frac{v_p(t) - v(t) - k(hv(t) + d_{\text{min}} - (p_p(t) - p(t) - l_p))}{h}, \end{aligned} \quad (17)$$

where k and h are constant parameters set as $k = 0.08$ and $h = 0.15$ for HEVs and $k = 0.08$ and $h = 0.2$ for PHEVs, respectively. The preceding vehicle is controlled using the same algorithm as that in Section 3 without the vehicle spacing cost term and the battery power use term. It is a kind of eco-driving method using the ROC approach.

4.2. Test Road Slope Profiles from a Digital Map and Calculation of the Road Slope and the Desired Battery SOC. The effectiveness of the proposed energy management system of the power-split HEV is evaluated using the slope information of a real road. It is a road from the Imajuku traffic light position to the Hatae traffic light position which is 6.2 [km] located at Route 202, Fukuoka, Japan. The maximum slope of this road is 3.65%, and the minimum slope of this road is -3.46% . This real terrain is typical in Japan where there are many hilly areas.

The effectiveness of the proposed energy management system of the power-split PHEV is evaluated using the

slope information of a real road. It is a road from the Oimatsuda Interchange to the Numazu Interchange of the Tomei Expressway which is 50.4 [km] located in Japan. The maximum slope of this road is 5.06%, and the minimum slope of this road is -4.79%.

We set the HEV desired battery SOC value according to the road elevation which represents the main part of the future road load information. It is reasonable to utilize the road elevation information since this future road load information is known already. The desired battery SOC is assumed using the function as

$$\begin{aligned} \text{SOC}_{d\text{HEV}}(p) &= k_{\text{SOCHEV}} \left(\frac{s_1}{1 + e^{(s_3(p-s_2))}} + \frac{s_4}{1 + e^{(s_6(p-s_5))}} + \dots \right) \\ &+ \text{SOC}_{k\text{HEV}}, \end{aligned} \quad (18)$$

where k_{SOCHEV} and $\text{SOC}_{k\text{HEV}}$ are constant parameters set as $k_{\text{SOCHEV}} = -2.5$ and $\text{SOC}_{k\text{HEV}} = 0.71$, respectively.

We set the PHEV desired battery SOC value according to the road elevation which represents the main part of the future road load information. The future road information can be modelled into the real-time optimal control algorithm to control the battery charge depletion for better fuel economy. The desired battery SOC is assumed using the function as

$$\begin{aligned} \text{SOC}_{d\text{PHEV}}(p) &= \frac{t}{t_{\text{sim}}} k_{\text{SOC}} \\ &+ k_{\text{SOC PHEV}} \left(\frac{s_1}{1 + e^{(s_3(p-s_2))}} + \frac{s_4}{1 + e^{(s_6(p-s_5))}} + \dots \right) \\ &+ \text{SOC}_{k\text{PHEV}}, \end{aligned} \quad (19)$$

where k_{SOC} , $k_{\text{SOC PHEV}}$, and $\text{SOC}_{k\text{PHEV}}$ are constant parameters set as $k_{\text{SOC}} = -0.7$, $k_{\text{SOC PHEV}} = -2$, and $\text{SOC}_{k\text{PHEV}} = 0.9$, respectively. t_{sim} is the total simulation time.

4.3. Simulation Conditions. In these simulations, HEV parameters are used from ADVISOR 2002 [24] Toyota Prius data. PHEV parameters are also used from ADVISOR 2002 [24] Toyota Prius data with a double battery capacity. The simulation parameters are shown in Table 1. The nonlinear real-time optimal control problem is solved using the numerical computation method: the continuation and generalized minimum residual (C/GMRES) method [34]. The C/GMRES method uses forward difference approach and discretizes the HEV/PHEV plant with a sampling interval h_t to implement the nonlinear real-time optimal control algorithm. The flowchart of the nonlinear model predictive control algorithm implementation is shown in Figure 3. A brief description of the solution of the real-time optimal control problem using the C/GMRES method is included in the Appendix. The nonlinear real-time optimal control algorithm is realized by utilizing the C MEX S-function

TABLE 1: Simulation parameters.

Parameters	Values
m	1504
C_D	0.3
g	9.8 [m/s ²]
V_{OCHEV}	307.9 [V]
V_{OCPHEV}	615.8 [V]
Q_{batt}	6 [Ah]
g_f	3.93
I_{MG1}	0.0226 [kgm ²]
I_w	3.3807 [kgm ²]
b	1576
h	1032000
l	-2.401
R	78
l_p	4.31 [m]
h_t	0.1 [s]
$\text{SOC}_{\text{minHEV}}$	0.6
$\text{SOC}_{\text{minPHEV}}$	0.15
τ_{MG2max}	305 [Nm]
τ_{MG1max}	55 [Nm]
τ_{brakemax}	2655 [Nm]
$P_{\text{battmaxHEV}}$	23.684 [kW]
$P_{\text{battminPHEV}}$	-47.367 [kW]
ω_{engmax}	418.8790 [rad/s]
ω_{MG1max}	575.9587 [rad/s]
$w_{y\text{HEV}}$	9000
$w_{d\text{HEV}}$	50000000
$w_{f\text{HEV}}$	10000
$w_{h\text{HEV}}$	1
$w_{j\text{HEV}}$	0.1
$w_{l\text{HEV}}$	90
$w_{x\text{PHEV}}$	19000
$w_{z\text{PHEV}}$	150
$w_{e\text{PHEV}}$	1
$w_{g\text{PHEV}}$	50000
$w_{l\text{PHEV}}$	0.1
$w_{k\text{PHEV}}$	1000
$w_{m\text{HEV}}$	0.00001
ρ	1.23 [kg/m ³]
A	1.746 [m ²]
μ	0.695
R_{battHEV}	1.0 [Ω]
R_{battPHEV}	2.0 [Ω]
r_w	0.287 [m]
I_{eng}	1.746 [m ²]
I_{MG2}	0.0226 [kgm ²]
a	40.88
c	-0.004051
k	365.7
S	30
T	10 [s]
d_{min}	1 [m]
SOC_d	0.7
$\text{SOC}_{\text{maxHEV}}$	0.8
$\text{SOC}_{\text{maxPHEV}}$	0.95
τ_{MG2min}	-305 [Nm]

TABLE 1: Continued.

Parameters	Values
τ_{MGlmin}	-55 [Nm]
$P_{battminHEV}$	-23.684 [kW]
v_d	60 [km/h]
$P_{battmaxPHEV}$	47.367 [kW]
ω_{MGlmin}	-575.9587 [rad/s]
w_{xHEV}	39000
w_{zHEV}	800
w_{eHEV}	260000
w_{gHEV}	50000
w_{iHEV}	0.1
w_{kHEV}	1000
w_{mHEV}	0
w_{yPHEV}	1000
w_{dPHEV}	550000
w_{fPHEV}	10000
w_{hPHEV}	1
w_{jPHEV}	0.1
w_{iPHEV}	90

builder in Matlab/Simulink. Direct control input torque of the engine, the two M/Gs, and the mechanical brake are given by the nonlinear real-time optimal controller. The fuel economy is calculated using the engine fuel consumption map which is obtained from ADVISOR 2002.

4.4. Simulation Results. Figure 4 shows the driving profile of the real-time optimal control algorithm for the HEV. The first column of Figure 4 is the road elevation. The next five columns show the vehicle control input, the optimized vehicle speed, and the vehicle spacing. Figure 5 shows the driving profile of the real-time optimal control algorithm for the PHEV. The vehicle tracking nonlinear real-time optimal control vehicle predicts the upcoming up-down hills, and avoids the abrupt acceleration or deceleration as shown in the ACC method at the link parts of different slopes. The vehicle spacing is kept above the minimum using both the real-time optimal control and the ACC. Instead of converging to a value, the vehicle tracking nonlinear real-time optimal control approach can make good use of the vehicle spacing range to get better fuel economy with the predicted preceding eco-driving vehicle information. In this way the vehicle tracking nonlinear real-time optimal control algorithm helps to improve the fuel economy.

Figure 6 shows the power-split profile of the vehicle using the vehicle tracking real-time optimal control algorithm for the HEV. The columns of Figure 6 from the top are the road elevation, the battery SOC, the speed of the engine and the two M/Gs, the torque of the engine and the two M/Gs, and the power of the engine and the two M/Gs. Figure 7 shows the power-split profile of the vehicle using the vehicle tracking ADVISOR rule-based algorithm for the HEV. Figure 8 shows the power-split profile of the vehicle using the real-time optimal control algorithm for the PHEV. Figure 9 shows the power-split profile of the vehicle using the vehicle tracking ADVISOR rule-based algorithm for the PHEV. As for the

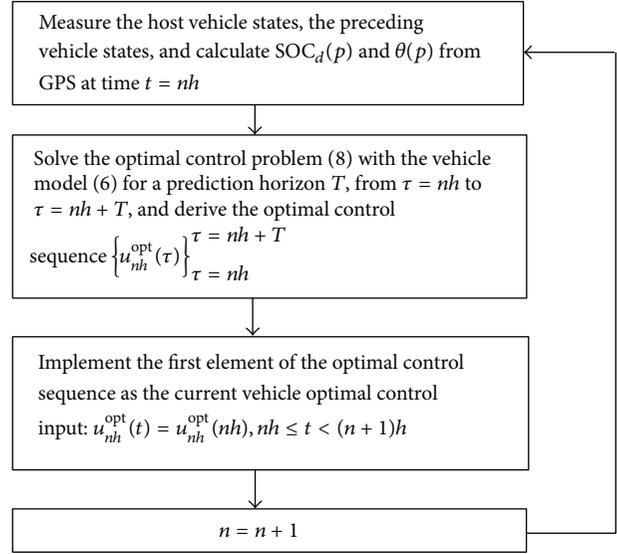


FIGURE 3: Flowchart of the nonlinear real-time optimal control algorithm.

power-split profile of the vehicle using the vehicle tracking real-time optimal control algorithm for the HEV, there is some causality between the road elevation and the battery SOC. The lowest point of the road elevation corresponds to the highest point of the battery SOC. The highest point of the road elevation corresponds to the lowest point of the battery SOC. By using the slope information in advance to better use the battery SOC range, the vehicle tracking nonlinear real-time optimal control algorithm helps to reduce the fuel consumption efficiently.

In Figure 9 the battery is depleted to its minimum allowable charge by the end of the trip. At first, the vehicle is driven by M/G2 until the battery charge is depleted to its minimum using the charge depleting mode, and then the control is switched to the charge sustaining mode. During the charge depleting mode in the above approach for the PHEV, the engine can be turned on if the driving power or torque requests exceed the capabilities of the battery or the motors. The battery SOC constraint is satisfied in the above simulations. The torque, speed, and power of the engine and M/Gs are more reasonable using the proposed real-time optimal control approach compared with the ADVISOR rule-based approach according to the commercially available Toyota Prius PHEV. There are a lot of spikes of the torque, speed, and power of the engine and M/Gs using the ADVISOR rule-based approach.

Overall, the real-time optimal control approach for the PHEV uses the M/Gs to drive the vehicle compared with the vehicle tracking real-time optimal control approach for the HEV, which helps to improve the fuel economy. The vehicle tracking real-time optimal control approach for the HEV uses the engine fuel energy to drive the vehicle instead of the free energy recovered by the battery, which results in worse fuel economy. Since the slopes in the real road for the HEV are short and gentle, the fuel economy improvement is

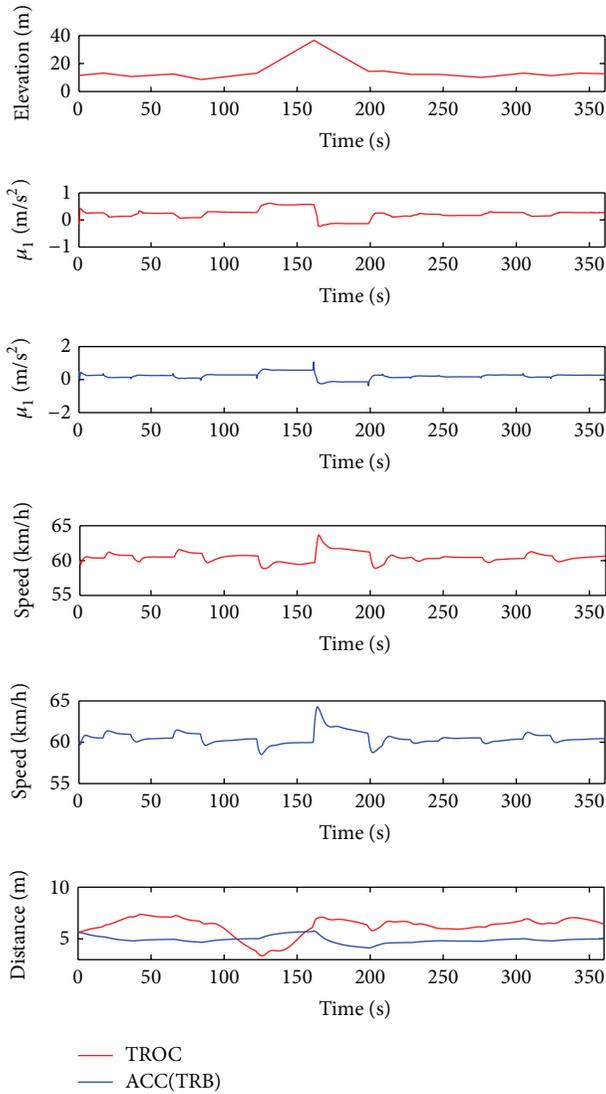


FIGURE 4: Driving profile of the vehicle tracking nonlinear real-time optimal control algorithm and the ACC algorithm for the HEV.

not significant using the vehicle tracking real-time optimal control approach algorithm for the HEV compared with that for the PHEV. Without slope previews, the engine and the M/Gs work abruptly, especially at the beginning of the simulation; and the link parts of different slopes using the ADVISOR rule-based approach. The battery SOC decreases continually. The vehicle does not get the regenerative braking energy properly.

A significant benefit of the power-split architecture is the fact that it decouples the engine crankshaft from the road and allows the electric machines to move the engine speed where fuel efficiency is maximized [35]. This is identified by the engine operating point distribution. Figure 10 shows the distribution of the engine operating points using the vehicle tracking nonlinear real-time optimal control algorithm for the HEV. Figure 11 shows the distribution of the engine

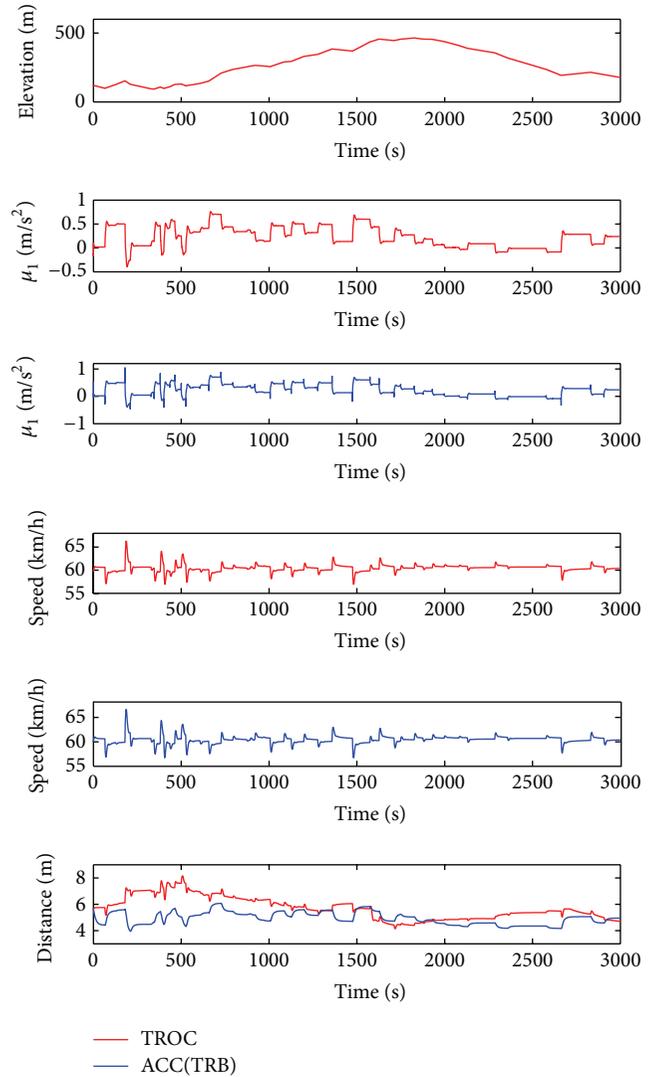


FIGURE 5: Driving profile of the vehicle tracking nonlinear real-time optimal control algorithm and the ACC algorithm for the PHEV.

operating points using the vehicle tracking nonlinear real-time optimal control algorithm for the PHEV.

The line at the top left corner is the engine max torque line. The engine operating points cannot go beyond the line.

As shown in Figures 10 and 11 the vehicle tracking nonlinear real-time optimal control algorithm operates the engine at fairly low speed and high torque, which means high engine efficiency and low brake specific fuel consumption values. The vehicle tracking nonlinear real-time optimal control algorithm forces the engine to work regularly near the engine low fuel consumption rate areas. In contrast, the ADVISOR rule-based approach operates the engine at fairly high speed and low torque, which means low engine efficiency and high brake specific fuel consumption values. By adapting the battery power to the future road load, the vehicle tracking nonlinear real-time optimal control algorithm develops the ability of the power-split architecture. We can see that the

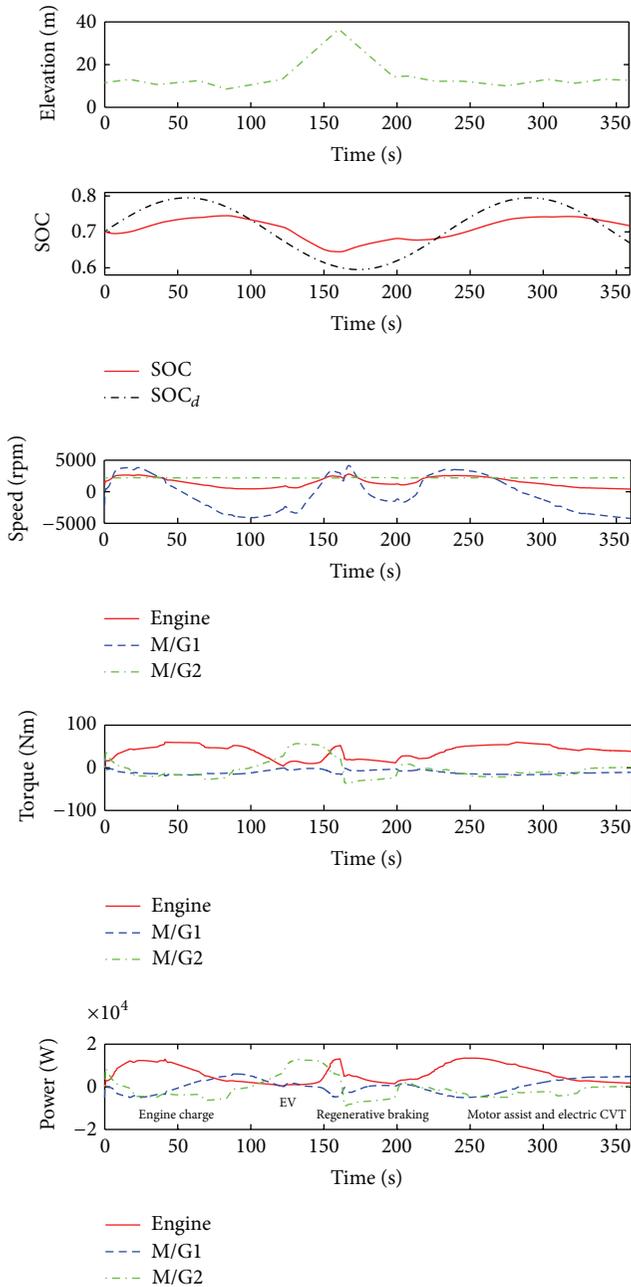


FIGURE 6: Power-split profile of the vehicle using the vehicle tracking nonlinear real-time optimal control algorithm for the HEV.

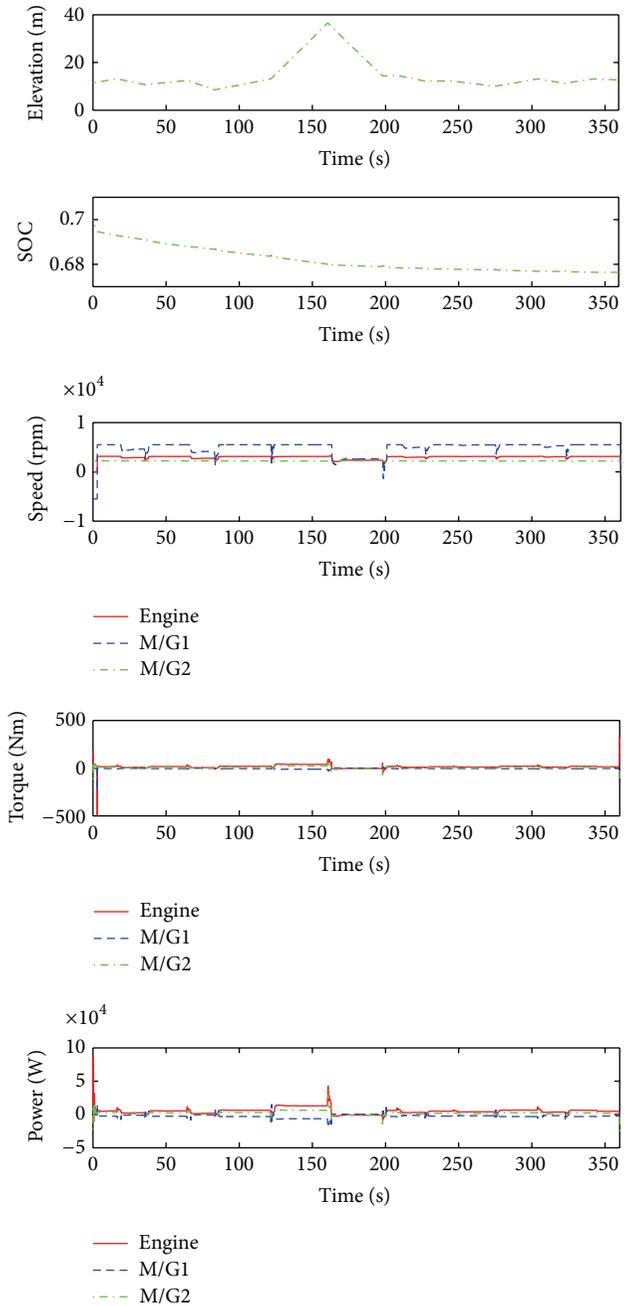


FIGURE 7: Power-split profile of the vehicle using the vehicle tracking ADVISOR rule-based algorithm for the HEV.

engine operating points of the vehicle tracking nonlinear real-time optimal control algorithm are distributed in better areas than those of the ADVISOR algorithm. Compared with the engine operating points using the vehicle tracking nonlinear real-time optimal control algorithm for the HEV, the engine operating points using the vehicle tracking nonlinear real-time optimal control algorithm for the PHEV are distributed closer to the left corner of the engine fuel consumption map that consumes less fuel. The vehicle tracking nonlinear real-time optimal control algorithm can make the engine work in better areas rather than those along the best efficiency line

of the engine using the CVT. The fuel efficiency depends on the real efficiency of the engine, which makes the point that a high efficiency area is more profitable.

The overall fuel economy results are presented in Table 2. The vehicle tracking nonlinear real-time optimal control approach for the HEV can improve fuel economy by 38.6% compared to the ADVISOR rule-based approach for the HEV. The vehicle tracking nonlinear real-time optimal control approach for the PHEV can improve fuel economy by 74.6% compared to the ADVISOR rule-based approach for the

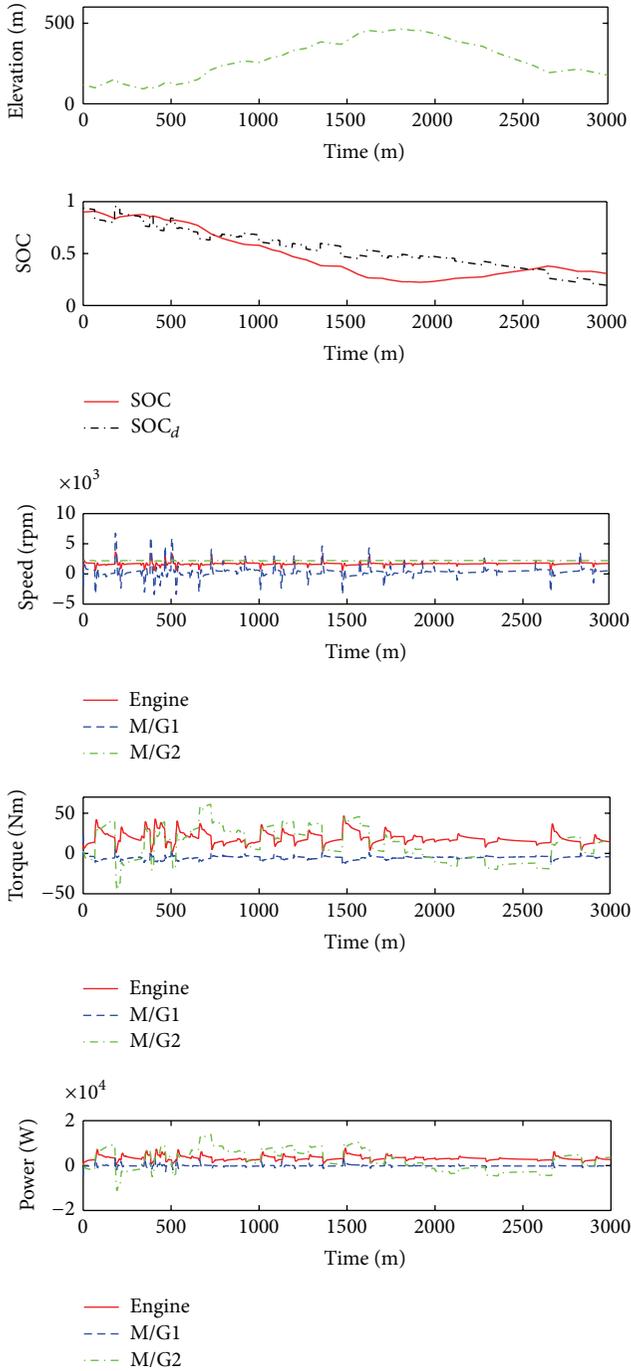


FIGURE 8: Power-split profile of the vehicle using the vehicle tracking nonlinear real-time optimal control algorithm for the PHEV.

PHEV. Since the vehicle tracking ADVISOR approach cannot avoid the acceleration or deceleration spikes at the link parts of different slopes, it gets worse fuel economy than that using the real-time optimal control approach. We can see from Figure 8 and Table 2 that better using of the battery SOC range results in better fuel economy. This leads to better fuel economy using the real-time optimal control approach for the PHEV than the real-time optimal control approach

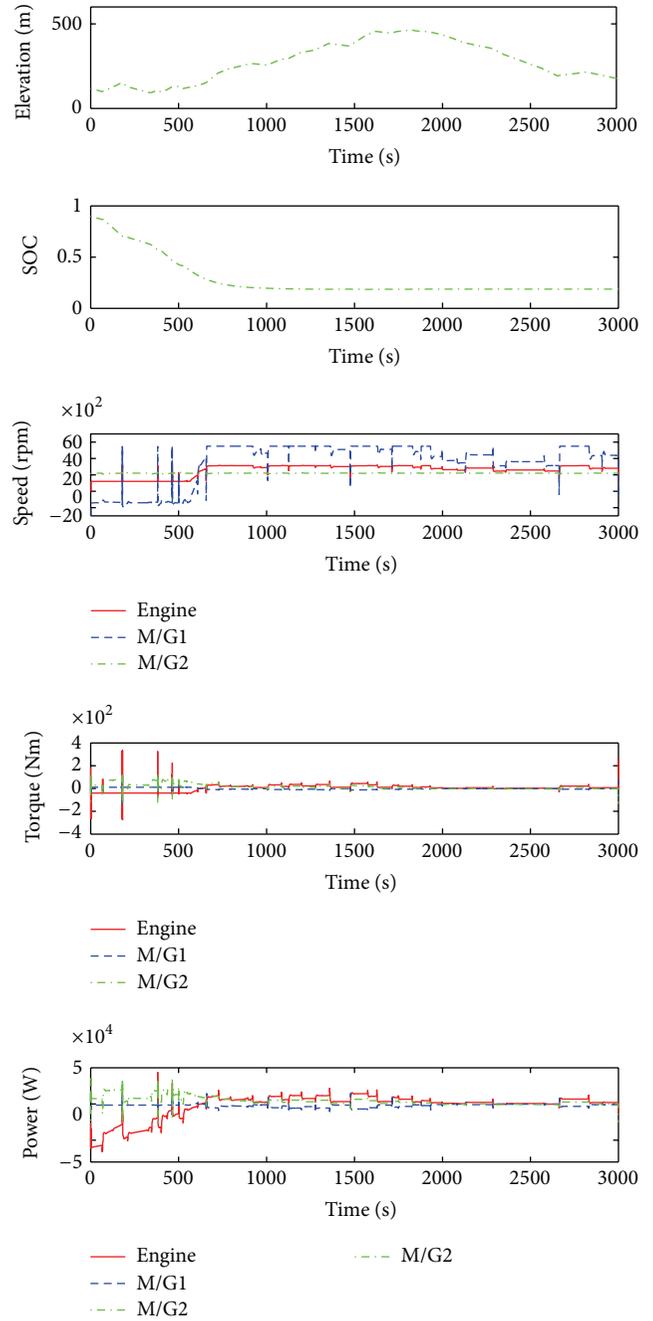


FIGURE 9: Power-split profile of the vehicle using the vehicle tracking ADVISOR rule-based algorithm for the PHEV.

for the HEV. By using the predicted road slope information freely, the real-time optimal control algorithm can adapt the HEV/PHEV battery SOC profile according to the known bounds of the parameters to get better fuel economy. Since the fuel economy is calculated by the high fidelity map of the real engine, which is the most accurate evaluating method in the computer simulation environment, these results are promising.

The proposed vehicle tracking nonlinear real-time optimal control algorithm is fast for computation. The computer

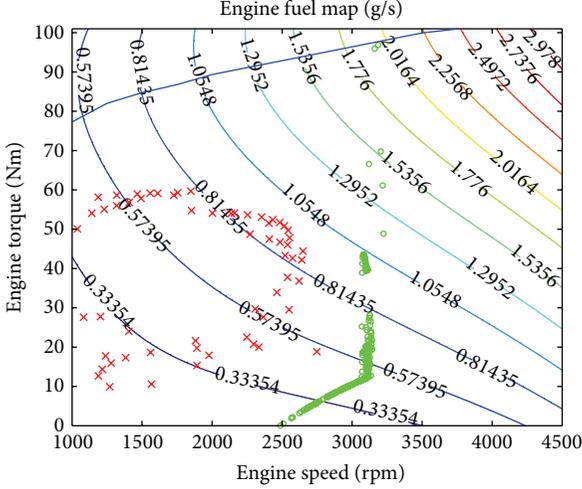


FIGURE 10: Engine operating point distribution using the vehicle tracking nonlinear real-time optimal control algorithm and the vehicle tracking ADVISOR rule-based algorithm for the HEV. The crosses and the circles denote the engine operating points of the vehicle tracking nonlinear real-time optimal control algorithm and the vehicle tracking ADVISOR rule-based algorithm, respectively.

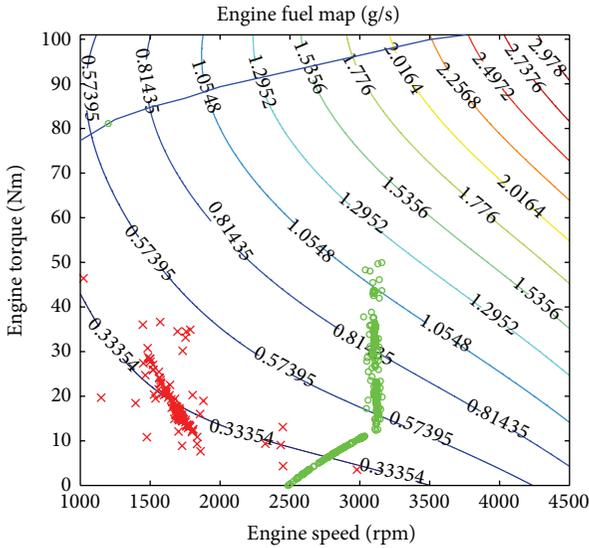


FIGURE 11: Engine operating point distribution using the vehicle tracking nonlinear real-time optimal control algorithm and the car tracking ADVISOR rule-based algorithm for the PHEV. The crosses and the circles denote the engine operating points of the vehicle tracking nonlinear real-time optimal control algorithm and the vehicle tracking ADVISOR rule-based algorithm, respectively.

simulation time for the HEV is 360 [s], and the computer simulation time for the PHEV is 3000 [s]. The computation time of the proposed vehicle tracking nonlinear real-time optimal control algorithm for the HEV is 29.0 [s]. The computation time of the proposed vehicle tracking nonlinear real-time optimal control algorithm for the PHEV is 402.7 [s]. The simulation is run in a Matlab/Simulink environment

TABLE 2: Fuel economy comparison results.

Method	Initial SOC	Final SOC	Fuel economy [km/l]
TROC (HEV)	0.700	0.717	29.1 (+38.6%)
TRB (HEV)	0.700	0.677	21.0 (+0.00%)
TROC (PHEV)	0.900	0.307	45.4 (+74.6%)
TRB (PHEV)	0.900	0.188	26.0 (+0.00%)

using a laptop with an Intel processor at 2.27 [GHz] processing speed and 2 [GB] of RAM. The sampling interval is 100 [ms]. The computation time per sampling interval of the proposed vehicle tracking nonlinear real-time optimal control algorithm for the HEV is 8.1 [ms]. The computation time per sampling interval of the proposed vehicle tracking nonlinear real-time optimal control algorithm for the PHEV is 13.4 [ms]. So we can conclude that the proposed vehicle tracking nonlinear real-time optimal control algorithm has the potential for real-time vehicle control.

5. Conclusions

A real-time optimal control approach for the energy management problem of a power-split HEV/PHEV system during car following was presented. The nonlinear system model with road slope information was developed. The validity of the proposed algorithm was demonstrated by the significant fuel economy improvements. In the future, the traffic control signals and congestion will be modelled to get better fuel economy. A non-fragile output tracking controller [36–38] for a fleet of vehicles, which guarantees the tracking error dynamics to be robustly stable, especially to design, is our top priority.

Appendix

Solution of the Real-Time Optimal Control Problem

A brief description of the solution of the real-time optimal control problem is provided as follows.

To implement the real-time optimal control algorithm, the horizon T is divided into N steps, and the optimal control problem is discretized. The general discretized optimal control problem is formulated as

$$\min_u J = \sum_{i=0}^{N-1} L(x_i(\tau | t), u_i(\tau | t)) \Delta\tau (t)$$

subject to

$$\begin{aligned} x_{i+1}(\tau | t) &= x_i(\tau | t) + f(x_i(\tau | t), u_i(\tau | t)) \Delta\tau (t) \\ G(x_i(\tau | t), u_i(\tau | t)) &\leq 0, \end{aligned} \quad (\text{A.1})$$

where u is the control input and x is the state, L is the cost function. $f(x, u)$ is the state equation. $G(x, u)$ is the inequality constraint.

The inequality constraint in the optimal control problem is converted to an equality constraint by introducing a dummy input u_d for computation simplicity as follows:

$$C(x(t), u(t)) = u^2(t) + u_d^2(t) - u_{\max}^2 = 0, \quad (\text{A.2})$$

where u_{\max} denotes the upper bound of the control input.

To solve this optimal control problem with the calculus of variation method [32], the Hamiltonian function is defined by

$$H(x, u, \lambda, \psi) = L(x, u) + \lambda^T f(x, u) + \psi^T C(x, u), \quad (\text{A.3})$$

where λ denotes the costate and ψ denotes the Lagrange multiplier associated with the equality constraint.

The first-order necessary conditions for the optimal control input u , the multiplier ψ , and the costate λ are obtained using the calculus of variation as

$$\begin{aligned} x_{i+1}(t) &= x_i(t) + f(x_i(t), u_i(t)) \Delta\tau(t), \quad x_0(t) = x(t), \\ \lambda_i(t) &= \lambda_{i+1}(t) + H_x(x_i(t), u_i(t), \lambda_{i+1}(t), \psi_i(t)) \Delta\tau(t), \\ \lambda_N(t) &= 0, \\ H_u(x_i(t), u_i(t), \lambda_{i+1}(t), \psi_i(t)) &= 0, \\ C(x(t), u(t)) &= 0, \end{aligned} \quad (\text{A.4})$$

where x_0 is the initial state.

To solve this optimal control problem, the continuation and GMRES (C/GMRES) method is employed for computation cost reduction. The necessary conditions of optimality for the constrained control input can be expressed as the following equation:

$$F(U(\tau | t), x(\tau | t), t)$$

$$:= \begin{bmatrix} H_u(u_0(\tau | t), x_0(\tau | t), \lambda_1(\tau | t), \psi_0(\tau | t)) \\ C(u_0(\tau | t), x_0(\tau | t)) \\ \vdots \\ H_u(u_{N-1}(\tau | t), x_{N-1}(\tau | t), \lambda_N(\tau | t), \psi_{N-1}(\tau | t)) \\ C(u_{N-1}(\tau | t), x_{N-1}(\tau | t)) \end{bmatrix} = 0,$$

$$U(t) := [u_0^T(\tau | t), \psi_0^T(\tau | t), \dots, u_{N-1}^T(\tau | t), \psi_{N-1}^T(\tau | t)]^T. \quad (\text{A.5})$$

$F(U(t), x(t), t) = 0$ is identical to

$$F(U(0), x(0), 0) := 0, \quad (\text{A.6})$$

$$\dot{F}(U, x, t) = -A_s F(U(t), x(t), t),$$

where A_s is a stable matrix introduced to stabilize $F = 0$. If F_U is nonsingular, a differential equation for $U(t)$ can be obtained as

$$\dot{U} = -F_U^{-1} (A_s F - F_x \dot{x} - F_t). \quad (\text{A.7})$$

The above differential equation can be solved by the GMRES method. The presented approach is also a kind of continuation method. The solution curve $U(t)$ is traced by integrating the above differential equation. Because there is no need to calculate the Jacobians and the linear equation iteratively, C/GMRES method assures the real-time optimal control ability because of small computational cost. The detailed description of the solution for the real-time optimal control algorithm can be found in [34].

Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

References

- [1] A. Sciarretta and L. Guzzella, "Control of hybrid electric vehicles," *IEEE Control Systems Magazine*, vol. 27, no. 2, pp. 60–70, 2007.
- [2] S. J. Moura, H. K. Fathy, D. S. Callaway, and J. L. Stein, "A stochastic optimal control approach for power management in plug-in hybrid electric vehicles," *IEEE Transactions on Control Systems Technology*, vol. 19, no. 3, pp. 545–555, 2011.
- [3] L. Serrao, *A comparative analysis of energy management strategies for hybrid electric vehicles [Ph.D. thesis]*, The Ohio State University, 2009.
- [4] J. Liu and H. Peng, "Modeling and control of a power-split hybrid vehicle," *IEEE Transactions on Control Systems Technology*, vol. 16, no. 6, pp. 1242–1251, 2008.
- [5] N. Kim, S. Cha, and H. Peng, "Optimal control of hybrid electric vehicles based on Pontryagin's minimum principle," *IEEE Transactions on Control Systems Technology*, vol. 19, no. 5, pp. 1279–1287, 2011.
- [6] C. Zhang, A. Vahidi, P. Pisu, X. Li, and K. Tennant, "Role of terrain preview in energy management of hybrid electric vehicles," *IEEE Transactions on Vehicular Technology*, vol. 59, no. 3, pp. 1139–1147, 2010.
- [7] M. Koot, J. T. B. A. Kessels, B. de Jager, W. P. M. H. Heemels, P. P. J. van den Bosch, and M. Steinbuch, "Energy management strategies for vehicular electric power systems," *IEEE Transactions on Vehicular Technology*, vol. 54, no. 3, pp. 771–782, 2005.
- [8] H. Borhan, A. Vahidi, A. M. Phillips, M. L. Kuang, I. V. Kolmanovskiy, and S. Di Cairano, "MPC-based energy management of a power-split hybrid electric vehicle," *IEEE Transactions on Control Systems Technology*, vol. 20, no. 3, pp. 593–603, 2012.
- [9] Y. Deguchi and T. Kawabe, "HEV charge/discharge control system based on navigation information," SAE Technical Paper 2004-21-0028, 2004.
- [10] S. Stockar, V. Marano, M. Canova, G. Rizzoni, and L. Guzzella, "Energy-optimal control of plug-in hybrid electric vehicles for real-world driving cycles," *IEEE Transactions on Vehicular Technology*, vol. 60, no. 7, pp. 2949–2962, 2011.
- [11] Q. Gong and Y. Li, "Trip based optimal power management of plug-in hybrid electric vehicle with advanced traffic modeling," SAE Technical Paper 2008-01-1316, 2008.
- [12] Q. Gong, Y. Li, and Z.-R. Peng, "Trip-based optimal power management of plug-in hybrid electric vehicles," *IEEE Transactions on Vehicular Technology*, vol. 57, no. 6, pp. 3393–3401, 2008.

- [13] L. Guzzella and A. Sciarretta, *Vehicle Propulsion Systems: Introduction to Modeling and Optimization*, Springer, New York, NY, USA, 2005.
- [14] A. Sciarretta, M. Back, and L. Guzzella, "Optimal control of parallel hybrid electric vehicles," *IEEE Transactions on Control Systems Technology*, vol. 12, no. 3, pp. 352–363, 2004.
- [15] G. Paganelli, S. Delprat, T. M. Guerra, J. Rimaux, and J. J. Santin, "Equivalent consumption minimization strategy for parallel hybrid powertrains," in *Proceedings of the 55th Vehicular Technology Conference*, pp. 2076–2081, May 2002.
- [16] C. Zhang and A. Vahidi, "Route preview in energy management of plug-in hybrid vehicles," *IEEE Transactions on Control Systems Technology*, vol. 20, no. 2, pp. 546–553, 2012.
- [17] K. Yu, M. Mukai, and T. Kawabe, "Performance of an eco-driving nonlinear MPC system for a power-split HEV during car following," *SICE Journal of Control, Measurement, and System Integration*, vol. 7, no. 1, pp. 55–62, 2014.
- [18] K. Yu, M. Mukai, and T. Kawabe, "Model predictive control of a power-split hybrid electric vehicle system," *Artificial Life and Robotics*, vol. 17, no. 2, pp. 221–226, 2012.
- [19] K. Yu, M. Mukai, and T. Kawabe, "A battery management system using nonlinear model predictive control for a hybrid electric vehicle," in *Proceedings of the 7th IFAC Symposium on Advances in Automotive Control*, September 2013.
- [20] E. Hellström, M. Ivarsson, J. Åslund, and L. Nielsen, "Look-ahead control for heavy trucks to minimize trip time and fuel consumption," *Control Engineering Practice*, vol. 17, no. 2, pp. 245–254, 2009.
- [21] A. Vahidi and A. Eskandarian, "Research advances in intelligent collision avoidance and adaptive cruise control," *IEEE Transactions on Intelligent Transportation Systems*, vol. 4, no. 3, pp. 132–153, 2003.
- [22] D. Swaroop, J. K. Hedrick, C. C. Chien, and P. Ioannou, "Comparison of spacing and headway control laws for automatically controlled vehicles," *Vehicle System Dynamics*, vol. 23, no. 8, pp. 597–625, 1994.
- [23] J. Kim, N. Kim, S. Hwang, Y. Hori, and H. Kim, "Motor control of input-split hybrid electric vehicles," *International Journal of Automotive Technology*, vol. 10, no. 6, pp. 733–742, 2009.
- [24] K. B. Wipke, M. R. Cuddy, and S. D. Burch, "ADVISOR 2.1: a user-friendly advanced powertrain simulation using a combined backward/forward approach," *IEEE Transactions on Vehicular Technology*, vol. 48, no. 6, pp. 1751–1761, 1999.
- [25] C.-C. Lin, H. Peng, J. W. Grizzle, and J.-M. Kang, "Power management strategy for a parallel hybrid electric truck," *IEEE Transactions on Control Systems Technology*, vol. 11, no. 6, pp. 839–849, 2003.
- [26] M. Ehsani, Y. Gao, and A. Emadi, *Modern Electric, Hybrid Electric, and Fuel Cell Vehicles: Fundamentals, Theory, and Design*, CRC, Boca Raton, Fla, USA, 2nd edition, 2009.
- [27] H. N. Gupta, *Fundamentals of Internal Combustion Engines*, Prentice-Hall, New Delhi, India, 2006.
- [28] D. Rotenberg, A. Vahidi, and I. Kolmanovsky, "Ultracapacitor assisted powertrains: modeling, control, sizing, and the impact on fuel economy," *IEEE Transactions on Control Systems Technology*, vol. 19, no. 3, pp. 576–589, 2011.
- [29] K. Ahn, S. Cho, and S. W. Cha, "Optimal operation of the power-split hybrid electric vehicle powertrain," *Proceedings of the Institution of Mechanical Engineers D: Journal of Automobile Engineering*, vol. 222, no. 5, pp. 789–800, 2008.
- [30] M. Schulz, "Set values for a power-split hybrid electric vehicle through numerical optimization," *International Journal of Alternative Propulsion*, vol. 1, no. 1, pp. 6–31, 2006.
- [31] K. Ahn and P. Y. Papalambros, "Engine optimal operation lines for power-split hybrid electric vehicles," *Proceedings of the Institution of Mechanical Engineers D: Journal of Automobile Engineering*, vol. 223, no. 9, pp. 1149–1162, 2009.
- [32] D. E. Kirk, *Optimal Control Theory: An Introduction*, Prentice Hall, Englewood Cliffs, NJ, USA, 1970.
- [33] J. M. Maciejowski, *Predictive Control with Constraints*, Prentice Hall, Englewood Cliffs, NJ, USA, 2002.
- [34] T. Ohtsuka, "A continuation/GMRES method for fast computation of nonlinear receding horizon control," *Automatica*, vol. 40, no. 4, pp. 563–574, 2004.
- [35] K. Muta, M. Yamazaki, and J. Tokieda, "Development of new-generation hybrid system THS II-Drastic improvement of power performance and fuel economy," SAE Technical Paper 2004-01-0064, 2004.
- [36] F. B. Li, P. Shi, L. G. Wu, and X. Zhang, "Fuzzy-model-based D-stability and non-fragile control for discrete-time descriptor systems with multiple delays," *IEEE Transactions on Fuzzy Systems*, 2014.
- [37] L. Wu, X. Yang, and F. Li, "Nonfragile output tracking control of hypersonic air-breathing vehicles with an LPV model," *IEEE/ASME Transactions on Mechatronics*, vol. 18, no. 4, pp. 1280–1288, 2013.
- [38] F. Li and X. Zhang, "A delay-dependent bounded real lemma for singular LPV systems with time-variant delay," *International Journal of Robust and Nonlinear Control*, vol. 22, no. 5, pp. 559–574, 2012.

Research Article

Robust Control for Uncertain Linear System Subject to Input Saturation

Qingyun Yang and Mou Chen

College of Automation Engineering, Nanjing University of Aeronautics and Astronautics, Nanjing 210016, China

Correspondence should be addressed to Qingyun Yang; yang_980060@163.com

Received 3 March 2014; Accepted 23 May 2014; Published 16 June 2014

Academic Editor: Yuxin Zhao

Copyright © 2014 Q. Yang and M. Chen. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

A robust control scheme using composite nonlinear feedback (CNF) technology is proposed to improve tracking control performance for the uncertain linear system with input saturation and unknown external disturbances. A disturbance observer is presented to estimate the unknown disturbance generated by a linear exogenous system. The designed gain matrix of the disturbance observer is determined by solving linear matrix inequalities (LMIs). Based on the output of the designed disturbance observer, a robust CNF controller including a linear feedback control item and a nonlinear item is developed to follow the desired tracking signals. The linear feedback controller is designed using LMIs and the stability of the closed-loop system is proved via rigorous Lyapunov analysis. Finally, the extensive simulation results are presented to illustrate the effectiveness of the proposed control scheme.

1. Introduction

As is well known, almost all practical control systems have limitations on the amplitudes or rates of the control input [1]. Therefore, the input saturation which can cause the nonlinearity usually appears in most of the physical systems in our real life, such as aircraft, robot [2, 3], and industry control systems [4–6]. The input saturation problem is of great importance because it may lead to performance degradation and even destroy the stability of the control systems if they are ignored in the process of controller design [7, 8]. In general, it is hard to overcome the effect of input saturation through traditional linear control technologies because of the nonlinear characteristic of input saturation. Meanwhile, the linear system usually possesses unmodelled dynamics, modeling error, system parameter perturbations, and other uncertainties [9]. Generally speaking, the control performance of linear systems is severely affected by uncertainties. Thus, the task that designs high performance feedback control schemes for systems with input saturation and parametric uncertainties is theoretically challenging and critical for practical applications [10, 11].

Over the past years, several research methods on the input saturation problem have been reported in the literature, for

example, antiwindup schemes in [12, 13], predictive control in [14, 15], positively invariant sets method in [16, 17], low gain technology in [18, 19], variable structure control in [20, 21], and adaptive fuzzy control in [22–24]. Among these, a composite nonlinear feedback (CNF) control scheme, as an effective method to solve this problem, has been exclusively studied. The CNF method was proposed for a class of second-order linear systems in [25]. Then, a CNF control technique was developed for general single-input single-output (SISO) systems with measurement feedback and successfully applied to a hard disk drive (HDD) servo system in [26]. In [27], the design and implementation of a dual-stage actuated HDD servo system were studied via composite nonlinear control approach. Inspired by these works, the CNF control technique was extended to a general multi-input multioutput (MIMO) system under state feedback in [28] and a class of cascade nonlinear systems with input saturation in [29]. However, in the process of control design, the external disturbance has not been explicitly considered in above-mentioned literature. Considering the transient tracking performance and external disturbances, disturbance estimator was introduced into CNF control framework to propose a control strategy for servo system subject to actuator saturation and disturbances which was assumed to be an

unknown constant and applied to discrete-time systems in [30] and continuous systems in [31], respectively. In [32], the CNF control technique was extended to design a robust flight control system for an unmanned helicopter system with a wind gust disturbance. From above analysis, most of disturbances were assumed to be constant in these research works and parametric uncertainties are not explicitly considered in the control design. However, disturbances including external disturbances and parameter perturbations widely exist in practical systems, such as aircrafts, missiles, satellites, and many other systems [33, 34]. Thus, to solve this problem, disturbance observer-based control (DOBC) as a promising approach to handle system disturbances and to improve robustness can be employed.

The disturbance observer as an effective method which is extensively used to approximate unknown external disturbance has been attracting increasing attention [35, 36]. A two-stage design procedure was developed to improve disturbance attenuation ability of current linear/nonlinear controllers, where the disturbance observer design is separated from the controller design in [37]. In [38], a new DOBC was presented for a class of MIMO nonlinear systems to attenuate and reject the disturbances. A novel fuzzy-observer-design approach was presented for Takagi-Sugeno fuzzy models with unknown output disturbances in [39], where the disturbance was supposed to be an auxiliary state vector by an augmented fuzzy descriptor model and can be in any form. A disturbance observer-based multivariable control (DOMC) scheme was developed to control a two-input-two-output ball mill grinding circuit in [40]. Various control schemes based on the output of the disturbance observer can also be exclusively studied. In [41], a novel type of control scheme combining the DOBC with H_∞ control was proposed for a class of complex continuous models with disturbances. In [42], a sliding mode control (SMC) scheme was developed for a class of nonlinear systems based on disturbance observers. In [43], a nonlinear output disturbance observer based on the model of ocean wave was proposed to eliminate the disturbance of depth signal. A new DOBC technique for mismatched disturbances/uncertainties was presented in [44]. However, these research results did not consider the system subject to input saturation.

In [45], the system with input saturation and external disturbance has been studied, but the parametric uncertainties problem has not been considered, and the tracking signals are assumed to be constant. Thus, this paper is motivated by the robust control for the uncertain system with parametric uncertainties, input saturation, and external disturbance. A robust control design scheme based on disturbance observer will be proposed for the uncertain system subject to input saturation and unknown external disturbance. A disturbance observer is developed to estimate disturbances generated by an exogenous system via solving linear matrix inequality (LMI). Then, based on the output of the disturbance observer, a robust control scheme is proposed and the stability of the closed-loop system is proved by the Lyapunov function method. The outline of this paper is as follows. Section 2 gives the description and formulation of the problem. In Section 3, the design of disturbances observer is presented. In

Section 4, the robust CNF control method will be described and the developed control method is applied to design a tracking controller for a control system. Finally, simulation results will be given in Section 5 to illustrate the effectiveness of the proposed control scheme, followed by drawing some concluding remarks in Section 6.

2. Problem Formulation

Considering a class of linear systems with parametric uncertainties, unknown disturbances and input saturation are described as

$$\begin{aligned}\dot{x} &= (A + \Delta A)x + (B_1 + \Delta B_1) \text{sat}(u) + B_2 d, \\ y &= C_1 x, \\ z &= C_2 x + D_2 d,\end{aligned}\tag{1}$$

where $x \in R^n$, $u \in R^m$, $y \in R^p$, $z \in R^m$, and $d \in R^l$ are the state, control input, measurement output, controlled output, and external disturbance of system, respectively. d is norm bounded by a nonnegative scalar η , that is, $\|d\| \leq \eta$, and $A, B_1, B_2, C_1, C_2, D_2$ are appropriate dimensional constant matrices. The function $\text{sat}(\cdot)$ represents the input saturation of system defined as

$$\text{sat}(u_i) = \text{sign}(u_i) \min\{u_{\max i}, |u_i|\}, \quad i = 1, 2, \dots, m,\tag{2}$$

where $u_{\max i}$ represents the saturation level of the i th input and is known. ΔA and ΔB_1 representing the parametric uncertainties of system (1) are assumed to be in the following form:

$$[\Delta A \quad \Delta B_1] = DF(t) [E_1 \quad E_2],\tag{3}$$

where D, E_1 , and E_2 are appropriate dimensional constant matrices. $F(t)$ is an unknown, real, and possibly time-varying matrix with Lebesgue measurable elements satisfying

$$F^T(t)F(t) \leq I, \quad \forall t.\tag{4}$$

To continue the composite nonlinear feedback control design, the following assumptions and the lemma for the given system (1) are required [31].

Assumption 1. (A, B_1) is stabilizable.

Assumption 2. (A, C_1) is detectable.

Assumption 3. (A, B_1, C_2) is invertible and has no invariant zero at $s = 0$.

Assumption 4. Control gain matrix B_1 is row full rank.

Lemma 5. Assume that U and V are vectors or matrices with appropriate dimension; then, for any positive constant α , the following inequality holds:

$$U^T V + V^T U \leq \alpha U^T U + \alpha^{-1} V^T V.\tag{5}$$

Remark 1 (see [31]). Note that Assumption 1 means that there exists state feedback matrix K which satisfies that $A + B_1K$ is an asymptotically stable matrix. Assumption 2 denotes that the states variables can be detected by the output y of system (1). Assumption 3 implies that the matrix $C_2(A + B_1K)^{-1}B_1$ is invertible and will be used in the control design. For the convenience of the robust controller design, Assumption 4 is adopted to avoid the control singularity. Thus, all these assumptions are fairly standard for the tracking control.

In this paper, the control objective is that the robust CNF controller based on disturbance observer will be designed for uncertain systems (1) with input saturation and disturbances such that the closed-loop system is asymptotically stable and the controlled output z can well track the reference signal r .

3. Design of Disturbance Observer

In this section, a disturbance observer is developed to estimate the unknown disturbance of the system (1). To design the robust controller, suppose that the disturbance d is generated by a linear exogenous system [46]:

$$\begin{aligned}\dot{w} &= W_1 w, \\ d &= V_1 w,\end{aligned}\quad (6)$$

where $w \in R^p$, $d \in R^l$, W_1 , and V_1 are matrices with corresponding dimensions.

The disturbance observer is formulated as

$$\begin{aligned}\dot{v} &= (W_1 + LB_2V_1)(v - Lx) + L(Ax + B_1 \text{sat}(u)), \\ \hat{w} &= v - Lx, \\ \hat{d} &= V_1 \hat{w},\end{aligned}\quad (7)$$

where \hat{d} is the estimate of d and v is the auxiliary design vector of the disturbance observer. $L \in R^{p \times n}$ is a designed gain matrix and will be given by solving linear matrix inequalities (LMIs). The estimation error is defined as $\tilde{w} = w - \hat{w}$; based on (6) and (7), it is shown that the error dynamic satisfies

$$\begin{aligned}\dot{\tilde{w}} &= W_1 w - (W_1 + LB_2V_1)(v - Lx) \\ &\quad - L(Ax + B_1 \text{sat}(u)) + L\dot{x} \\ &= (W_1 + LB_2V_1)\tilde{w} + L\Delta B_1 \text{sat}(u) + L\Delta Ax.\end{aligned}\quad (8)$$

In this case, it is obvious that the designed observer gain matrix L not only needs to satisfy the desired stability of the disturbance observer (7), that is, $W_1 + LB_2V_1 < 0$, but also achieves robustness performance under the uncertainties $L\Delta B_1 \text{sat}(u)$ and $L\Delta Ax$.

Remark 2 (see [42]). As it is known, a wide class of real engineering disturbances can be presented by this disturbance model such as unknown constant load disturbances and harmonic disturbances. For example, unknown constant disturbance can be presented with $W_1 = 0$ and $V_1 = 1$, and a

harmonic disturbance with known frequency ω but unknown phase and magnitude can be represented with

$$W_1 = \begin{bmatrix} 0 & \omega \\ \omega & 0 \end{bmatrix}, \quad V_1 = [1 \ 0]. \quad (9)$$

4. Robust CNF Control Design Using the Disturbance Observer

In this section, we will design robust CNF control law using the disturbance observer. The state feedback robust CNF control law can be designed by the following step-by-step procedure.

Step 1. A linear state feedback control law with a disturbance compensation term is designed as

$$u_L = Kx + K_d w + \Lambda r + \Lambda_r \dot{r}, \quad (10)$$

where K is the designed state feedback matrix, and satisfies that $A + B_1K$ is an asymptotically stable matrix. $K_d w$ is the disturbance compensation term and r is the tracking reference signal differing from previous CNF method in which the tracking reference signal must be constant. Next, Λ is chosen as

$$\Lambda = -[C_2(A + B_1K)^{-1}B_1]^{-1}. \quad (11)$$

It is apparent that (11) is well defined when Assumption 3 is given. Considering Assumption 4, the matrix K_d is given by

$$\begin{aligned}K_d &= B_1^T (B_1 B_1^T)^{-1} \Lambda_d V_1 W_1, \\ \Lambda_d &= -(A + B_1K)^{-1} B_2.\end{aligned}\quad (12)$$

At the same time, Λ_r is chosen as

$$\Lambda_r = B_1^T (B_1 B_1^T)^{-1} \Lambda_e r, \quad (13)$$

where

$$\Lambda_e = -(A + B_1K)^{-1} B_1 \Lambda. \quad (14)$$

Step 2. The nonlinear feedback law u_N is constructed as

$$u_N = \begin{cases} -\frac{B_1^T (B_1 B_1^T)^{-1} P_2^{-1} (x - x_e) K_a}{\|x - x_e\|^2}, & \|x - x_e\| \geq \varepsilon \\ 0, & \|x - x_e\| < \varepsilon, \end{cases} \quad (15)$$

where $P_2 > 0$ is a positive-definite matrix, K_a is a designed matrix, ε is a minimal positive design constant, and x_e is defined as

$$x_e = \Lambda_e r + \Lambda_d \hat{d}. \quad (16)$$

In (15), K_a is designed as

$$K_a = \frac{1}{2q} \left((\alpha_5^{-1} + \alpha_6^{-1}) u_{\max}^2 E_2^T E_2 + (\alpha_4^{-1} + \alpha_4^{-1}) x_e^T E_1^T E_1 x_e \right), \quad (17)$$

where α_4 , α_5 , and α_6 will be defined in Theorem 6.

Step 3. The linear and nonlinear feedback laws designed in above steps are now combined as a robust CNF controller:

$$u = u_L + u_N = Kx + K_d w + \Lambda r + \Lambda_r \dot{r} + u_N. \quad (18)$$

If the disturbance is replaced by the estimated one, the CNF controller is given by

$$u = Kx + K_d \hat{w} + \Lambda r + \Lambda_r \dot{r} + \hat{u}_N, \quad (19)$$

$$\hat{u}_N = \begin{cases} -\frac{B_1^T (B_1 B_1^T)^{-1} P_2^{-1} (x - \hat{x}_e) K_a}{\|x - \hat{x}_e\|^2}, & \|x - \hat{x}_e\| \geq \epsilon \\ 0, & \|x - \hat{x}_e\| < \epsilon, \end{cases} \quad (20)$$

where $\hat{x}_e = \Lambda_e r + \Lambda_d \hat{d}$ is the estimation of x_e , and K_a is rewritten as

$$K_a = \frac{1}{2q} \left((\alpha_5^{-1} + \alpha_6^{-1}) u_{\max}^2 E_2^T E_2 + (\alpha_2^{-1} + \alpha_4^{-1}) \hat{x}_e^T E_1^T E_1 \hat{x}_e \right). \quad (21)$$

This completes the robust CNF controller design procedure.

The main objective of the designed robust CNF controller is to ensure not only the asymptotical stability of the closed-loop system and the disturbance observer estimate error, but also that the controlled output z can track the reference signal r as smooth as possible. Thus, the stability and tracking condition is given in the following theorem.

Theorem 6. *Considering the given uncertain system (1) with external disturbance and input saturation and provided that the following conditions are satisfied:*

- (1) for any $\tau \in (0, 1)$, let $\rho_\tau > 0$ be the largest positive scalar such that for all $x \in X_\tau$, the following property holds:

$$\left\| [K, K_1] \begin{pmatrix} x \\ w \end{pmatrix} \right\| \leq (1 - \tau) u_{\max}, \quad (22)$$

where

$$X_\tau = \left\{ \begin{pmatrix} x \\ w \end{pmatrix} \mid \begin{pmatrix} x \\ w \end{pmatrix}^T \begin{pmatrix} P_1 & 0 \\ 0 & P_2 \end{pmatrix} \begin{pmatrix} x \\ w \end{pmatrix} \leq \rho_\tau \right\} \quad (23)$$

and $K_1 = -K \Lambda_d V_1$, $P_1, P_2 > 0$;

- (2) the initial conditions x_0 satisfy

$$\begin{pmatrix} x_0 - \hat{x}_{e0} \\ w_0 - \hat{w}_0 \end{pmatrix} \in X_\tau; \quad (24)$$

- (3) the reference signal r satisfies

$$\|Mr\| + \|\Lambda_r \dot{r}\| \leq \tau u_{\max} - \eta \|K_2\| - \|K_d \hat{w}\|, \quad (25)$$

where

$$\begin{aligned} M &= K \Lambda_e + \Lambda, \\ K_2 &= K \Lambda_d; \end{aligned} \quad (26)$$

- (4) for given positive constants α_2, α_4 , and α_6 , there exist constants $\alpha_1 > 0, \alpha_3 > 0$, and $\alpha_5 > 0$ and some matrices $X \in R^{n \times n} > 0, Y \in R^{m \times n}, P_1 \in R^{m \times m} > 0$, and $T \in R^{m \times n}$ such that the following LMI holds:

$$\begin{pmatrix} \Gamma_{11} & \Gamma_{12} & XE_1^T & XE_1^T & 0 \\ \Gamma_{12}^T & \Gamma_{22} & 0 & 0 & TD \\ E_1 X & 0 & -\alpha_1 I & 0 & 0 \\ E_1 X & 0 & 0 & -\alpha_3 I & 0 \\ 0 & D^T T^T & 0 & 0 & -(\alpha_1 + \alpha_3 + \alpha_4)^{-1} I \end{pmatrix} < 0, \quad (27)$$

where

$$\begin{aligned} \Gamma_{11} &= AX + XA^T + B_1 Y + Y B_1^T + (\alpha_1 + \alpha_2 + \alpha_5) DD^T, \\ \Gamma_{12} &= B_2 V_1, \end{aligned} \quad (28)$$

$$\Gamma_{22} = P_1 W_1 + W_1^T P_1 + T B_2 V_1 + V_1^T B_2^T T^T,$$

then the disturbance observer approximation error is asymptotically stable and the controlled output z can track the reference r asymptotically under the developed CNF control law (19), where $K = YX^{-1}$, $L = P_1^{-1} T$.

Proof. Let us define a new state variable $\bar{x} = x - \hat{x}_e$, and $\bar{d} = d - \hat{d}$. Invoking the definition of \bar{w} , the CNF control law of (19) can be rewritten as

$$\begin{aligned} u &= [K, -K \Lambda_d V_1] \begin{pmatrix} \bar{x} \\ \bar{w} \end{pmatrix} \\ &+ [K \Lambda_e + \Lambda, K \Lambda_d] \begin{pmatrix} r \\ d \end{pmatrix} + K_d \hat{w} + \Lambda_r \dot{r} + \hat{u}_N. \end{aligned} \quad (29)$$

Considering (22) and (26), we obtain

$$u = [K, K_1] \begin{pmatrix} \bar{x} \\ \bar{w} \end{pmatrix} + [M, K_2] \begin{pmatrix} r \\ d \end{pmatrix} + K_d \hat{w} + \Lambda_r \dot{r} + \hat{u}_N. \quad (30)$$

Invoking (1) and the the definition of variables \bar{x} , the time derivative of \bar{x} can be written as

$$\dot{\bar{x}} = \dot{x} - \dot{\hat{x}}_e = (A + \Delta A) x + (B_1 + \Delta B_1) \text{sat}(u) + B_2 d - \dot{\hat{x}}_e. \quad (31)$$

Next, we note that

$$(A + B_1 K) x_e = (A + B_1 K) (\Lambda_e r + \Lambda_d \hat{d}) = -B_1 \Lambda r - B_2 d. \quad (32)$$

According to (32), we have

$$(A + B_1 K) x_e + B_1 \Lambda r + B_2 d = 0. \quad (33)$$

Based on the the definition of variables \hat{x}_e and substituting (33) into (31), we obtain

$$\begin{aligned} \dot{\bar{x}} &= Ax + B_1 \text{sat}(u) - (A + B_1 K) x_e - B_1 \Lambda r \\ &- B_1 K_d \hat{w} - \Lambda_r \dot{r} + \Delta Ax + \Delta B_1 \text{sat}(u). \end{aligned} \quad (34)$$

Then, according to $\bar{x} = x - \hat{x}_e$, we have

$$\begin{aligned} \dot{\bar{x}} &= (A + B_1K)\bar{x} + B_1 \text{sat}(u) + A\hat{x}_e - B_1K\bar{x} - (A + B_1K)x_e \\ &\quad - B_1\Lambda r - B_1K_d\hat{w} - B_1\Lambda_r\dot{r} + \Delta Ax + \Delta B_1 \text{sat}(u) \\ &= (A + B_1K)\bar{x} + B_1 \text{sat}(u) + (A + B_1K)\hat{x}_e \\ &\quad - (A + B_1K)x_e - B_1Kx \\ &\quad - B_1\Lambda r - B_1K_d\hat{w} - B_1\Lambda_r\dot{r} + \Delta Ax + \Delta B_1 \text{sat}(u). \end{aligned} \quad (35)$$

Considering the definition of variables \tilde{d} and \bar{x} and substituting $\hat{x}_e = \Lambda_e r + \Lambda_d \tilde{d}$ into (35) yield

$$\begin{aligned} \dot{\bar{x}} &= (A + B_1K)\bar{x} + B_1 \text{sat}(u) - (A + B_1K)\Lambda_d \tilde{d} \\ &\quad + B_1K\Lambda_d \tilde{d} - B_1K\bar{x} \\ &\quad - (B_1K\Lambda_e r + B_1\Lambda)r - B_1K\Lambda_d \tilde{d} - B_1K_d\hat{w} - B_1\Lambda_r\dot{r} \\ &\quad + \Delta Ax + \Delta B_1 \text{sat}(u). \end{aligned} \quad (36)$$

Considering the definition of variables \bar{x} , we have

$$\begin{aligned} \dot{\bar{x}} &= (A + B_1K + \Delta A)\bar{x} + B_1 \text{sat}(u) \\ &\quad - (A + B_1K)\Lambda_d \tilde{d} + B_1K\Lambda_d \tilde{d} \\ &\quad - B_1K\bar{x} - (B_1K\Lambda_e r + B_1\Lambda)r - B_1K\Lambda_d \tilde{d} \\ &\quad - B_1K_d\hat{w} - B_1\Lambda_r\dot{r} - \Lambda_r\dot{r} + \Delta A\hat{x}_e + \Delta B_1 \text{sat}(u) \\ &= (A + B_1K + \Delta A)\bar{x} - (A + B_1K)\Lambda_d \tilde{d} \\ &\quad + B_1 \left(\text{sat}(u) - [K, -K\Lambda_d V_1] \begin{pmatrix} \bar{x} \\ \tilde{w} \end{pmatrix} \right. \\ &\quad \left. - [K\Lambda_e + \Lambda, K\Lambda_d] \begin{pmatrix} r \\ d \end{pmatrix} - K_d\hat{w} - \Lambda_r\dot{r} \right) \\ &\quad + \Delta A\hat{x}_e + \Delta B_1 \text{sat}(u). \end{aligned} \quad (37)$$

Substituting (11), (22), and (26) into (37) gives

$$\begin{aligned} \dot{\bar{x}} &= (A + B_1K + \Delta A)\bar{x} + B_2V_1\tilde{w} \\ &\quad + B_1\gamma + \Delta A\hat{x}_e + \Delta B_1 \text{sat}(u), \end{aligned} \quad (38)$$

where

$$\sigma = \text{sat}(u) - [K, K_1] \begin{pmatrix} \bar{x} \\ \tilde{w} \end{pmatrix} - [H, K_2] \begin{pmatrix} r \\ d \end{pmatrix} - K_d\hat{w} - \Lambda_r\dot{r}. \quad (39)$$

Note that for all $\begin{pmatrix} \bar{x} \\ \tilde{w} \end{pmatrix} \in X_\tau$ and $\|Mr\| + \|\Lambda_r\dot{r}\| \leq \tau u_{\max} - \eta\|K_2\| - \|K_d\hat{w}\|$, we have

$$\begin{aligned} &\left\| [K, K_1] \begin{pmatrix} \bar{x} \\ \tilde{w} \end{pmatrix} + [M, K_2] \begin{pmatrix} r \\ d \end{pmatrix} + K_d\hat{w} + \Lambda_r\dot{r} \right\|, \\ &\leq \left\| [K, K_1] \begin{pmatrix} \bar{x} \\ \tilde{w} \end{pmatrix} \right\| + \|Mr\| + \|\Lambda_r\dot{r}\| \\ &\quad + \eta\|K_2\| + \|K_d\hat{w}\| \leq u_{\max}. \end{aligned} \quad (40)$$

Thus, the value of σ can be determined via (30) and (39) for three different values of saturation function:

$$\begin{aligned} \hat{u}_N < \sigma < 0, & \quad u < -u_{\max}, \\ \sigma = \hat{u}_N, & \quad |u| \leq u_{\max}, \\ 0 < \sigma < \hat{u}_N, & \quad u > u_{\max}. \end{aligned} \quad (41)$$

From above analysis, we can obtain that

$$\sigma = q\hat{u}_N, \quad (42)$$

where $q \in [0, 1]$.

Substituting (42) into system (38), we have

$$\begin{aligned} \dot{\bar{x}} &= (A + B_1K + \Delta A)\bar{x} + B_2V_1\tilde{w} \\ &\quad + qB_1\hat{u}_N + \Delta A\hat{x}_e + \Delta B_1 \text{sat}(u). \end{aligned} \quad (43)$$

Then, the error dynamic equation (8) can be rewritten as

$$\dot{\tilde{w}} = (W_1 + LB_2V_1)\tilde{w} + L\Delta B_1 \text{sat}(u) + L\Delta A\bar{x} + L\Delta A\hat{x}_e. \quad (44)$$

Choose the Lyapunov function as

$$V = \bar{x}^T P_2 \bar{x} + \tilde{w}^T P_1 \tilde{w}. \quad (45)$$

Invoking (43) and (44), the time derivative of V along the trajectory of the system (45) is

$$\begin{aligned} \dot{V} &= \bar{x}^T \left((A + B_1K + \Delta A)^T P_2 + P_2 (A + B_1K + \Delta A) \right) \bar{x} \\ &\quad + 2q\bar{x}^T P_2 B_1 \hat{u}_N + \bar{x}^T P_2 B_2 V_1 \tilde{w} + \tilde{w}^T V_1^T B_2^T P_2 \bar{x} \\ &\quad + \bar{x}^T P_2 \Delta A \hat{x}_e + \hat{x}_e^T \Delta A^T P_2 \bar{x} \\ &\quad + \tilde{w}^T \left((W_1 + LB_2V_1)^T P_1 + P_1 (W_1 + LB_2V_1) \right) \tilde{w} \\ &\quad + \tilde{w}^T P_1 L \Delta A \bar{x} + \bar{x}^T \Delta A^T L^T P_1 \tilde{w} \\ &\quad + \tilde{w}^T P_1 L \Delta A \hat{x}_e + \hat{x}_e^T \Delta A^T L^T P_1 \tilde{w} \\ &\quad + 2\bar{x}^T P_2 \Delta B_1 \text{sat}(u) + 2\tilde{w}^T P_1 L \Delta B_1 \text{sat}(u). \end{aligned} \quad (46)$$

Recalling (3), it obtains that

$$\begin{aligned} \dot{V} &= \bar{x}^T \left((A + B_1K)^T P_2 + P_2 (A + B_1K) \right) \bar{x} \\ &\quad + \bar{x}^T \left(E_1^T F^T D^T P_2 + P_2 D F E_1 \right) \bar{x} \\ &\quad + 2q\bar{x}^T P_2 B_1 \hat{u}_N + \bar{x}^T P_2 B_2 V_1 \tilde{w} + \tilde{w}^T V_1^T B_2^T P_2 \bar{x} \\ &\quad + \bar{x}^T P_2 D F E_1 \hat{x}_e + \hat{x}_e^T E_1^T F^T D^T P_2 \bar{x} \\ &\quad + \tilde{w}^T \left((W_1 + LB_2V_1)^T P_1 + P_1 (W_1 + LB_2V_1) \right) \tilde{w} \\ &\quad + \tilde{w}^T P_1 L D F E_1 \bar{x} + \bar{x}^T E_1^T F^T D^T L^T P_1 \tilde{w} \\ &\quad + \tilde{w}^T P_1 L D F E_1 \hat{x}_e + \hat{x}_e^T E_1^T F^T D^T L^T P_1 \tilde{w} \\ &\quad + 2\bar{x}^T P_2 D F E_2 \text{sat}(u) + 2\tilde{w}^T P_1 L D F E_2 \text{sat}(u). \end{aligned} \quad (47)$$

Using Lemma 5, we have

$$\begin{aligned}
& \bar{x}^T (E_1^T F^T D^T P_2 + P_2 D F E_1) \bar{x} \\
& \leq \alpha_1 \bar{x}^T P_2 D D^T P_2 \bar{x} + \alpha_1^{-1} \bar{x}^T E_1^T E_1 \bar{x}, \\
& \bar{x}^T P_2 D F E_1 \hat{x}_e + \hat{x}_e^T E_1^T F^T D^T P_2 \\
& \leq \alpha_2 \bar{x}^T P_2 D D^T P_2 \bar{x} + \alpha_2^{-1} \hat{x}_e^T E_1^T E_1 \hat{x}_e, \\
& \bar{w}^T P_1 L D F E_1 \bar{x} + \bar{x}^T E_1^T F^T D^T L^T P_1 \bar{w} \\
& \leq \alpha_3 \bar{w}^T P_1 L D D^T L^T P_1 \bar{w} + \alpha_3^{-1} \bar{x}^T E_1^T E_1 \bar{x}, \\
& \bar{w}^T P_1 L D F E_1 \hat{x}_e + \hat{x}_e^T E_1^T F^T D^T L^T P_1 \bar{w} \\
& \leq \alpha_4 \bar{w}^T P_1 L D D^T L^T P_1 \bar{w} + \alpha_4^{-1} \hat{x}_e^T E_1^T E_1 \hat{x}_e, \\
& 2\bar{x}^T P_2 D F E_2 \text{sat}(u) \\
& \leq \alpha_5 \bar{x}^T P_2 D D^T P_2 \bar{x} + \alpha_5^{-1} \text{sat}(u)^T E_2^T E_2 \text{sat}(u), \\
& 2\bar{w}^T P_1 L D F E_2 \text{sat}(u) \\
& \leq \alpha_6 \bar{w}^T P_1 L D D^T L^T P_1 \bar{w} + \alpha_6^{-1} \text{sat}(u)^T E_2^T E_2 \text{sat}(u).
\end{aligned} \tag{48}$$

Substituting (48) into (47) and considering (20) yield

$$\begin{aligned}
\dot{V} & \leq \bar{x}^T \left((A + B_1 K)^T P_2 + P_2 (A + B_1 K) \right) \bar{x} \\
& + \alpha_1 \bar{x}^T P_2 D D^T P_2 \bar{x} \\
& + \alpha_1^{-1} \bar{x}^T E_1^T E_1 \bar{x} - 2qK_a + \bar{x}^T P_2 B_2 V_1 \bar{w} + \bar{w}^T V_1^T B_2^T P_2 \bar{x} \\
& + \alpha_2 \bar{x}^T P_2 D D^T P_2 \bar{x} + \alpha_2^{-1} \hat{x}_e^T E_1^T E_1 \hat{x}_e \\
& + \bar{w}^T \left((W_1 + L B_2 V_1)^T P_1 + P_1 (W_1 + L B_2 V_1) \right) \bar{w} \\
& + \alpha_3 \bar{w}^T P_1 L D D^T L^T P_1 \bar{w} + \alpha_3^{-1} \bar{x}^T E_1^T E_1 \bar{x} \\
& + \alpha_4 \bar{w}^T P_1 L D D^T L^T P_1 \bar{w} + \alpha_4^{-1} \hat{x}_e^T E_1^T E_1 \hat{x}_e \\
& + \alpha_5 \bar{x}^T P_2 D D^T P_2 \bar{x} + \alpha_6 \bar{w}^T P_1 L D D^T L^T P_1 \bar{w} \\
& + (\alpha_5^{-1} + \alpha_6^{-1}) \text{sat}(u)^T E_2^T E_2 \text{sat}(u).
\end{aligned} \tag{49}$$

Considering the definition of $\text{sat}(u)$, we have

$$\| \text{sat}(u)^T E_2^T E_2 \text{sat}(u) \| \leq u_{\max}^2 E_2^T E_2. \tag{50}$$

Substituting (21) and (50) into (49) yields

$$\begin{aligned}
\dot{V} & \leq \bar{x}^T \left((A + B_1 K)^T P_2 + P_2 (A + B_1 K) \right) \bar{x} \\
& + (\alpha_1 + \alpha_2 + \alpha_5) \bar{x}^T P_2 D D^T P_2 \bar{x} + (\alpha_1^{-1} + \alpha_3^{-1}) \bar{x}^T E_1^T E_1 \bar{x} \\
& + \bar{x}^T P_2 B_2 V_1 \bar{w} + \bar{w}^T V_1^T B_2^T P_2 \bar{x}
\end{aligned}$$

$$\begin{aligned}
& + \bar{w}^T \left((W_1 + L B_2 V_1)^T P_1 + P_1 (W_1 + L B_2 V_1) \right) \bar{w} \\
& + (\alpha_3 + \alpha_4 + \alpha_6) \bar{w}^T P_1 L D D^T L^T P_1 \bar{w}.
\end{aligned} \tag{51}$$

Equation (51) can be rewritten as

$$\dot{V} \leq \begin{pmatrix} \bar{x} \\ \bar{w} \end{pmatrix}^T \bar{\Gamma} \begin{pmatrix} \bar{x} \\ \bar{w} \end{pmatrix}, \tag{52}$$

where

$$\bar{\Gamma} = \begin{bmatrix} \bar{\Gamma}_{11} & \bar{\Gamma}_{12} \\ \bar{\Gamma}_{12}^T & \bar{\Gamma}_{22} \end{bmatrix},$$

$$\begin{aligned}
\bar{\Gamma}_{11} & = (A + B_1 K)^T P_2 + P_2 (A + B_1 K) \\
& + (\alpha_1 + \alpha_2 + \alpha_5) P_2 D D^T P_2 + (\alpha_1^{-1} + \alpha_3^{-1}) E_1^T E_1
\end{aligned} \tag{53}$$

$$\bar{\Gamma}_{12} = P_2 B_2 V_1,$$

$$\begin{aligned}
\bar{\Gamma}_{22} & = (W_1 + L B_2 V_1)^T P_1 + P_1 (W_1 + L B_2 V_1) \\
& + (\alpha_3 + \alpha_4 + \alpha_6) P_1 L D D^T L^T P_1.
\end{aligned}$$

Let $P_2^{-1} = X$, $K = YX^{-1}$, and $L = P_1^{-1}T$. Both sides of (53), multiplying by $\text{diag}(P_2^{-1}, I)$, yield

$$\tilde{\Gamma} = \begin{bmatrix} \tilde{\Gamma}_{11} & \tilde{\Gamma}_{12} \\ \tilde{\Gamma}_{12}^T & \tilde{\Gamma}_{22} \end{bmatrix}, \tag{54}$$

where

$$\begin{aligned}
\tilde{\Gamma}_{11} & = AX + XA^T + B_1 Y + YB_1^T \\
& + (\alpha_1 + \alpha_2 + \alpha_5) DD^T + (\alpha_1^{-1} + \alpha_3^{-1}) XE_1^T E_1 X, \\
\tilde{\Gamma}_{12} & = B_2 V_1, \\
\tilde{\Gamma}_{22} & = P_1 W_1 + W_1^T P_1 + TB_2 V_1 + V_1^T B_2^T T^T \\
& + (\alpha_3 + \alpha_4 + \alpha_6) TDD^T T^T.
\end{aligned} \tag{55}$$

Equation (51) can be rewritten as

$$\tilde{\Gamma} = \begin{bmatrix} \Gamma_{11} & \Gamma_{12} \\ \Gamma_{12}^T & \Gamma_{22} \end{bmatrix} + \bar{E} \Delta \bar{F}, \tag{56}$$

where

$$\begin{aligned}
\bar{E} & = \begin{bmatrix} XE_1^T & XE_1^T & 0 \\ 0 & 0 & TD \end{bmatrix}, \\
\Delta & = \begin{bmatrix} \alpha_1^{-1} I & 0 & 0 \\ 0 & \alpha_3^{-1} I & 0 \\ 0 & 0 & (\alpha_3 + \alpha_4 + \alpha_6) I \end{bmatrix}, \\
\bar{F} & = \begin{bmatrix} E_1 X & 0 \\ E_1 X & 0 \\ 0 & D^T T^T \end{bmatrix}.
\end{aligned} \tag{57}$$

Considering (27) and using the Schur complement theorem, we have

$$\bar{\Gamma} \leq 0. \quad (58)$$

Thus, combining (52)–(56) with (57), we obtain

$$\dot{V} \leq 0, \quad \forall \begin{pmatrix} \bar{x} \\ \bar{w} \end{pmatrix} \in X_\tau. \quad (59)$$

From (59), we can know that X_τ is an invariant set of the closed-loop system (43) and (44) and all the trajectories of the closed-loop system starting from inside X_τ will converge to the origin; meanwhile, the disturbance observer estimate error is asymptotically stable. Thus, we have

$$\lim_{t \rightarrow \infty} \begin{pmatrix} \bar{x} \\ \bar{w} \end{pmatrix} = 0 \implies \lim_{t \rightarrow \infty} \bar{w} = \lim_{t \rightarrow \infty} \bar{x} = 0 \implies \lim_{t \rightarrow \infty} x = x_e. \quad (60)$$

Furthermore, if $D_2 = C_2(A + B_1K)^{-1}B_2$, we obtain

$$\lim_{t \rightarrow \infty} z = \lim_{t \rightarrow \infty} C_2x_e = r. \quad (61)$$

This completes the proof of Theorem 6. \square

Remark 3. It can be seen from Theorem 6 that the closed-loop system and the disturbance observer estimate error for the studied plant in (1) under the developed robust CNF controller of (20) and disturbance observer (7) are asymptotically stable.

Remark 4. To handle the nonlinear terms $\text{sat}(u)^T E_2^T E_2 \text{sat}(u)$ and $\hat{x}_e^T E_1^T E_1 \hat{x}_e$ in (49), the nonlinear feedback law \hat{u}_N is designed as the form of (20). It can be seen from (60) and (61) that $x = x_e$ when $t \rightarrow \infty$. From above analysis, we can obtain that $z = r$. Thus, $x - \hat{x}_e = 0$ means that the controlled output z can track the reference r asymptotically under the control of the CNF control law of (20). Therefore, $\hat{u}_N = 0$ is reasonable when the difference between x and \hat{x}_e is small enough; that is, $x - \hat{x}_e < \epsilon$.

Remark 5. It can be seen from [26] that the CNF control can actually improve the transient performance of output response of the closed-loop system by introducing nonlinear feedback portion in which the desired trajectory is normally assumed to be constant. However, as the desired tracking trajectory used in this paper is time-varying, which differs from [26], our control objective is to ensure that output of the closed-loop system can track the time-variant trajectory in the presence of input saturation, external disturbance, and uncertainties and the tracking errors are asymptotically stable under the control of the proposed robust CNF controller. Thus, the improvement of the transient performance is not investigated in the paper, and this study is our future research work.

5. Simulation Results

In this section, the extensive simulation results are given to demonstrate the effectiveness of the proposed robust CNF control techniques by using two simulation examples.

5.1. Numerical Example. Consider an uncertain system [47] characterized by (1) with

$$\begin{aligned} A &= \begin{bmatrix} 0.1 & -0.1 \\ 0.1 & -3 \end{bmatrix}, & B_1 &= \begin{bmatrix} 5 & 0 \\ 0 & 1 \end{bmatrix}, \\ B_2 &= \begin{bmatrix} 1 \\ 1 \end{bmatrix}, & C_1 = C_2 = D &= \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \\ E_1 &= \text{diag}(0.2, 0.2), & E_2 &= \text{diag}(0.3, 0.3), \\ F &= \text{diag}(0.5 \sin(t), 0.5 \cos(t)). \end{aligned} \quad (62)$$

The references signals are $r = (3, 0.12)^T$. The system disturbance d is generated by a linear exogenous system described by (3) with

$$W_1 = \begin{pmatrix} 0 & 1.5 \\ -1.5 & 0 \end{pmatrix}, \quad V_1 = (1 \ 0). \quad (63)$$

Here, the given disturbance represents an external harmonic disturbance with known frequency but without any information of its magnitude and phase. Choosing $\alpha_2 = \alpha_4 = \alpha_6 = 1$ and solving LMI (27) give

$$\begin{aligned} K &= \begin{bmatrix} -1.0561 & -0.0355 \\ -0.0108 & -0.1691 \end{bmatrix}, & L &= \begin{bmatrix} -0.1297 & -0.0613 \\ -0.0652 & -0.0079 \end{bmatrix}, \\ P_2 &= \begin{bmatrix} 2.9955 & 0.0353 \\ 0.0353 & 4.3282 \end{bmatrix}, & \alpha_1 &= 3.0227, \\ \alpha_3 &= 3.0227, & \alpha_5 &= 3.0227. \end{aligned} \quad (64)$$

The initial state values are $x_0 = [0, 0]^T$ and $u_{\max} = [5, 2]^T$, the initial generated disturbance value is $d_0 = 0.12$, and the disturbance observer initial value is $\hat{d}_0 = 0.2$. The CNF controller is designed according to (19).

The simulation results for the system using the developed CNF controller are presented in Figures 1, 2, 3, 4, and 5. Figure 1 indicates that the output of disturbance can effectively approximate the unknown external harmonic disturbance. It is shown in Figures 2 and 3 that the control output z can track the references r asymptotically under the control of (19). Figures 4 and 5 show that the control input does not exceed the limitation of input. Thus, the developed composite nonlinear feedback control (CNF) scheme is valid for the uncertain linear system with input saturation and unknown external disturbances.

5.2. Chaotic System. A chaotic system with disturbance is described as follows:

$$\dot{x} = (A + \Delta A)x + B_2d, \quad (65)$$

where $x = [x_1, x_2, x_3, x_4]^T$ is the system state, $d = \sin(10t)$ is the external disturbance, $\Delta A = DF(t)E_1$ represents the system uncertainties, and $D = I_4$, $E_1 = 0.2I_4$, and

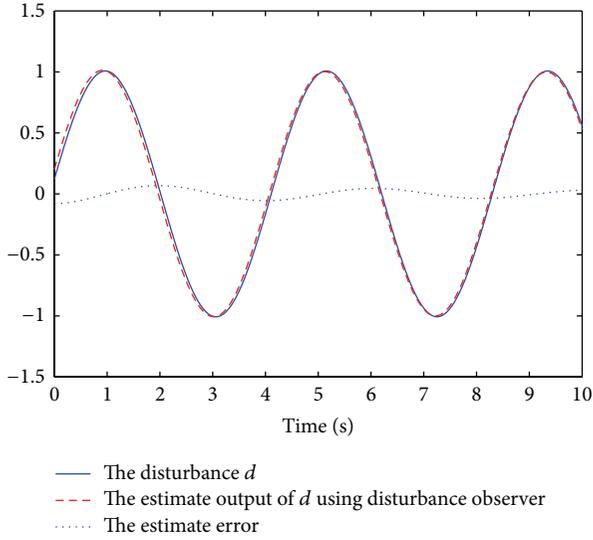


FIGURE 1: The disturbance d and the approximation output of \hat{d} .

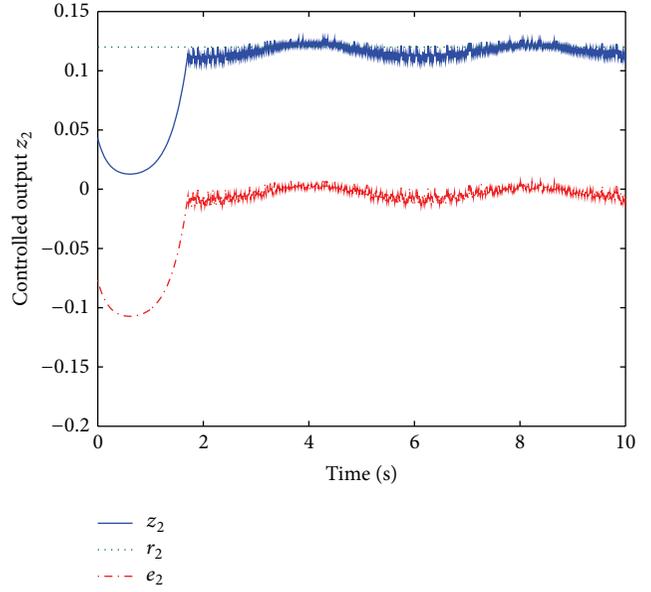


FIGURE 3: Controlled output z_2 and tracking error e_2 .

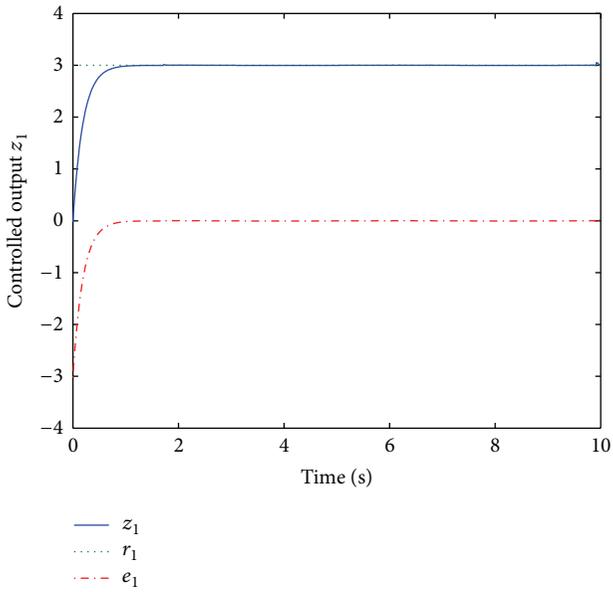


FIGURE 2: Controlled output z_1 and tracking error e_1 .

$F = \text{diag}(0, 10 \cos(t), 100 \sin(t), 0)$; the system matrices are given by

$$A = \begin{bmatrix} 1 & 0 & 1 & 0 \\ -20 & 0 & 0 & -20 \\ -2 & 6 & -66 & 0 \\ 0 & 1.5 & -1 & 50 \end{bmatrix}, \quad B_2 = [1, 1, 1, 1]^T. \quad (66)$$

The system state responses without control are shown in Figure 6. It can be seen that it is a typical chaotic system.

To control the chaotic system (65), a controller u is introduced. Thus, the system (65) can be transformed into another system as the form of plant (1) used to synchronize with the chaotic system, and the references signals r of this system are obtained from system (65), where the system matrixes A, B_2

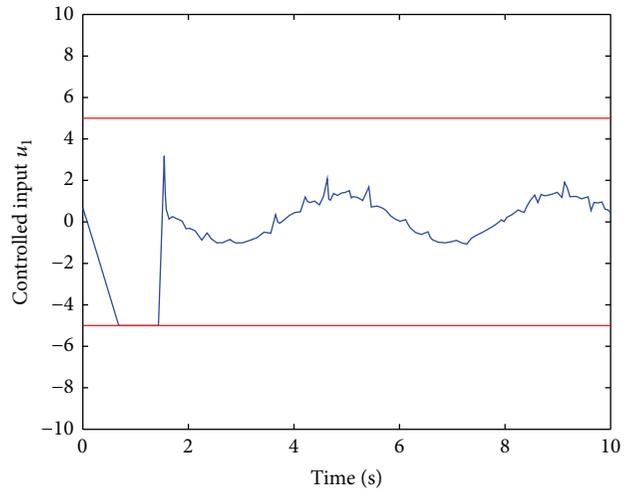


FIGURE 4: Control input u_1 .

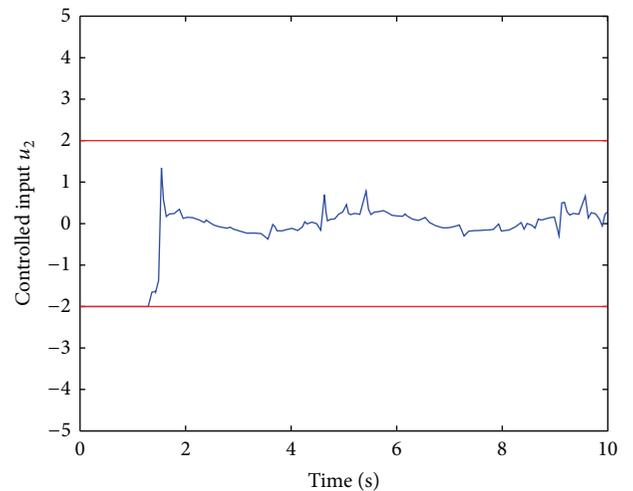


FIGURE 5: Control input u_2 .

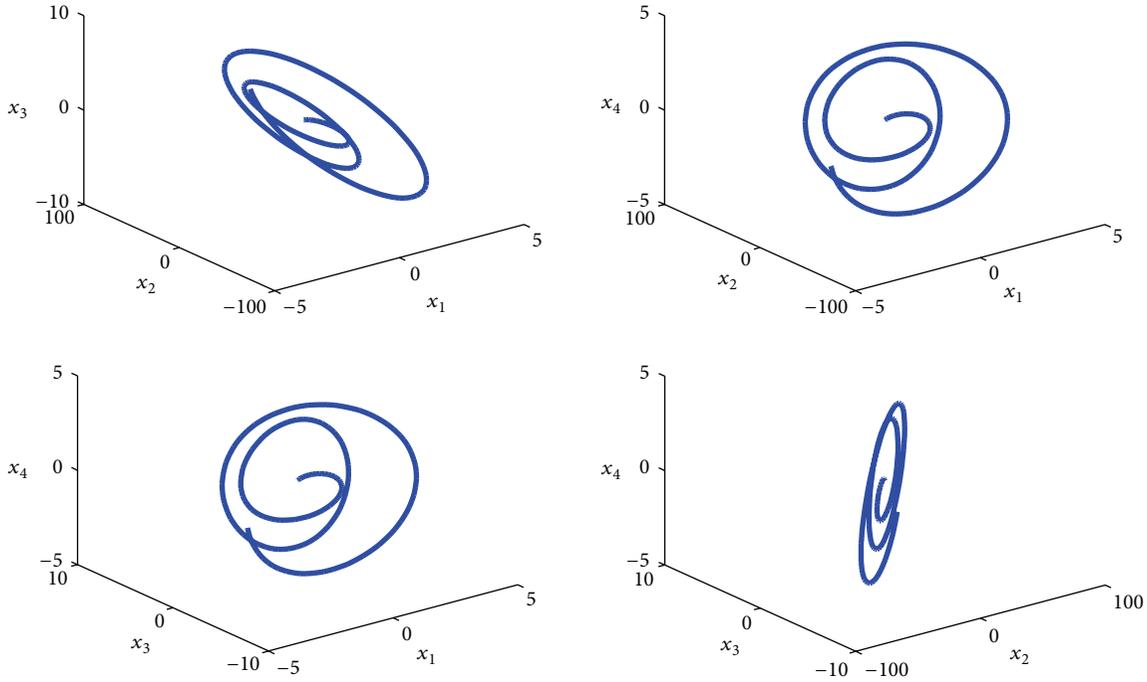


FIGURE 6: The states response of chaotic system.

are the same as those of the chaotic system (65), $B_1 = \text{diag}(10, 10, 10, 10)$ and $C_1 = C_2 = \text{diag}(1, 1, 1, 1)$. It is obvious that (A, B_1) is stabilizable. The system disturbance d is equal to that of last section and uncertainties ΔB_1 are given by $E_2 = 0.3I_4$.

Choosing $\alpha_2 = \alpha_4 = \alpha_6 = 1$ and solving LMI (27) give

$$K = \begin{bmatrix} -1.6871 & 0.8152 & 0.0741 & -0.6320 \\ 0.7644 & -0.9456 & -0.1257 & 0.7273 \\ 0.0847 & -0.1188 & -0.1133 & 0.2485 \\ -0.6227 & 0.7553 & 0.3379 & -1.5494 \end{bmatrix},$$

$$L = \begin{bmatrix} -0.0704 & -0.0704 & -0.0704 & -0.0704 \\ -0.0029 & -0.0029 & -0.0029 & -0.0029 \end{bmatrix},$$

$$P_2 = \begin{bmatrix} 0.5327 & -0.1736 & -0.0169 & 0.1476 \\ -0.1736 & 0.4143 & 0.0183 & -0.1619 \\ -0.0169 & 0.0183 & 0.3090 & -0.0714 \\ 0.1476 & -0.1619 & -0.0714 & 0.5091 \end{bmatrix}, \quad (67)$$

$$\alpha_1 = 2.6417, \quad \alpha_3 = 2.6417, \quad \alpha_5 = 2.6417.$$

The initial state values of the synchronize system are $x_0 = [1, 1, 1, 1]^T$ and $u_{\max} = 1$, the initial generated disturbance value is $\hat{d}_0 = 1$, and the disturbance observer initial value is $\hat{d}_0 = 1.16$. The CNF controller is designed according to (19).

The simulation results for the the synchronization of chaotic circuit system and designed system using the developed CNF controller are presented in Figures 7, 8, and 9. Figure 7 indicates that the output of disturbance can effectively approximate the unknown external harmonic disturbance. It is shown in Figure 8 that the control output z can track the references r asymptotically under the control of (19) and the tracking errors are asymptotically stable. Therefore,

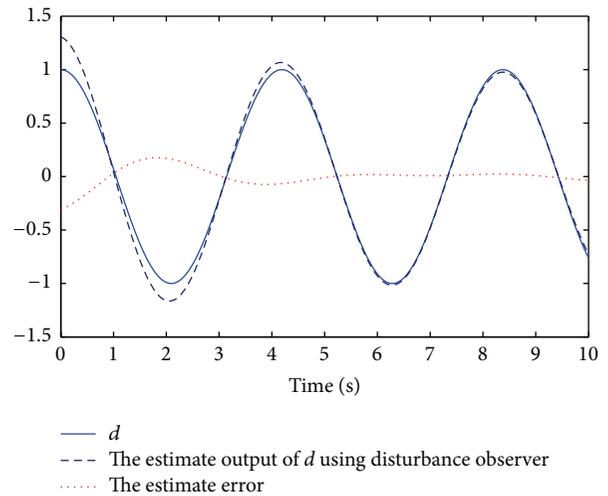


FIGURE 7: The disturbance d and the approximation output of \hat{d} .

the outputs of the chaotic system and designed system are asymptotically synchronized. Figure 9 shows that the control input does not exceed the limitation of input.

It can be shown from these simulation results of the numerical example and uncertain system that the disturbance observer can well estimate the system disturbance, and the closed-loop system for the linear system with input saturation and parametric uncertainties under the the designed robust control scheme using the disturbance observer is asymptotically stable. Thus, the proposed robust control method is valid.

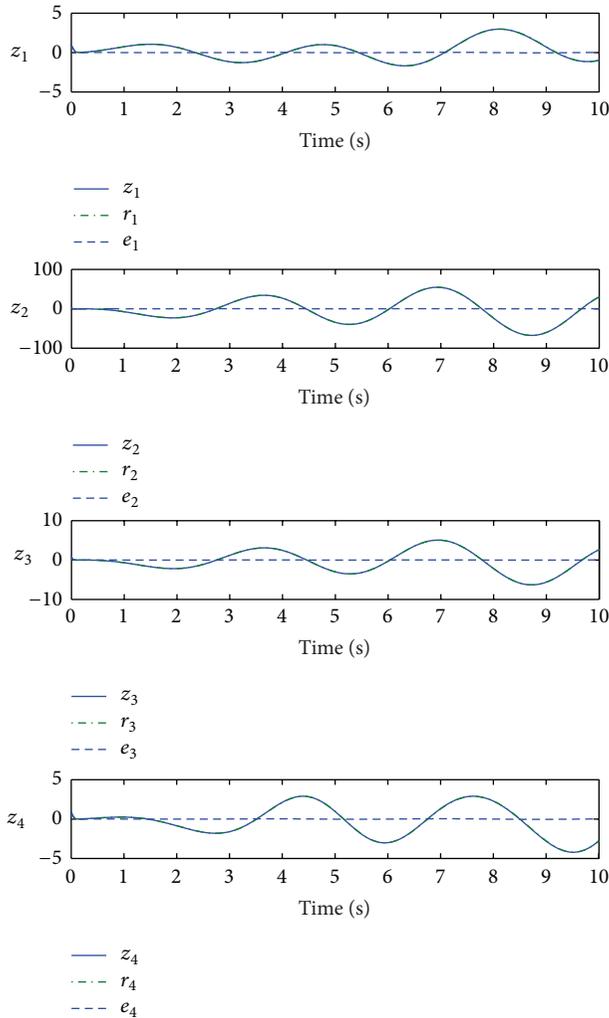


FIGURE 8: Controlled output z , reference r , and tracking error e .

6. Conclusion

In this paper, a CNF control scheme based on the disturbance observer has been proposed to achieve satisfactory tracking control performance for the linear system subject to input saturation, parametric uncertainties, and unknown external disturbance. The disturbance observer has been designed to approximate the system disturbance generated by a linear exogenous system. Based on the output of the disturbance observer, a CNF controller has been developed for the uncertain system subject to input saturation; then, the stability of the closed-loop system under the designed controller has been rigorously proved. Finally, the control method has been applied to the uncertain linear system to illustrate the effectiveness of the proposed control scheme. The simulation results have suggested that the designed CNF control scheme is valid. The direction of future research is to make further improvement of transient tracking performance and extend our results to other MIMO systems, such as near space vehicles (NSV), helicopters, and aircrafts.

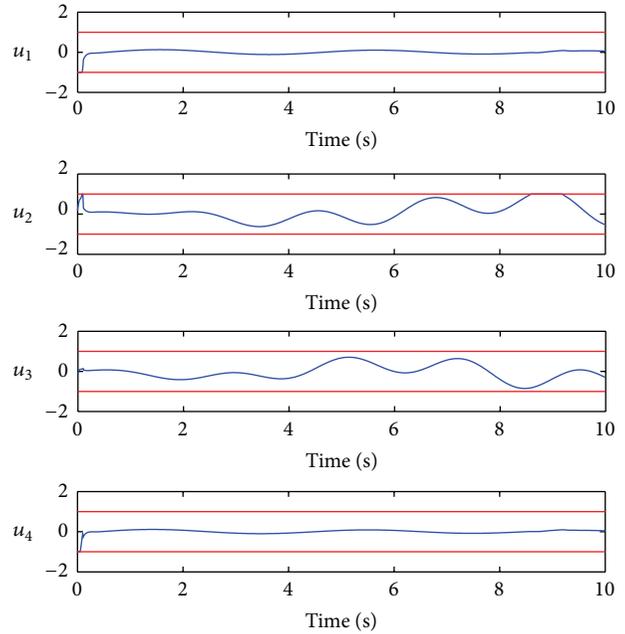


FIGURE 9: Control input u .

Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

Acknowledgments

This work is partially supported by National Natural Science Foundation of China (Grant no. 61174102), Jiangsu Natural Science Foundation of China (Grant nos. SBK20130033 and SBK2011069), Program for New Century Excellent Talents in University of China (Grant no. NCET-11-0830), and Specialized Research Fund for the Doctoral Program of Higher Education (Grant no. 20133218110013).

References

- [1] Y.-Y. Cao, Z. Lin, and T. Hu, "Stability analysis of linear time-delay systems subject to input saturation," *IEEE Transactions on Circuits and Systems I: Fundamental Theory and Applications*, vol. 49, no. 2, pp. 233–240, 2002.
- [2] Z. Li, S. S. Ge, and A. Ming, "Adaptive robust motion/force control of holonomic-constrained nonholonomic mobile manipulators," *IEEE Transactions on Systems, Man, and Cybernetics B: Cybernetics*, vol. 37, no. 3, pp. 607–616, 2007.
- [3] Z. Li, S. S. Ge, M. Adams, and W. S. Wijesoma, "Robust adaptive control of uncertain force/motion constrained nonholonomic mobile manipulators," *Automatica*, vol. 44, no. 3, pp. 776–784, 2008.
- [4] C. Dong, Y. Hou, Y. Zhang, and Q. Wang, "Model reference adaptive switching control of a linearized hypersonic flight vehicle model with actuator saturation," *Proceedings of the Institution of Mechanical Engineers I: Journal of Systems and Control Engineering*, vol. 224, no. 3, pp. 289–303, 2010.

- [5] Y. Zhao, W. Sun, and H. Gao, "Robust control synthesis for seat suspension systems with actuator saturation and time-varying input delay," *Journal of Sound and Vibration*, vol. 329, no. 21, pp. 4335–4353, 2010.
- [6] W. Sun, Z. Zhao, and H. Gao, "Saturated adaptive robust control for active suspension systems," *IEEE Transactions on Industrial Electronics*, vol. 60, no. 9, pp. 3889–3896, 2013.
- [7] T. Hu and Z. Lin, *Control Systems with Actuator Saturation: Analysis and Design*, Birkhäuser, Boston, Mass, USA, 2001.
- [8] T. Hu, A. R. Teel, and L. Zaccarian, "Anti-windup synthesis for linear control systems with input saturation: achieving regional, nonlinear performance," *Automatica*, vol. 44, no. 2, pp. 512–519, 2008.
- [9] Y.-J. Liu, S. Tong, and C. L. P. Chen, "Adaptive fuzzy control via observer design for uncertain nonlinear systems with unmodeled dynamics," *IEEE Transactions on Fuzzy Systems*, vol. 21, no. 2, pp. 275–288, 2013.
- [10] B. Zhou, H. Gao, Z. Lin, and G.-R. Duan, "Stabilization of linear systems with distributed input delay and input saturation," *Automatica*, vol. 48, no. 5, pp. 712–724, 2012.
- [11] Y. Li, S. Tong, and T. Li, "Direct adaptive fuzzy backstepping control of uncertain nonlinear systems in the presence of input saturation," *Neural Computing and Applications*, vol. 23, no. 5, pp. 1207–1216, 2013.
- [12] G. Grimm, J. Hatfield, I. Postlethwaite, A. R. Teel, M. C. Turner, and L. Zaccarian, "Antiwindup for stable linear systems with input saturation: an LMI-based synthesis," *IEEE Transactions on Automatic Control*, vol. 48, no. 9, pp. 1509–1525, 2003.
- [13] Y.-Y. Cao, Z. Lin, and D. G. Ward, "An antiwindup approach to enlarging domain of attraction for linear systems subject to actuator saturation," *IEEE Transactions on Automatic Control*, vol. 47, no. 1, pp. 140–145, 2002.
- [14] Y. I. Lee and B. Kouvaritakis, "Robust receding horizon predictive control for systems with uncertain dynamics and input saturation," *Automatica*, vol. 36, no. 10, pp. 1497–1504, 2000.
- [15] Z.-J. Li, W. Tan, S.-C. Nian, and J.-Z. Liu, "A stabilizing model predictive control for linear systems with input saturation," in *Proceedings of the International Conference on Machine Learning and Cybernetics*, pp. 671–675, August 2006.
- [16] H. Richter, B. D. O'Dell, and E. A. Misawa, "Robust positively invariant cylinders in constrained variable structure control," *IEEE Transactions on Automatic Control*, vol. 52, no. 11, pp. 2058–2069, 2007.
- [17] B. Zhou, D. Li, and Z. Lin, "Control of discrete-time periodic linear systems with input saturation via multi-step periodic invariant set," in *Proceedings of the 10th World Congress on Intelligent Control and Automation (WCICA '12)*, pp. 1372–1377, July 2012.
- [18] S. Tong, T. Wang, Y. Li, and B. Chen, "A combined backstepping and stochastic small-gain approach to robust adaptive fuzzy output feedback control," *IEEE Transactions on Fuzzy Systems*, vol. 21, no. 2, pp. 314–327, 2013.
- [19] Y. Li, S. Tong, Y. Liu, and T. Li, "Adaptive fuzzy robust output feedback control of nonlinear systems with unknown dead zones based on small-gain approach," *IEEE Transactions on Fuzzy Systems*, vol. 22, no. 1, pp. 164–176, 2014.
- [20] Q. Hu, G. Ma, and L. Xie, "Robust and adaptive variable structure output feedback control of uncertain systems with input nonlinearity," *Automatica*, vol. 44, no. 2, pp. 552–559, 2008.
- [21] Z. Zhu, Y. Xia, and M. Fu, "Adaptive sliding mode control for attitude stabilization with actuator saturation," *IEEE Transactions on Industrial Electronics*, vol. 58, no. 10, pp. 4898–4907, 2011.
- [22] S. Tong and Y. Li, "Adaptive fuzzy output feedback tracking backstepping control of strict-feedback nonlinear systems with unknown dead zones," *IEEE Transactions on Fuzzy Systems*, vol. 20, no. 1, pp. 168–180, 2012.
- [23] S. Tong and Y. Li, "Adaptive fuzzy output feedback control of MIMO nonlinear systems with unknown dead-zone inputs," *IEEE Transactions on Fuzzy Systems*, vol. 21, no. 1, pp. 134–146, 2013.
- [24] Y. Li, T. Li, and X. Jing, "Indirect adaptive fuzzy control for input and output constrained nonlinear systems using a barrier Lyapunov function," *International Journal of Adaptive Control and Signal Processing*, vol. 28, no. 2, pp. 184–199, 2014.
- [25] Z. Lin, M. Pachter, and S. Banda, "Toward improvement of tracking performance—nonlinear feedback for linear systems," *International Journal of Control*, vol. 70, no. 1, pp. 1–11, 1998.
- [26] B. M. Chen, T. H. Lee, K. Peng, and V. Venkataramanan, "Composite nonlinear feedback control for linear systems with input saturation: theory and an application," *IEEE Transactions on Automatic Control*, vol. 48, no. 3, pp. 427–439, 2003.
- [27] K. Peng, B. M. Chen, T. H. Lee, and V. Venkataramanan, "Design and implementation of a dual-stage actuated HDD servo system via composite nonlinear control approach," *Mechatronics*, vol. 14, no. 9, pp. 965–988, 2004.
- [28] Y. He, B. M. Chen, and C. Wu, "Composite nonlinear control with state and measurement feedback for general multivariable systems with input saturation," *Systems & Control Letters*, vol. 54, no. 5, pp. 455–469, 2005.
- [29] W. Lan, B. M. Chen, and Y. He, "Composite nonlinear feedback control for a class of nonlinear systems with input saturation," in *Proceedings of the 16th Triennial World Congress of International Federation of Automatic Control (IFAC '05)*, pp. 622–627, Prague, Czech, July 2005.
- [30] K. Peng, G. Cheng, B. M. Chen, and T. H. Lee, "Improvement of transient performance in tracking control for discrete-time systems with input saturation and disturbances," *IET Control Theory and Applications*, vol. 1, no. 1, pp. 65–74, 2007.
- [31] G. Cheng and K. Peng, "Robust composite nonlinear feedback control with application to a servo positioning system," *IEEE Transactions on Industrial Electronics*, vol. 54, no. 2, pp. 1132–1140, 2007.
- [32] G. Cai, B. M. Chen, X. Dong, and T. H. Lee, "Design and implementation of a robust and nonlinear flight control system for an unmanned helicopter," *Mechatronics*, vol. 21, no. 5, pp. 803–820, 2011.
- [33] Y.-J. Liu, S.-C. Tong, and T.-S. Li, "Observer-based adaptive fuzzy tracking control for a class of uncertain nonlinear MIMO systems," *Fuzzy Sets and Systems*, vol. 164, no. 1, pp. 25–44, 2011.
- [34] H. R. Karimi, M. Zapateiro, and N. Luo, "Adaptive synchronization of master-slave systems with mixed neutral and discrete time-delays and nonlinear perturbations," *Asian Journal of Control*, vol. 14, no. 1, pp. 251–257, 2012.
- [35] W.-H. Chen, "Nonlinear disturbance observer-enhanced dynamic inversion control of missiles," *Journal of Guidance, Control, and Dynamics*, vol. 26, no. 1, pp. 161–166, 2003.
- [36] S. Tong, Y. Li, Y. Li, and Y. Liu, "Observer-based adaptive fuzzy backstepping control for a class of stochastic nonlinear strict-feedback systems," *IEEE Transactions on Systems, Man, and Cybernetics B: Cybernetics*, vol. 41, no. 6, pp. 1693–1704, 2011.

- [37] W.-H. Chen, "Disturbance observer based control for nonlinear systems," *IEEE Transactions on Mechatronics*, vol. 9, no. 4, pp. 706–710, 2004.
- [38] L. Guo and W.-H. Chen, "Disturbance attenuation and rejection for systems with nonlinearity via DOBC approach," *International Journal of Robust and Nonlinear Control*, vol. 15, no. 3, pp. 109–125, 2005.
- [39] Z. Gao, X. Shi, and S. X. Ding, "Fuzzy state/disturbance observer design for T-S fuzzy systems with application to sensor fault estimation," *IEEE Transactions on Systems, Man, and Cybernetics B: Cybernetics*, vol. 38, no. 3, pp. 875–880, 2008.
- [40] X. S. Chen, J. Yang, S. H. Li, and Q. Li, "Disturbance observer based multi-variable control of ball mill grinding circuits," *Journal of Process Control*, vol. 19, no. 7, pp. 1205–1213, 2009.
- [41] X. Wei and L. Guo, "Composite disturbance-observer-based control and H_∞ control for complex continuous models," *International Journal of Robust and Nonlinear Control*, vol. 20, no. 1, pp. 106–118, 2010.
- [42] M. Chen and W.-H. Chen, "Sliding mode control for a class of uncertain nonlinear system based on disturbance observer," *International Journal of Adaptive Control and Signal Processing*, vol. 24, no. 1, pp. 51–64, 2010.
- [43] H.-H. Zhang, D. Wu, C.-C. Li, and Z.-P. Yan, "Research on depth control based on output disturbance observer for UUVs maneuvering near the surface," in *Proceedings of the 9th IEEE International Conference on Mechatronics and Automation (ICMA '12)*, pp. 2564–2568, August 2012.
- [44] J. Yang, S. Li, and X. Yu, "Sliding-mode control for systems with mismatched uncertainties via a disturbance observer," *IEEE Transactions on Industrial Electronics*, vol. 60, no. 1, pp. 160–169, 2013.
- [45] Q. Yang and M. Chen, "Composite nonlinear control for near space vehicles with input saturation based on disturbance observer," in *Proceedings of the IEEE 32nd Chinese Control Conference (CCC '13)*, pp. 2763–2768, 2013.
- [46] M. Chen and W.-H. Chen, "Disturbance-observer-based robust control for time delay uncertain systems," *International Journal of Control, Automation and Systems*, vol. 8, no. 2, pp. 445–453, 2010.
- [47] J.-H. Kim and Z. Bien, "Robust stability of uncertain linear systems with saturating actuators," *IEEE Transactions on Automatic Control*, vol. 39, no. 1, pp. 202–207, 1994.

Research Article

Support Vector Regression-Based Adaptive Divided Difference Filter for Nonlinear State Estimation Problems

Hongjian Wang, Jinlong Xu, Aihua Zhang, Cun Li, and Hongfei Yao

College of Automation, Harbin Engineering University, Harbin 150001, China

Correspondence should be addressed to Jinlong Xu; xujinlong_1983@sina.cn

Received 2 March 2014; Accepted 4 May 2014; Published 25 May 2014

Academic Editor: Weichao Sun

Copyright © 2014 Hongjian Wang et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

We present a support vector regression-based adaptive divided difference filter (SVRADDf) algorithm for improving the low state estimation accuracy of nonlinear systems, which are typically affected by large initial estimation errors and imprecise prior knowledge of process and measurement noises. The derivative-free SVRADDf algorithm is significantly simpler to compute than other methods and is implemented using only functional evaluations. The SVRADDf algorithm involves the use of the theoretical and actual covariance of the innovation sequence. Support vector regression (SVR) is employed to generate the adaptive factor to tune the noise covariance at each sampling instant when the measurement update step executes, which improves the algorithm's robustness. The performance of the proposed algorithm is evaluated by estimating states for (i) an underwater nonmaneuvering target bearing-only tracking system and (ii) maneuvering target bearing-only tracking in an air-traffic control system. The simulation results show that the proposed SVRADDf algorithm exhibits better performance when compared with a traditional DDF algorithm.

1. Introduction

The problem of state estimation for nonlinear systems has been a subject of considerable research interest in recent years, but there is still no single solution that outperforms all other approaches. Most proposed estimators are nonlinear extensions of the dominated Kalman filter (see [1]), and each approach provides a suboptimal trade-off between properties such as numerical robustness, computational burden, and estimation accuracy. The extended Kalman filter (EKF), which linearizes both nonlinear terms of a current estimated state trajectory, is based on a first-order Taylor series and displays poor performance if the system is highly nonlinear. The limitations of EKFs are enumerated in [2]. Another improved algorithm is the iterated extended Kalman filter (IEKF), which linearizes the nonlinear model around an updated state rather than the predicted state (see [3]). Although IEKFs have been proven to perform better than EKFs in addition to globally guaranteeing convergence, the algorithm still requires a Jacobian matrix just like EKFs. However, no solution exists for the Jacobian matrix in nonlinear systems

for some situations, which limits the potential application of both EKFs and IEKFs.

In recent years, a new class of filter known as sigma-point Kalman filter (SPKF) has attracted a great deal of attention. In SPKFs, the algorithm propagates a cluster of points centered on the current state instead of linearizing the system dynamics to improve the approximations of the conditional mean and covariance. Unscented Kalman filters (UKF) and divided difference filters (DDF) are two kinds of SPKFs.

UKFs use a deterministic sampling technique to pick a minimal set of sample points around the mean to catch the higher order statistics of the system so as to better estimation accuracy and convergence characteristics (see [4]). In [5, 6], a UKF for a class of nonlinear discrete-time systems with correlated noises was designed to deal with the problem of nonlinear filtering failure found in conventional UKFs when system noise is correlated with measurement noise. The proposed UKF breaks the limitation of conventional UKFs that requires system noise and measurement noise to be uncorrelated Gauss white noises, thus extending the potential

application of conventional UKFs. In [7], a UKF filtering algorithm with colored measurement noise was proposed. The algorithm was first derived on the basis of augmented measurement information and minimum mean square error estimation, and a filtering recursive formula of UKF with colored noise then added by applying an unscented transformation to calculate the posterior mean and covariance of the nonlinear state within the optimal framework. The proposed UKF effectively dealt with the fact that traditional UKFs fail when measurement noise is colored. In [8], a UKF was applied to multiple target tracking, with the proposed UKF shown to have improved performance versus previous EKF approaches. In [9], the paper discussed an adaptive multiuser receiver for CDMA systems in which the scaled unscented filter (SUF) and the square root unscented filter (SURF) were used for joint estimation and tracking of the code delays and multipath coefficients of the received CDMA signals. The proposed channel estimators were more near-far resistant than in conventional EKFs and presented lower complexity than conventional particle filter- (PF-) based methods. Computer simulation results demonstrated the superior performance of the proposed channel estimators, and the proposed estimators were shown to exhibit lower complexity relative to the PF-based method. Although UKFs have undergone a significant amount of meaningful theory innovation and are now used in many fields, [10–12] show that UKF accuracy is lower than that of DDF, while also having a higher computational cost.

The divided difference filter (DDF) first proposed by Nøgaard et al. (see [13]) linearizes the nonlinear terms based on Stirling's interpolation polynomial approximations formula rather than Taylor's approximation of nonlinear terms in an EKF. Conceptually, the implementation principle resembles that of an EKF; however, the DDF is significantly simpler as it does not need to calculate the Jacobian matrix and no derivatives are required. The DDF that Nøgaard et al. developed works on general discrete-time nonlinear models in which the noises are not assumed to be additive. In [14], the paper further formulated a DDF in terms of the innovation vector approach, the additive process, and the measurement noise sources. In [15], the paper proposed a new filter named the maximum likelihood-based iterated divided difference filter (MLIDDF), which improved the low state estimation accuracy of nonlinear state estimation that results from large initial estimation errors and the nonlinearity of the measurement equations. Simulation results showed that the MLIDDF algorithm possessed better state estimation accuracy and a faster convergence rate. In [16], the authors proposed a novel adaptive version of the DDF that was applicable to nonlinear systems with a linear output equation. In order to make the filter robust to modeling errors, upper bounds on the state covariance matrix were derived. The parameters of the upper bound were then estimated using a combination of offline tuning and online optimization with a linear matrix inequality constraint, which ensured that the predicted output error covariance was larger than the observed output error covariance. Simulation results demonstrated the superior performance of the proposed filter as compared to the standard DDF. Reference [17] presented

an ensemble-based approach that handled nonlinearity based on a simplified divided difference approximation through Stirling's interpolation formula. The algorithm used Stirling's interpolation formula to evaluate the statistics of the background ensemble during the prediction step, employing the formula in an ensemble square root filter (EnSRF) at the filtering step to update the background for analysis. In this sense, the algorithm is a hybrid of Stirling's interpolation formula and the EnSRF method, while the computational cost of the algorithm is less than that of EnSRF.

Different studies have focused on the application of DDFs to nonlinear state estimation problems. In [18], time delay and channel gain estimation for multipath fading code division multiple access (CDMA) signals using a DDF were investigated, and the simulation results showed that the DDF was simpler to implement and more resilient to near-far interference in CDMA networks compared with an EKF. In [19, 20], the relative kinematic states of a reentry vehicle obtained from noisy seeker measurements using a DDF were examined. The results were compared to those obtained using an EKF and a UKF and showed that the DDF was more accurate than estimators based on a Taylor approximation like the EKF. Reference [21] investigated the possibility of using a DDF for estimating the internal variables of a synchronous generator, such as the rotor angle where the data acquired is from a phasor measurement. The effectiveness of the method was tested on a single machine infinite bus system, a nine-bus system, and a 68-bus New England-New York interconnected system. In [22], a DDF using orientation estimation was considered. The fourth element of the quaternion error vector was removed from the system states to alleviate estimated error covariance matrix divergence. The measurement system was a MARG sensor, which consisted of a triaxial rate gyro, a triaxial accelerometer, and a triaxial magnetometer. The nonlinear measurement model was obtained based on the principals of operation of the magnetometer and accelerometer and the properties of the quaternion vector space. The performance of three filters, DDF, EKF, and UKF, was compared with different sampling frequencies. The work showed that the tested DDF and the UKF were more robust than the EKF under the same initial angle-error conditions. The DDF also performed better than the UKF, although the computational load for the UKF was less. In [23], a DDF-based data fusion algorithm was presented, which utilized the complementary noise profile of rate gyros and gravimetric inclinometers to extend their limits and achieve more accurate attitude estimates. In [24], a DDF-based ballistic target tracking system for the reentry phase was proposed. The paper compared DDF, EKF, and UKF algorithms using a Monte Carlo simulation approach, with the simulation results showing that the DDF outperformed both the EKF and UKF in terms of estimation accuracy and filtering credibility. In [25], a DDF with quaternion-based dynamic process modeling was applied to global positioning system (GPS) navigation to increase navigation estimation accuracy at high-dynamic regions while preserving precision at low-dynamic regions. Some properties and performance metrics were assessed and compared to those using EKF and UKF approaches.

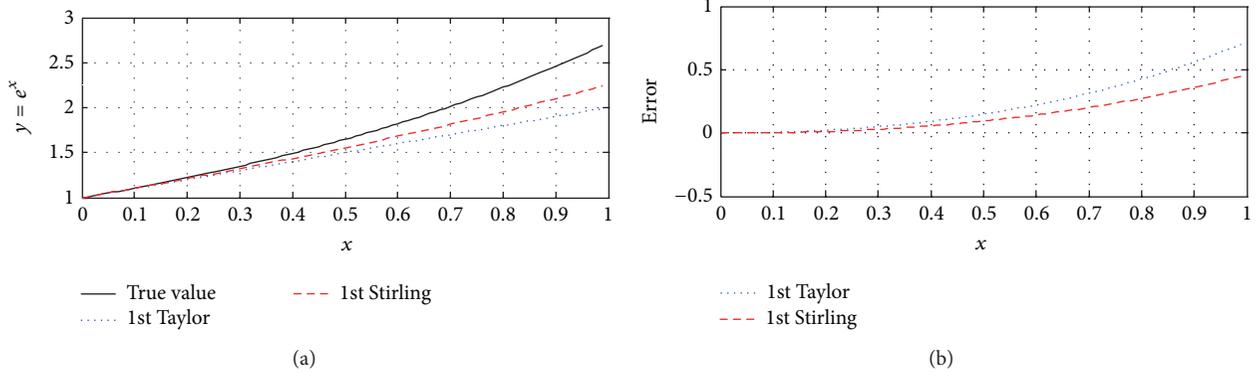


FIGURE 1: Comparison of first-order Stirling series with first-order Taylor series results.

Despite their recent popularity, DDF algorithms require that both the system model and the stochastic information must be accurate. However, this condition cannot be satisfied in many practical situations, which forces the filter to adapt itself to changing conditions. One of the problems with this requirement is that any change in the process introduces measurement noise covariance. In this work, we make use of the theoretical and actual covariance of the innovation sequence, employing SVR to generate the scale factor to tune the noise covariance at each sampling instant when the measurement update step is executed to adapt the filtering algorithm. This paper is organized as follows. Section 2 briefly introduces DDF theory and the proposed SVR-based adaptive strategy. Passive target tracking is then carried out to evaluate the performance of DDF and SVRADDF algorithms using a Monte Carlo simulation in Section 3. Finally, conclusions are provided in Section 4.

2. Development of the Support Vector Regression-Based Adaptive Divided Difference Filter

2.1. Divided Difference Filter. Consider the nonlinear function

$$\mathbf{y} = \mathbf{f}(\mathbf{x}), \quad (1)$$

where $\mathbf{x} \in \mathbf{R}^{n_x}$ and $\mathbf{y} \in \mathbf{R}^{n_y}$. If the function is analytic, then the first-order Taylor series expanded about some point $\mathbf{x} = \bar{\mathbf{x}}$ becomes

$$\begin{aligned} \mathbf{y} = & \left(\mathbf{f}(\bar{\mathbf{x}} + \Delta\mathbf{x}) = \mathbf{f}(\bar{\mathbf{x}}) + \mathbf{D}_{\Delta\mathbf{x}}\mathbf{f} = \mathbf{f}(\bar{\mathbf{x}}) \right. \\ & \left. + \left(\Delta x_1 \frac{\partial}{\partial x_1} + \Delta x_2 \frac{\partial}{\partial x_2} + \cdots + \Delta x_n \frac{\partial}{\partial x_n} \right) \mathbf{f}(\mathbf{x}) \right) \Big|_{\mathbf{x}=\bar{\mathbf{x}}} \end{aligned} \quad (2)$$

with (2) truncated after the first-order term. Note that (2) can achieve a better local approximation if more terms are included. However, such an expanded Taylor series requires derivatives and cannot be fulfilled in some situations. Stirling's interpolation formula is based on a finite number of

evaluations of the function and does not require derivatives, with the first-order approximation yielding

$$\begin{aligned} \mathbf{y} = \mathbf{f}(\bar{\mathbf{x}} + \Delta\mathbf{x}) &= \mathbf{f}(\bar{\mathbf{x}}) + \tilde{\mathbf{D}}_{\Delta\mathbf{x}}\mathbf{f}(\bar{\mathbf{x}}) \\ &= \mathbf{f}(\bar{\mathbf{x}}) + \frac{1}{h} \left(\sum_{i=1}^n \Delta x_i \mu_i \delta_i \right) \mathbf{f}(\bar{\mathbf{x}}), \end{aligned} \quad (3)$$

where h denotes a selected interval length and δ and μ are determined by

$$\begin{aligned} \delta_i \mathbf{f}(\bar{\mathbf{x}}) &= \mathbf{f}\left(\bar{\mathbf{x}} + \frac{h}{2} \mathbf{e}_i\right) - \mathbf{f}\left(\bar{\mathbf{x}} - \frac{h}{2} \mathbf{e}_i\right), \\ \mu_i \mathbf{f}(\bar{\mathbf{x}}) &= \frac{1}{2} \left(\mathbf{f}\left(\bar{\mathbf{x}} + \frac{h}{2} \mathbf{e}_i\right) + \mathbf{f}\left(\bar{\mathbf{x}} - \frac{h}{2} \mathbf{e}_i\right) \right), \end{aligned} \quad (4)$$

with \mathbf{e}_i being the i th unit vector.

Figure 1 compares the results found by using (2) and (3). The function example is $f(x) = e^x$, where $h = 0.56$. From the figure, we can see that Stirling's interpolation provides better accuracy than the Taylor series under the same order approximations.

We now assume that the variable \mathbf{x} has a Gaussian density with mean $\bar{\mathbf{x}}$ and covariance $\mathbf{P}_{\mathbf{x}}$. We can introduce a transformation matrix $\mathbf{S}_{\mathbf{x}}$ which we select as a square Cholesky factor of $\mathbf{P}_{\mathbf{x}}$, such that $\mathbf{P}_{\mathbf{x}} = \mathbf{S}_{\mathbf{x}}\mathbf{S}_{\mathbf{x}}^T$. To illustrate how others can be derived, we introduce the linear transformation of \mathbf{x} :

$$\mathbf{z} = \mathbf{S}_{\mathbf{x}}^{-1}\mathbf{x}. \quad (5)$$

This linear transformation results in a stochastic decoupling of \mathbf{x} as the elements of \mathbf{z} become mutually uncorrelated (see [13]). This changes (3) to

$$\mathbf{y} = \mathbf{f}(\mathbf{S}_{\mathbf{x}}\mathbf{z}) = \tilde{\mathbf{f}}(\mathbf{z}) = \tilde{\mathbf{f}}(\bar{\mathbf{z}}) + \tilde{\mathbf{D}}_{\Delta\mathbf{z}}\tilde{\mathbf{f}}, \quad (6)$$

where $\tilde{\mathbf{D}}_{\Delta\mathbf{z}}\tilde{\mathbf{f}}$ is determined by

$$\tilde{\mathbf{D}}_{\Delta\mathbf{z}}\tilde{\mathbf{f}} = \frac{1}{h} \left(\sum_{i=1}^n \Delta z_i \mu_i \delta_i \right) \mathbf{f}(\bar{\mathbf{z}}). \quad (7)$$

The mean \bar{y} , covariance \mathbf{P}_{yy} , and cross covariance \mathbf{P}_{xy} of y are obtained from

$$\begin{aligned}
\bar{y} &= \mathbf{E}[y] = \mathbf{E}[\tilde{\mathbf{f}}(\bar{\mathbf{z}}) + \tilde{\mathbf{D}}_{\Delta z} \tilde{\mathbf{f}}] = \mathbf{E}[\tilde{\mathbf{f}}(\bar{\mathbf{z}})] = \mathbf{f}(\bar{\mathbf{x}}), \\
\mathbf{P}_{yy} &= \mathbf{E}[(y - \bar{y})(y - \bar{y})^T] \\
&= \mathbf{E}[(\tilde{\mathbf{D}}_{\Delta z} \tilde{\mathbf{f}})(\tilde{\mathbf{D}}_{\Delta z} \tilde{\mathbf{f}})^T] \\
&= \frac{1}{4h^2} \sum_{i=1}^n (\mathbf{f}(\bar{\mathbf{z}} + h\mathbf{e}_i) - \mathbf{f}(\bar{\mathbf{z}} - h\mathbf{e}_i)) \\
&\quad \times (\mathbf{f}(\bar{\mathbf{z}} + h\mathbf{e}_i) - \mathbf{f}(\bar{\mathbf{z}} - h\mathbf{e}_i))^T \\
&= \frac{1}{4h^2} \sum_{i=1}^n (\mathbf{f}(\bar{\mathbf{x}} + h\mathbf{s}_{x,i}) - \mathbf{f}(\bar{\mathbf{x}} - h\mathbf{s}_{x,i})) \\
&\quad \times (\mathbf{f}(\bar{\mathbf{x}} + h\mathbf{s}_{x,i}) - \mathbf{f}(\bar{\mathbf{x}} - h\mathbf{s}_{x,i}))^T, \\
\mathbf{P}_{xy} &= \mathbf{E}[(\mathbf{x} - \bar{\mathbf{x}})(y - \bar{y})^T] \\
&= \mathbf{E}[\mathbf{S}_x \Delta \mathbf{z} (\tilde{\mathbf{D}}_{\Delta z} \tilde{\mathbf{f}})^T] \\
&= \frac{1}{2h} \sum_{i=1}^n \mathbf{s}_{x,i} (\mathbf{f}(\bar{\mathbf{x}} + h\mathbf{s}_{x,i}) - \mathbf{f}(\bar{\mathbf{x}} - h\mathbf{s}_{x,i}))^T,
\end{aligned} \tag{8}$$

where $\mathbf{s}_{x,i}$ is the i th column of the matrix \mathbf{S}_x .

Consider the following nonlinear dynamic system with states to be estimated:

$$\begin{aligned}
\mathbf{x}_{k+1} &= \mathbf{f}(\mathbf{x}_k) + \boldsymbol{\omega}_k \\
\mathbf{y}_k &= \mathbf{h}(\mathbf{x}_k) + \boldsymbol{\nu}_k,
\end{aligned} \tag{9}$$

where $\boldsymbol{\omega}_k$ and $\boldsymbol{\nu}_k$ are assumed to be independent and identically distributed and independent of current and past states, such that $\boldsymbol{\omega}_k \sim \mathbf{N}(\mathbf{0}, \mathbf{Q}_k)$ and $\boldsymbol{\nu}_k \sim \mathbf{N}(\mathbf{0}, \mathbf{R}_k)$.

The DDF takes the same predictor-corrector structures in the EKF and can be described as follows.

Step 1 (initialization). Suppose the state distribution at k instant is $\mathbf{x}_k \sim \mathbf{N}(\hat{\mathbf{x}}_k, \mathbf{P}_k)$, where $\hat{\mathbf{x}}_k$ and \mathbf{P}_k are obtained by

$$\begin{aligned}
\hat{\mathbf{x}}_k &= \mathbf{E}[\mathbf{x}_k], \\
\mathbf{P}_k &= \mathbf{E}[(\mathbf{x} - \hat{\mathbf{x}}_k)(\mathbf{x} - \hat{\mathbf{x}}_k)^T].
\end{aligned} \tag{10}$$

Step 2 (square Cholesky factorizations). Consider the following:

$$\begin{aligned}
\mathbf{P}_k &= \mathbf{S}_x \mathbf{S}_x^T \\
\mathbf{S}_{x\hat{x}} &= \frac{1}{2h} \left\{ \mathbf{f}_i(\hat{\mathbf{x}}_k + h\mathbf{s}_{x,j}) - \mathbf{f}_i(\hat{\mathbf{x}}_k - h\mathbf{s}_{x,j}) \right\},
\end{aligned} \tag{11}$$

where $\mathbf{s}_{x,i}$ is the j th column of \mathbf{S}_x .

Step 3 (state and covariance propagation). One has

$$\hat{\mathbf{x}}_{k+1}^- = \mathbf{f}(\hat{\mathbf{x}}_k) \tag{12}$$

$$\begin{aligned}
\mathbf{P}_{k+1}^- &= \mathbf{S}_{x\hat{x}} (\mathbf{S}_{x\hat{x}})^T + \mathbf{Q}_k \\
&= \mathbf{S}_x^- (\mathbf{S}_x^-)^T
\end{aligned} \tag{13}$$

$$\mathbf{S}_{y\hat{x}} = \frac{1}{2h} \left\{ \mathbf{h}_i(\hat{\mathbf{x}}_{k+1}^- + \mathbf{s}_{x,j}^-) - \mathbf{h}_i(\hat{\mathbf{x}}_{k+1}^- - \mathbf{s}_{x,j}^-) \right\}, \tag{14}$$

where $\hat{\mathbf{x}}_{k+1}^-$ is the predicted state and \mathbf{P}_{k+1}^- is the predicted covariance matrix.

Step 4 (observation and innovation covariance propagation). Consider

$$\hat{\mathbf{y}}_{k+1}^- = \mathbf{h}(\hat{\mathbf{x}}_{k+1}^-) \tag{15}$$

$$\mathbf{P}_{k+1}^{\mathcal{Y}\mathcal{Y}} = (\mathbf{S}_{y\hat{x}}) (\mathbf{S}_{y\hat{x}})^T + \mathbf{R}_k \tag{16}$$

$$\mathbf{P}_{k+1}^{\mathcal{X}\mathcal{Y}} = \mathbf{S}_x^- (\mathbf{S}_{y\hat{x}})^T, \tag{17}$$

where $\hat{\mathbf{y}}_{k+1}^-$ is the predicted observation vector, $\mathbf{P}_{k+1}^{\mathcal{Y}\mathcal{Y}}$ is the innovation covariance matrix, and $\mathbf{P}_{k+1}^{\mathcal{X}\mathcal{Y}}$ is the cross correlation matrix.

Step 5 (update). Consider the following:

$$\begin{aligned}
\boldsymbol{\kappa}_{k+1} &= \mathbf{P}_{k+1}^{\mathcal{X}\mathcal{Y}} (\mathbf{P}_{k+1}^{\mathcal{Y}\mathcal{Y}})^{-1} \\
\mathbf{P}_{k+1}^+ &= \mathbf{P}_{k+1}^- - \boldsymbol{\kappa}_{k+1} \mathbf{P}_{k+1}^{\mathcal{Y}\mathcal{Y}} \boldsymbol{\kappa}_{k+1}^T \\
\mathbf{v}_{k+1} &= \mathbf{y}_{k+1} - \hat{\mathbf{y}}_{k+1}^- \\
\hat{\mathbf{x}}_{k+1}^+ &= \hat{\mathbf{x}}_{k+1}^- + \boldsymbol{\kappa}_{k+1} \mathbf{v}_{k+1},
\end{aligned} \tag{18}$$

where $\boldsymbol{\kappa}_{k+1}$ is the gain, \mathbf{P}_{k+1}^+ is the updated covariance matrix, \mathbf{v}_{k+1} is the innovation vector, and $\hat{\mathbf{x}}_{k+1}^+$ is the updated estimated state.

2.2. Adaptive DDF Algorithm. As stated previously, the DDF algorithm assumes a complete prior knowledge of the process and the measurement noise statistics \mathbf{Q}_k and \mathbf{R}_k . However, \mathbf{Q}_k and \mathbf{R}_k are unknown in most applications, and incorrect priori noise statistics can lead to performance degradation or even divergence for the solution. One of the effective ways to overcome this weakness is to use an algorithm to adapt the noise statistics. In this paper, we propose using a support vector regression adaptive scheme of the DDF to adjust \mathbf{Q}_k and \mathbf{R}_k , respectively.

2.2.1. Support Vector Regression. The principles underlying support vector regression (SVR) developed by Vapnik and are presented in several works (see [26, 27]).

Given the train set $TS = \{(\mathbf{x}_1, y_1), \dots, (\mathbf{x}_m, y_m)\} \in (\mathbf{R}^n \times \mathbf{R})^m$, where $\mathbf{x}_i \in \mathbf{R}^n$, $y_i \in \mathbf{R}$, $i = 1, 2, \dots, m$, the SVR problem can be defined as solving for the nonlinear function $g(\mathbf{x})$ about $\mathbf{x} \in \mathbf{R}^n$ to construct a relationship between the output and an arbitrary input \mathbf{x} :

$$y(\mathbf{x}, \boldsymbol{\omega}) = \sum_{i=1}^m \omega_i g(\mathbf{x}_i) + b = (\boldsymbol{\omega} \cdot \mathbf{g}(\mathbf{x})) + b. \quad (19)$$

In [26], the author shows that changing the regression estimate minimizes the risk functional by using the following form:

$$\begin{aligned} y(\mathbf{x}, \boldsymbol{\alpha}) &= \sum_{i=1}^m (\alpha_i^* - \alpha_i) (g(\mathbf{x}_i) \cdot g(\mathbf{x})) + b \\ &= \sum_{i=1}^m (\alpha_i^* - \alpha_i) K(\mathbf{x}_i, \mathbf{x}) + b, \end{aligned} \quad (20)$$

where α_i^* and α_i are Lagrange multipliers that satisfy the condition $\alpha_i, \alpha_i^* \geq 0$, $\alpha_i \alpha_i^* = 0$ and $K(\mathbf{x}_i, \mathbf{x})$ is a kernel function that satisfies Mercer's condition. In this paper, we use a translation invariant Gaussian kernel, that is, $\exp(-\|\mathbf{x}_i - \mathbf{x}\|^2 / 2\sigma^2)$.

2.2.2. Adaptive Scheme Based on SVR (\mathbf{Q}_k Is Fixed). The covariance matrix \mathbf{R}_k represents the accuracy of the measurement instrument. If we assume that the noise covariance \mathbf{Q}_k is completely known, then we can derive the SVR algorithm to estimate the measurement noise covariance \mathbf{R}_k by defining an adaptive factor Δr_k to get the form:

$$\mathbf{R}_k = \Delta r_k \mathbf{R}, \quad (21)$$

where \mathbf{R} is the constant noise covariance matrix.

This work uses such an SVR algorithm to derive the adaptive factor at time instant k , so as to estimate the value of \mathbf{R}_k during the algorithm's execution.

The innovation sequence \mathbf{v}_{k+1} has a theoretical covariance $\mathbf{P}_{k+1}^{\text{yy}}$, as (16) shows. The actual residual covariance $\mathbf{P}_{k+1}^{\text{yy}}$ can be approximated using its sample covariance by averaging inside a moving window of size N , where

$$\mathbf{C}_k = \frac{1}{N} \sum_{i=k-N+1}^k \mathbf{v}_{k+1} \mathbf{v}_{k+1}^T. \quad (22)$$

If \mathbf{C}_k differs from $\mathbf{P}_{k+1}^{\text{yy}}$, then the diagonal elements of \mathbf{R}_k can be adjusted to minimize the difference as much as possible. The size of this difference is given by

$$\mathbf{DOM}_k = \text{diag}(\mathbf{P}_{k+1}^{\text{yy}} - \mathbf{C}_k) \in \mathbf{R}^n, \quad (23)$$

where the function diag denotes the diagonal elements of the matrix. If the elements of \mathbf{DOM}_k are rounded with zero, the covariance matrix \mathbf{R}_k is likely accurate; otherwise there may be a large deviation between values.

To adjust \mathbf{R}_k , we define the SVR train set $TS_{\Delta r}$ as $TS_{\Delta r} = \{(\mathbf{DOM}_{k-N+1}, \Delta r_{k-N+1}), \dots, (\mathbf{DOM}_k, \Delta r_k)\}$. With this form, we can then create a corresponding function about the

adaptive factor Δr_k and \mathbf{DOM}_k based on (19), such that when \mathbf{DOM}_{k+1} comes, then Δr_{k+1} can be resolved by the function.

To avoid having any element within the train set be too large to affect the accuracy of Δr_k , we use a normalized function of the train set TS ; that is, $TS'_{\Delta r} = \{(\mathbf{DOM}'_{k-N+1}, \Delta r'_{k-N+1}), \dots, (\mathbf{DOM}'_k, \Delta r'_k)\}$, where \mathbf{DOM}'_k and r'_k can be expressed as follows:

$$\begin{aligned} \mathbf{DOM}'_k &= \frac{1}{N} \left\{ \sum_{i=1}^N \mathbf{DOM}_{k-i+1, j} \right\} \quad j = 1, 2, \dots, n, \\ \Delta r'_k &= \frac{1}{N} \sum_{i=1}^N \Delta r_{k-i+1}. \end{aligned} \quad (24)$$

2.2.3. Adaptive Scheme Based on SVR (\mathbf{R}_k Is Fixed). Assuming that the noise covariance \mathbf{R}_k is completely known, we can derive an SVR algorithm to estimate the measurement noise covariance \mathbf{Q}_k . From (13), (14), and (16), we can deduce that a change in \mathbf{Q}_k will affect the covariance matrix $\mathbf{P}_{k+1}^{\text{yy}}$; if we increase \mathbf{Q}_k , then $\mathbf{P}_{k+1}^{\text{yy}}$ also increases. We can adjust \mathbf{Q}_k in the SVR by deliberately mismatching $\mathbf{P}_{k+1}^{\text{yy}}$ and \mathbf{C}_{k+1} .

We first define an adaptive factor Δr_k , where \mathbf{R}_k has the following form:

$$\mathbf{Q}_k = \Delta q_k \mathbf{Q}, \quad (25)$$

where \mathbf{Q} is the constant noise covariance matrix.

We can then define the SVR train set $TS_{\Delta q}$ as $TS_{\Delta q} = \{(\mathbf{DOM}_{k-N+1}, \Delta r_{k-N+1}), \dots, (\mathbf{DOM}_k, \Delta r_k)\}$. At this point, the solution process is the same as for solving Δr_k .

3. Monte Carlo Simulation Results and Discussion

In this section, we report the experimental results obtained by applying SVRADDF to the nonlinear state estimation of a nonmaneuvering target in an underwater tracking control scenario and a maneuvering target in an air-traffic control scenario. To demonstrate the performance of the SVRADDF algorithm, we compare its performance against a DDF algorithm.

3.1. Underwater Nonmaneuvering Target Bearing-Only Tracking Control Scenario. We consider a bearing-only tracking control scenario, where an underwater target executes a uniform motion in a horizontal plane but unknown velocity, while a passive sonar platform performs a uniform circular motion in a horizontal plane. Figure 2 shows a representative trajectory of the target and the passive sonar platform. The kinematics of the relative motion between the target and the platform can be modeled using the following linear process equation:

$$\mathbf{x}_{k+1} = \begin{bmatrix} 1 & 0 & T & 0 \\ 0 & 1 & 0 & T \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \mathbf{x}_k + \mathbf{w}_k. \quad (26)$$

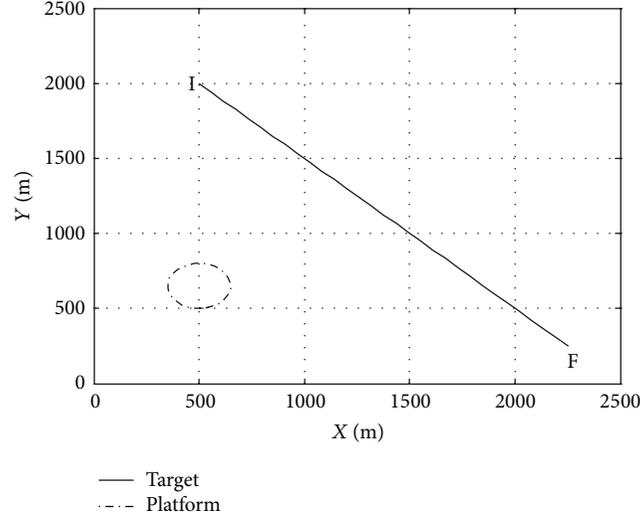


FIGURE 2: Target trajectory (I-initial position and F-final position) and sonar trajectory.

Here, the state of the equation is $\mathbf{x} = [x \ y \ \dot{x} \ \dot{y}]$, where x and y denote position and \dot{x} and \dot{y} denote velocity in the x and y directions, respectively, and T is the time interval between two consecutive measurements and the process noise $\mathbf{w}_k \sim N(0, \mathbf{Q})$ with a nonsingular covariance where

$$\mathbf{Q} = q_1 \times \begin{bmatrix} \frac{T^2}{2} \\ \frac{T^2}{2} \\ T \\ T \end{bmatrix}. \quad (27)$$

In (27), the parameter q_1 is related to process noise intensities. The measurement equation is written as follows:

$$\mathbf{y}_k = \theta_k = \tan^{-1} \left(\frac{y_k}{x_k} \right) + v_k, \quad (28)$$

where the measurement noise $v_k \sim N(0, R)$ with a nonsingular covariance.

Given the following initial conditions:

$$\begin{aligned} T &= 1 \text{ s} \\ q_1 &= 0.0001 \text{ m}^2 \text{ s}^{-3} \\ R &= 0.02 \text{ mrad}, \end{aligned} \quad (29)$$

the true initial state is

$$\mathbf{x}_0 = [0 \text{ m} \ 1500 \text{ m} \ 0 \text{ m s}^{-1} \ 0 \text{ m s}^{-1}]^T \quad (30)$$

and the associated covariance is

$$\mathbf{P}_0 = \text{diag} [100 \text{ m}^2 \ 2000 \text{ m}^2 \ 1 \text{ m}^2 \text{ s}^{-2} \ 1 \text{ m}^2 \text{ s}^{-2}]. \quad (31)$$

The initial estimate state $\hat{\mathbf{x}}_0$ was chosen randomly from $N(\mathbf{x}_0, \mathbf{P}_0)$ in each run, and the total number of scans per run was 1000.

To provide a fair comparison, we performed 50 independent Monte Carlo runs. To track the underwater target, we used both the SVRADDf and the DDF algorithms and compared their performance. The adaptive factor was set to $\Delta q_k = \Delta r_k = 0.1$. Both of the filters were initialized with the same initial conditions for each run.

Performance metrics: to compare the nonlinear performance of the filters, we used the root mean square error (RMSE) of the target position and velocity. The RMSE yields a combined measure of the bias and variance of a filter estimate. The RMSE of the position at time k was found using

$$\text{RMSE}_{\text{pos}}(k) = \sqrt{\frac{1}{N} \sum_{i=1}^N ((x_k^i - \hat{x}_k^i)^2 + (y_k^i - \hat{y}_k^i)^2)}, \quad (32)$$

where (x_k^i, y_k^i) and $(\hat{x}_k^i, \hat{y}_k^i)$ are the true and estimated positions, respectively, in the i th Monte Carlo run. The form for the RMSE of the velocity is similar.

Figures 3 and 4 show the estimated RMSE in target position and velocity. The SVRADDf uses SVR to adjust the adaptive factor during algorithm execution, which leads to a marginally better performance compared to the DDF, as seen in the figures.

3.2. Maneuvering Target Tracking in the Air-Traffic Control Scenario. A typical air-traffic control scenario was considered next, where an aircraft executes a maneuvering turn in a horizontal plane at a constant and known turn rate Ω . Figure 5 shows a representative trajectory of the aircraft.

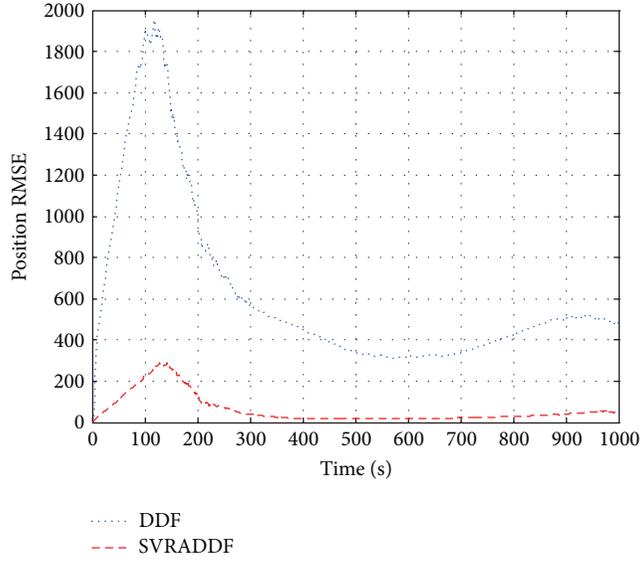


FIGURE 3: RMSE in position for DDF and SVRADDF.

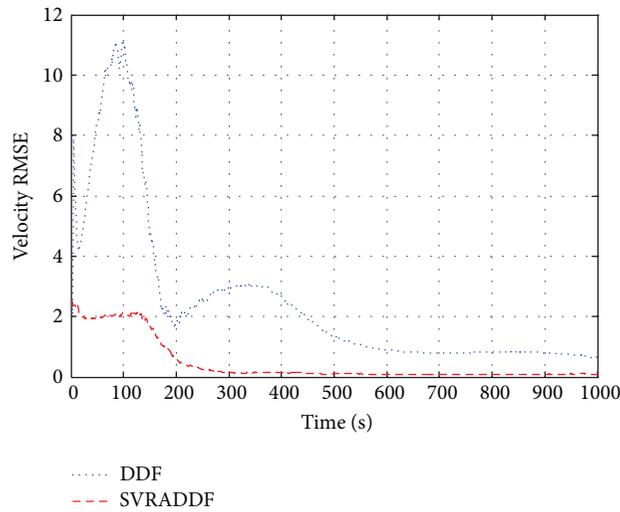


FIGURE 4: RMSE in velocity for DDF and SVRADDF.

The kinematics of the turning motion can be modeled using the following nonlinear process equation:

$$\mathbf{x}_k = \begin{bmatrix} 1 & \frac{\sin \Omega T}{\Omega} & 0 & -\frac{1 - \cos \Omega T}{\Omega} \\ 0 & \cos \Omega T & 0 & -\sin \Omega T \\ 0 & \frac{1 - \cos \Omega T}{\Omega} & 1 & \frac{\sin \Omega T}{\Omega} \\ 0 & \sin \Omega T & 0 & \cos \Omega T \end{bmatrix} \times \mathbf{x}_{k-1} + \mathbf{w}_{k-1}. \quad (33)$$

The state of the aircraft is given by $\mathbf{x} = [x \ \dot{x} \ y \ \dot{y}]^T$, where x and y denote position and \dot{x} and \dot{y} denote velocity in the x and y directions, respectively, and T is the time

interval between two consecutive measurements and the process noise $\mathbf{w}_k \sim N(0, \mathbf{Q})$ with a nonsingular covariance, where

$$\mathbf{Q} = q_1 \times \begin{bmatrix} \frac{T^3}{3} & \frac{T^2}{2} & 0 & 0 \\ \frac{T^2}{2} & T & 0 & 0 \\ 0 & 0 & \frac{T^3}{3} & \frac{T^2}{2} \\ 0 & 0 & \frac{T^2}{2} & T \end{bmatrix}. \quad (34)$$

The parameter q_1 related to process noise intensities. A passive radar is fixed at the origin and equipped to measure

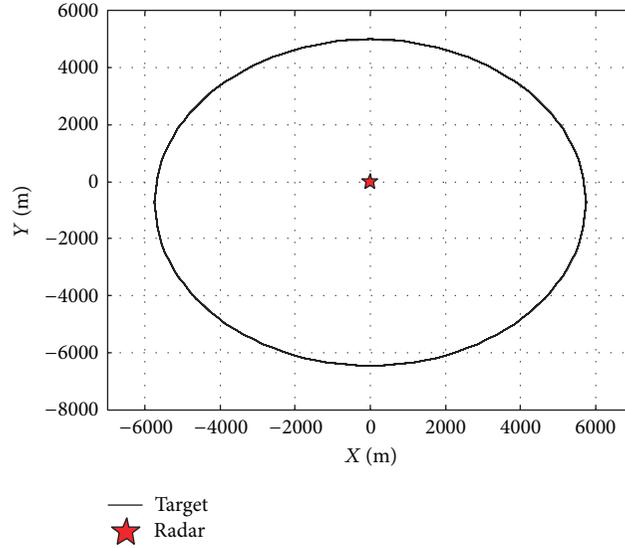


FIGURE 5: Aircraft trajectory and radar location.

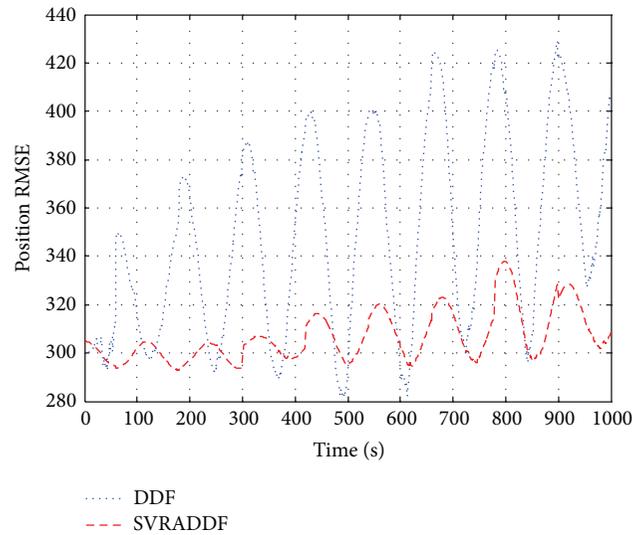


FIGURE 6: RMSE in position for DDF and SVRADDF.

the bearing θ . The measurement equation is written as follows:

$$y_k = \theta_k = \tan^{-1} \left(\frac{y_k}{x_k} \right) + v_k, \quad (35)$$

where the measurement noise $v_k \sim N(0, R)$.

The parameters used in this simulation were $T = 1 \text{ s}$, $\Omega = -3^\circ \text{ s}^{-1}$, $q_1 = 0.1 \text{ m}^2 \text{ s}^{-3}$, and $R = \sqrt{10} \text{ mard}$. The true initial state of the aircraft was $\mathbf{x}_0 = [1000 \text{ m } 300 \text{ m s}^{-1} 1000 \text{ m } 0 \text{ m s}^{-1}]^T$ and the associated covariance matrix was $\mathbf{P}_0 = \text{diag}[100 \text{ m}^2 \ 10 \text{ m}^2 \text{ s}^{-2} \ 100 \text{ m}^2 \ 10 \text{ m}^2 \text{ s}^{-2}]$. The initial state $\hat{\mathbf{x}}_0$ for the filters was chosen randomly from $N(\mathbf{x}_0, \mathbf{P}_0)$ in each Monte Carlo run, and the simulation time per run was 1000.

For a fair comparison, we performed 100 independent Monte Carlo runs for each filter. To track the maneuvering aircraft, we used both the SVRADDF and the DDF algorithms and compared their performance. The adaptive factor was set to $\Delta q_k = 0.3$ and $\Delta r_k = 0.5$. Both filters were initialized with the same initial conditions for each run.

Figures 6 and 7 show the estimate RMSE in target position and velocity for the SVRADDF and DDF filters. Not surprisingly, both filters exhibit divergence due to a mismatch between the initial filter design assumption and the Gaussian noise nature of the problem. The SVRADDF filter exhibits marginally better performance compared to the DDF since it was able to adjust the statistical properties of the noise during execution.

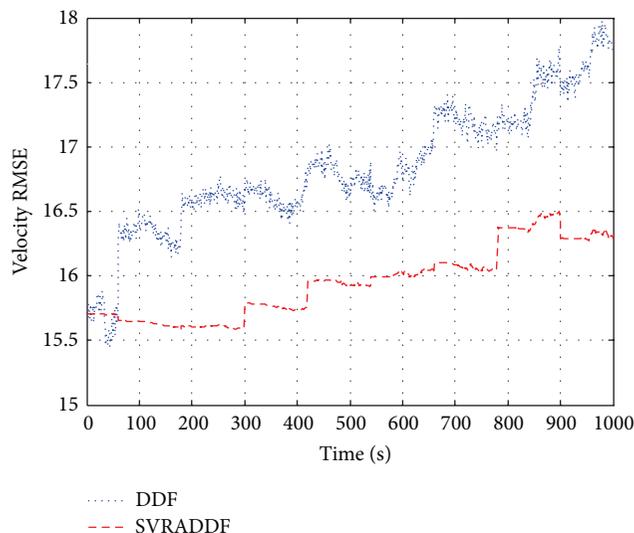


FIGURE 7: RMSE in velocity for DDF and SVRADDF.

4. Conclusions

This work has proposed and developed an innovation-based SVRADDF algorithm. The algorithm introduces an adaptive factor estimated using SVR, which allows for estimation of the noise statistical characteristics of nonlinear stochastic systems. The SVRADDF algorithm avoids instability and divergence in the solution which is caused by incorrect statistical characteristics of the noise. Monte Carlo simulation results of an underwater nonmaneuvering bearing-only target tracking system and a maneuvering target bearing-only tracking system in an air-traffic control setting showed that the SVRADDF algorithm provides better state estimation accuracy than a traditional DDF algorithm.

Although the SVRADDF algorithm has showed better performance under Monte Carlo simulation, there are still several challenging issues to be considered for future study. To improve its feasibility and effectiveness in the complex environment, more comprehensive and detailed studies are still needed to solve under the nonlinear and non-Gaussian noise conditions. Since the algorithm has only been tested under Monte Carlo simulation, the following work might be testing the algorithm under trial data to approve its better performance online.

Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

Acknowledgments

This research work is supported by the National Natural Science Foundation of China (Grant no. E091002/50979017), Ph.D. Programs Foundation of Ministry of Education of China and Basic Technology (Grant no. 20092304110008), and Harbin Science and Technology Innovation Talents of

Special Fund Project (Outstanding Subject Leaders) (no. 2012RFXXG083).

References

- [1] Y. Bar-Shalom and T. Kirubarajan, *Estimation with Applications to Tracking and Navigation*, John Wiley & Sons, New York, NY, USA, 2001.
- [2] S. J. Julier, J. K. Uhlmann, and H. F. Durrant-Whyte, "A new approach for filtering nonlinear systems," in *Proceedings of the American Control Conference*, pp. 1628–1632, June 1995.
- [3] B. M. Bell and F. W. Cathey, "Iterated Kalman filter update as a Gauss-Newton method," *IEEE Transactions on Automatic Control*, vol. 38, no. 2, pp. 294–297, 1993.
- [4] K.-H. Kim, G.-I. Jee, C.-G. Park, and J.-G. Lee, "The stability analysis of the adaptive fading extended kalman filter using the innovation covariance," *International Journal of Control, Automation and Systems*, vol. 7, no. 1, pp. 49–56, 2009.
- [5] X.-X. Wang, L. Zhao, Q.-X. Xia, W. Cao, and L. Li, "Design of unscented Kalman filter with correlative noises," *Control Theory & Applications*, vol. 27, no. 10, pp. 1362–1368, 2010.
- [6] X.-X. Wang, L. Zhao, Q. Pan, Q.-X. Xia, and W. Hong, "Design of UKF with correlative noises based on minimum mean square error estimation," *Control and Decision*, vol. 25, no. 9, pp. 1393–1398, 2010.
- [7] X.-X. Wang, Y. Liang, Q. Pan, C.-H. Zhao, and H.-Z. Li, "Unscented Kalman filter for nonlinear systems with colored measurement noise," *Acta Automatica Sinica*, vol. 38, no. 6, pp. 986–998, 2012.
- [8] W. F. Leven and A. D. Lanterman, "Unscented Kalman filters for multiple target tracking with symmetric measurement equations," *IEEE Transactions on Automatic Control*, vol. 54, no. 2, pp. 370–375, 2009.
- [9] J.-S. Kim, D.-R. Shin, and E. Serpedin, "Adaptive multiuser receiver with joint channel and time delay estimation of CDMA signals based on the square-root unscented filter," *Digital Signal Processing*, vol. 19, no. 3, pp. 504–520, 2009.
- [10] W. Fan and Y. Li, "Accuracy analysis of sigma-point Kalman filters," in *Proceedings of the Chinese Control and Decision*

- Conference (CCDC '09)*, pp. 2883–2888, Guilin, China, June 2009.
- [11] T. Lefebvre, H. Bruyninckx, and J. de Schutter, “Kalman filters for non-linear systems: a comparison of performance,” *International Journal of Control*, vol. 77, no. 7, pp. 639–653, 2004.
 - [12] J. Dunik and M. Simandl, “Performance analysis of derivative-free filters,” in *Proceedings of the 44th IEEE Conference on Decision and Control*, pp. 1941–1946, Seville, Spain, 2005.
 - [13] M. Nørgaard, N. K. Poulsen, and O. Ravn, “New developments in state estimation for nonlinear systems,” *Automatica*, vol. 36, no. 11, pp. 1627–1638, 2000.
 - [14] L. Deok-Jin, *Nonlinear Bayesian filtering with applications to estimation and navigation [Ph.D. thesis]*, 2005.
 - [15] C. Wang, J. Zhang, and J. Mu, “Maximum likelihood-based iterated divided difference filter for nonlinear systems from discrete noisy measurements,” *Sensors*, vol. 12, no. 7, pp. 8912–8929, 2012.
 - [16] N. Subrahmanya and Y. C. Shin, “Adaptive divided difference filtering for simultaneous state and parameter estimation,” *Automatica*, vol. 45, no. 7, pp. 1686–1693, 2009.
 - [17] X. Luo, I. Hoteit, and I. M. Moroz, “On a nonlinear Kalman filter with simplified divided difference approximation,” *Physica D: Nonlinear Phenomena*, vol. 241, no. 6, pp. 671–680, 2012.
 - [18] Z. Ali, M. A. Landolsi, and M. Deriche, “Multiuser parameter estimation using divided difference filter in CDMA systems,” *International Journal of Advanced Science and Technology*, vol. 5, pp. 35–50, 2009.
 - [19] P. G. Bhale, P. N. Dwivedi, P. Kumar, and A. Bhattacharyaa, “Estimation of ballistic coefficient of reentry vehicle with divided difference filtering using noisy RF seeker data,” in *Proceedings of the IEEE International Conference on Industrial Technology (ICIT '06)*, pp. 1087–1092, Mumbai, India, December 2006.
 - [20] M. Jing, W. Chang-yuan, Z. Juan, and C. Fang, “State estimation of reentry target based on divided difference filter,” *Journal of Xi'an Technological University*, vol. 31, no. 4, pp. 377–381, 2011.
 - [21] P. Tripathy, S. C. Srivastava, and S. N. Singh, “A divide-by-difference-filter based algorithm for estimation of generator rotor angle utilizing synchrophasor measurements,” *IEEE Transactions on Instrumentation and Measurement*, vol. 59, no. 6, pp. 1562–1570, 2010.
 - [22] M. Ahmadi, A. Khayatian, and P. Karimaghvae, “Attitude estimation by divided difference filter in quaternion space,” *Acta Astronautica*, vol. 75, pp. 95–107, 2012.
 - [23] P. Setoodeh, A. Khayatian, and E. Farjah, “Attitude estimation by divided difference filter-based sensor fusion,” *The Journal of Navigation*, vol. 60, no. 1, pp. 119–128, 2007.
 - [24] C. Wu and C. Han, “Second-order divided difference filter with application to ballistic target tracking,” in *Proceedings of the 7th World Congress on Intelligent Control and Automation (WCICA '08)*, pp. 6336–6341, Chongqing, China, June 2008.
 - [25] D.-J. Jwo, M.-Y. Hsieh, and S.-Y. Lai, “GPS navigation processing using the quaternion-based divided difference filter,” *GPS Solutions*, vol. 14, no. 3, pp. 217–228, 2010.
 - [26] V. N. Vapnik, *The Nature of Statistical Learning Theory*, Springer, Berlin, Germany, 1995.
 - [27] L. Guozheng, W. Meng, and Z. Huajun, *An Introduction to Support Vector Machines and Other Kernel-Based Learning Methods*, Publishing House of Electronics Industry, Beijing, China, 2004.

Research Article

A Multiagent Transfer Function Neuroapproach to Solve Fuzzy Riccati Differential Equations

Mohammad Shazri Shahrir,^{1,2} N. Kumaresan,² Kuru Ratnavelu,² and M. Z. M. Kamali³

¹ Telekom Research & Development Sdn Bhd, 63000 Cyberjaya, Selangor, Malaysia

² Institut Sains Matematik (ISM), Universiti Malaya, Kuala Lumpur, Malaysia

³ Pusat Asasi Sains, Universiti Malaya, Kuala Lumpur, Malaysia

Correspondence should be addressed to N. Kumaresan; drnk2008@gmail.com

Received 18 February 2014; Revised 25 April 2014; Accepted 25 April 2014; Published 21 May 2014

Academic Editor: Weichao Sun

Copyright © 2014 Mohammad Shazri Shahrir et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

A numerical solution of fuzzy quadratic Riccati differential equation is estimated using a proposed new approach for neural networks (NN). This proposed new approach provides different degrees of polynomial subspaces for each of the transfer function. This multitude of transfer functions creates unique “agents” in the structure of the NN. Hence it is named as multiagent neuroapproach (multiagent NN). Previous works have shown that results using Runge-Kutta 4th order (RK4) are reliable. The results can be achieved by solving the 1st order nonlinear differential equation (ODE) that is found commonly in Riccati differential equation. Multiagent NN shows promising results with the advantage of continuous estimation and improved accuracy that can be produced over Mabood et al. (2013), RK-4, and the existing neuromethod (NM). Numerical examples are discussed to illustrate the proposed method.

1. Introduction

In optimal control theory, solving the Riccati differential equation for state space representation of a dynamical system is a central issue. Chen et al. [1] have shown that the stochastic LQR problem is well posed if there are solutions to the Riccati equation and then optimal feedback control can be obtained. In this paper, a new multiagent NN approach is proposed to solve nonlinear Riccati differential equation that is related to the LQR. Multiagent NN approach improved the results generated by Mabood et al. [2], RK-4, group method of data handling (GMDH), and neuromethod (NM).

Fundamental theories of Riccati equation can be applied to stochastic processes and diffusion problems [3], robust stabilization, network synthesis, and financial mathematics [4, 5]. Apart from the more traditional methods like RK-4 and forward Euler method, there are other nontraditional approaches for solving the aforementioned problems such as the unconditionally stable scheme by Dubois and Saïdi [6]. El-Tawil et al. [7] also have presented the usage of Adomian

decomposition method (ADM) to solve the nonlinear Riccati differential equation in an analytical form. In this method, the ODE can be decomposed to a set of Adomian polynomials.

Tan and Abbasbandy [8] employed the analytic technique called homotopy analysis method (HAM) to solve a quadratic Riccati equation. This technique has been derived from perturbation theory. A modified variational iteration method was used to solve quadratic Riccati equation by Geng [9]. HAM method is the generalization of other methods. But it has some limitations to solve the differential equations.

Multiagent NN has its origins from group method of data handling (GMDH) neural networks. GMDH polynomial neural network was created by the Ukrainian scientist Aleksey Ivakhnenko. GMDH design and weights (parameters of polynomials) adjustment resembles the Kolmogorov-Gabor polynomial by using low order polynomials for every pair of the input values [10]. There are shortcomings of GMDH, firstly the inability to control contribution of subspaces to the final solution and secondly the choice of degree of polynomial sometimes that does not fit with the complexity

of the problem in hand. To overcome the limitations and shortcoming of the existing methods, the new multiagent NN approach is proposed to solve nonlinear fuzzy Riccati differential equation.

The general concept of solving ODE using neural networks is to fit the derivative model of the neural network (as given in (1)) for the given ODE:

$$\min_{w \in (-\infty, +\infty)} \frac{dy}{dx} - NN'(w), \quad (1)$$

where w represents the weights in the neural network.

One of the advantages of neural network is its nonparametric nature. Another advantage is its ability to estimate unseen or untrained points. Previous studies of the first order differential equation (FODE) and second order differential equation (SODE) solution via neural networks [10–12] have been promising. It has also been shown in [10] that neural network methods are more stable and accurate than Euler method, first order implicit method, and second order implicit method. There is no report on NN method [13–16] to solve nonlinear ODE.

Traditionally, the input is given in one form to the NN. In this paper, we propose a multiagent approach (multiagent NN) in which different forms of inputs are used and each transfer function is an agent. Each agent has different degrees of polynomial. Our paper will investigate and show the fact that the fuzzy Riccati differential equation can be solved for fuzzy control system using multiagent NN. In Section 2, the statement of the problem in hand and the corresponding Riccati differential equation are derived. In Section 3, a proposed new approach of multiagent transfer function neuromethod is defined. In Sections 4 and 5, the reader can find the results of fuzzy Riccati differential equation using multiagent NN and conclusion.

2. Statement of the Problem

Any first order ordinary differential equation that is quadratic in the unknown function is called the Riccati equation. It is usually written as

$$\dot{K}_i(t) = q_0(t) + q_1(t)K_i(t) + q_2(t)K_i^2(t). \quad (2)$$

Given the linear time-invariant fuzzy system that can be expressed in the form

$$R^i: \text{If } x_j \text{ is } T_{ji}(\mu_{ji}, \sigma_{ji}), \quad i = 1, 2, \dots, r, \quad j = 1, 2, \dots, n, \quad (3)$$

then

$$\dot{x}(t) = A_i x(t) + B_i u(t), \quad x(0) = x_0, \quad (4)$$

where R^i indicates the i th rule of the fuzzy model, μ_{ji} and σ_{ji} are the mean and standard deviation of the Gaussian membership function, $x(t) \in \mathbb{R}^n$ is a generalized state space vector, $u(t) \in \mathbb{R}^m$ is a control variable, $A_i \in \mathbb{R}^{n \times n}$ and $B_i \in \mathbb{R}^{n \times m}$ are known as coefficient matrices associated with

$x(t)$ and $u(t)$, respectively, x_0 is given initial state vector, and $m \leq n$.

If all state variables are measurable, then a linear state feedback control law is given as

$$u(t) = -R^{-1}B_i^T \lambda(t). \quad (5)$$

Equation (5) can be obtained to the system from (4) and

$$\lambda(t) = K_i(t)x(t), \quad (6)$$

where $K_i(t) \in \mathbb{R}^{n \times n}$ is a symmetric matrix such that $K_i(t_f) = S$.

To minimize both the state and control signals of the feedback control system, a quadratic performance index is minimized:

$$J = \frac{1}{2}x^T(t_f)Sx(t_f) + \frac{1}{2} \int_0^{t_f} [x^T(t)Qx(t) + u^T(t)Ru(t)] dt, \quad (7)$$

where T represents the transpose operator, $S \in \mathbb{R}^{n \times n}$ and $Q \in \mathbb{R}^{n \times n}$ are symmetric and positive definite (semidefinite), and $R \in \mathbb{R}^{m \times m}$ is a symmetric and positive definite weighting matrix for $u(t)$.

Based on standard procedure, J can be minimized by minimizing the Hamiltonian equation

$$H(x(t), u(t), \lambda(t)) = \frac{1}{2}x^T(t)Qx(t) + \frac{1}{2}u^T(t)Ru(t) + \lambda^T(t)[A_i(x) + B_i(t)]. \quad (8)$$

The necessary condition for optimality $(\partial H / \partial u)(x, u, \lambda, t) = 0$ implies that $Ru(t) + B_i^T \lambda(t) = 0$ and

$$\frac{\partial H}{\partial x} = \dot{\lambda}(t), \quad (9)$$

$$\rightarrow \dot{\lambda}(t) = -Qx(t) - A_i^T \lambda(t),$$

$$\frac{\partial H}{\partial \lambda} = \dot{x}(t), \quad (10)$$

$$\rightarrow \dot{x}(t) = A_i x(t) - B_i u(t),$$

and, from (5), we have

$$\dot{x}(t) = Ax(t) - B_i R^{-1} B_i^T \dot{\lambda}(t). \quad (11)$$

Equations (9) and (11) can be written in a matrix form as follows:

$$\begin{bmatrix} \dot{x}(t) \\ \dot{\lambda}(t) \end{bmatrix} = \begin{bmatrix} A_i & -B_i R^{-1} B_i^T \\ -Q & -A_i^T \end{bmatrix} \begin{bmatrix} x(t) \\ \lambda(t) \end{bmatrix}, \quad (12)$$

where $x(0) = x_0$.

From (6), we have

$$\dot{\lambda}(t) = \dot{K}_i x(t) + K_i \dot{x}(t). \quad (13)$$

Substituting (9) and (11) into (13) will generate the following equation:

$$\left[\dot{K}_i(t) + K_i(t) A_i + A_i^T K_i(t) + Q - K_i(t) B_i R^{-1} B_i^T K_i(t) \right] x(t) = 0. \quad (14)$$

Since (14) holds for all nonzero $x(t)$, the term premultiplying $x(t)$ must be zero. Therefore, we obtain the following fuzzy Riccati differential equation for the fuzzy linear system (4):

$$\left[\dot{K}_i(t) + K_i(t) A_i + A_i^T K_i(t) + Q - K_i(t) B_i R^{-1} B_i^T K_i(t) \right] = 0. \quad (15)$$

Assume $m = n = 1$; then (15) becomes

$$\dot{K}_i(t) + 2A_i K_i(t) + Q - B_i^2 R K_i^2(t) = 0. \quad (16)$$

Fuzzy systems are control systems where the feedback loop is based on fuzzy logic. Fuzzy-feedback system mechanism is based on imprecise logic or “fuzzy logic.” The following are the rules applied to the fuzzy system:

Fuzzy systems are control systems where the feedback loop is based on fuzzy logic. Fuzzy-feedback system mechanism is based on imprecise logic or “fuzzy logic.” The following are the rules applied to the fuzzy system:

Fuzzy Rule 1:

$$K = y, \quad Q = 1, \quad A_1 = 0, \quad B_1^2 R = 1: \quad (17)$$

$$\frac{dy}{dx} = -y^2(t) + 1, \quad y(0) = 0,$$

Fuzzy Rule 2:

$$K = y, \quad Q = 1, \quad A_2 = 1, \quad B_2^2 R = 1: \quad (18)$$

$$\frac{dy}{dt} = 2y(t) - y^2(t) + 1, \quad y(0) = 0.$$

This paper investigates the performance of solving the aforementioned equations generated from Fuzzy Rules 1 and 2.

3. Neural Networks

By mapping the output of the neural network to the expected value and using the characteristics of NM (19), the numerical integration can be acquired [11]:

$$\frac{d\sigma(z_i)}{dx} = \sigma(z_i)(1 - \sigma(z_i)). \quad (19)$$

Multiagent NN differs from NM. In the present NM, only one form of input is used in the transfer functions whereas multiagent NN has given different form of inputs in the transfer functions.

Neural methods, being NM or multiagent NN, use neural networks to estimate solution for ODE. Traditionally neural network has shown performance capability as a *function estimator* [10] and *regressor* [10].

In the neuroapproach, to solve ODE [12], $\sigma(z_i) = 1/(1 + e^{-\alpha z_i})$ which is the transfer function in the neural network process. Part of NM process is to find the derivative of neural networks using (19). In our proposed multiagent NN, different agents of σ have been used in the transfer functions.

3.1. Multiagent NN Algorithm

Step 1. Feed the input vector t_j .

Step 2. Initialize the weight matrix w_{ij} and bias u_i .

Step 3. Compute $z_i = \sum_{j=1}^n w_{ij} t_j^p$, where p is the power of the input.

Step 4. Compute $\sigma_i(z_i) = z_i$.

Step 5. Initialize weight vector v_i hidden to output layer.

Step 6. Calculate $N_{ij} = \sum_{i=1}^n v_i \sigma(z_i)$.

Step 7. Compute purelin function of N_{ij} .

Step 8. Minimize the following fitness function:

$$E_r = \sum_{i,j=1}^n \left(\frac{d^2 y}{dx^2} - \theta \left(y, \frac{dy}{dx} \right) \right)^2, \quad (20)$$

where $y = A_i + t_j N_{ij}$ is the trial function for first order and r is the epoch. The proposed multiagent NN method converges when the error tends to zero as given in (21). The convergence of method usually takes around 10–15 mins on an i7 Processor, 4-Gigabyte RAM system. The method does not need to be recomputed for the untrained points within the bounds of the model or problem defined. Consider

$$\lim_{r \rightarrow \infty} E_r \rightarrow 0. \quad (21)$$

The k th order derivative of neural networks is shown in

$$\frac{d^k N}{dx_j^k} = \sum_{i=1}^H v_i w_{ij}^k \sigma_i^k. \quad (22)$$

Example of derived fitness function of second order in Matlab (to solve $y'' = 2$) is as follows:

$$\text{Error} = 2.*(\text{Weight1t2}.*\text{Weight2t3}'*(\text{SigOut1}_')) + ((\text{Input}(:,\text{cnt}).*((\text{Weight2t3}'.*\text{Weight1t2}.*\text{Weight1t2})*\text{SigOut1}_')) - 2.$$

The above fitness function can easily be derived using chain rule:

$$\text{Error} = (\text{Out2}_ + ((\text{Input}(:,\text{cnt}).*(\text{Weight1t2}*(\text{Weight2t3}.*\text{SigOut1}_')))))) - \text{TargetFunction}.$$

The fitness function in (20) is simply defined by the derivatives of neural networks in the given differential equation.

TABLE 1: Solution of fuzzy Riccati differential equation by various methods for Fuzzy Rule 1.

x	RK-4	Mabood et al. [2]	NM	GMDH	Multiagent NN	Analytical
0	0	0	0	0	0	0
0.1	0.1	0.099668	0.095826	0.024865	0.099668	0.099668
0.2	0.199	0.197376	0.199513	0.055562	0.197375	0.197375
0.3	0.29504	0.291315	0.298841	0.093447	0.291313	0.291313
0.4	0.386335	0.379949	0.391187	0.139876	0.379949	0.379949
0.5	0.47141	0.462092	0.472797	0.19621	0.462121	0.462117
0.6	0.549187	0.53691	0.544233	0.26381	0.537077	0.53705
0.7	0.619026	0.603815	0.606699	0.344037	0.604513	0.604368
0.8	0.680707	0.662245	0.661873	0.438244	0.664640	0.664037
0.9	0.734371	0.711287	0.710854	0.547766	0.718390	0.716298
1	0.780441	0.749123	0.75499	0.673906	0.767899	0.761594
SAE	0.101748	0.019995	0.059152	1.936046	0.009178	

TABLE 2: Solution of fuzzy Riccati differential equation by various methods for Fuzzy Rule 2.

x	RK-4	Mabood et al. [2]	NM	GMDH	Multiagent NN	Analytical
0	0	0	0	0	0	0
0.1	0.1	0.110328	0.109895	0.062432	0.110295	0.110295
0.2	0.219	0.242273	0.237926	0.139493	0.241976	0.241977
0.3	0.358004	0.396175	0.416948	0.234568	0.395089	0.395105
0.4	0.516788	0.570231	0.652779	0.351046	0.56766	0.567812
0.5	0.693439	0.759555	0.930998	0.492328	0.755134	0.756014
0.6	0.884041	0.955094	1.200455	0.661821	0.949964	0.953566
0.7	1.082696	1.142444	1.396588	0.862943	1.141423	1.152949
0.8	1.282012	1.300569	1.519028	1.099099	1.315723	1.346364
0.9	1.474059	1.400444	1.616554	1.373660	1.456545	1.526911
1	1.651586	1.403645	1.694985	1.689896	1.546032	1.689498
SAE	0.478868	0.477505	1.008306	1.774001	0.25181	

4. Results and Discussion

The following are the generated results from the FODE of Fuzzy Rule 1 and Fuzzy Rule 2. The methods used to solve both of the FODEs are Runge-Kutta 4th order (RK-4), Mabood et al. [2], neuromethod (NM), GMDH, and the proposed new method (multiagent NN).

Revisiting Fuzzy Rule 1, $K = y, Q = 1, A_i = 0,$ and $B_i^2 R = 1$:

$$\frac{dy}{dx} = -y^2(t) + 1, \quad y(0) = 0. \tag{23}$$

Revisiting Fuzzy Rule 2, $K = y, Q = 1, A_i = 1,$ and $B_i^2 R = 1$:

$$\frac{dy}{dt} = 2y(t) - y^2(t) + 1, \quad y(0) = 0. \tag{24}$$

4.1. Multiagent NN Solution. In multiagent NN, one input layer, one hidden layer containing seven neurons, and one output layer are taken for training the NN (Figure 1). In the hidden layer, linear functions are assigned to the neurons. Initially the weights from input layer to hidden layer and from

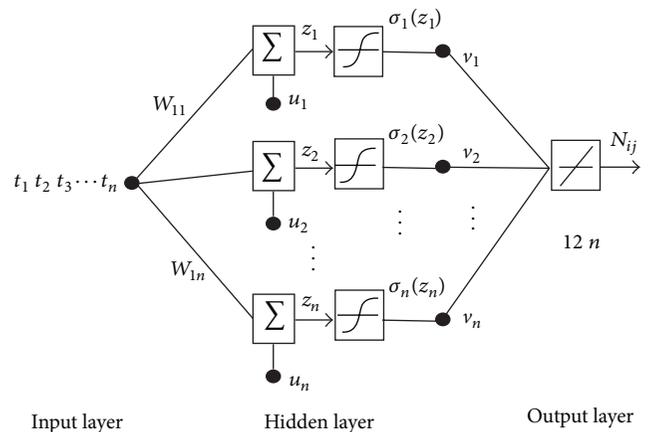


FIGURE 1: A neural network with arbitrary number of hidden nodes.

hidden layer to output layer are taken randomly. Then the weights are updated using gradient decent method while NN is under training the inputs. The NN solutions are compared with the solutions obtained by the existing methods. All the solutions are presented in Tables 1 and 2. All the solutions curves are displayed in Figures 2 and 3.

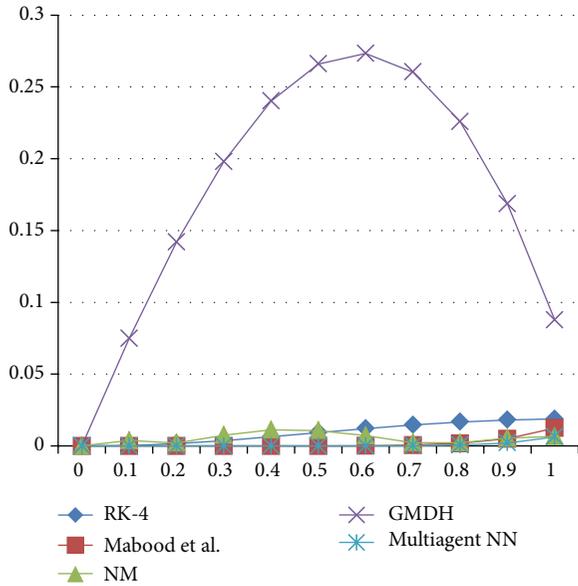


FIGURE 2: Solution curves by various methods for Fuzzy Rule 1.

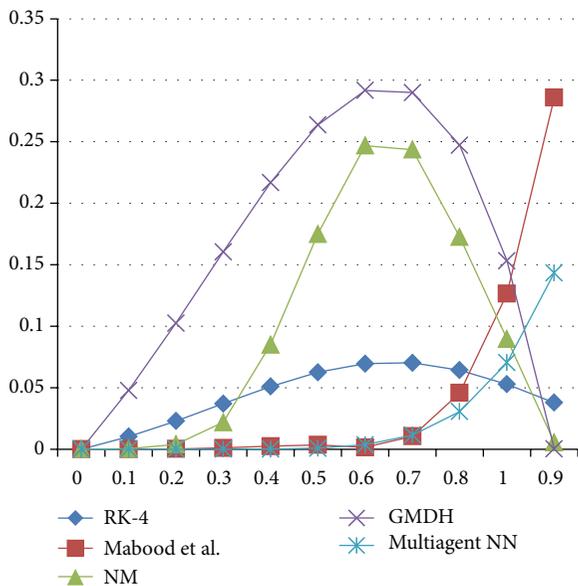


FIGURE 3: Solution curves by various methods for Fuzzy Rule 2.

For Fuzzy Rule 1, in Table 1, nontraditional methods (except GMDH) in all cases outperformed traditional method (RK-4).

In the case of Fuzzy Rule 2, only 2 methods, Mabood et al. [2] (only 0.3% improvement against RK-4) and multiagent NN, outperformed RK-4.

For Rule 1, multiagent NN improvement against RK-4 is 91%, Mabood et al. [2] is 54%, neuromethod (NM) is 90%, and GMDH is 99.5%.

For Rule 2, multiagent NN improvement against RK-4 is 47%, Mabood et al. [2] is 47%, neuromethod (NM) is 75%, and GMDH is 85.8%.

5. Conclusions

Multiagent neural network method has shown improved performance against RK-4, Mabood et al. [2], and neuromethod (NM). Added advantage of using neuromethod is its ability of continuous estimation. Because of this ability of continuous estimation, points, which are not explicitly trained, can be estimated without going through another phase of computational processing.

Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

Acknowledgments

The authors are very much thankful to the referees for their valuable comments and suggestions to improve this paper.

References

- [1] S. Chen, X. Li, and X. Y. Zhou, "Stochastic linear quadratic regulators with indefinite control weight costs," *SIAM Journal on Control and Optimization*, vol. 36, no. 5, pp. 1685–1702, 1998.
- [2] F. Mabood, A. I. M. Ismail, and I. Hashim, "Application of optimal homotopy asymptotic method for the approximate solution of Riccati equation," *Sains Malaysiana*, vol. 42, no. 6, pp. 863–867, 2013.
- [3] W. T. Reid, *Riccati Differential Equations*, vol. 86 of *Mathematics in Science and Engineering*, Academic Press, New York, NY, USA, 1972.
- [4] B. D. O. Anderson and J. B. Moore, *Optimal Control-Linear Quadratic Methods*, Prentice-Hall, Upper Saddle River, NJ, USA, 1989.
- [5] I. Lasiecka and R. Triggiani, *Differential and Algebraic Riccati Equations with Application to Boundary/Point Control Problems: Continuous Theory and Approximation Theory*, vol. 164 of *Lecture Notes in Control and Information Sciences*, Springer, Berlin, Germany, 1991.
- [6] F. Dubois and A. Saïdi, "Unconditionally stable scheme for Riccati equation," in *Contrôle des Systèmes Gouvernés par des Équations aux Dérivées Partielles (Nancy, 1999)*, vol. 8 of *ESAIM Proceedings*, pp. 39–52, Société de Mathématiques Appliquées et Industrielles, Paris, France, 2000.
- [7] M. A. El-Tawil, A. A. Bahnasawi, and A. Abdel-Naby, "Solving Riccati differential equation using Adomian's decomposition method," *Applied Mathematics and Computation*, vol. 157, no. 2, pp. 503–514, 2004.
- [8] Y. Tan and S. Abbasbandy, "Homotopy analysis method for quadratic Riccati differential equation," *Communications in Nonlinear Science and Numerical Simulation*, vol. 13, no. 3, pp. 539–546, 2008.
- [9] F. Geng, "A modified variational iteration method for solving Riccati differential equations," *Computers and Mathematics with Applications*, vol. 60, no. 7, pp. 1868–1872, 2010.
- [10] S. Ferrari and R. F. Stengel, "Smooth function approximation using neural networks," *IEEE Transactions on Neural Networks*, vol. 16, no. 1, pp. 24–38, 2005.
- [11] N. Kumaresan, "Solution of generalized matrix Riccati differential equation for indefinite stochastic linear quadratic singular

- fuzzy system with cross-term using neural networks,” *Neural Computing and Applications*, vol. 21, no. 3, pp. 497–503, 2012.
- [12] I. E. Lagaris, A. C. Likas, and D. G. Papageorgiou, “Neural-network methods for boundary value problems with irregular boundaries,” *IEEE Transactions on Neural Networks*, vol. 11, no. 5, pp. 1041–1049, 2000.
- [13] M. S. Bin Shahrir, K. Ratnavely, and N. Kumaresan, “A Neural Network (NN) approach to solving a static-non-exchange scattering of electron-hydrogen,” in *Proceeding of the 5th IEEE International Conference on Computational Intelligence, Modelling and Simulation*, pp. 14–16, 2013.
- [14] J. Fojdl and R. W. Brause, “The performance of approximating ordinary differential equations by neural nets,” in *Proceedings of the 20th IEEE International Conference on Tools with Artificial Intelligence (ICTAI '08)*, pp. 457–464, November 2008.
- [15] S. Chandana and R. V. Mayorga, “Can we work around numerical methods? An insight,” in *Proceedings of the 21st National Conference on Artificial Intelligence (AAAI '06)*, vol. 2, pp. 1859–1860, July 2006.
- [16] Z. Zeng, Y.-N. Wang, and H. Wen, “Numerical integration based on a neural network algorithm,” *Computing in Science and Engineering*, vol. 8, no. 4, pp. 42–48, 2006.

Research Article

Sign Stability for Switched Linear Systems and Its Application in Flight Control

Qing Wang,¹ Tong Wang,¹ Yanze Hou,² and Chaoyang Dong³

¹ Department of Automation Science and Electrical Engineering, Beijing University of Aeronautics and Astronautics, Beijing 100191, China

² Institute of Manned Space System Engineering, China Academy of Space Technology, Beijing 100094, China

³ Department of Aeronautic Science and Engineering, Beijing University of Aeronautics and Astronautics, Beijing 100191, China

Correspondence should be addressed to Tong Wang; xenon@buaa.edu.cn

Received 8 December 2013; Revised 21 April 2014; Accepted 23 April 2014; Published 13 May 2014

Academic Editor: Yuxin Zhao

Copyright © 2014 Qing Wang et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

The sign stability concept in ecological systems is introduced into the analysis and synthesis of switched linear system to explore new control design technique. The necessary and sufficient condition for sign stability of a switched linear system under arbitrary switching is achieved via the notion of complete isogenous sign stable set (CISSS). A new approach for the stabilization of switched system is presented. Although the controllers are devised for each subsystem, respectively, the switched system is sign stabilized by the constitution of CISSS. The provided method has natural robustness and more design freedoms than the familiar Lyapunov function method, which bears relative conservativeness as the requirement of solving LMIs. The presented technique is validated by an example of flight control within a large-scale flight envelop. Simulation results indicate that the proposed method can stabilize the flight attitude under large variations of system parameters and external perturbations.

1. Introduction

The sign stability (or qualitative stability) concept is first proposed in bionomics and utilized to analyze the interactions of different species in a large-scale ecosystem which lacks exact model but presents high robustness under various perturbations [1, 2]. The sign stability approach is also applied to population biology and economics in respect that these systems are also short of quantitative mathematical models. This qualitative analysis technique provides an avenue to research the linear system stability by the Jacobian matrix with only signs, and still attracts growing attention in system science community [3, 4].

For a given matrix, the signs (+, −, or 0) of its elements are taken to make up a new matrix named the sign-pattern. A matrix is called sign stable if arbitrary matrix which has the same sign-pattern is Hurwitz stable, regardless of the elements' magnitudes. Hence, the sign stability of a matrix equates to that of the corresponding sign-pattern. Jeffries developed necessary and sufficient conditions for sign stability of matrix and proposed an approach named the color

test to verify an arbitrary matrix in ecological terms [5, 6]. Yedavalli translated the color test conditions in matrix theory notation and devised a programmable set of conditions for the color test in an irreducible matrix [7]. It is worth noticing that the conclusions in this work are based on the conditions and criterions for the matrix sign stability presented in the aforementioned literatures.

Based on theoretical researches, the sign stability technique is applied in control engineering, especially the robust controllers design for aerospace flight control system. A three-axis attitude stabilization controller for an axisymmetric satellite is provided in [7]. The closed-loop system matrix is designed to possess the specific sign-pattern and sign stability property which will bring the system robust stability under arbitrarily large variations in the spin angular velocity. In [8], the effect of the elements' signs on the matrix properties such as eigenvalues and condition number is shown. Efforts are made to identify target closed-loop systems that incorporate the desirable features of ecological systems, and an algorithm for the design of controller is

given. The control design procedure is illustrated with two applications in the aerospace field: satellite attitude control and aircraft lateral dynamics control.

However, the existing applications of sign stability only consider linear time-invariant systems which cannot approximate the flight dynamics within the full flight envelope. One of the alternative solutions is the switched linear system which is a good approximation of complicated system characteristics. Flight-control-oriented analysis and synthesis of switched systems are studied in the prior literatures and show large potential in engineering practice [9–12]. For example, the gain schedule control that is widely applied in flight control systems can be abstracted as a switching control law if variations of system parameters are regarded as a series of switching with the transition of the schedule variable. Hence, the aim of this paper is to generalize the sign stability concept to switched linear system and develop a new and practicable control scheme.

Compared to the prior references, the significant contribution of this work is the conclusion on sign stability analysis of switched linear system, which is presented by a necessary and sufficient condition. The system model treated with sign stability approach is extended from LTI system to a hybrid system. The mix of continuous linear systems and discrete signals make switched system suitable to describe a wide range of engineering systems such as power systems, automotive engine control, flight control, and networked control systems. As a result, the application areas of sign stability approach are enlarged remarkably.

Another important contribution is that a new stabilization technique of switched system is presented. As for the asymptotic stability of switched system under arbitrary switching, the common Lyapunov function is a necessary and sufficient condition and a usual method. In [13, 14], the existing condition of common quadratic Lyapunov function for switched linear system is discussed. Furthermore, multiple Lyapunov functions method [15], dwell-time method [16], and average dwell-time method [17] are proposed to adapt different design situations. In recent years, Zhao and his coworkers propose the mode-dependent average dwell-time method in order to decrease conservativeness [18, 19]. To utilize the above methods, the LMI technique is widely used, whereas the feasible solution of LMI is only a sufficient condition for the existence of common Lyapunov function, and an expectant controller may not exist or may not be found even if one does exist [20]. That leads to uncertainty of the controller solvability, even though modern mathematic tools are utilized. In contrast, the sign stability theory offers a fire-new approach with which designers can configure target sign-patterns of closed-loop switched subsystems via state feedback and guarantee the stability under arbitrary switching. The primary advantage of the sign stability technique is that there is no need to solve LMIs. That provides more design freedoms than Lyapunov function approach. Besides, the provided control synthesis technique possesses natural robustness because the sign stability is independent of the elements' magnitudes. This property allows the sign stabilization controllers to fit discretionarily

large sign-preserving parameter perturbations, especially in the flight control system.

With this consideration, the paper is organized as follows. In Section 2, the basic concepts and conditions for the sign stability of matrix are reviewed. The definition of sign stability for switched linear system is presented to formulate the main issue of this paper. In Section 3, the main result of this paper is given in the form of necessary and sufficient condition for the sign stability of switched system along with a new approach for the stabilization problem. In Section 4, the proposed technique is illustrated by an application to flight control of the HiMAT vehicle. Finally, Section 5 provides a detailed conclusion of this work.

2. Problem Formulation

In this section, preliminary knowledge about the sign stability is reviewed, and then the definition of sign stability for switched linear system is proposed to formulate the main problem of this paper.

2.1. Preliminary Knowledge. A switched linear autonomous system is usually described as

$$\dot{x}(t) = A_{\sigma(t)}x(t), \quad (1)$$

in which $\sigma(t) : [0, t) \rightarrow \Omega = \{1, 2, \dots, n\}$ is the switching signal and $\{A_i, i \in \Omega\}$ is the set of switching subsystem matrices. Then the aim of this section is to establish the concept of sign stability for the switched system (1). Above all, basic knowledge and results are needed.

For the $n \times n$ real matrix $A = (a_{ij})$ and $B = (b_{ij})$, if $\text{sgn } a_{ij} = \text{sgn } b_{ij}$, for all i, j , it is defined that the two matrices have the same sign-pattern. The sign-pattern can be represented as a matrix, of which all the entries simply consist of signs $+$, $-$, or 0 . The $\text{sgn}(A)$ is defined as the sign-pattern matrix of A . Similarly, $\text{sgn}\{A_i, i \in \Omega\}$ is defined as the sign-pattern set of the matrix set $\{A_i, i \in \Omega\}$; that is, $\text{sgn}\{A_i\} = \{\text{sgn}(A_i), i \in \Omega\}$. For two sets of matrices $\{A_i, i \in \Omega\}$ and $\{B_j, j \in \Delta\}$, if for all $i \in \Omega$ $\text{sgn}(A_i) \in \text{sgn}\{B_j\}$ and for all $j \in \Delta$ $\text{sgn}(B_j) \in \text{sgn}\{A_i\}$, it is defined that the two matrix sets have the same sign-pattern set.

Definition 1. The $n \times n$ real matrix $A = (a_{ij})$ is sign stable if each matrix $B = (b_{ij})$ of the same sign-pattern as A is Hurwitz stable.

The necessary and sufficient condition for the sign stability of matrix is provided in referenced researches [5, 21–23]. The main necessary conditions are listed in Lemma 2 as the foundation of main results in this paper.

Lemma 2. *The following are necessary conditions for sign stability of matrix $A = (a_{ij})$.*

- (1) For all i , $a_{ii} \leq 0$.
- (2) $a_{ii} < 0$ for at least one i .
- (3) For all $i \neq j$, $a_{ij}a_{ji} \leq 0$.

(4) $a_{ij}a_{jk} \cdots a_{qr}a_{ri} = 0$ for any sequences of three or more distinct indices i, j, k, \dots, q, r .

(5) $\det A \neq 0$.

It is worth announcing that the color test, given by Jeffries in [5], is confirmed as the criterion for the sign stability of any matrix or sign-pattern in this work.

2.2. Problem Statement. Compared with the definition of sign stability for matrix, sign stability of switched linear system is defined as below.

Definition 3. The switched linear system (1) is sign stable under arbitrary switching if each switched system $\dot{x}(t) = A_{\sigma(t)}x(t)$ which has the same sign-pattern set as $\{A_{\sigma(t)}\}$ is asymptotically stable under arbitrary switching.

From Definition 3 it is seen that, for a given switched system, that is, sign stable, it is obviously asymptotically stable under arbitrary switching. Besides, the determinant of sign stability is the sign-pattern set, and the concept of sign stability for the sign-pattern set $\text{sgn}\{A_{\sigma}\}$ is similar to that for the switched system. We can also know that a sign stable switched system will be achieved by assigning arbitrary values for elements of sign-patterns in a sign-pattern set which is sign stable. In addition, the number of subsystems will not affect the stability of the switched system.

Remark 4. In a sign stable switched system, there may be several subsystems of which the matrices have the same sign-pattern. In other words, the switched system derived from a given sign stable sign-pattern set is stable even if more than one matrices are built according to the same sign-pattern.

Remark 5. The subset of an arbitrary sign stable switched system (or sign-pattern set) is also sign stable. The conclusion is obtained under arbitrary switching laws.

With the concept given in Definition 3, we are now interested in the properties of a sign stable switched linear system, and it is also important to develop a technique to constitute a sign stable switched system. Towards the above targets, the main results are provided in the next section. To accomplish the proof of the main theorem, a primary conclusion given by [24] is used and presented here as a lemma.

Lemma 6. *The following statements are equivalent:*

(1) *the switched linear system*

$$\dot{x}(t) = A_{\sigma(t)}x(t), \quad (2)$$

where $A_{\sigma(t)} \in \{A_1, A_2, \dots, A_N\}$ is asymptotically stable under arbitrary switching;

(2) *the linear time-variant system*

$$\dot{x}(t) = F(t)x(t),$$

$$F(t) = \left\{ \sum_{i=1}^N \alpha_i(t) A_i \mid \alpha_i(t) \geq 0, \sum_{i=1}^N \alpha_i(t) = 1 \right\} \quad (3)$$

is asymptotically stable.

3. Main Results

3.1. Necessary Condition. From Definition 3 and Lemma 6, the sign stability of the switched linear system (2) under arbitrary switching implies that each sign-pattern in $\text{sgn}\{A_{\sigma}\}$ is sign stable. The necessary conditions of a sign stable switched linear system are given in the following theorem, where a_{ij}^k denotes the $i \times j$ th element of the k th subsystem matrix of switched system (2).

Theorem 7. *The following are necessary conditions for sign stability of switched linear system (2) under arbitrary switching.*

- (1) *For all A^k , which is the k th subsystem matrix of switched system (2), the necessary conditions in Lemma 2 are satisfied.*
- (2) *For all $p \neq q$, for all $i \neq j$, one of the following three conditions is satisfied: (i) $a_{ij}^p = a_{ij}^q = 0$; (ii) $a_{ji}^p = a_{ji}^q = 0$; (iii) $a_{ij}^p a_{ij}^q \geq 0$.*

Proof. For (1), noting that each subsystem of switched system (2) is sign stable, the conclusion is obvious.

For (2), with no loss of generality, it is supposed that there are 2 subsystems in switched system (2). From Lemma 6, the linear time-variant system (3) is asymptotically stable, and the system matrix can be written as

$$F(t) = \left\{ \alpha_1(t) A^1 + \alpha_2(t) A^2 \mid \alpha_1(t), \alpha_2(t) \geq 0, \alpha_1(t) + \alpha_2(t) = 1 \right\}. \quad (4)$$

It is easily understood that $F(t)$ is sign stable because the elements of A^1 and A^2 may take arbitrary values. Define f_{ij} as the $i \times j$ th element of $F(t)$, and it holds that

$$f_{ij}(t) = \alpha_1(t) a_{ij}^1 + \alpha_2(t) a_{ij}^2. \quad (5)$$

Due to condition (3) of Lemma 2, the following inequality holds for all $i \neq j$:

$$f_{ij}(t) f_{ji}(t) = \left(\alpha_1(t) a_{ij}^1 + \alpha_2(t) a_{ij}^2 \right) \left(\alpha_1(t) a_{ji}^1 + \alpha_2(t) a_{ji}^2 \right) \leq 0. \quad (6)$$

(i) For the situation that $f_{ij}(t) f_{ji}(t) = 0$, since $\alpha_1(t)$ and $\alpha_2(t)$ may take arbitrary values, it must hold that $a_{ij}^1 = a_{ij}^2 = 0$ or $a_{ji}^1 = a_{ji}^2 = 0$. That is, for all $p \neq q$, for all $i \neq j$, $a_{ij}^p = a_{ij}^q = 0$ or $a_{ji}^p = a_{ji}^q = 0$ is satisfied.

(ii) For the situation that $f_{ij}(t)f_{ji}(t) < 0$, with no loss of generality, supposing $\alpha_1(t) = \alpha_2(t) \equiv 0.5$ gives

$$(a_{ij}^1 + a_{ij}^2)(a_{ji}^1 + a_{ji}^2) < 0, \quad \forall i \neq j. \quad (7)$$

That is,

$$a_{ij}^1 a_{ji}^1 + a_{ij}^2 a_{ji}^1 + a_{ij}^1 a_{ji}^2 + a_{ij}^2 a_{ji}^2 < 0, \quad \forall i \neq j. \quad (8)$$

Considering that all the elements permit arbitrary sign-preserving variations, let

$$|a_{ij}^1| = |a_{ji}^2|, \quad |a_{ij}^2| = |a_{ji}^1| \quad (9)$$

which leads to

$$\begin{aligned} a_{ij}^1 a_{ji}^1 &= \operatorname{sgn}(a_{ji}^1 a_{ij}^2) a_{ij}^1 a_{ij}^2, \\ a_{ij}^2 a_{ji}^1 &= \operatorname{sgn}(a_{ji}^1 a_{ij}^2) (a_{ij}^2)^2, \\ a_{ij}^1 a_{ji}^2 &= \operatorname{sgn}(a_{ji}^1 a_{ij}^2) (a_{ij}^1)^2, \\ a_{ij}^2 a_{ji}^2 &= \operatorname{sgn}(a_{ji}^1 a_{ij}^2) a_{ij}^1 a_{ij}^2, \end{aligned} \quad (10)$$

with $\operatorname{sgn}(a)$ denoting the sign of a .

From condition (3) of Lemma 2, it is known that a_{ij}^1 and a_{ji}^1 have the opposite signs (including 0), so do a_{ij}^2 and a_{ji}^2 . It implies that $a_{ij}^1 a_{ji}^2$ and $a_{ji}^1 a_{ij}^2$ have the same sign. That is, $\operatorname{sgn}(a_{ij}^1 a_{ji}^2) = \operatorname{sgn}(a_{ji}^1 a_{ij}^2)$. Then (8) can be rewritten as

$$\operatorname{sgn}(a_{ij}^1 a_{ji}^2) \left(a_{ij}^1 a_{ij}^2 + (a_{ij}^2)^2 + (a_{ij}^1)^2 + a_{ij}^1 a_{ij}^2 \right) < 0, \quad \forall i \neq j. \quad (11)$$

That is,

$$\operatorname{sgn}(a_{ij}^1 a_{ji}^2) (a_{ij}^1 + a_{ij}^2)^2 < 0, \quad \forall i \neq j. \quad (12)$$

To ensure (12), a_{ij}^1 and a_{ji}^2 must have the opposite signs. Therefore, a_{ij}^1 and a_{ij}^2 have the same sign (including 0). That is, for all $p \neq q$, for all $i \neq j$, $a_{ij}^p a_{ij}^q \geq 0$.

Synthesizing situations (i) and (ii), condition (2) is proved. \square

It can be concluded by Theorem 7 that, for all the subsystem matrices of a sign stable switched system, the elements in the same position must satisfy specific conditions. According to this, the definition of an isogenous sign-pattern set and correlative corollary are given as below.

Definition 8. A sign-pattern set is called an isogenous set if the elements in the same position of all the sign-patterns satisfy one of the following terms: (i) they have the same sign (including 0); (ii) they are all 0; (iii) the elements in the opposite position are all 0.

Corollary 9. The sign-pattern set of a sign stable switched system is an isogenous set.

Remark 10. We also say that two sign-patterns are isogenous with each other if they conform to the condition of Definition 8. Apparently, if arbitrary two sign-patterns of a known set are isogenous, this set is an isogenous set. Corollary 9 can be acquired from Theorem 7 and Definition 8 easily. However, it should be noticed that the isogenous set is only a necessary condition for sign stable switched system. For example, an isogenous set by 3×3 is in (13). However, the set is not sign stable as the first sign-pattern is not sign stable obviously:

$$\left\{ \begin{bmatrix} - & + & + \\ - & 0 & - \\ 0 & + & - \end{bmatrix}, \begin{bmatrix} - & 0 & + \\ - & - & - \\ 0 & 0 & - \end{bmatrix}, \begin{bmatrix} 0 & + & - \\ - & - & 0 \\ 0 & + & - \end{bmatrix} \right\}. \quad (13)$$

Considering an arbitrary isogenous sign-pattern set, we define the joining operation of all the sign-patterns to develop a special sign-pattern named the original sign-pattern of the set. The joining operation should keep the following rules.

- (1) If elements in a certain position of all the sign-patterns have the same nonzero sign or 0, then the corresponding element of the original sign-pattern is the same nonzero sign.
- (2) If elements in a certain position of all the sign-patterns are all 0, then the corresponding element of the original sign-pattern is 0.
- (3) If elements in a certain position of all the sign-patterns have different signs, then the corresponding element of the original sign-pattern is written as * to denote arbitrary signs.

The original sign-pattern obtained from a known isogenous set may be a new sign-pattern or an existent one which is already in the set. The original sign-pattern of set (13) is given in as follows as an illustration:

$$\begin{bmatrix} - & + & * \\ - & - & - \\ 0 & + & - \end{bmatrix}. \quad (14)$$

Furthermore, it is obvious that the new original sign-pattern and the known isogenous set can be joined to build a new set which is still isogenous. Actually, new sign-patterns can be derived to extend the isogenous set by replacing some nonzero elements of the original sign-pattern by 0. To demonstrate the aforementioned operation, the set which is developed from (13) is given below with the first matrix as the original sign-pattern and the last two as accessorial ones:

$$\left\{ \begin{bmatrix} - & + & * \\ - & - & - \\ 0 & + & - \end{bmatrix}, \begin{bmatrix} - & + & + \\ - & 0 & - \\ 0 & + & - \end{bmatrix}, \begin{bmatrix} - & 0 & + \\ - & - & - \\ 0 & 0 & - \end{bmatrix}, \right. \\ \left. \begin{bmatrix} 0 & + & - \\ - & - & 0 \\ 0 & + & - \end{bmatrix}, \begin{bmatrix} - & + & - \\ - & - & - \\ 0 & + & 0 \end{bmatrix}, \begin{bmatrix} - & + & 0 \\ 0 & - & - \\ 0 & + & - \end{bmatrix} \right\}. \quad (15)$$

Of course, we can continue to add new isogenous sign-patterns of (15) on the basis of the original sign-pattern and list all the other ones. However, among all these sign-patterns,

the sign stable ones are what we focused on and can be selected out. That leads to another important definition in this paper.

Definition 11. A sign-pattern set is named a complete isogenous sign stable set (CISSS) if it contains a sign stable original sign-pattern and all of the derivative isogenous sign-patterns that are sign stable.

Remark 12. Definition 11 indicates that we can obtain a CISSS from an isogenous set, for example, (15), if the original sign-pattern and some other ones are sign stable. Unfortunately, as the original sign-pattern of (15) is not sign stable, the subset of (15) cannot be a CISSS. It can be seen that a sign stable original sign-pattern is necessary for a CISSS. This notion can be interpreted by another example. Set (16) is an isogenous set, but not a CISSS even though all the sign-patterns are sign stable. Actually, the original sign-pattern given by (17) is not included in (16), and it is more important that the original sign-pattern is not sign stable:

$$\left\{ \begin{bmatrix} - & 0 \\ - & - \end{bmatrix}, \begin{bmatrix} - & - \\ 0 & - \end{bmatrix} \right\}, \tag{16}$$

$$\begin{bmatrix} - & - \\ - & - \end{bmatrix}. \tag{17}$$

Remark 13. It also should be emphasized that the selection of the uncertain element * does not affect the sign stability, which is related to the products of the off-diagonal pairs. Since the sign * must appear in the form of an off-diagonal pair “0 and *,” leading the product to 0 at all time, the above pair can still exist in CISSS.

3.2. Necessary and Sufficient Condition. In what follows, the main result of this work is proposed.

Theorem 14. *The necessary and sufficient condition for sign stability of switched linear system (2) under arbitrary switching is that the sign-pattern set of (2) is a subset of a CISSS.*

Proof. (1) The necessity can be easily proved by Corollary 9 and Definition 11.

(2) If the sign-pattern set $\text{sgn}\{A_\sigma\}$ of (2) is a subset of a CISSS, then all the sign-patterns of $\text{sgn}\{A_\sigma\}$ are isogenous and sign stable. For another arbitrary matrix set $\{\bar{A}_i\}$ that has the same sign-pattern set $\text{sgn}\{A_\sigma\}$ and $\{\alpha_i \mid \alpha_i \geq 0, \sum_{i=1}^N \alpha_i = 1\}$, let

$$A^* = \sum_{i=1}^N \alpha_i \bar{A}_i. \tag{18}$$

According to Definition 11, it is certain that $\text{sgn}(A^*)$ belongs to the foregoing CISSS. Hence, A^* is sign stable and thus Hurwitz stable. By Lemma 6, the switched system $\dot{x}(t) = \bar{A}_{\sigma(t)}x(t)$ is asymptotically stable. At last, by Definition 3, the sufficiency is proved. \square

With Theorem 14, it is easy to judge whether a given switched system is sign stable. To construct a sign stable

switched system, the original sign-pattern and the CISSS are required first. Furthermore, a sign stable original sign-pattern with the least element “0” is the basis to constitute a CISSS.

3.3. Simple Examples. Two simple examples are given to illuminate the constitution of CISSS.

Example 1. For 2×2 matrices, the sets Δ_1, Δ_2 given by (19) are CISSS with the first sign-pattern as the original sign-pattern of each set, respectively:

$$\begin{aligned} \Delta_1 &= \left\{ \begin{bmatrix} - & + \\ - & - \end{bmatrix}, \begin{bmatrix} 0 & + \\ - & - \end{bmatrix}, \begin{bmatrix} - & + \\ - & 0 \end{bmatrix}, \begin{bmatrix} - & + \\ 0 & - \end{bmatrix}, \begin{bmatrix} - & 0 \\ - & - \end{bmatrix}, \begin{bmatrix} - & 0 \\ 0 & - \end{bmatrix} \right\}, \\ \Delta_2 &= \left\{ \begin{bmatrix} - & - \\ + & - \end{bmatrix}, \begin{bmatrix} 0 & - \\ + & - \end{bmatrix}, \begin{bmatrix} - & - \\ + & 0 \end{bmatrix}, \begin{bmatrix} - & - \\ 0 & - \end{bmatrix}, \begin{bmatrix} - & 0 \\ + & - \end{bmatrix}, \begin{bmatrix} - & 0 \\ 0 & - \end{bmatrix} \right\}. \end{aligned} \tag{19}$$

It has been summarized that the original sign-pattern is needed to have as few “0” elements as possible. Hence, all the diagonal elements are chosen to be “-,” and the off-diagonal elements are selected to possess as many pairs of opposite signs as possible. For 2×2 matrices, it is not necessary for us to configure any “0” element in the original sign-pattern as shown in Example 1, whereas, for 3×3 matrix $A_{3 \times 3} = (a_{ij})_{3 \times 3}$, according to condition (4) of Lemma 2, there must be at least two “0” elements in the off-diagonal positions to guarantee $a_{12}a_{23}a_{31} = 0$ and $a_{13}a_{32}a_{21} = 0$. It then can be concluded that, in the 3×3 matrix below, there must be at least one 0 in the positions represented by & and one 0 in the positions represented by #, respectively:

$$\begin{bmatrix} - & \& \# \\ \# & - & \& \\ \& \# & - \end{bmatrix}. \tag{20}$$

Example 2. Consider the 3×3 sign stable original sign-pattern given in (21), where each * presents an uncertain sign that may be +, -, or 0. The aim is to discover the corresponding CISSS:

$$\begin{bmatrix} - & 0 & 0 \\ * & - & + \\ * & - & - \end{bmatrix}. \tag{21}$$

Utilizing the criterion in [5] to select sign stable sign-patterns in the isogenous set of (21), the CISSS is acquired and shown in (22), with each * denoting an arbitrary sign. As each sign-pattern in (22) contains 9 (= 3×3) sign-patterns, there is a total of 54 (= 6×9) sign-patterns in the CISSS:

$$\begin{aligned} &\left\{ \begin{bmatrix} - & 0 & 0 \\ * & - & + \\ * & - & - \end{bmatrix}, \begin{bmatrix} - & 0 & 0 \\ * & 0 & + \\ * & - & - \end{bmatrix}, \begin{bmatrix} - & 0 & 0 \\ * & - & + \\ * & - & 0 \end{bmatrix}, \right. \\ &\left. \begin{bmatrix} - & 0 & 0 \\ * & - & 0 \\ * & - & - \end{bmatrix}, \begin{bmatrix} - & 0 & 0 \\ * & - & + \\ * & 0 & - \end{bmatrix}, \begin{bmatrix} - & 0 & 0 \\ * & - & 0 \\ * & 0 & - \end{bmatrix} \right\}. \end{aligned} \tag{22}$$

In fact, in the above process of finding CISSS, the two pairs of “0 and *” do not need to be changed. The first term

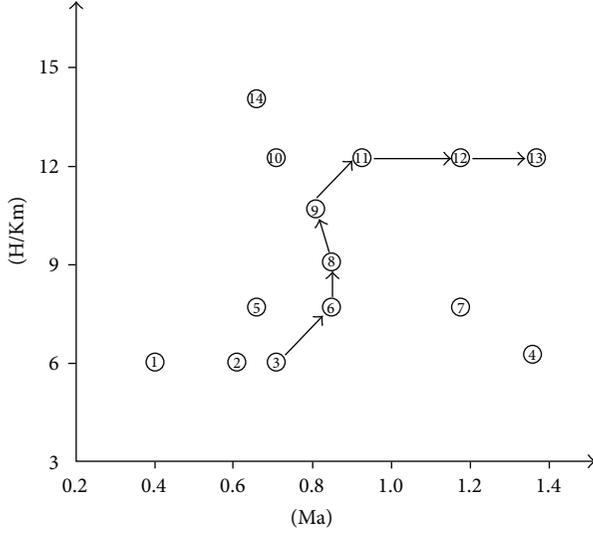


FIGURE 1: Partial flight envelope and operating points.

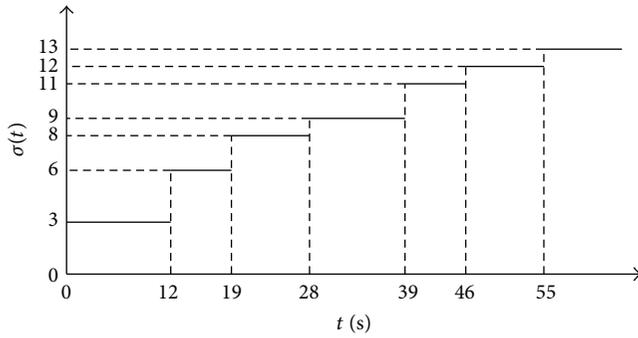


FIGURE 2: Switching law of the flight trajectory.

of the diagonal is also unchangeable. As a result, the only way is to transform the 2×2 block on the lower-right. Indeed, it can be seen that the lower-right blocks are the same as the set Δ_1 of Example 1.

4. Example

The controller design problem for the highly maneuverable technology (HiMAT) vehicle is presented to demonstrate the main results in this paper. As declared in [10], a switched system can be applied to describe flight conditions with fast varying parameters. Controllers for the subsystems can be synthesized via a switching law to adapt the parameter variations among different operating points. In this section, the sign stability approach is utilized to design a switching control scheme for the unstable longitudinal dynamics. The considered flight envelope and 14 operating points within it are depicted in Figure 1. We suppose that the dynamics in the contiguous region of each operating point can be approximated by the corresponding linear subsystem. These linear models are given in [25].

The linearization model for each operating point is regarded as a subsystem of the switched system. The switched system is then given by

$$\dot{x}(t) = A_\sigma x(t) + B_\sigma u(t), \quad \sigma \in \{1, 2, \dots, 14\}, \quad (23)$$

where $x(t) \in R^3$ is the state vector consisting of three state variables α (angle of attack), q (pitch rate), and ϕ (pitch angle), and $u(t) \in R^4$ is the control vector consisting of δ_e (elevator), δ_v (elevon), δ_c (canard), and δ_s (symmetric aileron).

The system matrices and control matrices for several operating points are shown to illustrate the sign-patterns:

$$A_1 = \begin{bmatrix} -0.8435 & 0.97505 & -0.0048 \\ 8.7072 & -1.1643 & 0.0026 \\ 0 & 1 & 0 \end{bmatrix},$$

$$B_1 = \begin{bmatrix} -0.1299 & -0.092 & -0.0107 & -0.0827 \\ -7.6833 & -4.7974 & 4.8178 & -5.7416 \\ 0 & 0 & 0 & 0 \end{bmatrix},$$

$$A_6 = \begin{bmatrix} -1.8997 & 0.98312 & -0.00073 \\ 11.720 & -2.6316 & 0.00088 \\ 0 & 1 & 0 \end{bmatrix}, \quad (24)$$

$$B_6 = \begin{bmatrix} -0.2436 & -0.1708 & -0.00497 & -0.1997 \\ -46.206 & -31.604 & 22.396 & -31.179 \\ 0 & 0 & 0 & 0 \end{bmatrix},$$

$$A_{13} = \begin{bmatrix} -1.2206 & 0.99411 & -0.00084 \\ -64.071 & -1.8876 & 0.00046 \\ 0 & 1 & 0 \end{bmatrix},$$

$$B_{13} = \begin{bmatrix} -0.0662 & -0.0315 & -0.0141 & -0.0749 \\ -27.333 & -13.163 & 11.058 & -26.878 \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

Furthermore, the sign-pattern set of the open-loop system matrices is given by the following:

$$\left\{ \left[\begin{array}{ccc} - & + & - \\ + & - & + \\ 0 & + & 0 \end{array} \right], \left[\begin{array}{ccc} - & + & - \\ - & - & + \\ 0 & + & 0 \end{array} \right] \right\}. \quad (25)$$

As explained in Section 3, a subset of a CISSS is required for the closed-loop switched system matrices. According to the structure of $\{B_\sigma\}$, the original sign-pattern is chosen as follows:

$$\left[\begin{array}{ccc} - & * & * \\ 0 & - & - \\ 0 & + & 0 \end{array} \right]. \quad (26)$$

Although the complete form of the CISSS derived from (26) has more sign-patterns, we can regard (26) as the target sign-pattern of closed-loop subsystems directly. Compared to the sign-patterns in (25), it is only needed to change two signs in each subsystem matrix. A state feedback control gain

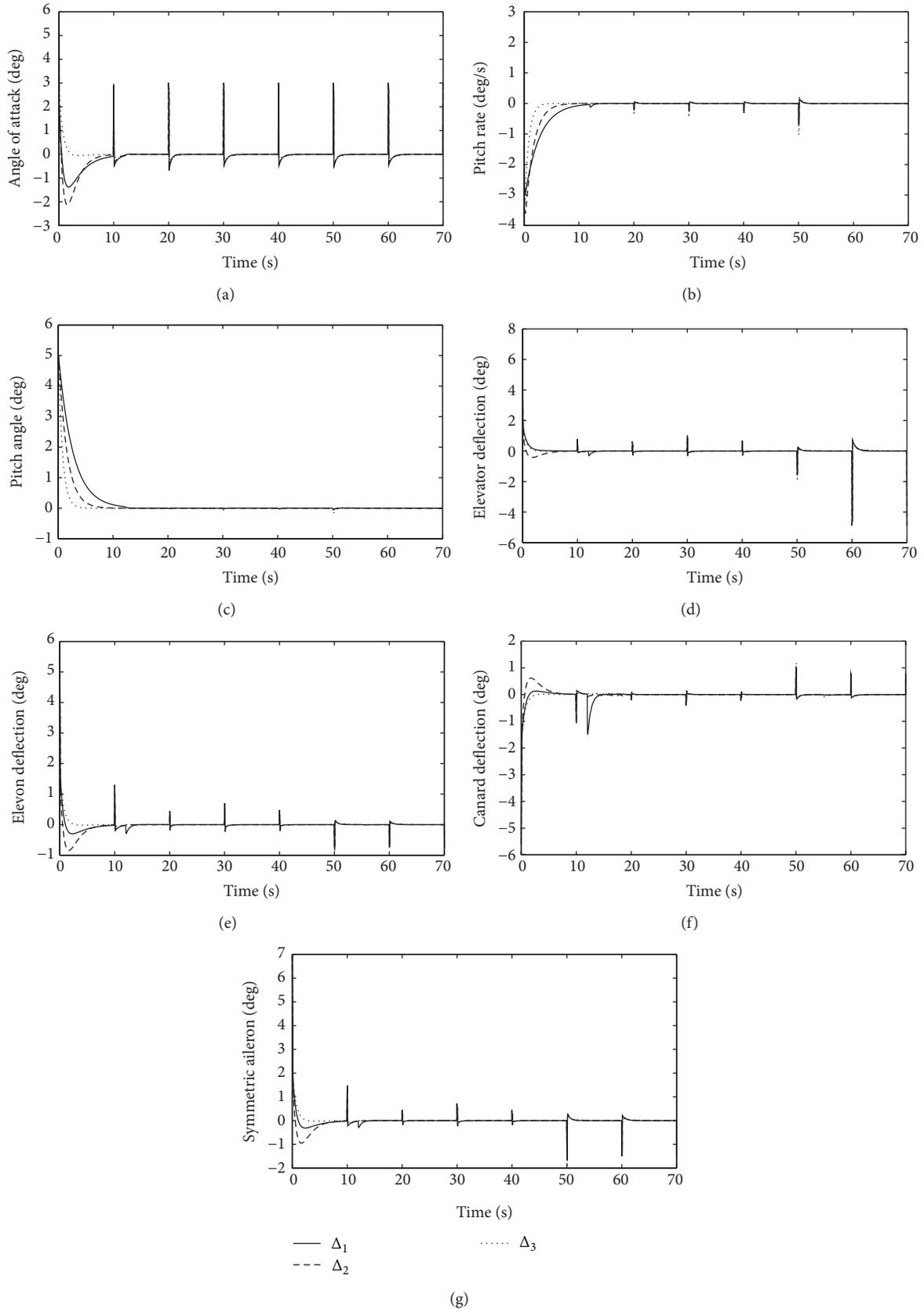


FIGURE 3: Time histories of system states and control inputs under perturbations: (a) angle of attack, (b) pitch rate, (c) pitch angle, (d) elevator, (e) elevon, (f) canard, and (g) symmetric aileron.

is designed for each operating point in the envelope, and the results for several points are given as follows:

$$\begin{aligned}
 A_{1c} &= \begin{bmatrix} -0.9555 & 0.8066 & -0.1616 \\ 0 & -13.6041 & -11.7700 \\ 0 & 1 & 0 \end{bmatrix}, \\
 K_1 &= \begin{bmatrix} 0.2618 & 0.6935 & 0.6468 \\ 0.4362 & 0.4339 & 0.4005 \\ -0.3589 & -0.4278 & -0.4323 \\ 0.5007 & 0.5171 & 0.4875 \end{bmatrix}, \\
 A_{6c} &= \begin{bmatrix} -1.9582 & -0.6336 & -0.3349 \\ 0 & -18.4138 & -67.8792 \\ 0 & 1 & 0 \end{bmatrix}, \\
 K_6 &= \begin{bmatrix} -0.0031 & 2.6675 & 0.6795 \\ 0.1577 & 2.4566 & 0.4644 \\ -0.0794 & 11.6769 & -0.3393 \\ 0.1636 & 2.4505 & 0.4557 \end{bmatrix}, \\
 A_{13c} &= \begin{bmatrix} -1.0714 & 0.9009 & 0.4040 \\ 0 & -42.1910 & -28.9097 \\ 0 & 1 & 0 \end{bmatrix}, \\
 K_{13} &= \begin{bmatrix} -1.6277 & 0.6241 & -1.3513 \\ -0.2504 & 0.3005 & -1.6873 \\ 0.2634 & -0.2534 & -11.2900 \\ -0.4974 & 0.6134 & -1.3687 \end{bmatrix}.
 \end{aligned} \tag{27}$$

It can be seen that the closed-loop system matrices A_{1c} , A_{6c} , and A_{13c} have different sign-patterns. However, these sign-patterns (also including those of other operating points) belong to the CISSS derived from the original sign-pattern (26). Based on Theorem 14, the closed-loop switched system is sign stable under arbitrary switching.

The simulation across different regions of the flight envelope is taken to validate the proposed technique. The flight trajectory travels through vicinities of 7 operating points ($3 \rightarrow 6 \rightarrow 8 \rightarrow 9 \rightarrow 11 \rightarrow 12 \rightarrow 13$; see Figure 1). The switching law is depicted in Figure 2.

To validate the stability within the full flight envelope, gust disturbance is considered in the form of perturbations acted on the angle of attack, which is a group of pulse signals with a period of 10 seconds, a width of 1 second, and a magnitude of 3 degrees. The initial values of the angle of attack, pitch rate, and pitch angle are set to 3 degrees, 3 degrees per second, and 5 degrees, respectively. To illustrate the robustness of the proposed approach, multiplicative perturbations on the closed-loop subsystem matrices are introduced. Each element in the matrices is multiplied by an arbitrary positive value, which will preserve all the signs

of the products. The perturbation matrices are given by the following:

$$\begin{aligned}
 \Delta_1 &= \begin{bmatrix} 1.5 & 1.2 & 0.9 \\ 0.8 & 1.8 & 1.3 \\ 1.4 & 0.6 & 0.7 \end{bmatrix}, \\
 \Delta_2 &= \begin{bmatrix} 1.2 & 1.6 & 1.3 \\ 0.5 & 1.2 & 1.1 \\ 1.5 & 0.8 & 1.4 \end{bmatrix}, \quad \Delta_3 = \begin{bmatrix} 0.5 & 0.6 & 1.1 \\ 1.4 & 2.2 & 1.8 \\ 0.5 & 1.6 & 2.4 \end{bmatrix}.
 \end{aligned} \tag{28}$$

The resulted closed-loop subsystem matrices are obtained by multiplying each element of the original matrix with the corresponding element in the perturbation matrix. Thus, it leads to 3 closed-loop switched systems with different multiplicative perturbations. The time histories of the system states and control surface deflections are depicted in Figure 3.

As depicted in Figure 3, the system states converge quickly under gust disturbances and keep stable while the switching occurs along the flight trajectory. There are vibrations in the deflections of the control surfaces at the time of the first switching, nevertheless the amplitudes are acceptable. It is concluded that the devised approach is robust for the sign-preserving parameter perturbations on the closed-loop systems.

5. Conclusion

A new concept of sign stable switched linear system is established to develop novel techniques for flight control. The main result is provided in the form of necessary and sufficient condition for the sign stability of switched system under arbitrary switching. Hence, the application areas of sign stability approach are enlarged remarkably. A new robust stabilization approach for switched linear system is proposed via the notion of CISSS. The sign stabilization controllers are devised for each subsystem, respectively; therefore the design process possesses more freedoms. Compared with the traditional Lyapunov function method, the proposed technique has natural robustness and decreased conservativeness. The aforementioned points are verified by a stabilization control problem of aircraft with a large-scale flight envelope.

Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

Acknowledgment

This work is supported by the National Natural Science Foundation of China (61273083).

References

- [1] R. M. May, "Will a large complex system be stable?" *Nature*, vol. 238, no. 5364, pp. 413–414, 1972.
- [2] L. Edelstein-Keshet, *Mathematical Models in Biology*, McGraw-Hill, New York, NY, USA, 1988.

- [3] D. O. Logofet, "Stronger-than-Lyapunov notions of matrix stability, or how "flowers" help solve problems in mathematical ecology," *Linear Algebra and its Applications*, vol. 398, pp. 75–100, 2005.
- [4] D. A. Grundy, D. D. Olesky, and P. van den Driessche, "Constructions for potentially stable sign patterns," *Linear Algebra and its Applications*, vol. 436, no. 12, pp. 4473–4488, 2012.
- [5] C. Jeffries, V. Klee, and P. van den Driessche, "When is a matrix sign stable?" *Canadian Journal of Mathematics*, vol. 29, no. 2, pp. 315–326, 1977.
- [6] C. Jeffries, "Qualitative stability and digraphs in model ecosystems," *Ecology*, vol. 55, no. 6, pp. 1415–1419, 1974.
- [7] R. K. Yedavalli, "Robust control design for linear systems using an ecological sign-stability approach," *Journal of Guidance, Control, and Dynamics*, vol. 32, no. 1, pp. 348–352, 2009.
- [8] R. K. Yedavalli and N. Devarakonda, "Sign-stability concept of ecology for control design with aerospace applications," *Journal of Guidance, Control, and Dynamics*, vol. 33, no. 2, pp. 333–346, 2010.
- [9] Y. Hou, Q. Wang, and C. Dong, "Gain scheduled control: switched polytopic system approach," *Journal of Guidance, Control, and Dynamics*, vol. 34, no. 2, pp. 623–628, 2011.
- [10] Y. Hou, C. Dong, and Q. Wang, "Stability analysis of switched linear systems with locally overlapped switching law," *Journal of Guidance, Control, and Dynamics*, vol. 33, no. 2, pp. 396–403, 2010.
- [11] T. Shimomura, "Hybrid control of gain-scheduling and switching: a design example of aircraft control," in *Proceedings of the American Control Conference*, pp. 4639–4644, Denver, Colo, USA, June 2003.
- [12] B. Lu, F. Wu, and S. Kim, "Switching LPV control of an F-16 aircraft via controller state reset," *IEEE Transactions on Control Systems Technology*, vol. 14, no. 2, pp. 267–277, 2006.
- [13] K. S. Narendra and J. Balakrishnan, "A common Lyapunov function for stable LTI systems with commuting A -matrices," *Institute of Electrical and Electronics Engineers. Transactions on Automatic Control*, vol. 39, no. 12, pp. 2469–2471, 1994.
- [14] D. Liberzon, J. P. Hespanha, and A. S. Morse, "Stability of switched systems: a Lie-algebraic condition," *Systems & Control Letters*, vol. 37, no. 3, pp. 117–122, 1999.
- [15] M. S. Branicky, "Multiple Lyapunov functions and other analysis tools for switched and hybrid systems," *Institute of Electrical and Electronics Engineers. Transactions on Automatic Control*, vol. 43, no. 4, pp. 475–482, 1998.
- [16] A. S. Morse, "Supervisory control of families of linear set-point controllers. I. Exact matching," *Institute of Electrical and Electronics Engineers. Transactions on Automatic Control*, vol. 41, no. 10, pp. 1413–1431, 1996.
- [17] J. P. Hespanha and A. S. Morse, "Stability of switched systems with average dwell-time," in *Proceedings of the 38th IEEE Conference on Decision and Control (CDC)*, pp. 2655–2660, Phoenix, Ariz, USA, December 1999.
- [18] X. Zhao, L. Zhang, P. Shi, and M. Liu, "Stability and stabilization of switched linear systems with mode-dependent average dwell time," *Institute of Electrical and Electronics Engineers. Transactions on Automatic Control*, vol. 57, no. 7, pp. 1809–1815, 2012.
- [19] Q. Lu, L. Zhang, H. R. Karimi, and Y. Shi, " H_∞ control for asynchronously switched linear parameter-varying systems with mode-dependent average dwell time," *IET Control Theory & Applications*, vol. 7, no. 5, pp. 673–683, 2013.
- [20] O. Mason, R. N. Shorten, and D. Leith, "Evaluation of piecewise quadratic methods applied to the wind-turbine benchmark example," Tech. Rep. NUM-SS-2001-11, Signals and Systems Group, NUI Maynooth, 2001.
- [21] V. Klee and P. van den Driessche, "Linear algorithms for testing the sign stability of a matrix and for finding Z -maximum matchings in acyclic graphs," *Numerische Mathematik*, vol. 28, no. 3, pp. 273–285, 1977.
- [22] J. Quirk and R. Ruppert, "Qualitative economics and the stability of equilibrium," *The Review of Economic Studies*, vol. 32, no. 4, pp. 311–326, 1965.
- [23] R. May, *Stability and Complexity in Model Ecosystems*, Princeton University Press, Princeton, NJ, USA, 1973.
- [24] H. Lin and P. J. Antsaklis, "Stability and stabilizability of switched linear systems: a survey of recent results," *Institute of Electrical and Electronics Engineers. Transactions on Automatic Control*, vol. 54, no. 2, pp. 308–322, 2009.
- [25] G. L. Hartmann, M. F. Barrett, and C. S. Greene, "Control design for an unstable vehicle," Tech. Rep. NASA CR-170393, 1979.

Research Article

Reliability Modeling and Evaluation of Electric Vehicle Motor by Using Fault Tree and Extended Stochastic Petri Nets

Bing Wang,^{1,2} Guangdong Tian,² Yanping Liang,¹ and Tiangang Qiang²

¹ School of Electrical and Electronic Engineering, Harbin University of Science and Technology, Harbin, Heilongjiang 150080, China

² Transportation College, Northeast Forestry University, Harbin, Heilongjiang 150040, China

Correspondence should be addressed to Guangdong Tian; tiangd2013@gmail.com

Received 20 March 2014; Accepted 6 April 2014; Published 30 April 2014

Academic Editor: Weichao Sun

Copyright © 2014 Bing Wang et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Performing reliability analysis of electric vehicle motor has an important impact on its safety. To do so, this paper proposes its reliability modeling and evaluation issues of electric vehicle motor by using fault tree (FT) and extended stochastic Petri nets (ESPN). Based on the concepts of FT and ESPN, an FT based ESPN model for reliability analysis is obtained. In addition, the reliability calculation method is introduced and this work designs a hybrid intelligent algorithm integrating stochastic simulation and NN, namely, NN based simulation algorithm, to solve it. Finally, taking an electric vehicle motor as an example, its reliability modeling and evaluation issues are analyzed. The results illustrate the proposed models and the effectiveness of proposed algorithms. Moreover, the results reported in this work could be useful for the designers of electric vehicle motor, particularly, in the process of redesigning the electric vehicle motor and scheduling its reliability growth plan.

1. Introduction

Along with increasing problems of energy and environment, more and more countries develop related policies to handle these changing and disturbing issues, that is, developing new-energy and low-carbon vehicles, implementing the remanufacturing, reuse, and recycling of waste products, and performing green transportation technologies [1–5]. Electric vehicle, as an important and green transportation tool, has been widely attracted by more and more researchers [6]. Motor is one of the key components of electric vehicle. Its reliability has an important impact on the system safety. Designers have well-recognized the importance of electric vehicle reliability, but to our best knowledge a detailed reliability analysis is still missing. Although the faults have been reduced in the last few years by some measures, the faults still affects the safety of vehicles, and faults of mechanical system occupy a large proportion of all the faults.

In the present literature, most of the current researches have discussed electric and electronic system issue and reliability prediction analysis of the electric vehicle. For

example, P. Liu and H. P. Liu present a permanent-magnet synchronous motor drive system of electric vehicles [7]. Peng et al. discuss driving and control problems of torque for direct-wheel-driven electric vehicle with motors in serial [8]. Quinn et al. present the effect of communication architecture on reliability [9]. Zhu et al. present a grey prediction model of motor reliability of electric vehicle [10]. They propose a grey prediction model of electric vehicle motor based on particle swarm optimization [11]. In addition, Zhu et al. discuss the reliability modeling method of solar array based on the fault tree (FT) analysis method [11].

It can be seen from the above literatures that the current research on reliability modeling of electric vehicle motor is limited to FT analysis method. There is no doubt that FT analysis has been widely employed as a powerful technique to evaluate the safety and reliability of complex systems by many scholars [12, 13]. However, FT analysis has some limitations in reliability analysis. Firstly, in FT analysis, the probabilities of basic events must be known before analysis. Thus based on this assumption, the reliability analysis of the system is only a probability decision-making process

and cannot achieve the real-time description of reliability information [14, 15]. Secondly, it is not easy for FT analysis to conduct further quantitative analysis automatically due to the lack of effective means of mathematical expression. Thirdly, FT analysis cannot find the dynamic description of fault information of the system precisely and cannot describe the propagation process of fault information.

The Petri net is one of the mathematical modeling approaches for the description of distributed systems, which consists of places, transitions, and directed arcs [16, 17]. Many extensions to the Petri nets have been successfully developed and applied in analyzing fault diagnosis, automated manufacturing systems, and product disassembly [18–21]. The extended stochastic Petri net is a high level one; it has been used to establish models of reconfigurable manufacturing system and network attack due to its better information expression ability and dynamic description performance of process features [22, 23]. Although some prior works [22–24] have proposed to use extended stochastic Petri nets to solve reliability issues of nitric acid reactor feed and reconfigurable manufacturing systems, they merely analyze the average failure rate/life time of system; the real-time probability analysis/reliability issues of the system are not yet addressed by using this method. Moreover, for reliability issue of electric vehicle motor, we cannot find reference to handle this issue by the extended stochastic Petri nets method to the best knowledge of the authors. To do so, this work addresses reliability modeling and evaluation of electric vehicle motor by using extended stochastic Petri nets based on fault tree for the first time. Namely, the aim of this work is to find a new way to analyze reliability of electric vehicle motor.

The remainder of this paper is organized as follows: reliability model and establishment method of FT and extended stochastic Petri nets of the mechanical system are given in Section 2. Section 3 presents the reliability analysis method and algorithm. In Section 4, taking an electric vehicle motor as an example, its reliability modeling and evaluation are presented. Section 5 concludes our work and describes some future research issues.

2. FT Based Extended Stochastic Petri Nets (ESPN) Models for Reliability Analysis

Reliability model is basis and premise for reliability analysis and evaluation; thus we first introduce the concept and establishment process of extended stochastic Petri nets model based on FT for reliability analysis. To easily establish it, the following method is proposed in this work. Namely, the FT model for reliability analysis is established based on its related concept, and then transformation rules of elements of FA to ESPN are defined. Finally, FT based ESPN model for reliability analysis is established.

2.1. FT Model for Reliability Analysis. FT is the most usual model of reliability analysis. Many references describe it in detail [25, 26]. In this work, we only present its basic elements and schematic diagram.

2.1.1. Basic Elements of FT. Usually, FT is composed of a series of events and logic gates. The main events include the following.

Top event: it is the most undesirable system failure event and the object of the analysis. It is denoted by \square .

Middle event: it is the subsystem or component failure event and the cause of the top event. It is denoted by \square .

Basic event: it is the primary failure event and the cause of the top event or middle events. It is denoted by \circ .

The main logic gates include the following.

Logic OR gate: it indicates that output event occurs if either one of the input events occurs. It is denoted by \cup .

Logic AND gate: it indicates that output event occurs only if all of the input events occur. It is denoted by \cap .

There may be many other types of events and logic gates involved in complex system reliability analysis. However, for the sake of concisions, we only list the most commonly used ones here. For other types of events and logic gates, please refer to [25, 26].

2.1.2. FT Model of a Mechanical System. Based on the presented basic elements and logical relationship of fault occurrence, the schematic diagram of FT for reliability analysis of a mechanical system is presented next, as shown in Figure 1.

As shown in Figure 1, this FT is composed of 1 top event, 2 middle events, and 4 basic events. They are T, M1 and M2, and B1, B2, B3, and B4.

2.2. ESPN Model for Reliability Analysis

2.2.1. Concept of ESPN. The Petri net is a graphic modeling method, which is widely used in modeling and analyzing discrete event systems such as semiconductor manufacturing, transportation, and automated manufacturing systems. An ESPN is a high level one. It is a type of improved stochastic Petri nets with arbitrary distribution. Before giving the formal definition, we present the definition of PN introduced by Petri in 1962 [27–30].

A PN is a five-tuple (P, T, I, O, m) , where

$P = \{p_1, p_2, \dots, p_n\}$, $n > 0$, and is a finite set of places pictured by circles;

$T = \{t_1, t_2, \dots, t_s\}$, $s > 0$, and is a finite set of transitions pictured by bars, with $P \cup T \neq \Phi$, $P \cap T = \Phi$;

$I: P \times T \rightarrow N$ and is an input function that defines the set of directed arcs from P to T , where $N = \{0, 1, 2, \dots\}$;

$O: T \times P \rightarrow N$ and is an output function that defines the set of directed arcs from T to P ;

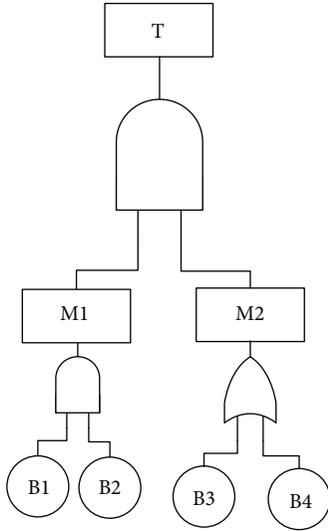


FIGURE 1: FT model for reliability analysis of a mechanical system.

$m_i : P \rightarrow N$ and is a marking whose i th component represents the number of tokens in the i th place. An initial marking is denoted by m_0 . The tokens are pictured by dots.

A simple PN and its elements are shown in Figure 2. The four-tuple (P, T, I, O) is called a PN structure that defines a directed graph structure. A PN models system dynamics using tokens and their firing rules.

Assume that every transition in a PN is associated with an exponentially distributed random delay from the enabling to the firing of the transition; then this PN is transformed into the stochastic Petri net (SPN), while every transition in a PN is associated with arbitrary distribution random delay; this PN is called ESPN. It is defined as follows [31].

An ESPN is a six-tuple (P, T, I, O, m, H) , where

$P = \{p_1, p_2, \dots, p_n\}$, $n > 0$, and is a finite set of places;

$T = \{t_1, t_2, \dots, t_s\}$, $s > 0$, and is a finite set of transitions, with $P \cup T = \Phi$, $P \cap T = \Phi$;

$I : P \times T \rightarrow N$ and is an input function that defines the set of directed arcs from P to T , where $N = \{0, 1, 2, \dots\}$;

$O : T \times P \rightarrow N$ and is an output function that defines the set of directed arcs from T to P ;

$m_i : P \rightarrow N$ and is a marking whose i th component represents the number of tokens in the i th place. An initial marking is denoted by m_0 .

$H : T \rightarrow R$ is a vector whose component is a firing time delay with an extended/arbitrary distribution function.

Note that the firing time delay is the life time of corresponding component in ESPN model for reliability analysis.

2.2.2. *Elements Transformation Rules of FT to ESPN.* To obtain the ESPN, we introduce the transformation rules of

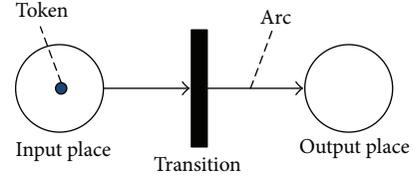


FIGURE 2: A simple PN and its elements.

elements of FT to ones of ESPN. Based on different logical relations, the transformation of AND/OR gates of FT to AND/OR transitions of ESPN is shown in Figure 3.

As shown in Figure 3, the top/middle/basic event, logical gate, and logical relation line of FT are transformed to place, transition, and arc of ESPN, respectively.

2.2.3. *The Establishment of ESPN Based on FT.* Based on the FT and elements transformation rules of FT to ESPN, the FT based ESPN model for reliability analysis is obtained, as shown in Figure 4.

From Figures 4 and 1, compared to FT model, we can see that the building FT model needs 6 types of elements, that is, 3 types of events, 2 types of logic gates, and a kind of relation line, while the building ESPN model only needs 3 types of elements, that is, place, transition, and arc. Namely, the ESPN model is more concise than FT one. In addition, when each transition is associated with its corresponding life distribution function, it can achieve the real-time description of reliability analysis. Also, this model can achieve the dynamic delivery and propagation of reliability/fault information due to the introduction of transition and directed arc. Overall, the results denote that using ESPN method to establish the product reliability model is more convenient, concise, and effective than FT model.

3. Method and Algorithm for Reliability Evaluation

3.1. Reliability Evaluation Method

3.1.1. *Reliability Evaluation Parameters and Calculation Method.* In this paper, the following two evaluation parameters of a system are adapted, that is, reliability degree of $R(x)$ and average life $E(x)$.

Let random variable X denote the life for a specified system; then the unreliability $F(x)$ of the system at time x is the probability of random variable $X \leq x$; namely,

$$F(x) = \Pr(X \leq x). \quad (1)$$

The reliability of the system $R(x)$ is

$$R(x) = 1 - F(x) = \Pr(X \geq x). \quad (2)$$

Let probability density function of the system be $f_s(x)$; then the average life $E(x)$ is

$$E(x) = \int_{-\infty}^{+\infty} x f_s(x) du. \quad (3)$$

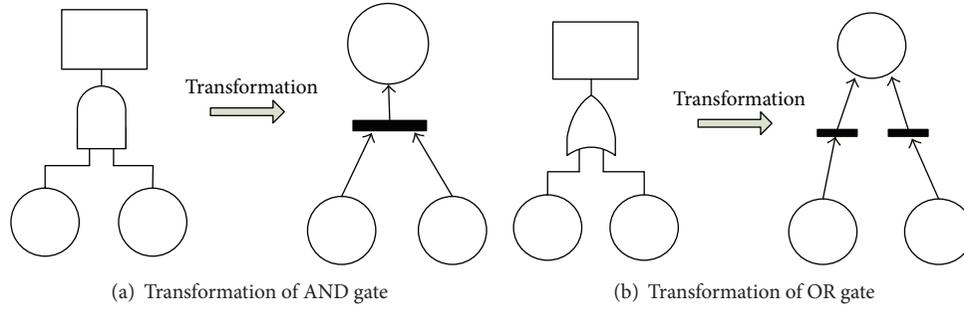


FIGURE 3: Transformation of AND/OR gate.

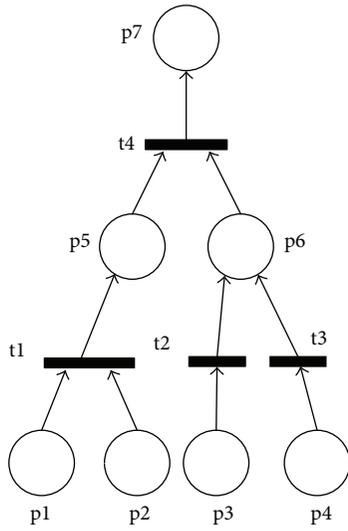


FIGURE 4: ESPN model for reliability analysis of a mechanical system.

In addition, for computer system reliability, the life calculation method of AND/OR transition in ESPN is pretended next.

Let a specified system consist of n components and let the life of these n components be x_1, x_2, \dots , and x_n , respectively.

For AND transitions, the system life x is expressed as

$$x = \min \{x_1, x_2, \dots, x_n\}. \quad (4)$$

For OR transitions, the system life x is expressed as

$$x = \max \{x_1, x_2, \dots, x_n\}. \quad (5)$$

For example, in Figure 4, let the life of corresponding components of places p_1 and p_2 be x_1 and x_2 , respectively; then the life of corresponding system x_5 of places p_5 is $x_5 = \min(x_1, x_2)$. Let the life of corresponding components of places p_3 and p_4 be x_3 and x_4 , respectively; then the life of corresponding system x_6 of places p_6 is $x_6 = \max(x_3, x_4)$.

For other systems, its life calculation can be obtained by the integration of AND and OR transition calculation. Such as, a system as shown in Figure 4, its system life $x_7 = \min\{x_5, x_6\} = \min\{\min(x_1, x_2), \max(x_3, x_4)\}$.

3.2. Algorithm for Reliability Evaluation. Stochastic simulation is an effective means to assess and calculate stochastic and probabilistic functions. It has effectively solved many stochastic programming problems [32–35]. While neural networks (NN) have been successfully used to solve many complex industrial evaluation and optimization problems due to their strong nonlinear fitting ability [36–38], we propose to use a stochastic simulation algorithm based on NN to solve the reliability of the proposed FT based ESPN model.

3.2.1. Stochastic Simulation of Reliability Function

Step 1. Initialize associated fault probability (life) distribution functions of each transition in ESPN model, and set the number of simulation cycles M .

Step 2. Generate the life time x_1, x_2, \dots , and x_k from their life distribution functions $f_{r1}(x), f_{r2}(x), \dots$, and $f_{rk}(x)$.

Step 3. Based on transition transmission rules, that is, (4) and (5), a system life value x is obtained from bottom to top in ESPN model.

Step 4. Repeat Steps 1–3 for M times; namely, M samples of system life are obtained.

Step 5. Calculation of the average value of M samples of system life obtained in Step 4; that is, average life $E(x)$ of the system is obtained.

Step 6. Given a time x' , record the number of $x > x'$ in M samples as M' ; then the system reality degree $R(x)$ is obtained; that is, $R(x) = M'/M$.

3.2.2. Neural Networks (NN). NN is treated as a nonlinear mapping system consisting of neurons (processing units), which are linked by weighted connections. It usually consists of three layers: input, hidden, and output layers. There is an activation function in the hidden layer. It is defined as the sigmoid function in this paper [36–39].

Firstly, the method to determine the number of neurons of the input, hidden, and output layers is presented as follows.

The number of input neurons of an NN structure is the number of bottom places in ESPN model, namely, the number of basic events in FT model; in other words, the number of input neurons is the number of life distributions of basic fault component of a system in this paper.

The number of output neurons is 1 representing one system life function. In terms of the NN structure, the main problem is to determine the best number of hidden neurons. The number can be infinite in theory, but finite in practice due to two reasons. Too many hidden neurons increase the training time and response time of the trained NN. On the other hand, too few hidden neurons make the NN lack of generalization ability. Therefore, it can usually be determined by the following formula; namely, $s = \sqrt{u + v} + b$, where u and v are the number of input neurons and output neurons, respectively, and b is a constant from 1 to 10 [40]. Based on it, in terms of a system shown in Figure 4, u is set to be 4 since this system is composed of 4 bottom places, $v = 1$; thus s is a constant from 3 to 12.

Secondly, backpropagation is the most commonly used method to calculate values for the weight and bias terms of an NN model. In this method, all weights are adjusted according to the calculated error term using a gradient method. Learning in an NN, that is, the calculation of the weights of the connections, is achieved by minimizing the error between its output and the actual output over a number of available training data points. In this paper, the error term is controlled by the following MATLAB function, namely, *net.trainParam.goal*. It denotes the mean squared error between the output of the neural network and the actual output over a number of available training data points.

Thus the NN algorithm is presented as follows.

Step 1. Initialize the number of neurons at the input, hidden, and output layers, and initialize weight vector w .

Step 2. Calculate the output of the hidden layer and the output of output layer, and adjust the corresponding weights w .

Step 3. Calculate the error term, namely, training performance goal. If it is larger than the given error, go to Step 2, and otherwise, end.

3.2.3. NN Based Simulation Algorithm. Based on the presented stochastic simulation and NN, the steps of the stochastic evaluation algorithm based on NN are presented as follows.

Step 1. Initialize the parameters of an NN structure and the number of training data points K .

Step 2. Establish an FT based ESPN model for reliability analysis based on fault logic relationships of each component in a system.

Step 3. Based on the relationship among the system life and the corresponding each component life in FT based ESPN model, generate the input-output data for NN training by the stochastic simulation technology.

Step 4. Train the NN to approximate the uncertain function, namely, the transition transmission rule/relation of the system life calculation, and obtain output data of a system life of the NN.

Step 5. Forecast outputs of system life value are obtained by the NN algorithm.

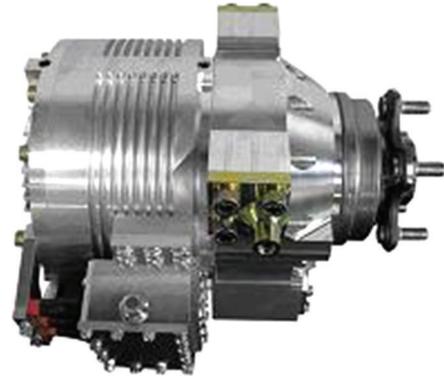


FIGURE 5: Motor of electric vehicle.

Step 6. Calculate reliability degree of $R(x)$ and average life $E(x)$ by stochastic simulation and obtained forecast outputs of system life value.

The above algorithm has been implemented in the MATLAB (R2009b) programming language.

4. Reliability Modeling and Evaluation of Electric Vehicle Motor

4.1. Reliability Modeling of Electric Vehicle Motor. As shown in Figure 5, it is a schematic graph of electric vehicle motor. It mainly consists of 3 parts, that is, stator, rotor, and axis. According to its components and their logical relation of fault occurrence, combined with the presented concept of FT, its FT model is shown in Figure 6.

Based on elements transformation rules of FA to ESPN, the FT based ESPN model of motor of electric vehicle is obtained and shown in Figure 7.

In addition, the paraphrase for the transitions and places of FT based ESPN model of electric vehicle motor is listed in Tables 1 and 2, respectively. Additionally, the life distribution types and parameters associated with each transition is listed in Table 2. Note that $\text{Norm}(\mu, \sigma)$ and $\text{Exp}(\lambda)$ denote normal and exponential distributions, respectively. The distribution parameters can be determined by the life test of motor and the unit of life time is hour [41].

4.2. Reliability Evaluation of Electric Vehicle Motor. The parameters of the NN based simulation algorithm are set as follows: for FT based ESPN model of electric vehicle motor, since there are 7 bottom places in this model, the number of input neurons is set to be $u = 7$. Thus the number of hidden neurons is 12 by letting $b = 12$. The given error term value is 0.000000004. Based on the stochastic simulation, 5000 input-output data points are generated and obtained. Moreover, they are separated into two groups: 3000 data points for training and 2000 points for testing. Note that the number of training data points is set to be 5000 due to two reasons. First, too few data points make the solution of models inaccurate. Second, too many points increase the training time and response time of the solution model. Note

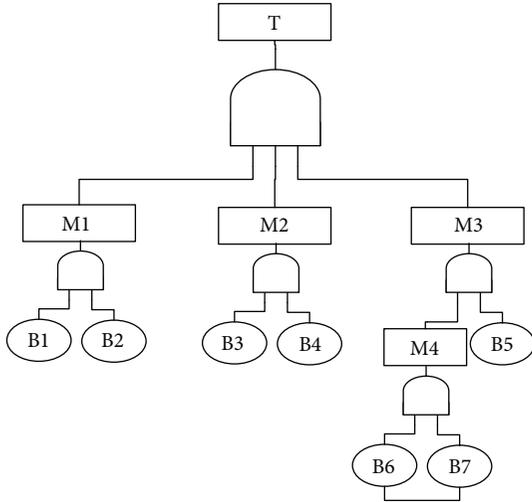


FIGURE 6: FT model of motor of electric vehicle.

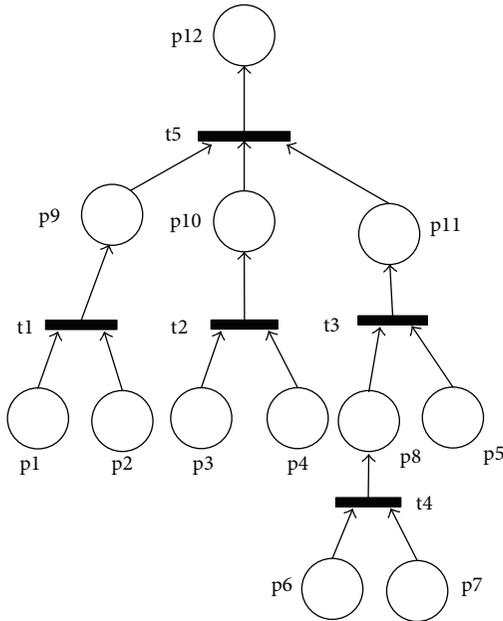


FIGURE 7: ESPN model of electric vehicle motor.

that the solution accuracy cannot be increased significantly with the number of data points [36, 37].

4.2.1. The Results of Average Life $E(x)$ of Electric Vehicle Motor. Based on the established FT based ESPN model and the above presented algorithm, after the proposed stochastic simulation algorithms based on NN are executed, the average life $E(x)$ of electric vehicle motor is obtained; that is, $E(x) = 4492.4$ hours. It denotes that the average use time of this electric vehicle motor is 4492.4 hours.

4.2.2. The Results of Reliability Degree $R(x)$ of Electric Vehicle Motor. Given the different running time x of the electric vehicle motor, after the proposed stochastic simulation

TABLE 1: Events and paraphrases of each place in FT and ESPN models.

FT	ESPN	Paraphrase
B1	p1	Failure of core of the stator
B2	p2	Failure of winding of the stator
B3	p3	Failure of core of the rotor
B4	p4	Failure of winding of the rotor
B5	p5	Failure of bearing
B6	p6	Failure of the shaft
B7	p7	Failure of the bond
M4	p8	Failure of the shaft assembly
M1	p9	Failure of the stator
M2	p10	Failure of the rotor
M3	p11	Failure of the axis system
T	p12	Failure of the motor system

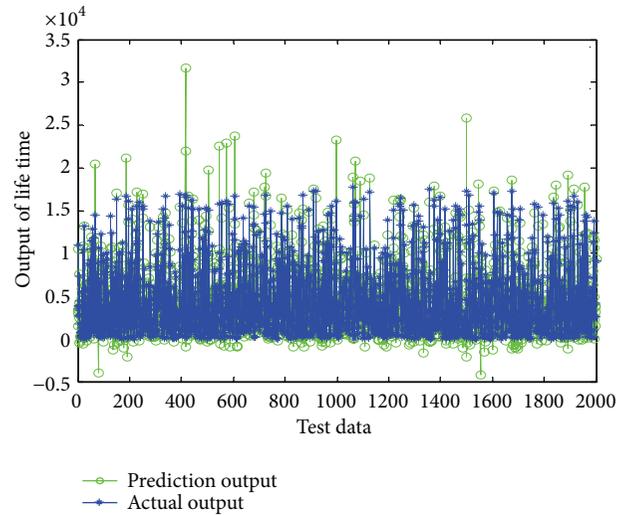


FIGURE 8: Prediction output of the proposed algorithm.

algorithms based on NN are executed, the corresponding reliability degree $R(x)$ is obtained, as listed in Table 3.

As seen in Table 3, for example, $R(800) = 0.871$; it denotes that the probability of this motor, which does not occur faulty, is 0.871 when it has run for 800-hour time.

In addition, in order to make future tests and observe the effectiveness of the proposed methods, the predicted outputs of NN method and their output error between forecast and actual outputs at test data points are shown in Figures 8 and 9.

From Figures 8 and 9, the predicted and output results of the proposed algorithm are highly close. It reveals that the proposed algorithm can accurately achieve reliability evaluation of the electric vehicle motor.

5. Conclusion

Electric vehicle motor is one of key components in electric vehicle and has a great impact on the vehicle safety; thus it is important to perform its reliability analysis. To do

TABLE 2: Paraphrases of each transition in ESPN models.

ESPN	Paraphrase	Life distribution type $f(x)$ associated with transition t
t_1	Faults of core and winding of the stator cause fault of stator	Life distribution of core of the stator x_1 is Norm(20000, 1100). Life distribution of winding of the stator x_2 is Exp(16000)
t_2	Faults of core and winding of the rotor cause fault of the rotor	Life distribution of core of the rotor x_3 is Exp(16000). Life distribution of winding of the rotor x_4 is Norm(18000, 1000)
t_3	Faults of bearing and shaft assembly cause fault of the axis system	Life distribution of bearing x_5 is Exp(12000). Life distribution of shaft assembly is obtained by x_6 and x_7
t_4	Faults of shaft and bond cause fault of shaft assembly	Life distribution of shaft x_6 is Norm(18000, 1000). Life distribution of bond x_7 is Norm(16000, 1100)
t_5	Faults of the stator, rotor, and axis system cause the fault of the motor system	Life distributions of stator, rotor, and axis system are obtained by calculation of life of constituting its components

TABLE 3: Reliability degree of electric vehicle motor under different running time.

x	600	800	1000	1200	1400
$R(x)$	0.889	0.871	0.841	0.8111	0.778

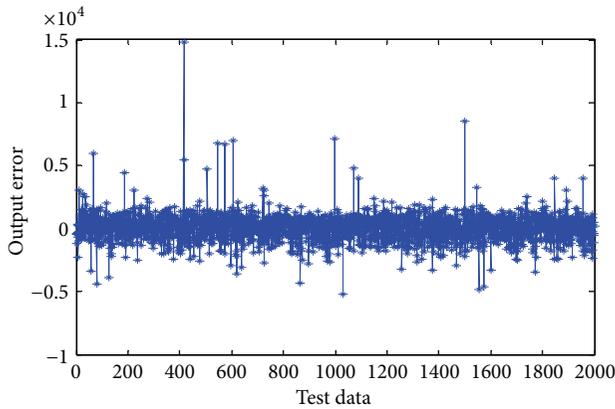


FIGURE 9: Output error of the proposed algorithm.

so, currently, researchers have discussed this problem by using FT analysis method. However, this method has many defects in analyzing the reliability of mechanical systems; for example, it cannot achieve the dynamic description of reliability and the building process of FT model needs a variety of elements. To deal with this problem, this paper proposes reliability modeling and evaluation issues of electric vehicle motor by using FT based extended stochastic Petri nets for the first time. Based on the concepts of FT and ESPN, combined with the defined transformation rules of their elements, an FT based ESPN model for reliability analysis of the mechanical system is obtained. In addition, the reliability calculation method is introduced on FT based ESPN model of the mechanical system and this work designs hybrid intelligent algorithms integrating stochastic simulation and NN, namely, NN based simulation algorithm, to solve it. Finally, taking an electric vehicle motor as an example, its reliability modeling and evaluation issues are analyzed. The results reveal that they are feasible when used to solve the proposed problems. The obtained results can be used to

guide decision makers in making better design when electric vehicle motor is developed and designed.

The future work is to find and use actual reliability test data to validate this method to provide the best decision support for reliability analysis of electric vehicle motor. This work merely analyzes the reliability of mechanical system of the electric vehicle motor. Thus, the reliability issue of integrated system integrating mechanical system and other software needs to be further discussed. In addition, advanced control and technology of electric vehicle motor should be further studied to improve its safety [42–44].

Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

Acknowledgments

This work is financially supported by the Fundamental Research Funds for the Central Universities of China (DL13BB12), and China Postdoctoral Science Foundation Project under Grant 2013M541329.

References

- [1] J. C. González Palencia, T. Furubayashi, and T. Nakata, “Energy use and CO₂ emissions reduction potential in passenger car fleet using zero emission vehicles and lightweight materials,” *Energy*, vol. 48, no. 1, pp. 548–565, 2012.
- [2] T. F. Go, D. A. Wahab, M. N. A. Rahman, R. Ramli, and C. H. Azhari, “Disassemblability of end-of-life vehicle: a critical review of evaluation methods,” *Journal of Cleaner Production*, vol. 19, no. 13, pp. 1536–1546, 2011.
- [3] G. Tian, Y. Liu, H. Ke, and J. Chu, “Energy evaluation method and its optimization models for process planning with stochastic characteristics: a case study in disassembly decision-making,” *Computers & Industrial Engineering*, vol. 63, no. 3, pp. 553–563, 2012.
- [4] G. Tian, M. Zhou, J. Chu, and Y. Liu, “Probability evaluation models of product disassembly cost subject to random removal time and different removal labor cost,” *IEEE Transactions on Automation Science and Engineering*, vol. 9, no. 2, pp. 288–295, 2012.

- [5] G. Tian, J. Chu, Y. Liu, H. Ke, X. Zhao, and G. Xu, "Expected energy analysis for industrial process planning problem with fuzzy time parameters," *Computers & Chemical Engineering*, vol. 35, no. 12, pp. 2905–2912, 2011.
- [6] M. Burgess, N. King, M. Harris, and E. Lewis, "Electric vehicle drivers' reported interactions with the public: driving stereotype change?" *Transportation Research F*, vol. 17, pp. 33–44, 2013.
- [7] P. Liu and H. P. Liu, "Application of z-source inverter for permanent-magnet synchronous motor drive system for electric vehicles," *Procedia Engineering*, vol. 15, pp. 309–314, 2011.
- [8] X. Peng, H. Zhe, G. Guifang, X. Gang, C. Binggang, and L. Zengliang, "Driving and control of torque for direct-wheel-driven electric vehicle with motors in serial," *Expert Systems with Applications*, vol. 38, no. 1, pp. 80–86, 2011.
- [9] C. Quinn, D. Zimmerle, and T. H. Bradley, "The effect of communication architecture on the availability, reliability, and economics of plug-in hybrid electric vehicle-to-grid ancillary services," *Journal of Power Sources*, vol. 195, no. 5, pp. 1500–1509, 2010.
- [10] X. H. Zhu, S. M. Cui, N. Shi, and Y. L. Min, "Grey prediction model of motor reliability of electric vehicle," *Electric Machines and Control*, vol. 16, no. 8, pp. 42–46, 2012.
- [11] X. H. Zhu, S. M. Cui, N. Shi, and Y. L. Min, "Grey prediction model of electric vehicle motor based on particle swarm optimization," *High Voltage Engineering*, vol. 38, no. 6, pp. 1391–1396, 2012.
- [12] K. Mokhtari, J. Ren, C. Roberts, and J. Wang, "Application of a generic bow-tie based risk analysis framework on risk management of sea ports and offshore terminals," *Journal of Hazardous Materials*, vol. 192, no. 2, pp. 465–475, 2011.
- [13] S.-R. Cheng, B. Lin, B.-M. Hsu, and M.-H. Shu, "Fault-tree analysis for liquefied natural gas terminal emergency shutdown system," *Expert Systems with Applications*, vol. 36, no. 9, pp. 11918–11924, 2009.
- [14] A. Montes and I. H. Helvacioğlu, "An application of fuzzy fault tree analysis for spread mooring systems," *Ocean Engineering*, vol. 38, no. 2-3, pp. 285–294, 2011.
- [15] J. Wu, S. Yan, and L. Xie, "Reliability analysis method of a solar array by using fault tree analysis and fuzzy reasoning Petri net," *Acta Astronautica*, vol. 69, no. 11-12, pp. 960–968, 2011.
- [16] M. C. Zhou and F. Dicesare, *Petri Net Synthesis for Discrete Event Control of Manufacturing Systems*, Kluwer Academic, 1993.
- [17] M. C. Zhou, *Modeling, Simulation, and Control of Flexible Manufacturing Systems: A Petri Net Approach*, World Scientific, 1999.
- [18] H. S. Hu, M. C. Zhou, Z. W. Li, and Y. Tang, "An optimization approach to improved Petri net controller design for automated manufacturing systems," *IEEE Transactions on Automation Science and Engineering*, vol. 10, no. 3, pp. 772–782, 2013.
- [19] Z. Li and M. Zhou, "Elementary siphons of Petri nets and their application to deadlock prevention in flexible manufacturing systems," *IEEE Transactions on Systems, Man, and Cybernetics A: Systems and Humans*, vol. 34, no. 1, pp. 38–51, 2004.
- [20] M. Gao, M. Zhou, and Y. Tang, "Intelligent decision making in disassembly process based on fuzzy reasoning Petri nets," *IEEE Transactions on Systems, Man, and Cybernetics B: Cybernetics*, vol. 34, no. 5, pp. 2029–2034, 2004.
- [21] A. Mhalla, S. Collart Dutilleul, E. Craye, and M. Benrejeb, "Estimation of failure probability of milk manufacturing unit by fuzzy fault tree analysis," *Journal of Intelligent & Fuzzy Systems*, vol. 26, no. 2, pp. 741–750, 2014.
- [22] N. Xie and A. Li, "Modeling and analysis of reconfigurable manufacturing system by extended stochastic petri nets," *Jixie Gongcheng Xuebao/Chinese Journal of Mechanical Engineering*, vol. 42, no. 12, pp. 224–231, 2006.
- [23] G. Q. Huang, B. Zhang, and C. Z. Wang, "Network attack model based on hierarchical expanded stochastic Petri net," *Computer Engineering*, vol. 37, no. 22, pp. 12–18, 2011.
- [24] A. Adamyan and D. He, "System failure analysis through counters of Petri net models," *Quality and Reliability Engineering International*, vol. 20, no. 4, pp. 317–335, 2004.
- [25] B. Bertsche, *Reliability in Automotive and Mechanical Engineering*, Springer, Berlin, Germany, 2008.
- [26] D. X. Liu and S. V. Amari, "Fault tree analysis," in *Handbook of Performability Engineering*, K. B. Misra, Ed., Springer, London, UK, 2008.
- [27] W. Li, Y. Yao, Z. Ding, and Q. Pu, "Efficient mobile authentication scheme preserving user privacy for large-scale wireless networks," *Ad Hoc & Sensor Wireless Networks*, vol. 17, no. 3-4, pp. 313–339, 2013.
- [28] Z. Ding, J. Liu, J. Wang, and F. Wang, "An executable service composition code automatic creation tool based on Petri net model," *Computing and Informatics*, vol. 32, no. 5, pp. 968–986, 2013.
- [29] L. Li, Z. J. Sun, M. C. Zhou, and F. Qiao, "Adaptive dispatching rule for semiconductor wafer fabrication facility," *IEEE Transactions on Science and Engineering*, vol. 10, no. 2, pp. 354–364, 2013.
- [30] M. Zhou, D. Guo, and F. Dicesare, "Integration of Petri nets and moment generating function approaches for system performance evaluation," *Journal of Systems Integration*, vol. 3, no. 1, pp. 43–62, 1993.
- [31] D. L. Guo, F. DiCesare, and M. C. Zhou, "A moment generating function based approach for evaluating extended stochastic Petri nets," *IEEE Transactions on Automatic Control*, vol. 38, no. 2, pp. 321–327, 1993.
- [32] B. Liu, *Uncertainty Theory: An Introduction to Its Axiomatic Foundations*, vol. 154, Springer, Berlin, Germany, 2004.
- [33] G. D. Tian, M. C. Zhou, and J. W. Chu, "A chance constrained programming approach to determine the optimal disassembly sequence," *IEEE Transactions on Automation Science and Engineering*, vol. 10, no. 4, pp. 1004–1013, 2013.
- [34] G. Tian, Y. Liu, H. Zhang, J. Chu, and G. Xu, "Chance programming models for time-energy trade-off problem of product disassembly process with multiple stochastic variables," *Advanced Science Letters*, vol. 4, no. 4-5, pp. 1851–1854, 2011.
- [35] M. A. Naeem, D. J. Dias, R. Tibrewal, P. C. Chang, and M. K. Tiwari, "Production planning optimization for manufacturing and remanufacturing system in stochastic environment," *Journal of Intelligent Manufacturing*, vol. 24, no. 4, pp. 717–728, 2013.
- [36] V. N. Gaitonde and S. R. Karnik, "Minimizing burr size in drilling using artificial neural network (ANN)-particle swarm optimization (PSO) approach," *Journal of Intelligent Manufacturing*, vol. 23, no. 5, pp. 1783–1793, 2012.
- [37] G. D. Tian, T. G. Qiang, J. W. Chu, G. Xu, and W. Zhou, "Efficiency optimization for disassembly tools via using NN-GA approach," *Mathematical Problems in Engineering*, vol. 2013, Article ID 173736, 8 pages, 2013.
- [38] U. Zuperl and F. Cus, "Optimization of cutting conditions during cutting by using neural networks," *Robotics and Computer-Integrated Manufacturing*, vol. 19, no. 1-2, pp. 189–199, 2003.

- [39] V. S. Sharma, S. Dhiman, R. Sehgal, and S. K. Sharma, "Estimation of cutting forces and surface roughness for hard turning using neural networks," *Journal of Intelligent Manufacturing*, vol. 19, no. 4, pp. 473–483, 2008.
- [40] G. D. Tian, T. G. Qiang, and J. W. Chu, "Influence factor analysis and prediction models for component removal time in manufacturing," *Proceedings of the Institution of Mechanical Engineers B: Journal of Engineering Manufacture*, vol. 227, no. 10, pp. 1533–1540, 2013.
- [41] Y. L. Min, *Research on life prediction and reliability testing method of the electric vehicle drive motor [M.S. thesis]*, Harbin Institute of Technology, 2011.
- [42] W. Sun, H. Gao Sr., and O. Kaynak, "Finite frequency H_∞ control for vehicle active suspension systems," *IEEE Transactions on Control Systems Technology*, vol. 19, no. 2, pp. 416–422, 2011.
- [43] W. H. Sun, Z. Zhao, and H. J. Gao, "Saturated adaptive robust Control for active suspension systems," *IEEE Transactions on Industrial Electronics*, vol. 60, no. 9, pp. 3889–3896, 2013.
- [44] W. H. Sun, H. J. Gao, and O. Kaynak, "Adaptive backstepping control for active suspension systems with hard constraints," *IEEE Transactions on Mechatronics*, vol. 18, no. 3, pp. 1072–1079, 2013.

Research Article

Research on Optimal Control for the Vehicle Suspension Based on the Simulated Annealing Algorithm

Jie Meng,¹ Qingzhang Chen,¹ and Ren He²

¹ Department of Automotive Engineering, Changshu Institute of Technology, Changshu 215500, China

² School of Automotive and Traffic Engineering, Jiangsu University, Zhenjiang 212013, China

Correspondence should be addressed to Qingzhang Chen; chenqz1973@sina.com

Received 17 February 2014; Accepted 1 March 2014; Published 3 April 2014

Academic Editor: Hamid R. Karimi

Copyright © 2014 Jie Meng et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

A method is designed to optimize the weight matrix of the LQR controller by using the simulated annealing algorithm. This method utilizes the random searching characteristics of the algorithm to optimize the weight matrices with the target function of suspension performance indexes. This method improves the design efficiency and control performance of the LQR control, and solves the problem of the LQR controller when defining the weight matrices. And a simulation is provided for vehicle active chassis control. The result shows that the active suspension using LQR optimized by the genetic algorithm compared to the chassis controlled by the normal LQR and the passive one, shows better performance. Meanwhile, the problem of defining the weight matrices is greatly solved.

1. Introduction

The suspension system is an important component of the vehicle, and its performance is a major factor that affects the vehicle ride comfort, operation, and stability. The traditional passive suspension is generally composed of the elastic component and damping components with the fixed parameters. And it is generally designed to meet a particular road, so it limits the vehicle performance. Therefore, industrial countries have already begun to study the active suspension and the semiactive suspension systems based on the active vibration control since the 1970s. In recent years, with the rapid development of the electronic technology, testing techniques, and system dynamics theories, the vibration isolation for the vehicle suspension system has changed from the traditional passive to the active vibration control. There is a lot of research in the suspension control areas in China. The application of the modern control theory in the vehicle suspension control also has a variety of algorithms, such as the PID, skyhook, optimal, and sliding-mode controllers [1–5].

The conventional fuzzy control lacks the control precision, and it is difficult to establish fuzzy rules. And the high

complexity of the neural network, as a representative of the intelligent control technology, makes it impossible to analyze the various performance indexes accurately. Therefore, the system versatility cannot be satisfied.

Among the modern control theories, the optimal control is widely used, and its theoretical foundation is also perfect. As a kind of optimal control, the linear quadratic regulator (LQR) can provide the required optimal performance indexes, by determining the weighting matrix of the system state and control variables. And it can also give sufficient design space for designers. When applied to the vehicle active suspension design, the LQR control can put forward different objective functions for the different performance need.

That is to say, the LQR control can improve the vehicle performance by the comprehensive consideration of the various factors in the vehicle suspension system.

As to the LQR control, the choice of the weight matrix for the performance indexes has great influence on the control system performance. Usually, the weight matrix is determined manually according to the system's physical process, so it needs full understanding and sufficient experience of the system. Therefore, such "optimal" results are completely subjective and artificial. If the weight matrix is improperly

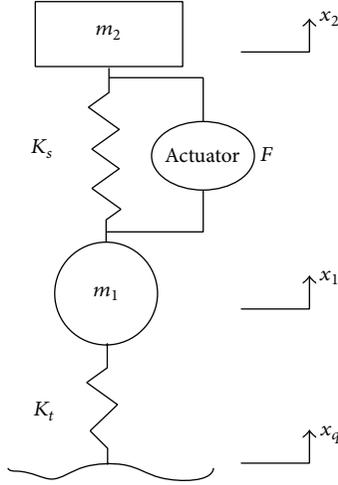


FIGURE 1: The quarter-vehicle body model of 2 DOF.

selected, the optimal solution of no practical significance. In addition, the choice of the weighted matrix relies too much on the designer's experience. The designers must adjust the weighted matrix step by step according to the system output until a satisfactory response value of output is acquired. Thus, the obtaining of feedback coefficient for the optimal control cannot guarantee the system to achieve the optimal aims [6–8].

At present, some random search methods such as the Genetic Algorithm (GA) and the Simulated Annealing (SA) have already caused the interest of many scholars. And they are already applied in the active suspension control field successfully. The simulated annealing algorithm is such a kind of the random search methods, and it is an approximation algorithm suitable and effective for solutions to the optimization problems in the mass general combinatorial fields. Compared with the previous approximate algorithm, the SA has the advantages of simple description, flexibility in use, universality, high efficiency, little constraint by the initial conditions, and so forth.

Due to the problems of the LQR in determining the weight matrix, a vehicle active suspension is selected as the control object, and the SA is applied in the LQR controller design optimized by the objective function of the active suspension performance index, aiming at improving the design efficiency and property.

2. Establishment of the Active Suspension Control System

For research convenience, a simplified quarter-vehicle model is selected as the research object, as Figure 1 shows.

In Figure 1, the symbol m_1 is the nonsprung mass and m_2 is the sprung mass. K_t means the tire stiffness, and K_s means the suspension stiffness. F is the active force, and X_q is the road input. X_1 is the displacement of the nonsprung mass

and X_2 is the displacement of the sprung mass. The dynamics equation can be got from Figure 1:

$$\begin{aligned} m_2 \ddot{x}_2 &= U - K_s (x_2 - x_1), \\ m_1 \ddot{x}_1 &= -U + K_s (x_2 - x_1) - K_t (x_1 - x_q). \end{aligned} \quad (1)$$

The state variables and the output variables are selected, respectively, as follows:

$$\begin{aligned} X &= (\dot{x}_2, \dot{x}_1, x_2, x_1, x_q)^T, \\ Y &= (\ddot{x}_2, x_2 - x_1, x_1 - x_q)^T. \end{aligned} \quad (2)$$

The control input is the active force F . And the filtering white noise is used to simulate the real road input:

$$\dot{x}_q(t) = -2\pi f_0 x_q(t) + 2\pi \sqrt{G_0 V} \omega(t). \quad (3)$$

In the formula, the symbol G_0 represents the pavement roughness coefficient (m^3/cycle). The symbol V means the vehicle speed. The $\omega(t)$ means the Gaussian white noise with zero mathematical expectations. The f_0 is the lower cutoff frequency.

Then, the system state equation can be achieved:

$$\begin{aligned} \dot{\mathbf{x}} &= \mathbf{A}\mathbf{x} + \mathbf{B}\mathbf{U}, \\ \mathbf{y} &= \mathbf{C}\mathbf{x} + \mathbf{D}\mathbf{U}. \end{aligned} \quad (4)$$

Annotation. Consider the following:

$$\mathbf{A} = \begin{bmatrix} 0 & 0 & -\frac{K_s}{m_2} & \frac{K_s}{m_2} & 0 \\ 0 & 0 & \frac{K_s}{m_1} & -\frac{K_t + K_s}{m_1} & \frac{K_t}{m_1} \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & -2\pi f_0 \end{bmatrix},$$

$$\mathbf{B} = \begin{bmatrix} 0 & \frac{1}{m_2} \\ 0 & \frac{-1}{m_1} \\ 0 & 0 \\ 0 & 0 \\ 2\pi \sqrt{G_0 V} & 0 \end{bmatrix}, \quad (5)$$

$$\mathbf{C} = \begin{bmatrix} 0 & 0 & -\frac{K_s}{m_2} & \frac{K_s}{m_2} & 0 \\ 0 & 0 & 1 & -1 & 0 \\ 0 & 0 & 0 & 1 & -1 \end{bmatrix}, \quad \mathbf{D} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix},$$

$$\mathbf{U} = \begin{bmatrix} \omega(t) \\ F \end{bmatrix}.$$

3. Optimal Controller Design for Active Suspension

The main performance for automobile suspension design includes three indexes: the body vertical vibration acceleration which represents the car ride, the suspension dynamic travel which represents the body posture and the suspension structure, and the dynamic load of the tyre which represents the tire grounding characteristic.

Therefore, the target performance index J can be set as the mean integral value of the weighted square sum of the body acceleration, the suspension dynamic travel, and the tire's dynamic displacement:

$$J = \lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T \left[q_1 (x_1 - x_q)^2 + q_2 (x_2 - x_1)^2 + q_3 \dot{x}_2^2 \right] dt. \quad (6)$$

In the equation, the symbols q_1 , q_2 , and q_3 represent the weight coefficients of the body acceleration, the suspension dynamic travel, and the tire's dynamic displacement, respectively. The choice of the three weight coefficients depends on the designer's tendency of the suspension performance.

Then, formula (4) is substituted into the performance index J and the following formula can be got:

$$J = \lim_{x \rightarrow \infty} \frac{1}{T} \int_0^T \left[\mathbf{x}^T \mathbf{Q} \mathbf{x} + \mathbf{u}^T \mathbf{R} \mathbf{u} + 2 \mathbf{x}^T \mathbf{N} \mathbf{u} \right] dt. \quad (7)$$

The symbols \mathbf{Q} , \mathbf{R} , and \mathbf{N} are noted as follows:

$$\mathbf{Q} = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & q_2 + \frac{K_s^2}{m_2^2} & -q_2 - \frac{K_s^2}{m_2^2} & 0 \\ 0 & 1 & -q_2 - \frac{K_s^2}{m_2^2} & q_1 + q_2 + \frac{K_s^2}{m_2^2} & -q_1 \\ 0 & 0 & 0 & -q_1 & q_1 \end{bmatrix}, \quad \mathbf{R} = \frac{1}{m_2^2},$$

$$\mathbf{N} = \frac{1}{m_2^2} \begin{bmatrix} 0 \\ 0 \\ -K_s \\ K_s \\ 0 \end{bmatrix}. \quad (8)$$

The feedback matrix K of the optimal control can be achieved from the Riccati matrix formula:

$$\mathbf{P} \mathbf{A} + \mathbf{A}^T \mathbf{P} - (\mathbf{P} \mathbf{B} + \mathbf{N}) \mathbf{R}^{-1} (\mathbf{B}^T \mathbf{P} + \mathbf{N}^T) + \mathbf{Q} = \mathbf{0}. \quad (9)$$

According to the feedback variables $\mathbf{X}(t)$ at any time, the optimal control force \mathbf{F} can be got by the following equation:

$$\mathbf{F} = -\mathbf{K} \mathbf{X}(t). \quad (10)$$

The \mathbf{K} value can be acquired through the design function for the optimal controller in Matlab:

$$[\mathbf{K}, \mathbf{S}, \mathbf{e}] = \text{LQR}(\mathbf{A}, \mathbf{B}, \mathbf{Q}, \mathbf{R}, \mathbf{N}). \quad (11)$$

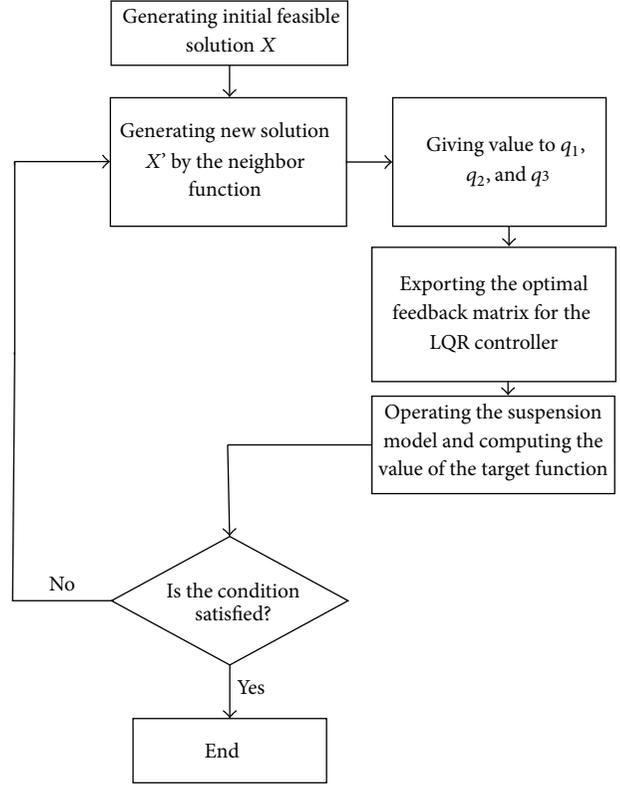


FIGURE 2: The optimization process for the LQR controller by the SA.

According to the above formulas, the optimal control design depends on the choice of weighting coefficients totally. Previous designs often rely on the designer's own experience through repeated trials. Though corresponding optimal results can be got, a large subjectivity is met.

Therefore, the simulated annealing algorithm is used to optimize the weights coefficient matrix in order to reduce the controller design time and avoid the subjectivity at the same time.

4. Optimum Controller Design Based on the Simulated Annealing Algorithm

The simulated annealing (SA) algorithm is put forward by Metropolis at the earliest. His starting point is based on the similarity of the solid material annealing process to the general combinatorial optimization problem. The SA algorithm is a general optimization algorithm with its annealing process composed of three processes: the heating process, the isothermal process, and the cooling process [9, 10].

The optimum controller design based on simulated annealing algorithm is a combination of the SA and LQR control method. The SA is used to optimize and solve the weight matrix, and the LQR is used to calculate the system performance index.

4.1. Design Model of the LQR Controller Optimized by the SA

4.1.1. *Designation of the Optimized Variable.* The weights coefficient is taken as the optimized variable of the SA:

$$X = [q_1, q_2, q_3]^T. \quad (12)$$

4.1.2. *Establishment of the Objective Function.* Because there is difference in the order of magnitude and the units, the optimization aim of the objective function is taken as follows:

$$\begin{aligned} \text{Minimum } f(\mathbf{X}) &= \frac{\text{RMS}[\ddot{x}_3(\mathbf{X})]}{\text{RMS}[\ddot{x}_{3\text{pass}}(\mathbf{X})]} \\ &+ \frac{\text{RMS}[(x_3 - x_2)(\mathbf{X})]}{\text{RMS}[(x_3 - x_2)_{\text{pass}}(\mathbf{X})]} \\ &+ \frac{\text{RMS}[(x_1 - x_q)(\mathbf{X})]}{\text{RMS}[(x_1 - x_q)_{\text{pass}}(\mathbf{X})]} \\ \text{s.t. } &\frac{\text{RMS}[\ddot{x}_3(\mathbf{X})]}{\text{RMS}[\ddot{x}_{3\text{pass}}(\mathbf{X})]} < 1 \\ &\frac{\text{RMS}[(x_3 - x_2)(\mathbf{X})]}{\text{RMS}[(x_3 - x_2)_{\text{pass}}(\mathbf{X})]} < 1 \\ &\frac{\text{RMS}[(x_1 - x_q)(\mathbf{X})]}{\text{RMS}[(x_1 - x_q)_{\text{pass}}(\mathbf{X})]} < 1 \end{aligned} \quad (13)$$

$$X = [q_1, q_2, q_3], \quad 0.1 < X_i < 10^6, \quad i = 1, 2, 3.$$

In the equations, the symbols are noted as follows:

- [RMS]: the root mean square value of relevant data;
- $\ddot{x}_{3\text{pass}}$: the vertical acceleration of the passive suspension;
- $(x_3 - x_2)_{\text{pass}}$: the dynamic displacement of the suspension;
- $(x_1 - x_q)_{\text{pass}}$: the dynamic displacement of the tyre;
- X : the coefficient matrix of the LQR.

4.2. *The Optimization Process of the SA.* The optimization process of the SA is shown in Figure 2.

- (1) Creation of the initial feasible solution: $X = [q_1, q_2, q_3]$. The initial temperature t_0 is determined, and random initial solution is selected together with the iterations times at any time which is the Metropolis chain length.
- (2) Then, new solutions are created through the neighborhood function according to the current solution X . This process continues until the new solutions are feasible, and then their optimization aims and comprehensive aims are evaluated.

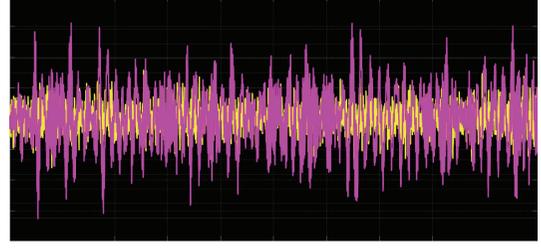


FIGURE 3: Vertical acceleration of the active and passive suspension.

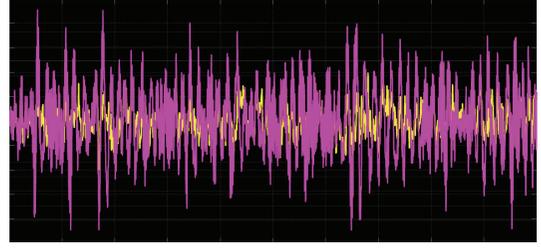


FIGURE 4: Dynamic travel of the active and passive suspension.

- (3) Make sure that the criteria are met at current temperature. If ok, then the process continues or returns to step (2).
- (4) Then, the cooling process is operated and $t_{k+1} = t_k \exp(-ck)$. The c is 0.9.
- (5) Make sure the algorithm meets the principle of termination. If it is ok, then the process is terminated or returns to step (2) again.

5. Simulation and the Results

Under the environment of the Matlab/Simulink, simulations are performed as to the active suspensions controlled by the LQR controller optimized by the SA and other controlling manners. The B-grade road in the standard file-ISO/TC108/SC2N67 is selected according to the previous model. A random road input is generated at the speed of 20 km/h, and G_0 is $5 \cdot 10^{-6} \text{ m}^3/\text{cycle}$. And m_1 is 50 kg, m_2 is 300 kg, m_3 is 50 kg, k_1 is 200000 N/m, and k_2 is 20000 N/m. The system simulation results are shown in Figures 3 to 5.

The LQR weights that the matrix gets from the simulation are [118.39 49.74 61.42].

In order to validate the superiority of the SALQR in the suspension control, the performance results of the suspension in Simulink simulation are compared and are shown in Table 1.

It can be seen from Figures 3, 4, and 5 and Table 1 the following.

- (1) The SALQR controller significantly reduces the vertical acceleration of the suspension. In this regard, its effect is very well.

TABLE 1: Performance contrast of the suspension of different control measures.

Performance indexes	Unit	The mean square root value		
		Passive suspension	Conventional LQR	SALQR
Vertical acceleration of suspension	m/s ²	2.8358	1.6342	0.3326
Dynamic travel of suspension	mm	40.7465	47.7899	35.4843
Dynamic deflection of the tyre	mm	11.8446	6.5684	24.7046

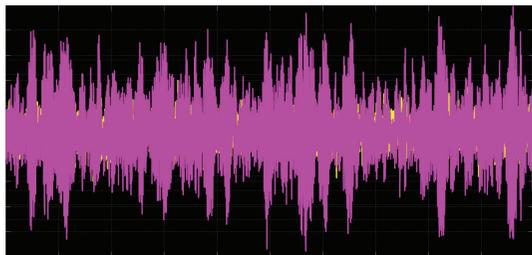


FIGURE 5: Dynamic deflection of the active and passive suspension.

- (2) In regard of the suspension's dynamic travel, the SALQR also gives good effect compared to the passive suspension and the conventional LQR suspension.
- (3) In the regard of the tire's dynamic deflection, the effect of the SALQR is no good compared to other controlling manners. But the weakness is not so significant to affect the whole performance of the suspension.

6. Conclusions

In view of the reliance on the designer's experience of the conventional LQR controller, a controller combining the SA and LQR is designed, and it has the following advantages.

- (1) The dependence on the designer's experience is reduced and the design convenience and efficiency are improved.
- (2) The controller performance is enhanced compared to the conventional LQR controller; therefore, the SA has a wide application prospect in the active control of automobile suspension.

Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

Acknowledgments

The authors gratefully acknowledge the financial support from the National Natural Science Foundation of China (no. 51275060) and the Natural Science Foundation of Jiangsu Province (no. BK2011367).

References

- [1] H. Zhang, J. Wang, and Y. Shi, "Robust H_∞ sliding-mode control for Markovian jump systems subject to intermittent observations and partially known transition probabilities," *Systems & Control Letters*, vol. 62, no. 12, pp. 1114–1124, 2013.
- [2] Y. G. Zhao and J. Chen, "Co-simulation on semi-active suspension of vehicles," *Noise and Vibration Control*, vol. 31, no. 5, pp. 104–107, 2011 (Chinese).
- [3] L. Chai, T. Sun, and J. Feng, "Design of the LQG controller for active suspension system based on analytic hierarchy process," *Automotive Engineering*, vol. 32, no. 8, pp. 712–718, 2010.
- [4] H. Zhang, Y. Shi, and J. Wang, "Observer-based tracking controller design for networked predictive control systems with uncertain Markov delays," *International Journal of Control*, vol. 86, no. 10, pp. 1824–1836, 2013.
- [5] R. Solea, A. Filipescu, and D. Cernega, "Lateral motion control of four-wheels steering vehicle using a sliding-mode controller," in *Proceedings of the 29th Chinese Control Conference (CCC '10)*, pp. 3699–3703, Beijing, China, July 2010.
- [6] L. Wang, *The Intelligence Algorithm and Its Application*, Tsinghua University press, Beijing, China, 2011.
- [7] J. Wang, J. Zhang, and X. Wei, "Multi-objective optimization design of gear reducer based on simulated annealing algorithms," *Transactions of the Chinese Society of Agricultural Machinery*, vol. 37, no. 10, pp. 120–123, 2006.
- [8] R. Zhou and Z. Guo, "Research on active suspension control technology," *Journal of Tongji University*, vol. 197, pp. 176–180, 2012.
- [9] D. Liu, Y. Tang, and H. Y. Gu, "The development of automotive active suspension control system," *Hydraulics Pneumatics & Seals*, vol. 5, no. 4, pp. 21–25, 2010.
- [10] W. J. Yan, D. Dong, and W. R. Wang, "On fuzzy strategy of nonlinear semi-active suspension system," *Control Engineering of China*, vol. 18, no. 6, pp. 941–946, 2011.

Research Article

A Direct Attitude Determination Approach Based on GPS Double-Difference Carrier Phase Measurements

Jianhua Cheng,^{1,2} Jing Wang,³ and Lin Zhao¹

¹ Marine Navigation Research Institute, College of Automation, Harbin Engineering University, Harbin 150001, China

² LASSENA Laboratory, Ecole de Technologie Supérieure, 1100 Notre-Dame Street West, Montreal, QC, Canada H3C 1K3

³ Beihang University, Beijing 100191, China

Correspondence should be addressed to Jing Wang; 050441jj@163.com

Received 26 November 2013; Accepted 18 February 2014; Published 25 March 2014

Academic Editor: Zheping Yan

Copyright © 2014 Jianhua Cheng et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

The principle of the traditional attitude solution approach based on GPS (Global Position System) is to get the attitude matrix according to the relationship of coordinates. During the progress, the error of baseline position assumed in ECEF (Earth-Centered Earth-Fixed) and the error of coordinate transform between body frame and reference frame (ENU, East-North-Up) have been included in the result, and finally the precision of attitude determination is reduced. This contribution presents a new approach of attitude determination, in which the attitude angles are calculated by the double-difference carrier phase equation of GPS according to the relationship of attitude matrix and attitude angles through least-squares estimate method. The new approach predigests the procedure of attitude determination which reduces the error assumed. According to the analysis the precision of attitude determination is higher than that of traditional method. It is shown it gets a precise attitude result with the direct attitude determination method in the simulation. A novel algorithm is proposed to solve some problems. Simulation results show the effectiveness of the proposed algorithm.

1. Introduction

GPS-based attitude determination is an important research field in recent years, where the key technologies involve the detection and reparation of cycle skip, integer ambiguities resolution, attitude solution, and so forth. At present, the common methods of attitude solution can be divided into two types [1, 2].

One type is to get the attitude matrix. Based on the relationship between baseline coordinates in the local orthogonal frame and the body frame, this type includes the TRIAD algorithm which estimates optimally according to the orthogonal property of the attitude matrix, the algorithms based on Wahba problem, for example, QUEST (QUaternion ESTimator), SVD (Singular Value Decomposition), FOAM (Fast Optimal Attitude Matrix), Euler-q, and so forth, and the least-square method using multiple epoch baseline coordinate [3]. Besides, resolving the equations which are established by the code or phase observations and the baseline coordinates can also estimate the attitude matrix.

The other type is to get the attitude angles. Setting one baseline along the azimuth axis of vehicle, the yaw and pitch angles can be obtained first, and then according to the rotating relationship of the second baseline, the roll angle can be calculated [4], that is the two-antenna attitude determination method. And the multiple-antenna method is rotating of the other baseline twice and getting the roll angle by the known yaw and pitch angles.

The method of getting attitude matrix includes two steps that estimations of baseline coordinate and attitude matrix parameters, and each of them will lead estimation errors in final results. On the other hand, the attitude matrix derived by the observed equations directly is usually not orthogonal. While, in the method of getting attitude angles the estimations of baseline coordinate and attitude matrix are drop, and procedure errors are except. But, when the z-axis component of all baselines in body frame is zero, the matrix is not full rank that results in unreliable results.

This contribution presents a new approach estimating the attitude angles by double-difference carrier phase equation

of GPS, which predigests the process calculating attitude angles by attitude matrix. The new method improves the precision and efficiency of estimation and avoids the problem of possible ambiguous values caused by traditional attitude-angle methods.

2. The Measurement Model of Attitude

Assume that attitude determination system consists of three antennas, namely, antenna 0 which is set as main antenna and antennas 1 and 2 which are set as slave antennas. All of the three antennas are tracking M satellites. The phase observations collected at receiver r tracking satellite i are modeled as follows:

$$\varphi_r^i = \lambda^{-1} (r_r^i - I_r^i + T_r^i) + f(\delta t_r - \delta t^i) + N_r^i + \varepsilon_{\varphi,r}^i, \quad (1)$$

where λ and f represent the wavelength and frequency of carrier wave, respectively; r_r^i is the real range between receiver and satellite (m); I_r^i is ionospheric effect (m); T_r^i is tropospheric effect (m); δt_r is receiver clock error (s); δt^i is satellite clock error (s); N_r^i is integer ambiguity; and $\varepsilon_{\varphi,r}^i$ is the measurement error.

2.1. Single-Difference Model. Single difference between the two slave receivers can be operated to eliminate many errors, such as the delays of ionospheric and tropospheric, the delays of instrumental and errors of satellite clock. Since the length of baseline is rather less than the distance between receiver and satellite, it is considered that the line-of-sight vectors of two endpoints of baseline are the same [5]. Signal is transmitted as shown in Figure 1.

According to the phase observation equation, when two antennas track the same satellite I, the single-difference equation can be derived as follows:

$$\varphi_{0r}^i = \lambda^{-1} r_{0r}^i + f\delta t_{0r} + N_{0r}^i + \varepsilon_{\varphi,0r}^i. \quad (2)$$

2.2. Double-Difference Model. Aiming to eliminate receiver clock error the differences between the observations of two satellites taken at the same time and different receivers are formed as follows:

$$\varphi_{0r}^{ij} = \varphi_{0r}^i - \varphi_{0r}^j = \lambda^{-1} r_{0r}^{ij} + N_{0r}^{ij} + \varepsilon_{\varphi,0r}^{ij}. \quad (3)$$

The purpose of relative positioning is to get the baseline vector b_{0r} . Therefore, it is necessary to contact the double-difference equation with baseline vector [6]. The line-of-sight vector is I_0^i , and the differences between the distance of main antenna to satellite i and that of slave antenna are r_{0r}^i , which is equal to the opposite number of the projection length of baseline vector on I_0^i , it is taken as

$$r_{0r}^i = -b_{0r} \cdot I_0^i. \quad (4)$$

Therefore, the relationship between double-difference distance and baseline vector can be expressed as

$$r_{0r}^{ij} = -b_{0r} \cdot I_0^i + b_{0r} \cdot I_0^j = -(I_0^i - I_0^j) \cdot b_{0r}. \quad (5)$$

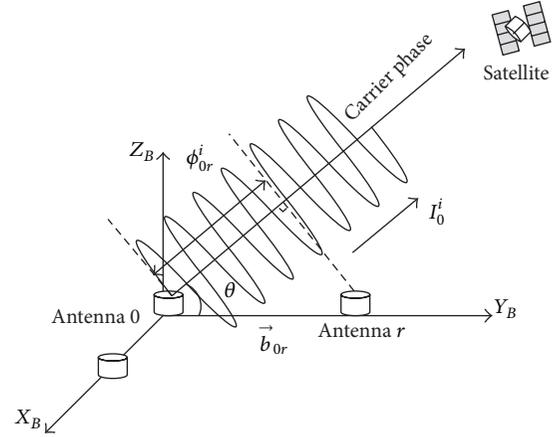


FIGURE 1: Single-difference observation of carrier phase measurements.

The double-difference equation of carrier phase is transformed into

$$\varphi_{0r}^{ij} = -\lambda^{-1} (I_0^i - I_0^j) \cdot b_{0r} + N_{0r}^{ij} + \varepsilon_{\varphi,0r}^{ij}, \quad (6)$$

where the relationship between double-difference φ_{0r}^{ij} and baseline vector b_{0r} is shown. φ_{0r}^{ij} is double-difference observation, which is a known quantity. b_{0r} is the baseline vector to be solved. N_{0r}^{ij} is double-difference integer ambiguity as an unknown quantity.

2.3. The Attitude Determination Model. In traditional process, the attitude determination needs two steps including the estimation of baseline coordinates and attitude matrix.

Aiming to simplify the calculations, avoiding the error during estimation, and improving the precision of solution, a direct method of gaining attitude angles using double-difference equation is introduced in the following [7].

Assume that GPS attitude determination system is connected rigidly with vehicle; that is, antenna coordinates in body frame are invariant and known which are, respectively, $r_{0,B}$, $r_{1,B}$, and $r_{2,B}$. The baseline vector b_i is equal to $r_{i,B} - r_{0,B}$, where i is equal to 1 or 2 as shown in Figure 2. In the reference frame, there is $r_i - r_0 = R^T(\varphi, \theta, \phi)b_i$.

According to the transformation relationship of coordinates, baseline vector can be expressed in body frame. Therefore, model (6) is transmitted into

$$\varphi_{i0}^{ku} = -\lambda^{-1} (I_0^k - I_0^u) \cdot A_R^E \cdot R^T(\varphi, \theta, \phi)b_i + N_{i0}^{ku} + \varepsilon_{\varphi,i0}^{ku}, \quad (7)$$

where R is rotation matrix from reference frame to body frame; A_R^E is transmission matrix from reference frame to ECEF.

After the double-difference integer ambiguity is determined, there are three attitude angles unknown in model (7). Then attitude angles can be estimated by observation equation.

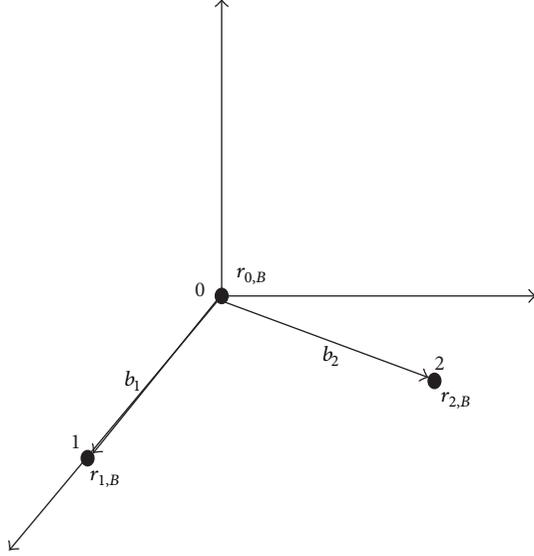


FIGURE 2: The relative position of antenna in the body frame.

M-1 double-difference equations can be established for every baseline:

$$\begin{bmatrix} \varphi_{i0}^{21} \\ \varphi_{i0}^{31} \\ \vdots \\ \varphi_{i0}^{M1} \end{bmatrix} = \lambda^{-1} \begin{bmatrix} -(I_0^2 - I_0^1)^T \\ -(I_0^3 - I_0^1)^T \\ \vdots \\ -(I_0^M - I_0^1)^T \end{bmatrix} A_R^E R^T(\varphi, \theta, \phi) b_i + \begin{bmatrix} N_{i0}^{21} \\ N_{i0}^{31} \\ \vdots \\ N_{i0}^{M1} \end{bmatrix}, \quad (8)$$

where satellite 1 is the reference satellite. According to relationship between attitude angles and the equation, linearization is operated. Then the final attitude angles can be gained through least-square estimation.

3. The Nonlinear Least-Square Solution

The form of rotation matrix R is as follows:

$$R = \begin{pmatrix} CrCy - SrSpSy & CrSy + SrSpCy & -SrCp \\ -CpSy & CpCy & Sp \\ SrCy + CrSpSy & SrSy - CrSpCy & CrCp \end{pmatrix}, \quad (9)$$

where S represents \sin ; C represents \cos ; y is the yaw angle of vehicle rotating around the z -axis of local frame; r is the roll angle of vehicle rotating around the y -axis of local frame; and p is the pitch angle of vehicle rotating around x -axis of local frame.

Because of the nonlinear relation between the model and attitude angle, attitude angle is determined by nonlinear least-square estimation. Assuming that state variables are three attitude angles as follows:

$$x = \begin{pmatrix} y \\ r \\ p \end{pmatrix}. \quad (10)$$

Firstly, with the high-order terms being neglected, the nonlinear equation is linearized through first-order Taylor expansion of every equation in the model. The nonlinear equations can be approximately transformed into linear equations expressed by the following matrix:

$$H \cdot \Delta x = \Delta y. \quad (11)$$

Observation matrix H can be expressed as follows:

$$H = \begin{pmatrix} \vdots \\ \cdots -(I_0^i - I_0^1)^T A_R^E \left(\frac{\partial R^T}{\partial x_i} \right) b_i \cdots \\ \vdots \end{pmatrix}, \quad (12)$$

where x_i ($i = 1, 2, 3$) is the attitude angle; $\partial R^T / \partial x_i$ is attitude matrix's partial derivative to each attitude, which can be expressed as follows:

$$\begin{aligned} \frac{\partial R^T}{\partial y} &= \begin{pmatrix} -CrSy - CySpSr & -CyCp & -SrSy + CySpCr \\ CyCr - SySpSr & -SyCp & CySr + SySpCr \\ 0 & 0 & 0 \end{pmatrix}, \\ \frac{\partial R^T}{\partial r} &= \begin{pmatrix} -CySr - SySpCr & 0 & CyCr - SySpSr \\ -SySr + CySpCr & 0 & SyCr + CySpSr \\ -CpCr & 0 & -CpSr \end{pmatrix}, \\ \frac{\partial R^T}{\partial p} &= \begin{pmatrix} -SyCpSr & SySp & SyCpCr \\ CyCpSr & -CySp & -CyCpCr \\ SpSr & Cp & -SpCr \end{pmatrix}. \end{aligned} \quad (13)$$

$\Delta y = \varphi_{i0}^{m1} - f(x_k)$, that is, the deviation of estimation state vector x_k to measured value φ_{i0}^{m1} .

Then the least-square solution is

$$\Delta x = (H^T H)^{-1} H^T \Delta y. \quad (14)$$

According to the equation of nonlinear least-square estimation and the initial attitude angle given above, the estimation of unknown parameters can be obtained. Namely, the attitude angle is estimated by double-difference observation equation of carrier phase.

4. Error Analysis

Compared with traditional approach of attitude determination, the solution error of model is analyzed.

4.1. Error of Traditional Attitude Determination Approach. Assume that the result is obtained from double-difference equation of carrier phase in both methods. In traditional attitude determination method, the baseline coordinates in local frame can be calculated through the equation firstly.

The phase observation error of all satellites is assumed as σ_φ , which is independent and its average value is zero. So the error of double-difference observation is $2\sigma_\varphi$. Uniting the double-difference equation of all visible satellites, solution is obtained through least-square estimation in overdetermined equation whose general expression is as follows:

$$Ax = b. \quad (15)$$

There is only one least-square solution expressed as follows:

$$A^H Ax = A^H b \implies x = (A^H A)^{-1} A^H b. \quad (16)$$

In (15), error of b is δb , and error of A is δA . The effects of δb and δA on solution x of the equation are proportional to the square of condition number of A , that means condition number of overdetermined equation will increase with a quadric function as the following formula shows:

$$\text{cond}(A^H A) = [\text{cond}(A)]^2. \quad (17)$$

According to the rule of error transmitting, the error of least-square solution is $4\sigma^2$. Here only the estimation of baseline coordinates in the local frame is finished. More resolution of baseline coordinates are needed in more observed epochs. Then the 9 parameters of three-dimensional attitude matrix can be estimated according to the transformation relationship between baseline coordinates in the body frame and local frame. Unit multipepoch baseline coordinates are as follows:

$$\begin{bmatrix} X^{BFS} \\ Y^{BFS} \\ Z^{BFS} \end{bmatrix} = R(r)_Y R(p)_X R(y)_Z \begin{bmatrix} X^{LLS} \\ Y^{LLS} \\ Z^{LLS} \end{bmatrix}, \quad (18)$$

where $(X^{BFS}, Y^{BFS}, Z^{BFS})$ are the baseline coordinates in the body frame; $(X^{LLS}, Y^{LLS}, Z^{LLS})$ are the baseline coordinates in the local frame.

So another least-square estimation is needed to solve the overdetermined equation. Then the estimation error increases to $16\sigma^4$ with a quadric function.

Many different theories are used to solve attitude matrix in the other methods, but the baseline coordinates in the local frame should be estimated at first. And the estimation error also exists in the process of attitude matrix determination. So the error of traditional attitude determination approach is in a high level.

4.2. Error of the Direct Attitude Determination Approach Based on Double-Difference Equation. Due to only one step of least-square estimation being operated in the direct attitude determination approach, the error of state vector estimated is as follows:

$$\begin{aligned} \text{cov}[\Delta x] &= E[\Delta x \Delta x^T] \\ &= (H^T H)^{-1} H^T E[\Delta y \Delta y^T] H (H^T H)^{-1}. \end{aligned} \quad (19)$$

Observation errors are mutually independent, whose variance is $4\sigma^2$ and their average is zero. Therefore, $E[\Delta y \Delta y^T] = 4I\sigma^2$, where I is unit matrix.

Dimensionless matrix \widehat{H} is defined as H/b_0 , then

$$\text{cov}[\Delta x] = (\widehat{H}^T \widehat{H})^{-1} \left(\frac{4\sigma_\varphi^2}{b_0^2} \right). \quad (20)$$

Therefore, the error of attitude angle estimated is concerned with the measurement error, measurement matrix, and length of baseline. Ignoring other errors, the longer the baseline is, the higher the precision is; according to (20), measurement matrix is only concerned with the geometric position of satellite. The better the geometric distribution is, the less the effect of measurement error on attitude is.

5. Simulation

5.1. Designation of Simulation Approach. Through setting the geodetic coordinate of main antenna, the body coordinate of baseline, and the true attitude angles, attitude determination algorithm can be verified by simulation system with the known satellite positions.

Simulation system consists of the following sections.

5.1.1. Obtainment of Satellites' Coordinates in ECEF. With the almanac files supplied by the website of navcen.uscg.gov and actually broadcasted by GPS satellites, the real-time coordinate of all satellites in orbit are calculated; then the coarse precision satellites' tracks can be obtained. After all elevations are calculated in the setting user position, visible satellites information can be obtained.

5.1.2. Simulation of Carrier Signal. After positions of main antenna and reference antenna are set, coordinates of all antennas in the local frame can be obtained according to the predetermined attitude angles. Namely, distance from satellite to antenna is obtained. With certain measurement noise added, carrier signal can be simulated. And the equation of phase observation can be established (assuming that integer ambiguity is determined).

5.1.3. Verification of Attitude Results. Parameters of simulation are known information, and according to different attitude solution approaches, equation is built to solve the attitude angles which will be compared with predetermined attitude angles to verify the algorithm.

5.2. Establishment of Attitude Simulation Model. Installing the array of the three antennas as Figure 3, the system determines attitude of vehicle.

Assume that reference antenna is at the rotation center of vehicle, phase error is set at (4, 5, 6) mm, and the other simulation conditions are set as Table 1 shown.

Elevation angle limited is set at 5° [8]; there are 8 visible satellites numbered as Table 2 shows, of which satellite number 22 is the reference satellite.

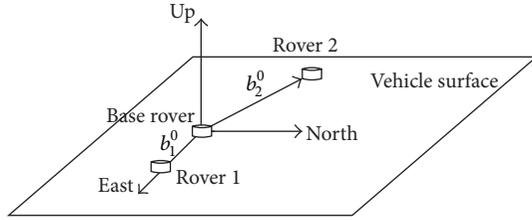


FIGURE 3: The distribution of antenna in the simulation model.

TABLE 1: Settings of simulation parameter.

Parameter	Value
Reference antenna coordinates in geographic frame	45°N, 120°E, 0 m
Antenna 1 coordinates in the body frame	(3, 0, 0) m
Antenna 2 coordinates in the body frame	(0, 3, 4) m
Predetermined attitude angles of simulation	Yaw: 90°, roll: 30°, pitch: 30°

TABLE 2: The visible satellites in the simulation position.

Sequence number	1	2	3	4	5	6	7	8
The visible satellite number	4	14	18	20	22	24	25	29

5.3. *Verification of Attitude Results.* Because that attitude solution of traditional approach needs multiepochs information, simulation is done for 10 minutes. Namely, 20 epochs is sampled. At the same time, the new approach is simulated. The two attitude determination results are compared.

In traditional attitude determination approach, the results estimated by least-square method once are listed as yaw is 90.0404°, roll is 30.0088°, and pitch is 29.9516°.

Attitude results estimated by nonlinear least-square iteration in the direct double-difference approach are listed as: yaw is 90.1092°, roll is 29.9925°, and pitch is 29.9831°. The estimation process is shown at Figure 4.

According to the result of direct approach, attitude angles become steady after 5 times cyclic calculations within only single epoch. While traditional method need more epochs for obtaining the final attitude solution. Precision of estimation is 10⁻¹ (°) that is higher than the result using traditional attitude determination approach.

6. Conclusions

The approach of solving attitude angles based on double-difference carrier phase measurements is proposed in this contribution, which is analyzed and compared with traditional approach, and some estimation procedures are avoided. The accuracy and low error in calculation of the new attitude determination model are verified. According to the attitude verification program based on MATLAB simulation, reality and reliability are certified and a certain attitude solution precision is achieved.

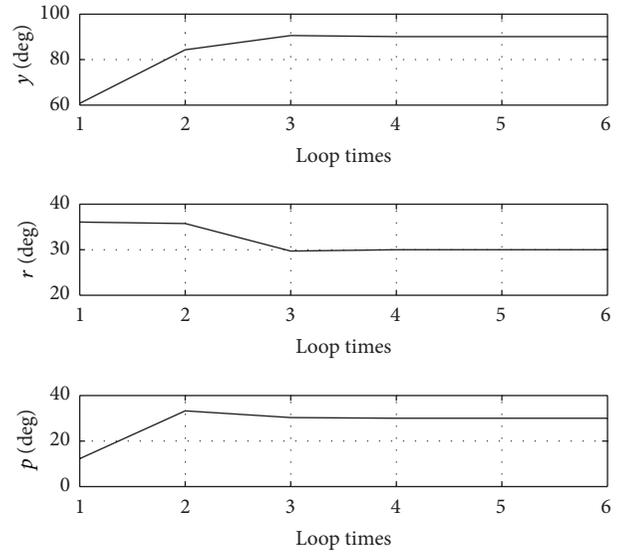


FIGURE 4: The process of attitude angles estimated using the direct approach based on carrier double-difference equation.

Conflict of Interests

The authors declare there is no conflict of interests regarding the publication of this paper.

Acknowledgments

Funding for this work was provided by the National Nature Science Foundation of China under Grant no. 61374007 and no. 61104036. The authors would like to thank all the editors and anonymous reviewers for improving this paper.

References

- [1] P. J. G. Teunissen, G. Giorgi, and P. J. Buist, “Testing of a new single-frequency GNSS carrier phase attitude determination method: land, ship and aircraft experiments,” *GPS Solutions*, vol. 15, no. 1, pp. 15–28, 2011.
- [2] G. Giorgi, T. P. Gourlay, P. J. G. Teunissen, L. Huisman, and K. Klaka, “Carrier phase ambiguity resolution for ship attitude determination and dynamic draught,” in *Proceedings of the 24th FIG International Congress*, pp. 11–16, Sydney, Australia, 2010.
- [3] J. C. Juang and G.-S. Huang, “Development of GPS-based attitude determination algorithms,” *IEEE Transactions on Aerospace and Electronic Systems*, vol. 33, no. 3, pp. 968–976, 1997.
- [4] G. Giorgi, P. J. G. Teunissen, S. Verhagen, and P. J. Buist, “Testing a new multivariate GNSS carrier phase attitude determination method for remote sensing platforms,” *Advances in Space Research*, vol. 46, no. 2, pp. 118–129, 2010.
- [5] S. Yoon and J. B. Lundberg, “An integer ambiguity resolution algorithm for real-time GPS attitude determination,” *Applied Mathematics and Computation*, vol. 129, no. 1, pp. 21–41, 2002.
- [6] G. Xie, *Principles of GPS and Receiver Design*, Publishing House of Electronics Industry, Beijing, China, 2011.

- [7] P. J. G. Teunissen, "A canonical theory for short GPS baselines—part I: the baseline precision," *Journal of Geodesy*, vol. 71, no. 6, pp. 320–336, 1997.
- [8] J. Liu, *The Principle and Method of GPS Navigation*, Publication of Science, Beijing, China, 2008.

Research Article

An Adaptive Unscented Kalman Filtering Algorithm for MEMS/GPS Integrated Navigation Systems

Jianhua Cheng,^{1,2} Daidai Chen,¹ Rene Jr. Landry,² Lin Zhao,¹ and Dongxue Guan¹

¹ Marine Navigation Research Institute, College of Automation, Harbin Engineering University, Harbin 150001, China

² LASSENA Laboratoire, Ecole de Technologie Supérieure, Université du Québec, Montréal, Canada H3C 1K3

Correspondence should be addressed to Daidai Chen; ins.dai@163.com

Received 13 November 2013; Accepted 15 February 2014; Published 20 March 2014

Academic Editor: Yuxin Zhao

Copyright © 2014 Jianhua Cheng et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

MEMS/GPS integrated navigation system has been widely used for land-vehicle navigation. This system exhibits large errors because of its nonlinear model and uncertain noise statistic characteristics. Based on the principles of the adaptive Kalman filtering (AKF) and unscented Kalman filtering (AUKF) algorithms, an adaptive unscented Kalman filtering (AUKF) algorithm is proposed. By using noise statistic estimator, the uncertain noise characteristics could be online estimated to adaptively compensate the time-varying noise characteristics. Employing the adaptive filtering principle into UKF, the nonlinearity of system can be restrained. Simulations are conducted for MEMS/GPS integrated navigation system. The results show that the performance of estimation is improved by the AUKF approach compared with both conventional AKF and UKF.

1. Introduction

Microelectromechanical systems (MEMS) and Global Positioning System (GPS) integrated navigation system have the advantages of small size, light weight, and low cost, but, because of its low accuracy, it can only be applied in low accuracy navigation fields such as unmanned aircrafts and land-vehicles [1, 2]. There are three main factors impacting its performance: (a) the inaccuracy of model parameters, (b) the uncertainty of measurement and observation noise statistic properties, and (c) the nonlinearity of model [3, 4].

The classical Kalman filter (KF) provides a recursive solution for estimation of linear dynamic systems. The optimality of the KF algorithm is mainly dependent on a priori statistic of the process and measurement noise and the linear system model. However, if the priori information is insufficient or biased, the precision of the estimated states will be degraded, even leading to divergences [5]. In the case of MEMS/GPS applications, the estimation system tends to be nonlinear as well as variational noise properties [6]. Meanwhile, for the vehicle navigation, there are many sudden motion changes. To deal with these problems, the implement of AKF appears to be one of suitable approaches [7].

By utilizing the innovation and residual information, the AKF could adapt the filter stochastic properties online to accommodate itself to changes in vehicle dynamics. Thus, this technique could reduce the reliance of filter on the prior statistical information and obtain the noise statistic parameters of the dynamic system. The essence of AKF is to adapt the filter weights, so as to restrain model errors and improve the accuracy of filters. It is showed that applying AKF to the INS/GPS integrated navigation system could obtain better estimated performance than by using conventional KF, especially less than 20% root mean square error in attitude estimation [8]. However, AKF is unreliable when it is applied into the nonlinear applications.

UKF, which is another extension of Kalman filter, could give reliable estimates even if the nonlinearities of system are quite severe [9]. The linearization procedure is avoided by introducing the unscented transformation (UT), which is a method to approximate the joint distribution of states and measurements variables. In UT, a number of sigma points are chosen which could maintain the desired mean and covariance of states. Theoretically, the performance of UKF could be close to that of the three-order Taylor series expansion approximation or better than it [10]. However,

many kinds of systems such as aircrafts, ships, and spacecrafts would be intensively disturbed by external environment. As a result, the statistical properties for system process and measurement noises cannot be predicted, and UKF cannot solve these problems effectively.

As a combination of AKF and UKF, the adaptive UKF has been developed and applied to nonlinear joint estimation of both time-varying states and parameters [11]. The adaptive principles have been employed to update the means and covariances of the process and measurement noises online. The contributions of model predicted states and measurement information for filtering are balanced. In this paper, a new AUKF algorithm is proposed for MEMS/GPS integrated navigation systems in vehicle applications. A noise estimator for UKF is designed to estimate and update the means and covariances of noises online. Then, the updated means and covariances are propagating through the UKF. The proposed AUKF has the adaptive ability to time-varying noises and the noise estimates are unbiased. The performance of AUKF applied in land navigation is evaluated by simulations and the results show that the integrated system exhibits excellent robustness and navigation performance.

2. Unscented Kalman Filtering Algorithm for Nonzero Mean Noise

In the integrated navigation field, almost all of the systems are nonlinear. The general nonlinear discrete system model is given as

$$\begin{aligned} x_k &= f_{k-1}(x_{k-1}) + w_{k-1}, \\ z_k &= h_k(x_k) + v_k, \end{aligned} \quad (1)$$

where $x_k \in R^n$ is the state vector, $z_k \in R^m$ is the measurement vector, and $f_k(\cdot) \in R^{n \times n}$ and $h_k(\cdot) \in R^{m \times n}$ are the state and measurement matrices of nonlinear system, respectively. w_k and v_k are the Gaussian white noise which are unrelated. The mean and covariance of w_k and v_k are given as follows:

$$\begin{aligned} E[w_k] &= q, & \text{cov}[w_k w_k^T] &= Q\delta_{kj}; \\ E[v_k] &= r, & \text{cov}[v_k v_k^T] &= R\delta_{kj}, \end{aligned} \quad (2)$$

where q and r are nonzero constant variables and δ_{kj} is a Kronecker delta function.

The initial state x_0 is uncorrelated with both the process noise and measurement noise. These initial states exhibit Gaussian normal distributions. The prior mean and covariance matrices of x_0 are defined by

$$\begin{aligned} \hat{x}_0 &= E(x_0), \\ P_0 &= \text{cov}(x_0) = E(x_0 - \hat{x}_0)(x_0 - \hat{x}_0)^T. \end{aligned} \quad (3)$$

Assuming that $\mu_k = w_k - q$ and $\eta_k = v_k - r$ and substituting them into (1) yield

$$\begin{aligned} x_k &= f_{k-1}(x_{k-1}) + q + \mu_{k-1}, \\ z_k &= h_k(x_k) + r + \eta_k, \end{aligned} \quad (4)$$

where the mean and covariance of μ_k and η_k are given as follows:

$$\begin{aligned} E[\mu_k] &= 0, & \text{cov}[\mu_k \mu_k^T] &= Q\delta_{kj}; \\ E[\eta_k] &= 0, & \text{cov}[\eta_k \eta_k^T] &= R\delta_{kj}. \end{aligned} \quad (5)$$

According to system model of (4), the recursive solution of UKF algorithm is noted as follows.

(1) *Sigma Points Sampling.* In order to guarantee positive semidefinite of state covariance, the modified sigma points sampling solution based on the scaling method is adopted [12].

Choose $2n + 1$ sigma points $\xi_{i,k/k-1}$ as follows:

$$\begin{aligned} \xi_{0,k/k-1} &= \bar{x}, \\ \xi_{i,k/k-1} &= \bar{x} + (\alpha\sqrt{(n+\lambda)P})_i, \quad i = 1, 2, \dots, n, \\ \xi_{i,k/k-1} &= \bar{x} - (\alpha\sqrt{(n+\lambda)P})_{i-L}, \quad i = n+1, n+2, \dots, 2n, \end{aligned} \quad (6)$$

where P is the covariance of the state vector x , $\alpha\sqrt{(n+\lambda)P}$ is the matrix square root of nP , and $(\alpha\sqrt{(n+\lambda)P})_i$ denotes the i th row items of $\alpha\sqrt{(n+\lambda)P}$.

(2) *Prediction.* Propagating these sigma points $\xi_{i,k/k-1}$ through nonlinear state function $f_k(\cdot) + q$, we obtain [13]

$$\gamma_{i,k/k-1} = f_{k-1}(\xi_{i,k/k-1}) + q, \quad i = 0, 1, \dots, 2n. \quad (7)$$

Then, computing the predicted state $\hat{x}_{k/k-1}$, the predicted covariance $P_{k/k-1}$ is as follows:

$$\begin{aligned} \hat{x}_{k/k-1} &= \sum_{i=0}^L W_i^m \gamma_{i,k/k-1} = \sum_{i=0}^L W_i^m f_{k-1}(\xi_{i,k/k-1}) + q, \\ P_{k/k-1} &= \sum_{i=0}^L W_i^c (\gamma_{i,k/k-1} - \hat{x}_{k/k-1})(\gamma_{i,k/k-1} - \hat{x}_{k/k-1})^T + Q, \end{aligned} \quad (8)$$

$$(9)$$

where W_i^m and W_i^c are associated weights:

$$\begin{aligned} W_0^m &= \frac{\lambda}{(n+\lambda)}, \\ W_0^c &= \frac{\lambda}{(n+\lambda)} + (1 - \alpha^2 + \beta), \\ W_i^m &= \frac{1}{2(n+\lambda)}, \quad i = 1, 2, \dots, 2n, \\ W_i^c &= \frac{1}{2(n+\lambda)}, \quad i = 1, 2, \dots, 2n. \end{aligned} \quad (10)$$

Parameter λ is a scaling parameter, which is defined by

$$\lambda = \alpha^2(n + \kappa) - n. \quad (11)$$

The parameters α , β , and κ are the positive constants in the sampling method.

Then, propagating the sigma points $\xi_{i,k/k-1}$, the measurement function $h_k(\cdot) + r$ yields

$$x_{i,k/k-1} = h_k(\xi_{i,k/k-1}) + r, \quad i = 0, 1, \dots, 2n. \quad (12)$$

Computing the predicted measurement vector $\widehat{z}_{k/k-1}$, the covariance of the measurement P_{z_k} and the cross-covariance of the state and measurement $P_{\widehat{x}_k \widehat{z}_k}$ are as follows:

$$\widehat{z}_{k/k-1} = \sum_{i=0}^L W_i^m x_{i,k/k-1} = \sum_{i=0}^L W_i^m h_k(\xi_{i,k/k-1}) + r, \quad (13)$$

$$P_{z_k} = \sum_{i=0}^L W_i^c (x_{i,k/k-1} - \widehat{z}_{k/k-1})(x_{i,k/k-1} - \widehat{z}_{k/k-1})^T + R, \quad (14)$$

$$P_{\widehat{x}_k \widehat{z}_k} = \sum_{i=0}^L W_i^c (y_{i,k/k-1} - \widehat{x}_{k/k-1})(x_{i,k/k-1} - \widehat{z}_{k/k-1})^T.$$

(3) *Updating*. Then, computing the filter gain K_k , the updated state $\widehat{x}_{k/k}$ and covariance $P_{k/k}$ are as follows:

$$\begin{aligned} K_k &= P_{\widehat{x}_k \widehat{z}_k} (P_{z_k})^{-1}, \\ \widehat{x}_{k/k} &= \widehat{x}_{k/k-1} + K_k (z_k - \widehat{z}_{k/k-1}), \\ P_{k/k} &= P_{k/k-1} - K_k P_{z_k} (K_k)^T. \end{aligned} \quad (15)$$

3. Noise Statistic Estimator

Aiming at the uncertainty of process and measurement noise statistic properties, the measurement information are used to real time estimate and update the means and covariances of noises. Assume that w_k and v_k are uncorrelated and obey the Gaussian distributions. Based on the maximum a posterior (MAP) principle, a noise statistic estimator is derived.

The noise parameters q , r and the noise matrices Q , R are unknown and need to be estimated according to the updated measurements. The MAP estimates of q , Q , r , R and the state x_k are denoted as \widehat{q} , \widehat{Q} , \widehat{r} , \widehat{R} , and $\widehat{x}_{j/k}$, respectively. The conditional distribution of interest based on the measurements is expressed as

$$J = p[X_k, q, Q, r, R | Z_k]. \quad (16)$$

Because $p[Z_k]$ is disrelated to other parameters except the estimate state, the problem in calculating (16) changes to calculate the joint conditional probability distribution:

$$\begin{aligned} J &= p[X_k, q, Q, r, R, Z_k] \\ &= p[X_k | q, Q, r, R] p[Z_k | X_k, q, Q, r, R] p[q, Q, r, R], \end{aligned} \quad (17)$$

where $p[q, Q, r, R]$ can be treated as a constant which is provided by the prior statistic information.

The original problem has changed to calculate the conditional probability distributions $p[X_k | q, Q, r, R]$ and $p[Z_k | X_k, q, Q, r, R]$.

According to the Gaussian distributions of μ_k , the conditional probability distribution $p[X_k | q, Q, r, R]$ could be expressed as

$$\begin{aligned} p[X_k | q, Q, r, R] &= p[x_0] \prod_{j=1}^k p[x_j | x_{j-1}, q, Q] \\ &= \frac{1}{(2\pi)^{n/2} |P_0|^{1/2}} \exp\left\{-\frac{1}{2} \|x_0 - \widehat{x}_0\|_{P_0^{-1}}^2\right\} \\ &\quad \times \left\{ \prod_{j=1}^k \frac{1}{(2\pi)^{n/2} |Q|^{1/2}} \right. \\ &\quad \left. \times \exp\left\{-\frac{1}{2} \|x_j - f_{j-1}(x_{j-1}) - q\|_{Q^{-1}}^2\right\} \right\} \\ &= C_1 |P_0|^{-1/2} |Q|^{-k/2} \\ &\quad \times \exp\left\{-\frac{1}{2} \|x_0 - \widehat{x}_0\|_{P_0^{-1}}^2\right. \\ &\quad \left. + \sum_{j=1}^k \|x_j - f_{j-1}(x_{j-1}) - q\|_{Q^{-1}}^2\right\}, \end{aligned} \quad (18)$$

where n is the dimension of system state, $C_1 = 1/(2\pi)^{n(k+1)/2}$ is a constant, and $|A|$ is the determinant of A and $\|u\|_A^2 = u^T A u$.

Moreover, assuming that the measurements z_1, z_2, \dots, z_k are known and unrelated to each other, the distribution of η_k is Gaussian, which can be expressed as

$$\begin{aligned} p[Z_k | X_k, q, Q, r, R] &= \prod_{j=1}^k p[z_j | x_j, r, R] \\ &= \prod_{j=1}^k \frac{1}{(2\pi)^{m/2} |R|^{1/2}} \exp\left\{-\frac{1}{2} \|z_j - h_j(x_j) - r\|_{R^{-1}}^2\right\} \\ &= C_2 |R|^{-k/2} \exp\left\{-\frac{1}{2} \sum_{j=1}^k \|z_j - h_j(x_j) - r\|_{R^{-1}}^2\right\}, \end{aligned} \quad (19)$$

where m denotes the dimension of measurements and $C_2 = 1/(2\pi)^{mk/2}$ is a constant.

Substituting (18) and (19) into (17) yields

$$\begin{aligned}
 J &= C_1 C_2 |P_0|^{-1/2} |Q|^{-k/2} |R|^{-k/2} p[q, Q, r, R] \\
 &= C |Q|^{-k/2} |R|^{-k/2} \exp \left\{ -\frac{1}{2} \left[\sum_{j=1}^k \|x_j - f_{j-1}(x_{j-1}) - q\|_{Q^{-1}}^2 \right. \right. \\
 &\quad \left. \left. + \sum_{j=1}^k \|z_j - h_j(x_j) - r\|_{R^{-1}}^2 \right] \right\}, \quad (20)
 \end{aligned}$$

where

$$C = \exp \left\{ -\frac{1}{2} \|x_0 - \hat{x}_0\|_{P_0^{-1}}^2 \right\} C_1 C_2 |P_0|^{-1/2} p[q, Q, r, R]. \quad (21)$$

Logarithm on both sides of (20) yields

$$\begin{aligned}
 \ln J &= -\frac{k}{2} \ln |Q| - \frac{k}{2} \ln |R| - \frac{1}{2} \sum_{j=1}^k \|x_j - f_{j-1}(x_{j-1}) - q\|_{Q^{-1}}^2 \\
 &\quad - \frac{1}{2} \sum_{j=1}^k \|z_j - h_j(x_j) - r\|_{R^{-1}}^2 + \ln C. \quad (22)
 \end{aligned}$$

By the logarithmic nature, J and $\ln J$ share the same extreme points. The partial derivative of J can be calculated by the following equations:

$$\begin{aligned}
 \frac{\partial \ln J}{\partial q} \Big|_{q=\hat{q}_k}^{x_{j-1}=\hat{x}_{j-1/k}, x_j=\hat{x}_{j/k}} &= 0, \\
 \frac{\partial \ln J}{\partial Q} \Big|_{Q=\hat{Q}_k}^{x_{j-1}=\hat{x}_{j-1/k}, x_j=\hat{x}_{j/k}} &= 0, \\
 \frac{\partial \ln J}{\partial r} \Big|_{r=\hat{r}_k}^{x_j=\hat{x}_{j/k}} &= 0, \\
 \frac{\partial \ln J}{\partial R} \Big|_{R=\hat{R}_k}^{x_j=\hat{x}_{j/k}} &= 0.
 \end{aligned} \quad (23)$$

Then, the noise statistic estimator can be derived, which is defined by

$$\hat{q}_k = \frac{1}{k} \sum_{j=1}^k [\hat{x}_{j/k} - f_{j-1}(\hat{x}_{j-1/k})], \quad (24)$$

$$\begin{aligned}
 \hat{Q}_k &= \frac{1}{k} \sum_{j=1}^k \left\{ [\hat{x}_{j/k} - f_{j-1}(\hat{x}_{j-1/k}) - q] \right. \\
 &\quad \left. \times [\hat{x}_{j/k} - f_{j-1}(\hat{x}_{j-1/k}) - q]^T \right\}, \quad (25)
 \end{aligned}$$

$$\hat{r}_k = \frac{1}{k} \sum_{j=1}^k [z_j - h_j(\hat{x}_{j/k})], \quad (26)$$

$$\hat{R}_k = \frac{1}{k} \sum_{j=1}^k \left\{ [z_j - h_j(\hat{x}_{j/k}) - r] [z_j - h_j(\hat{x}_{j/k}) - r]^T \right\}. \quad (27)$$

In (24) to (27), the smoothed estimates $\hat{x}_{j-1/k}$ and $\hat{x}_{j/k}$ can be replaced by the filtered estimate $\hat{x}_{j/j}$ or the predicted state $\hat{x}_{j/j-1}$ as approximating solutions.

4. Noise Statistic Estimator for UKF

From the consideration for the nonlinear purposes, the noise statistic estimator derived above should be modified. In the linear applications, the term of $f_{j-1}(x_{j-1})$ can be obtained by propagating each estimate \hat{x}_{j-1} through system model and $h_j(x_j)$ by propagating \hat{x}_j through measurement equation. However, for the nonlinear field, the scaled sigma points are inserted instead of the estimate. The predicted term for UKF could be expressed as a combination of all sigma points:

$$f_{j-1}(x_{j-1}) \Big|_{x_{j-1} \leftarrow \hat{x}_{j-1/j-1}} = \sum_{i=0}^L W_i^m f_{j-1}(\xi_{i,j-1/j-1}), \quad (28)$$

where $f_{j-1}(x_{j-1})$ is approximated by UT with a precision close to that using three-order Taylor series expansion method [14].

Similarly, $h_j(x_j)$ can be calculated by

$$h_j(x_j) \Big|_{x_j \leftarrow \hat{x}_{j/j-1}} = \sum_{i=0}^L W_i^m h_j(\xi_{i,j/j-1}). \quad (29)$$

Submitting (28) into (24) to (27) yields the noise statistic estimator for UKF:

$$\hat{q}_k = \frac{1}{k} \sum_{j=1}^k \left[\hat{x}_{j/j} - \sum_{i=0}^L W_i^m f_{j-1}(\xi_{i,j-1/j-1}) \right], \quad (30)$$

$$\hat{Q}_k = \frac{1}{k} \sum_{j=1}^k \left\{ [\hat{x}_{j/j} - \hat{x}_{j/j-1}] [\hat{x}_{j/j} - \hat{x}_{j/j-1}]^T \right\}, \quad (31)$$

$$\hat{r}_k = \frac{1}{k} \sum_{j=1}^k \left[z_j - \sum_{i=0}^L W_i^m h_j(\xi_{i,j/j-1}) \right], \quad (32)$$

$$\hat{R}_k = \frac{1}{k} \sum_{j=1}^k \left\{ [z_j - \hat{z}_{j/j-1}] [z_j - \hat{z}_{j/j-1}]^T \right\}. \quad (33)$$

The unbiased properties of the noise estimates for UKF are proved in the appendix.

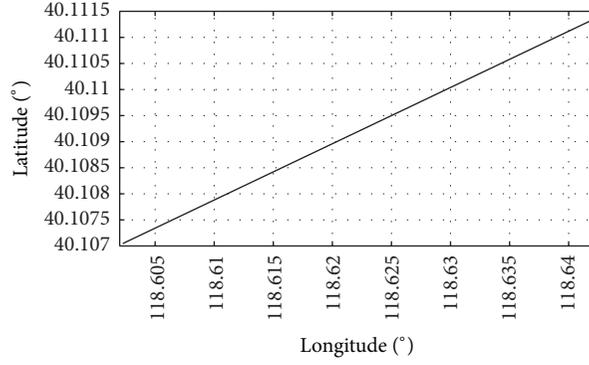


FIGURE 1: The track of land vehicle.

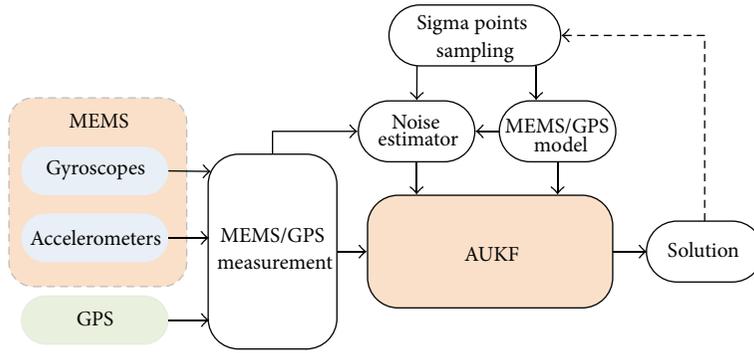


FIGURE 2: The architecture of MEMS/GPS integrated system with AUKF.

5. Recursive Equations of Adaptive Unscented Kalman Filtering Algorithm

Based on the UKF and its noise statistic estimator, the prediction and update steps of AUKF algorithm are as follows.

(1) *Prediction.* Propagating the sigma points $\xi_{i,k/k-1}$ through nonlinear state function $f_k(\cdot)$ yields

$$\gamma_{i,k/k-1} = f_{k-1}(\xi_{i,k/k-1}), \quad i = 0, 1, \dots, 2n. \quad (34)$$

Then, according to (30) and (31), estimate the process noise \hat{q}_k and covariance \hat{Q}_k , respectively.

With updated process noise parameters, compute the predicted state $\hat{x}_{k/k-1}$ and the predicted covariance $P_{k/k-1}$ as

$$\begin{aligned} \hat{x}_{k/k-1} &= \sum_{i=0}^L W_i^m \gamma_{i,k/k-1} = \sum_{i=0}^L W_i^m f_{k-1}(\xi_{i,k/k-1}) + \hat{q}_{k-1}, \\ P_{k/k-1} &= \sum_{i=0}^L W_i^m (\gamma_{i,k/k-1} - \hat{x}_{k/k-1}) \\ &\quad \times (\gamma_{i,k/k-1} - \hat{x}_{k/k-1})^T + \hat{Q}_{k-1}. \end{aligned} \quad (35)$$

(2) *Updating.* Propagating the sigma points $\xi_{i,k/k-1}$ through nonlinear state function $h_k(\cdot)$ yields

$$x_{i,k/k-1} = h_k(\xi_{i,k/k-1}), \quad i = 0, 1, \dots, 2n. \quad (36)$$

According to (32) and (33), estimate the measurement noise \hat{r}_k and its covariance \hat{R}_k , respectively.

With real-time measurement noise parameters, compute the predicted measurement $\hat{z}_{k/k-1}$ and the covariance $P_{\hat{z}_k}$ as

$$\begin{aligned} \hat{z}_{k/k-1} &= \sum_{i=0}^L W_i^m x_{i,k/k-1} = \sum_{i=0}^L W_i^m h_k(\xi_{i,k/k-1}) + \hat{r}_k, \\ P_{\hat{z}_k} &= \sum_{i=0}^L W_i^m (\chi_{i,k/k-1} - \hat{z}_{k/k-1}) (\chi_{i,k/k-1} - \hat{z}_{k/k-1})^T + \hat{R}_k. \end{aligned} \quad (37)$$

Estimate the cross-covariance $P_{\hat{x}_k \hat{z}_k}$ as

$$P_{\hat{x}_k \hat{z}_k} = \sum_{i=0}^L W_i^c (\gamma_{i,k/k-1} - \hat{x}_{k/k-1}) (\chi_{i,k/k-1} - \hat{z}_{k/k-1})^T. \quad (39)$$

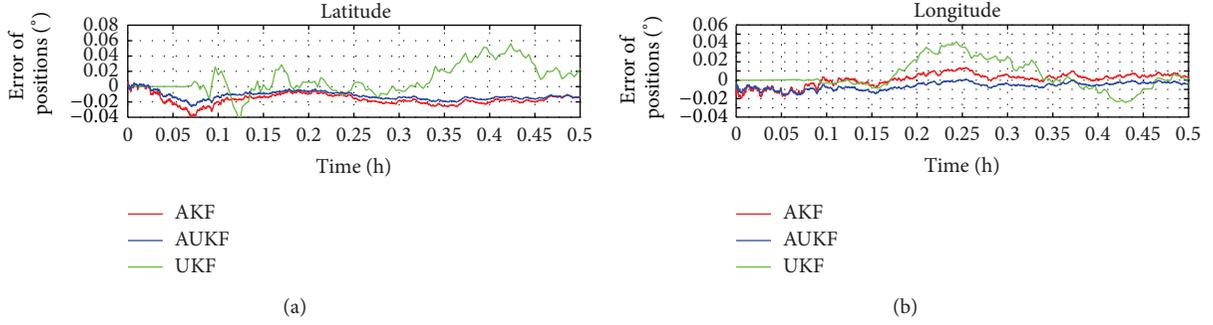


FIGURE 3: Comparison of position errors.

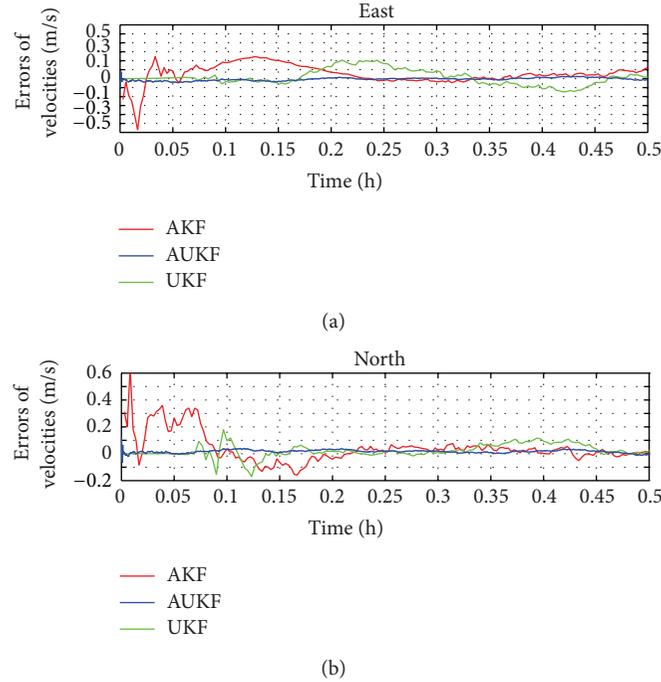


FIGURE 4: Comparison of velocity errors.

Then compute the filter gain K_k , the estimated state vector $\hat{x}_{k/k}$, and its covariance $P_{k/k}$:

$$\begin{aligned}
 K_k &= P_{\hat{x}_k \bar{z}_k} (P_{\bar{z}_k})^{-1}, \\
 \hat{x}_{k/k} &= \hat{x}_{k/k-1} + K_k (z_k - \hat{z}_{k/k-1}), \\
 P_{k/k} &= P_{k/k-1} - K_k P_{\bar{z}_k} (K_k)^T.
 \end{aligned} \tag{40}$$

6. MEMS/GPS Integrated Navigation for Land-Vehicle Using AUKF

Because of the highly nonlinear characteristic of MEMS/GPS, the conventional AKF based on small angle approximations is limited. Meanwhile, due to the time-varying noise stochastic properties for land-vehicle, the standard UKF in Section 1 cannot be directly applied to integrated navigation. On the other side, the modified AUKF based on a statistic estimator

in Section 4 appears appropriate for the MEMS/GPS integrated navigation. Simulations are conducted to compare the performances of AKF, UKF, and AUKF.

In the simulation, the parameters of sensor errors are shown in Table 1. The initial position of vehicle is east longitude 126° and north latitude 45° .

Figure 1 shows the trajectory of land-vehicle motion. The solid line in this figure illustrates the real simulated trajectory.

The architecture of MEMS/GPS integrated navigation with AUKF is shown in Figure 2.

As shown in Figures 3, 4, and 5, with the comparisons with AKF and UKF, the AUKF with noise estimator could obviously improve the accuracy of the velocity solutions. In addition, as shown in Figure 3, because the noise statistic estimators are designed in AKF and AUKF, the position errors of the two filters exhibit similar characteristics at most of the time in this simulation, and the performance of AUKF is slightly better. However, which is also seen in Figure 3, there

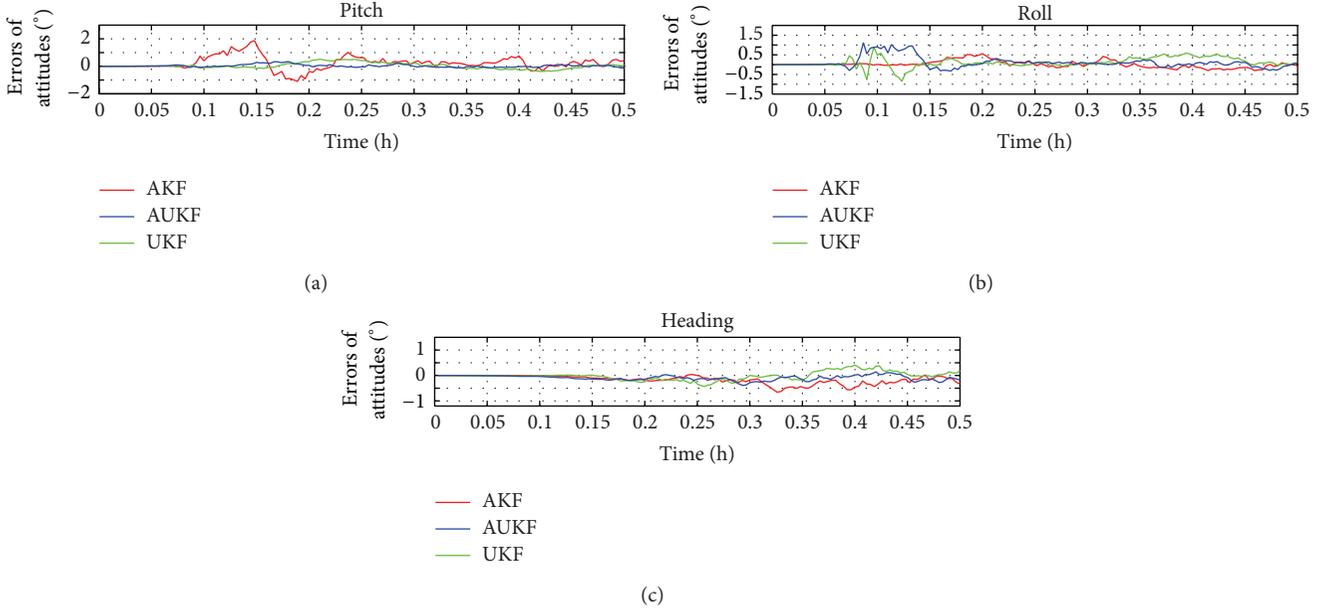


FIGURE 5: Comparison of attitude errors.

TABLE 1: Parameters of sensor errors.

Sensor	Characteristic	Value
MEMS Gyro	Drift	5°/h
MEMS Gyro	Measuring white noise	0.5°/h
MEMS accelerometer	Bias	2 mg
MEMS accelerometer	Measuring white noise	50 μg
GPS (position)	Measuring white noise	6 m
GPS (velocity)	Measuring white noise	0.01 m/s

are so notable vibrations for UKF in which stable estimate errors of positions could not be provided, which are mainly caused by the quickly changed system noises.

Figure 4 shows that the AUKF always has smaller velocity errors than AKF and UKF. And Figure 5 shows that at most of the time the AUKF scheme has a better performance on attitudes errors. The AKF and UKF solutions have large vibration errors in both Figures 4 and 5. That is because the velocity errors are related to the attitude errors, especially to the pitch and roll errors of AKF and UKF. Because of the strong nonlinear properties of system, it is difficult that AKF cannot be effectively operated for navigation. Meanwhile, for UKF solutions, the curves of speed errors are extremely similar to those of horizontal attitude errors.

The simulation results indicate that if the AUKF scheme is considered the filtering solution, there are only small variations impacting on the performance of MEMS/GPS integrated navigation system, and this system has an excellent robustness. However, long processing time would cause slight divergence of the attitude errors, sequentially the velocity and position errors.

7. Conclusions

This study has developed an AUKF approach to improve the navigation performance of MEMS/GPS integrated system for land-vehicle applications. By treating this problem within conventional UKF framework, the noise estimator is adopted and could effectively estimate the process and measurement noise characteristics online. The results indicate that the proposed AUKF algorithm could efficiently improve the navigation performance of land-vehicle integrated navigation system. By comparing with AKF and UKF methods, the AUKF solution has a more stable and superior performance.

Appendix

The process noise and measurement noise statistic properties are computed by estimator in (30) to (33). Their unbiased properties are proved as follows.

According to (8), the predicted state at time step j is given as

$$\hat{x}_{j/j-1} = \sum_{i=0}^L W_i^m f_{j-1}(\xi_{i,j-1/j-1}) + q. \quad (\text{A.1})$$

Hence the mean of the process noise can be expressed as

$$\begin{aligned} E[\hat{q}_k] &= \frac{1}{k} \sum_{j=1}^k \left[\hat{x}_{j/j} - \sum_{i=0}^L W_i^m f_{j-1}(\xi_{i,j-1/j-1}) \right] \\ &= \frac{1}{k} \sum_{j=1}^k [\hat{x}_{j/j} - \hat{x}_{j/j-1} + q]. \end{aligned} \quad (\text{A.2})$$

From (15), we have

$$\hat{x}_{j/j} = \hat{x}_{j/j-1} + K_j (z_j - \hat{z}_{j/j-1}). \quad (\text{A.3})$$

Substituting (A.3) into (A.2) yields

$$\begin{aligned} E[\hat{q}_k] &= \frac{1}{k} \sum_{j=1}^k [\hat{x}_{j/j} - \hat{x}_{j/j-1} + q] \\ &= \frac{1}{k} \sum_{j=1}^k [K_j (z_j - \hat{z}_{j/j-1}) + q]. \end{aligned} \quad (\text{A.4})$$

When the posteriori mean and covariance are known, the output residual vector of UKF is zero-mean Gaussian white noise and we see

$$E[z_j - \hat{z}_{j/j-1}] = 0. \quad (\text{A.5})$$

From (A.4) and (A.5), the mean of the process noise is

$$E[\hat{q}_k] = \frac{1}{k} \sum_{j=1}^k E[(z_j - \hat{z}_{j/j-1}) + q] = q. \quad (\text{A.6})$$

Thus, the estimate of the process noise noted by (24) is unbiased.

Similarly, the estimate of the measurement noise is

$$E[\hat{r}_k] = r. \quad (\text{A.7})$$

The unbiased properties for the estimates of noises are proved.

Conflict of Interests

The authors declare that they have no conflict of interests regarding the publication of this paper.

Acknowledgments

Funding for this work was provided by the National Nature Science Foundation of China under Grant nos. 61374007 and 61104036. The authors would like to thank all the editors and anonymous reviewers for improving this paper.

References

- [1] S. Mizukami, M. Sugiura, Y. Muramatsu, and H. Kumagai, "MEMS GPS/INS for micro air vehicle application," in *Proceedings of the 17th International Technical Meeting of the Satellite Division of the Institute of Navigation (ION GNSS '04)*, pp. 819–824, September 2004.
- [2] X. Niu, S. Nassar, and N. El-Sheimy, "An accurate land-vehicle MEMS IMU/GPS navigation system using 3D auxiliary velocity updates," *Navigation, Journal of the Institute of Navigation*, vol. 54, no. 3, pp. 177–188, 2007.
- [3] S. Y. Cho and B. D. Kim, "Adaptive IIR/FIR fusion filter and its application to the INS/GPS integrated system," *Automatica*, vol. 44, no. 8, pp. 2040–2047, 2008.
- [4] S. Y. Cho and W. S. Choi, "Robust positioning technique in low-cost DR/GPS for land navigation," *IEEE Transactions on Instrumentation and Measurement*, vol. 55, no. 4, pp. 1132–1142, 2006.
- [5] A. Gelb, *Applied Optimal Estimation*, The MIT press, 1974.
- [6] C. Hide, T. Moore, and M. Smith, "Adaptive Kalman filtering for low-cost INS/GPS," *Journal of Navigation*, vol. 56, no. 1, pp. 143–152, 2003.
- [7] C. W. Hu, Congwei, W. Chen, Y. Chen, and D. Liu, "Adaptive Kalman filtering for vehicle navigation," *Journal of Global Positioning Systems*, vol. 2, no. 1, pp. 42–47, 2003.
- [8] A. H. Mohamed and K. P. Schwarz, "Adaptive Kalman filtering for INS/GPS," *Journal of Geodesy*, vol. 73, no. 4, pp. 193–203, 1999.
- [9] D. Simon, *Optimal State Estimation: Kalman, H Infinity, and Nonlinear Approaches*, Wiley, 2006.
- [10] S. J. Julier and K. U. Jeffrey, "A general method for approximating nonlinear transformations of probability distributions," Tech. Rep., Robotics Research Group, University of Oxford, Oxford, UK, 1996.
- [11] Z. Jiang, Q. Song, Y. He, and J. Han, "A novel adaptive unscented kalman filter for nonlinear estimation," in *Proceedings of the 46th IEEE Conference on Decision and Control (CDC '07)*, pp. 4293–4298, New Orleans, LA, USA, December 2007.
- [12] S. J. Julier, "The scaled unscented transformation," in *Proceedings of the American Control Conference*, pp. 4555–4559, May 2002.
- [13] S. Särkkä, "On unscented Kalman filtering for state estimation of continuous-time nonlinear systems," *IEEE Transactions on Automatic Control*, vol. 52, no. 9, pp. 1631–1641, 2007.
- [14] L. Zhao, X.-X. Wang, H.-X. Xue, and Q.-X. Xia, "Design of unscented Kalman filter with noise statistic estimator," *Control and Decision*, vol. 24, no. 10, pp. 1483–1488, 2009.

Research Article

Dynamic Surface and Active Disturbance Rejection Control for Path Following of an Underactuated UUV

Juan Li, Haitao Gao, Jiajia Zhou, and Zheping Yan

College of Automation, Harbin Engineering University, Harbin 150001, China

Correspondence should be addressed to Jiajia Zhou; zhoujiajia@hrbeu.edu.cn

Received 20 December 2013; Accepted 6 February 2014; Published 19 March 2014

Academic Editor: Weichao Sun

Copyright © 2014 Juan Li et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

This paper addresses the problem of accurate path following control for an underactuated unmanned underwater vehicle (UUV) in the horizontal plane. For an underactuated UUV, the line-of-sight (LOS) guidance method is adopted to map 2D reference trajectory into a desired orientation, and through the tracking of heading to achieve path following, where the sideslip is introduced to modify the desired orientation. In this paper, we propose a method called dynamic surface and active disturbance rejection control (DS-ADRC) to solve the path following control problem. This controller can effectively avoid the phenomenon of explosion of terms in the conventional backstepping method, reduce the dependence on the UUV controller mathematical model, and enhance the antijamming ability. Simulation is carried out to verify the effectiveness of the proposed control method for an underactuated UUV. The results show that, even for this controller with disturbance, the cross-track error of UUV is gradually converged to zero and has some certain robustness.

1. Introduction

The high accuracy path following mission is a typical behavior of UUVs, and it is an important method for UUVs to complete other tasks (such as topography examination and long distance navigation) [1]. Under the influences of some uncertain disturbance such as the ocean current, the UUV's movements are very complicated with six degrees of freedom (DOF).

At present, the research on UUV's tracking control mainly focuses on the following three aspects: way-point tracking [2–4], path following [5], and trajectory tracking [6]. For the path following, many scholars at home and abroad have further studies and proposed some mature control technologies in those fields of unmanned aerial vehicle (UAV), unmanned boat, the surface ship, the autonomous underwater vehicle (AUV), and so on. The path following of UUVs is usually simplified as the horizontal and vertical planes, and the controllers are designed, respectively. This paper mainly studies the problem of horizontal path following. Lapierre and Soetanto [7] designed an AUV tracking controller based on the Lyapunov method and the backstepping technique, and the AUV tracking error is gradually converged to zero.

But it did not take the disturbance of the marine environment into account. Shi et al. [8] used the same method as Lapierre and Soetanto [7] and ensured the tracking error in the global asymptotic stability under the influence of the constant current. But there is a phenomenon called “explosion of terms” which may exist when we use the backstepping technique to obtain the repeated derivative of virtual control variation. Swaroop et al. [9] proposed a dynamic surface control (DSC) strategy for a class of nonlinear systems, which let the virtual control pass through a low-pass filter, where a new value was obtained to calculate the approximate derivation. This technique can avoid the phenomenon of “explosion of terms.” However, even a simple dynamic surface controller design requires the accurate mathematical model of controlled object. Li et al. [10] used the active disturbance rejection control (ADRC) technology to design a path following controller for an underactuated surface ship. It could deal with the problem of the transverse drift of the ship caused by the constant wind disturbance, but the path following responds slowly with the use of a linear PD control rate.

Usually, the accurate mathematical model of an UUV is difficult to obtain. Even with a precise mathematical model, it is also so complex that it needs to do some simplification

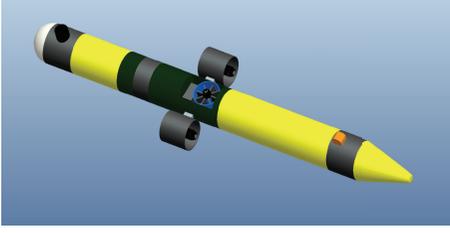


FIGURE 1: Minesniper MkII.

when doing control system design, which eventually leads to model error. Based on the DS-ADRC technique, this paper designs a horizontal path following controller for an underactuated UUV, which effectively avoids the “explosion of terms” phenomenon when using backstepping method and also reduces the high requirements of the dynamic surface control technology to the accurate mathematical model. It does real-time estimation to the internal and external disturbance in the loop and eventually compensates the estimation into the control system, which improves the control accuracy and has high stability.

2. Mathematical Models for MKII

This paper selects Minesniper MkII as a simulation object, which is shown in Figure 1 and based on [11]. The length of the UUV is 1.929 m and weight is 40 Kg. There are propellers equipped, respectively, on the left and right sides of the UUV, whose maximum speed is 2000 rpm, providing maximum thrust between ± 80 N. Through the left and right speed difference, it achieves the steering movement. Between the two longitudinal propellers, there is a vertical propeller, whose maximum speed is 100 rpm providing a maximum thrust for ± 20 N.

Here, we assume that the position vector of the UUV is $[x, y, \psi]^T$, velocity vector of the UUV is $[u, v, r]^T$ and the ocean current speed is $[u_c, v_c, 0]^T$, the UUV's relative speed is $[u_r, v_r, r]^T$, and the controlling force and moment are $[\tau_u, 0, \tau_r]^T$, and, by means of simplifying the 6-DOF model of the UUV, we get its horizontal mathematical model with ocean current as follows:

$$\begin{aligned}
 \dot{u}_r &= \frac{(-d_{11}u_r + \tau_u)}{m_{11}}, \\
 \dot{v}_r &= \frac{(Am_{66} - Bm_{26})}{(m_{22}m_{66} - m_{26}^2)}, \\
 \dot{r} &= \frac{(Bm_{22} - Am_{26})}{(m_{22}m_{66} - m_{26}^2)}, \\
 \dot{x} &= u \cos(\psi) - v \sin(\psi), \\
 \dot{y} &= u \sin(\psi) + v \cos(\psi), \\
 \dot{\psi} &= r,
 \end{aligned} \tag{1}$$

TABLE 1: UUV relevant parameters.

$m = 40$	$N_{\dot{v}} = 2.2$	$N_v = 36$
$X_{\dot{u}} = -1.42$	$X_{u_r} = 0.1$	$X_{ u_r u_r} = 8.2$
$Y_{\dot{v}} = -38.4$	$Y_{v_r} = 10$	$Y_{ v_r v_r} = 200$
$Y_{\dot{r}} = -2.5$	$Y_r = 5$	$I_z = 8.0$
$N_{\dot{r}} = -8.9$	$N_r = 5$	$N_{ r r} = 15$

where $A = -d_{22}v_r + (d_{26} - u_r c_{26} - mu_c)r$, $B = (d_{62} - u_r c_{62})v_r - d_{66}r + \tau_r$, and

$$\begin{aligned}
 m_{11} &= m - X_{\dot{u}}, & d_{11} &= X_{u_r} + X_{|u_r|u_r} |u_r|, \\
 m_{22} &= m - Y_{\dot{v}}, & d_{22} &= Y_{v_r} + Y_{|v_r|v_r} |v_r|, \\
 m_{26} &= -Y_{\dot{r}}, & d_{26} &= Y_r, \\
 m_{66} &= I_z - N_{\dot{r}}, & d_{62} &= N_{v_r}, \\
 c_{26} &= m - X_{\dot{u}}, & d_{66} &= N_r + N_{|r|r} |r|, \\
 c_{62} &= X_{\dot{u}} - Y_{\dot{v}}.
 \end{aligned} \tag{2}$$

According to [11, 12], the relevant parameters and hydrodynamic coefficients in the above equations are shown in Table 1.

3. The Line-of-Sight Guidance System

Considering the underactuated UUV in this paper, we select the LOS guidance as adopted in [11–13], which converts the two-dimensional desired location to the expected heading angle. And through tracking the expected heading angle, it achieves the path following control of the vehicle.

For the horizontal path following, it can divide the expecting path into a series of points: p_k and $k = 0, 1, \dots, n$, where $p_k = (x_k, y_k) \in R^2$. Taking the connected two points p_{k-1}, p_k on the expecting path and taking the p_{k-1} as original points to establish north-east coordination system, these have been shown in Figure 2. β_i is the included angle of the directed line segment $p_{k-1}p_k$ and the north coordinate axis, (x_t, y_t) is the real-time location coordinates of the UUV, $\psi(t)$ is the real-time heading angle of the UUV, $d(t)$ is the distance between location and the endpoint of the path, Δ is the selecting foresight vector quantity (Δ generally is selected 2–6 times long of the UUV [14]), and $\varepsilon(t)$ is the lateral error of path tracking made by the LOS guidance system. $\delta(t)$ is the included angle of current location to path endpoint ligature and the path of the UUV, $p_{\text{los}}(x_{\text{los}}, y_{\text{los}})$ is the current foresight point, ψ_d is the expecting angle, $\alpha(t)$ is the included angle of foresight vector quantity and the expected path, and R_{accept} is the switching condition of the two expected straight line segments.

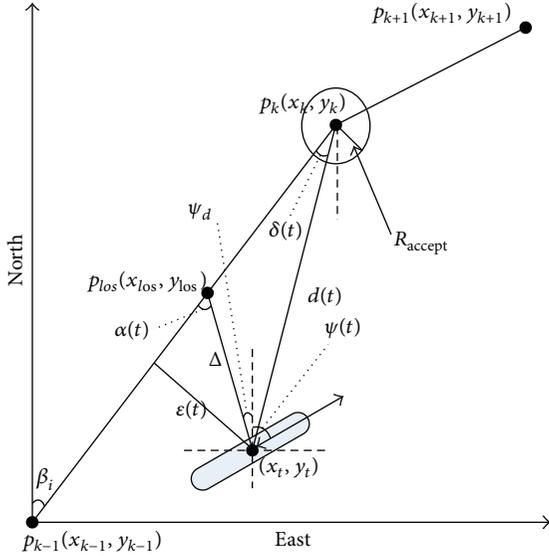


FIGURE 2: Line-of-sight-guidance.

Through the geometric relation, we can obtain

$$\begin{aligned}\beta_i &= a \tan 2(y_k - y_{k-1}, x_k - x_{k-1}), \\ \delta(t) &= \beta_i - a \tan 2(y_k - y_t, x_k - x_t), \\ d(t) &= \sqrt{(x_k - x_t)^2 + (y_k - y_t)^2}, \\ \varepsilon(t) &= d(t) * \sin(\delta(t)), \\ \psi_d &= \beta_i - \alpha(t).\end{aligned}\quad (3)$$

For the selection of $\alpha(t)$, there are some certain rules. When the current position of the UUV is far away from the desired path, that is, $\varepsilon(t) > \Delta$, there are no interaction points between foresight vector and path at this time; thus there is $|\alpha(t)| = \pi/2$. When the current position of the UUV is near to the desired path, we can select $\alpha(t) = a \sin(\varepsilon(t)/\Delta)$. So we can obtain formula (4) as follows:

$$\alpha(t) = \begin{cases} a \sin\left(\frac{\varepsilon(t)}{\Delta}\right), & |\varepsilon(t)| \leq \Delta, \\ \frac{\pi}{2} * \text{sign}(\varepsilon(t)), & \text{others.} \end{cases}\quad (4)$$

There are currents in the case of interference; if we get the desired heading angle as desired input for the UUV control system with the above method, eventually the heading tracking is achieved, but there will be a stable track error. In this paper, in order to eliminate it, we introduce a sideslip angle $\beta = a \tan(v/u)$, the specific expression, as follows:

$$\psi_d = \beta_i - \alpha(t) - \beta.\quad (5)$$

When $d(t) < R_{\text{accept}}$ (R_{accept} is the allowable maximum error), it can select straightway $p_k p_{k+1}$ as desired path.

4. Controller Design

The horizontal movement of the UUV can be divided into two subaspects as follows [15]: one aspect is the geometrical problem, that is, controlling the position of the UUV to the expecting path, and the other one is the dynamics aspect, that is, controlling the UUV's longitudinal velocity to an expecting one. As for the former, by using the LOS guidance system, it maps the expected position instructions to the expected heading angle instructions; thus we can achieve the UUV's path following. For the latter, it is mainly to control the longitudinal propellers.

The dynamic surface control (DSC) technique has efficiently avoided the "explosion of terms" phenomenon which is caused by repeatedly instructing on the virtual controlling, but it needs the precise model of the controlled object. The active disturbance rejection control (ADRC) technique does not depend on the precise model of the controlled object, but the feedback efficiency is not high enough, and its control signal may easily have high frequency oscillation. With the help of DSC and ADRC, this paper designs a longitudinal velocity controller and a heading controller, respectively, by using the DS-ADRC method.

4.1. Model Transformation. In order to use the ADRC technique more conveniently, it firstly converts the UUV mathematical model to a standard form according to the ADRC, and then the controllers are designed, respectively.

4.1.1. Model for Heading Control. From the model of (1), it can obtain the UUV heading control mathematical model as follows:

$$\begin{aligned}\dot{\psi} &= r, \\ \dot{r} &= \frac{(Bm_{22} - Am_{26})}{(m_{22}m_{66} - m_{26}^2)}.\end{aligned}\quad (6)$$

In the above model, ψ is heading angle, r is turning heading angular velocity, $A = -d_{22}v_r + (d_{26} - u_r c_{26} - mu_c)r$, $B = (d_{62} - u_r c_{62})v_r - d_{66}r + \tau_r$, $m_{22} = m - Y_{\dot{v}}$, $m_{26} = -Y_{\dot{r}}$, and $m_{66} = I_z - N_{\dot{r}}$; the specific parameters are all given in Table 1.

Assume $x_{11} = \psi$, $x_{12} = r$, and $u_1 = \tau_r$, so that we can simplify the above system as to the standard ADRC form as follows:

$$\begin{aligned}\dot{x}_{11} &= x_{12}, \\ \dot{x}_{12} &= f_1(\cdot) + b_{10}u_1, \\ y_1 &= x_{11}.\end{aligned}\quad (7)$$

Hereinto, u_1 , y_1 are the system input and output, that is, turning heading torque and heading angle, respectively, $f_1(\cdot)$ is the system nonlinear part, and b_{10} is the input gain coefficient.

Taking $u_{10} = f_1(\cdot) + b_{10}u_1$, so the above system can be converted into

$$\begin{aligned} \dot{x}_{11} &= x_{12}, \\ \dot{x}_{12} &= u_{10}, \\ y_1 &= x_{11}. \end{aligned} \quad (8)$$

4.1.2. Model for the Longitudinal Velocity Control. In the same way, from (1), we can obtain the longitudinal velocity controlling mathematical model as follows:

$$\dot{u}_r = \frac{(-d_{11}u_r + \tau_u)}{m_{11}}. \quad (9)$$

Here u_r is the UUV's relative velocity, τ_u is the longitudinal thrust, and d_{11} and $m_{11} = m - X_{\dot{u}}$ are the hydrodynamic coefficients. And the specific numerical values have been given in Table 1.

Supposing that $x_{21} = u_r$, so the above longitudinal velocity control model can be converted into the standard ADRC form as follows:

$$\begin{aligned} \dot{x}_{21} &= f_2(\cdot) + b_{20}u_2, \\ y_2 &= x_{21}. \end{aligned} \quad (10)$$

Hereinto, u_2 , y_2 are the system input and output, that is, the UUV's longitudinal thrust and the longitudinal relative velocity, respectively, $f_2(\cdot)$ is the nonlinear part, and $b_{20} = 1/m_{11}$, $f_2(\cdot) = -d_{11}u/m_{11}$. The specific parameters are shown in Table 1.

Taking $u_{20} = f_2(\cdot) + b_{20}u_2$, so the above system can be converted into

$$\begin{aligned} \dot{x}_{21} &= u_{20}, \\ y_2 &= x_{21}. \end{aligned} \quad (11)$$

4.2. DSC-ADRC Controller. The ADRC controller is made up of four parts [16], including arranging transition process, extended state observer (ESO) design, nonlinear feedback law design, and dynamic compensation (DC). Compared with ADRC, the DS-ADRC controller replaces the nonlinear feedback law design in ADRC with DSC; thus the specific design process is stated as follows.

Step 1. Arrange the transition process of expected signal using the tracking-differentiator (TD). And calculate the signal v_1 and its differential v_2 from the expected signal.

Step 2. Estimate the system total disturbance through the ESO in real time.

Step 3. Calculate the part of control input u_0 based on DS-ADRC.

Step 4. Compensate the estimation using ESO to u_0 . Then we can obtain the actual control input u for the UUV.

The schematic of the controller is shown in Figure 3.

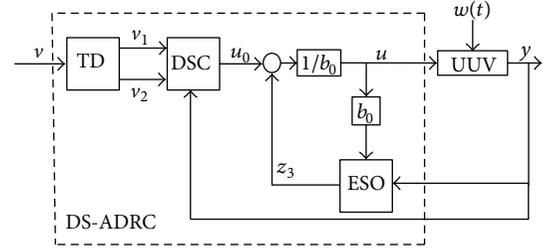


FIGURE 3: Design of the DS-ADRC controller.

4.2.1. Heading Controller Design

(1) TD Design. Based on the expected heading v_ψ , we arrange the transition process v_{11} and calculate its differential signal v_{12} at the same time. Thus its discrete model is written as follows:

$$\begin{aligned} fh &= fhan(v_{11}(k) - v_\psi(k), v_{12}(k), r, h), \\ v_{11}(k+1) &= v_{11}(k) + h \cdot v_{12}(k), \\ v_{12}(k+1) &= v_{12}(k) + h \cdot fh, \end{aligned} \quad (12)$$

where r is the regulation factor, h is the simulation step size, and $fhan(x_1, x_2, r, h)$ is the time-optimal feedback function whose specific algorithm is described as follows:

$$\begin{aligned} d &= rh, \\ d_0 &= hd, \\ y &= x_1 + hx_2, \\ a_0 &= \sqrt{d^2 + 8r|y|}, \\ a &= \begin{cases} x_2 + \frac{(a_0 - d)}{2} \text{sign}(y), & |y| > d_0, \\ x_2 + \frac{y}{h}, & |y| \leq d_0, \end{cases} \\ fhan &= \begin{cases} -r \text{sign}(a), & |a| > d, \\ -r \frac{a}{d}, & |a| \leq d. \end{cases} \end{aligned} \quad (13)$$

(2) ESO Design. The function of the ESO makes use of the vehicle's input and output to estimate the system's related states, including the total disturbance which acts on the system. For the heading control, the system input is the turning heading moment and its output is the heading angle; that is, $u_1 = \tau_r$ and $y_1 = \psi$, respectively. z_{11} , z_{12} , and z_{13} are the estimated values by the observer, which are induced by the heading angle, turning heading angular velocity, and

the system total disturbance, respectively. And its discrete algorithm is written as

$$\begin{aligned}
 e(k) &= z_{11}(k) - y_1(k), \\
 fe &= fal(e(k), 0.5, \delta), \\
 fe_1 &= fal(e(k), 0.25, \delta), \\
 z_{11}(k+1) &= z_{11}(k) + h(z_{12}(k) - \beta_{11}e(k)), \\
 z_{12}(k+1) &= z_{12}(k) + h(z_{13}(k) - \beta_{12}fe + b_{10}u_1(k)), \\
 z_{13}(k+1) &= z_{13}(k) + h(-\beta_{13}fe_1),
 \end{aligned} \tag{14}$$

where β_{11} , β_{12} , and β_{13} are the parameters which need to be designed and $fal(e, \alpha, \delta)$ is a continuous exponential function with linear segment near the original point, whose specific algorithm is stated as follows:

$$fal(e, \alpha, \delta) = \begin{cases} \frac{e}{\delta^{1-\alpha}}, & |e| \leq \delta, \\ |e|^\alpha \text{sign}(e), & |e| > \delta. \end{cases} \tag{15}$$

Hereinto, δ is the length of the linear segment.

For the stability analysis of ESO, please refer to [16].

(3) *DSC Design.* From the converted heading control model (8), it is shown that the state of the UUV is x_{1i} and the expected state is v_{1i} , calculated by TD.

Define the first dynamic surface as s_{11} as follows:

$$s_{11} = x_{11} - v_{11}. \tag{16}$$

Its first derivative is

$$\dot{s}_{11} = x_{12} - v_{12}. \tag{17}$$

Consider x_{12} as a virtual control and design a virtual stabilization function $\bar{\alpha}_{11}$:

$$\bar{\alpha}_{11} = v_{12} - k_{11}s_{11}. \tag{18}$$

where k_{11} is a positive parameter. Based on the dynamic state design method, a first-order low-pass filter can be introduced as follows:

$$\tau_{11}\dot{\alpha}_{11} + \alpha_{11} = \bar{\alpha}_{11}, \quad \alpha_{11}(0) = \bar{\alpha}_{11}(0), \tag{19}$$

where τ_{11} is the time constant of filter.

Define the second dynamic surface as s_{12} :

$$s_{12} = x_{12} - \alpha_{11}. \tag{20}$$

Its first derivative is

$$\dot{s}_{12} = \dot{x}_{12} - \dot{\alpha}_{11}. \tag{21}$$

In order to stabilize the system, we choose a variable x_{12} which satisfies

$$\dot{x}_{12} = \dot{\alpha}_{11} - k_{12}s_{12}. \tag{22}$$

Then (23) can be obtained as follows:

$$u_{10} = \dot{\alpha}_{11} - k_{12}s_{12}. \tag{23}$$

k_{12} is also a positive parameter.

The stabilization analysis of the DSC algorithm is discussed as follows. Defining $y = \alpha_{11} - \bar{\alpha}_{11}$, then there is $y = -\tau_{11}\dot{\alpha}_{11}$. Combining (17)–(22) we obtain (24) and (25), as follows:

$$\dot{s}_{11} = s_{12} + y - k_{11}s_{11}, \tag{24}$$

$$\dot{s}_{12} = -k_{12}s_{12}. \tag{25}$$

Then (26) can be obtained as follows:

$$\dot{y} = -\frac{y}{\tau_{11}} + k_{11}\dot{s}_{11} - \dot{v}_{12}. \tag{26}$$

Herein, we define $V_{1s} = s_{11}^2/2$, $V_{2s} = s_{12}^2/2$, and $V_y = y^2/2$. The derivative of V_{1s} can be calculated as

$$\begin{aligned}
 \dot{V}_{1s} &= s_{11}(s_{12} + y - k_{11}s_{11}), \\
 \dot{V}_{2s} &= -k_{12}s_{12}^2,
 \end{aligned} \tag{27}$$

$$\dot{V}_y = -\frac{y^2}{\tau_{11}} + k_{11}y\dot{s}_{11} - y\dot{v}_{12}.$$

Define a Lyapunov function as follows:

$$V = V_{1s} + V_{2s} + V_y. \tag{28}$$

Based on the assumption, v_{11} , v_{12} , and \dot{v}_{12} are all continuous variables. Now assuming

$$\begin{aligned}
 s_{11}^2 + y^2 + s_{12}^2 &\leq 2p, \quad \forall p < 0, \\
 k_{11}\dot{s}_{11} - \dot{v}_{12} &\leq q, \\
 k_{1i} &= 2 + a, \quad i = 1, 2,
 \end{aligned} \tag{29}$$

$$\tau_{11} = 1 + \left(\frac{q^2}{2\varepsilon}\right) + a,$$

then

$$\begin{aligned}
 \dot{V} &\leq -(2+a)(s_{11}^2 + s_{12}^2) \\
 &+ \left[\frac{(2s_{11}^2 + s_{12}^2 + y^2)}{2} + \left(1 + \frac{q^2}{2\varepsilon} + a\right)y^2 \right. \\
 &\quad \left. + \frac{q^2 y^2}{2\varepsilon} \right] + \frac{\varepsilon}{2} \\
 &\leq -2aV + \frac{\varepsilon}{2}.
 \end{aligned} \tag{30}$$

When $V = p$, which satisfies $a > \varepsilon/2p$, it can be concluded that $\dot{V} < 0$.

Through the above analysis, we can know that the DSC algorithm can guarantee all the states of the closed-loop

system final convergence. And with appropriate coefficients k_{1i} and τ_{1i} , it can get a good controlling quality.

(4) *DC Design.* Using the estimated value z_{13} to compensate u_{10} , which obtained from the DSC. In this way, the final control variable obtained as follows:

$$u_1 = u_{10} - \frac{z_{13}}{b_{10}} \quad \text{or} \quad u_1 = \frac{(u_{10} - z_{13})}{b_{10}}, \quad (31)$$

where the parameter b_{10} is an adjustable compensating factor which decides the compensation degree.

Through the above four steps, we can finally obtain the controller as follows:

$$\begin{aligned} e &= v_{11} - v_\psi, \\ fh &= fhan(e, v_{12}, r, h), \\ v_{11} &= v_{11} + hv_{12}, \\ v_{12} &= v_{12} + h \cdot fh, \\ e_1 &= z_{11} - y_1, \\ fe &= fal(e_1, 0.5, \delta), \\ fe_1 &= fal(e_1, 0.25, \delta), \\ z_{11} &= z_{11} + h(z_{12} - \beta_{11}e_1), \\ z_{12} &= z_{12} + h(z_{13} - \beta_{12} \cdot fe + b_{10}u_1), \\ z_{13} &= z_{13} + h(-\beta_{13} \cdot fe_1), \\ s_{11} &= x_{11} - v_{11}, \\ \bar{\alpha}_{11} &= v_{12} - k_{11}s_{11}, \\ \tau_{11}\dot{\alpha}_{11} + \alpha_{11} &= \bar{\alpha}_{11}, \\ \alpha_{11}(0) &= \bar{\alpha}_{11}(0), \\ s_{12} &= x_{12} - \alpha_{11}, \\ u_{10} &= \dot{\alpha}_{11} - k_{12}s_{12}, \\ u_1 &= \frac{(u_{10} - z_{13})}{b_{10}}. \end{aligned} \quad (32)$$

4.2.2. Longitudinal Velocity Controller Design. Similar to the derivation of heading controller, the longitudinal velocity controller can be deduced based on DS-ADRC as follows:

$$\begin{aligned} e &= v_{21} - v_u, \\ fh &= fhan(e, v_{22}, r, h), \\ v_{21} &= v_{21} + hv_{22}, \\ v_2 &= v_2 + h \cdot fh, \end{aligned}$$

$$\begin{aligned} e_{21} &= z_{21} - y_2, \\ fe &= fal(e_1, \alpha, \delta), \\ z_{21} &= z_{21} + h(z_{22} - \beta_{21}e_1 + b_{20}u_2), \\ z_{22} &= z_{22} + h(-\beta_{22} \cdot fe), \\ s_{21} &= x_{21} - v_{21}, \\ u_{20} &= v_{22} - k_{21}s_{21}, \\ u_2 &= \frac{(u_{20} - z_{22})}{b_{20}}. \end{aligned} \quad (33)$$

In (33), the definition of relevant parameters is similar to the definition in the heading controller design.

5. Simulation Results

A numerical example is given to illustrate the proposed path following control algorithm. In the simulation, our objective is to control the UUV to follow the path with speed at 1 m/s, and the sequence of points is $P = \{(5, 0), (50, 50), (50, 100), (5, 150), (5, 200), (50, 200), (50, 250)\}$. The UUV can initially rest at a random position with an unspecified attitude, supposing that it is $(x_0, y_0, \psi_0) = (0, 0, 0)$ and its velocity is $(u_0, v_0, r_0) = (0, 0, 0)$. Under the condition of constant current and other disturbances, we have a simulation comparison between the DSC and the DS-ADRC; the results are shown in Figures 4–11. Under the condition of constant current, the parameters of DSC are selected as $k_{11} = 0.8$, $\tau_{12} = 0.1$, $k_{12} = 1$, and $k_{21} = 1$, and with other disturbances the DSC parameters are $k_{11} = 0.5$, $\tau_{12} = 0.4$, $k_{12} = 1$, and $k_{21} = 2$, while for the above two conditions the parameters of the DS-ADRC both are selected as follows: $r = 0.8$, $h = 0.02$, $\delta = 0.1$, $\beta_{11} = 100$, $\beta_{12} = 300$, $\beta_{13} = 1000$, $\beta_{21} = 100$, $\beta_{22} = 1000$, $k_{11} = 1$, $k_{12} = 1$, $k_{13} = 1$, $k_{21} = 5$, $\tau_{11} = 0.1$, $b_{10} = 0.025$, and $b_{20} = 0.06$.

5.1. Path Following with Constant Current. In the north-east coordinates, we set the current velocity as 0.5 m/s and the direction as $\pi/4$. With the DSC and DS-ADRC, the relationship between UUV's "actual" path and its expected path is plotted in Figure 4. And Figures 5–7 are the UUV's cross-track error, its output of force, and moment, respectively.

Figures 4 and 5 show that the control effect of DSC and DS-ADRC is almost the same under the constant ocean current. However, from Figures 6 and 7, we can clearly see that the overshoot of DS-ADRC is smaller than the DSC and the setting time is shorter.

5.2. Path Following with Constant Current and Other Disturbances. Under the above constant current condition, we add a disturbance with which amplitude is 0.2 N·m and period is 20π , and the simulation results are as follows: Figure 8 is "actual" path and expected path under the DSC and DS-ADRC, respectively. And Figures 9–11 are the UUV's cross-track error, its output of force, and moment, respectively.

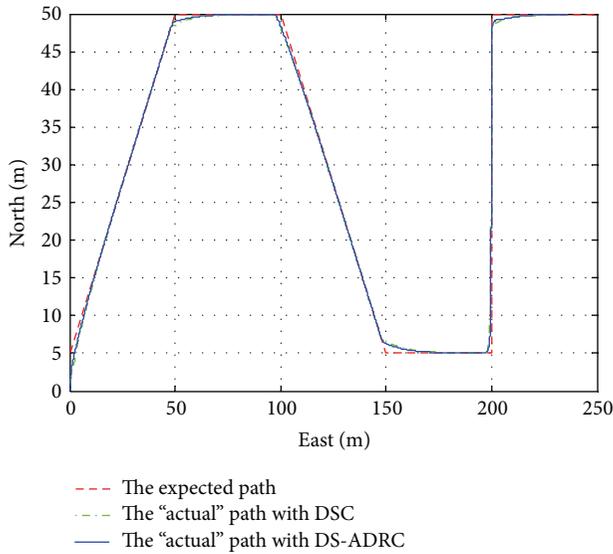


FIGURE 4: The responses of path following.

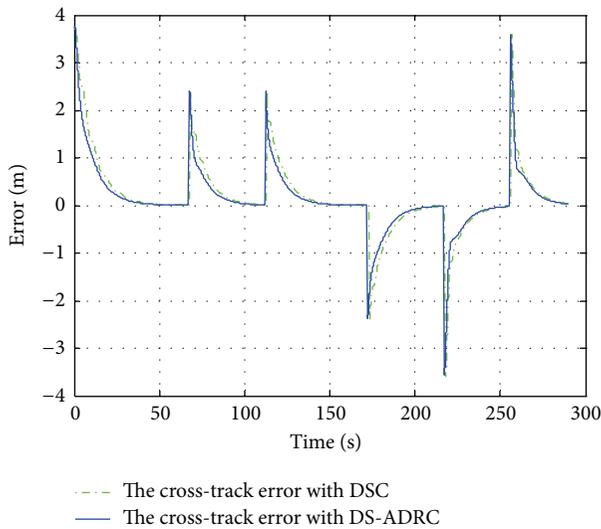


FIGURE 5: The cross-track error.

In the presence of other disturbances, it is interesting to note that the UUV's cross-track error of the proposed method in this paper is still gradually converged to zero according to Figures 8 and 9, which show strong antidisturbances ability, while the DSC cannot. When combined with Figures 10 and 11, we can see that the overshoot of DS-ADRC is smaller and its setting time is shorter, too.

Through the above two groups of simulation contrast, it can be seen that the effects of the DSC and the DS-ADRC are similar under the constant current, but when there are other external disturbances, the effect of the DS-ADRC is much better than the simple DSC, which show strong antidisturbance characteristic to the external unknown disturbances.

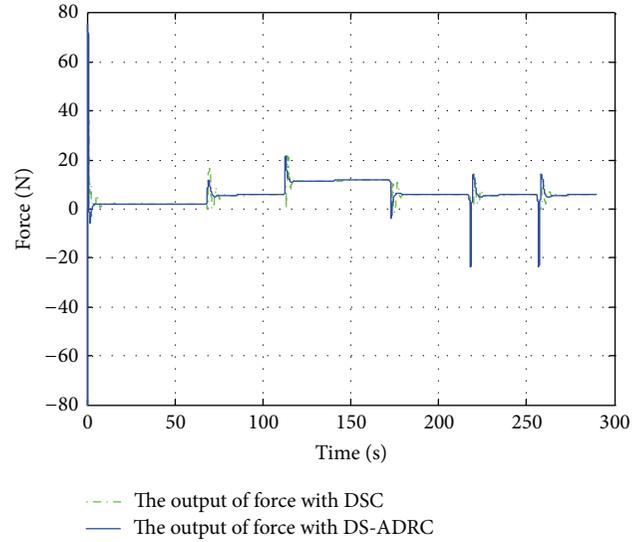


FIGURE 6: The output of force.

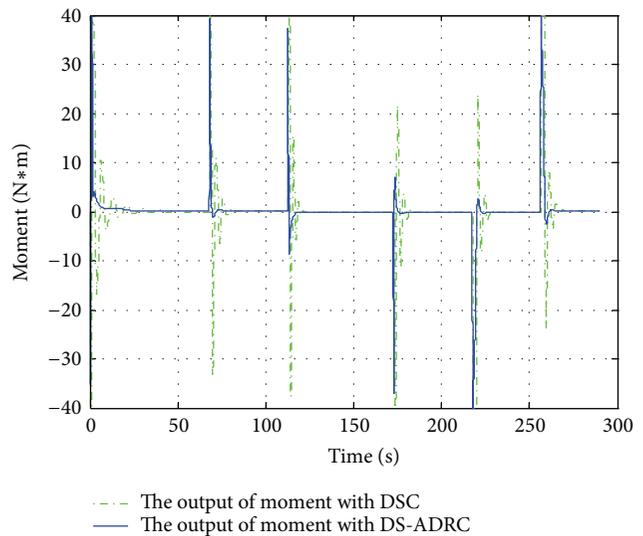


FIGURE 7: The output of moment.

6. Conclusion

This paper has presented a novel path following control method to UUV; it uses the line-of-sight guidance method to solve the real-time expectations of UUV's heading and has revised the method heading under the condition of current interference with introducing sideslip at the same time; it eliminates the stable cross-track error which is caused by the normal line-of-sight guidance method with current interference. Combining the advantages of the DSC technique and the ADRC technology, we, respectively, designed the UUV heading controller and the longitudinal velocity controller. This control method avoids the conventional dynamic surface control systems, relies on accurate mathematical models and improved antijamming capability. At the same time, the control method for a class of strict feedback forms is

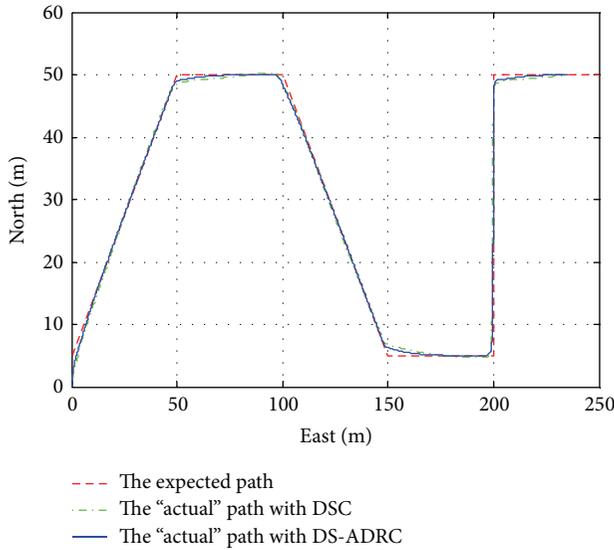


FIGURE 8: The responses of path following.

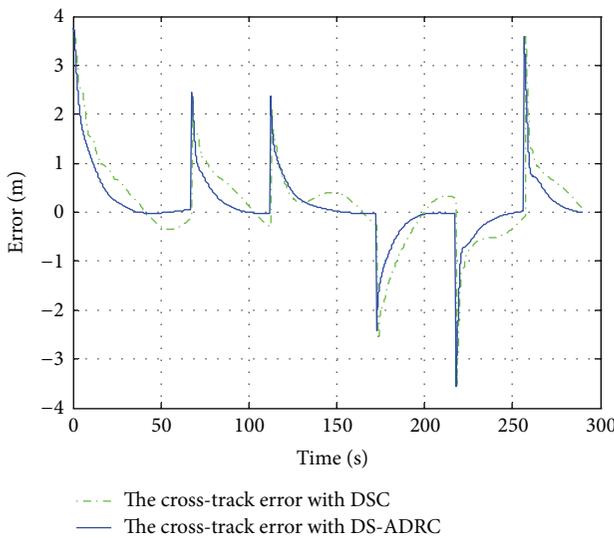


FIGURE 9: The cross-track error.

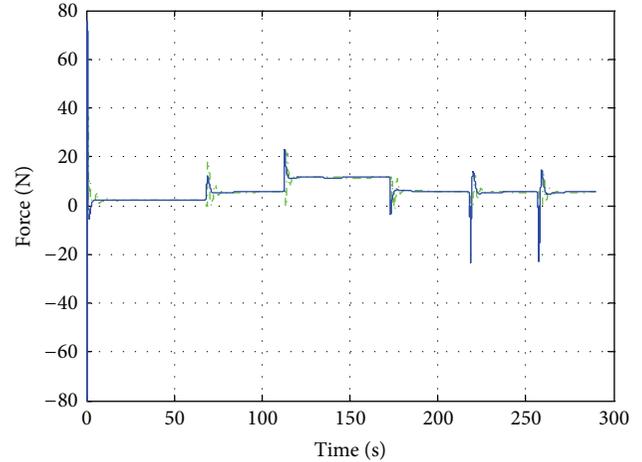


FIGURE 10: The output of force.

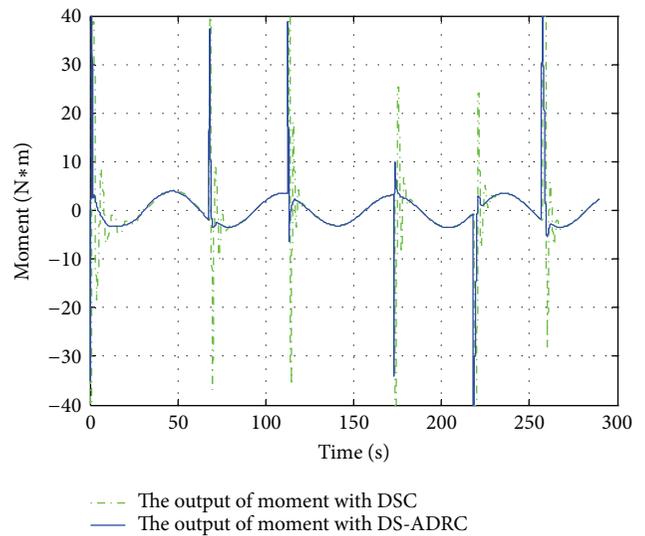


FIGURE 11: The output of moment.

applicable, which makes the design of the controller be in common use and be more conducive to the engineering practice. The simulation results also show that the control method has an excellent performance.

Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

Acknowledgments

This research work is supported by the National Natural Science Foundation of China (Grant No. 51309067/E091002),

the Fundamental Research Funds for the Center Universities (HEUCFX041402), and the National Defense Key Laboratory of Autonomous Underwater Vehicle Technology (9140C270208140C27004).

References

- [1] R. Coulson, J. C. Lambiotte, G. Grenon, T. Pantelakis, J. Curran, and E. An, "Development of a modular docking sub-system for 12" class autonomous underwater vehicles," in *Proceedings of the Bridges across the Oceans Conference (OCEANS '04. MTTs/IEEE TECHNO-OCEAN '04)*, pp. 1745–1752, November 2004.
- [2] E. Borhaug and K. Y. Fetteresen, "Adaptive way-point tracking control for underactuated autonomous vehicles," in *Proceedings of the 44th IEEE Conference on Decision and Control, and the*

- European Control Conference (CDC-ECC '05)*, pp. 4028–4034, December 2005.
- [3] J. Gao, D. M. Xu, and W. S. Yan, “Global K -exponential straight-line tracking control of an underactuated AUV in three dimensions using a cascaded approach,” *Control and Decision*, vol. 27, no. 9, pp. 1281–1287, 2012.
- [4] P. Herman, “Modified set-point controller for underwater vehicles,” *Mathematics and Computers in Simulation*, vol. 80, no. 12, pp. 2317–2328, 2010.
- [5] S.-R. Oh and J. Sun, “Path following of underactuated marine surface vessels using line-of-sight based model predictive control,” *Ocean Engineering*, vol. 37, no. 2-3, pp. 289–295, 2010.
- [6] M. Chyba, T. Haberkorn, R. N. Smith, and S. K. Choi, “Design and implementation of time efficient trajectories for autonomous underwater vehicles,” *Ocean Engineering*, vol. 35, no. 1, pp. 63–76, 2008.
- [7] L. Lapierre and D. Soetanto, “Nonlinear path-following control of an AUV,” *Ocean Engineering*, vol. 34, no. 11-12, pp. 1734–1744, 2007.
- [8] S. Shi, W. Yan, J. Gao et al., “Path following control of an AUV in the horizontal plane with constant ocean currents,” *Acta Armamentarii*, vol. 31, no. 3, pp. 375–379, 2010.
- [9] D. Swaroop, J. K. Hedrick, P. P. Yip, and J. C. Gerdes, “Dynamic surface control for a class of nonlinear systems,” *IEEE Transactions on Automatic Control*, vol. 45, no. 10, pp. 1893–1899, 2000.
- [10] R. Li, T. Li, R. Bu, Q. Zheng, and C. L. P. Chen, “Active disturbance rejection with sliding mode control based course and path following for underactuated ships,” *Mathematical Problems in Engineering*, vol. 2013, Article ID 743716, 9 pages, 2013.
- [11] J. E. Refsnes, A. J. Sorensen, and K. Y. Pettersen, “Model-based output feedback control of slender-body underactuated AUVs: theory and experiments,” *IEEE Transactions on Control Systems Technology*, vol. 16, no. 5, pp. 930–946, 2008.
- [12] J. E. G. Refsnes, “Nonlinear model-based control of slender body AUVs,” *Norwegian University of Science and Technology*, vol. 30, no. 226, pp. 229–231, 2007.
- [13] T. I. Fossen, *Handbook of Marine Craft Hydrodynamics and Motion Control*, Norwegian University of Science and Technology, 2011.
- [14] E. Fredriksen and K. Y. Pettersen, “Global k -exponential way-point manoeuvring of ships,” in *Proceedings of the 43rd IEEE Conference on Decision and Control*, pp. 5360–5367, 2004.
- [15] J. I. A. Heming, *Study of Spatial Target Tracking Nonlinear Control of Underactuated UUV Based on Backstepping*, Harbin Engineering University, 2012.
- [16] H. A. N. Jingqing, *Active Disturbance Rejection Control Technique—The Technique for Estimating and Compensating the Uncertainties*, National Defense Industry Press, 2007.