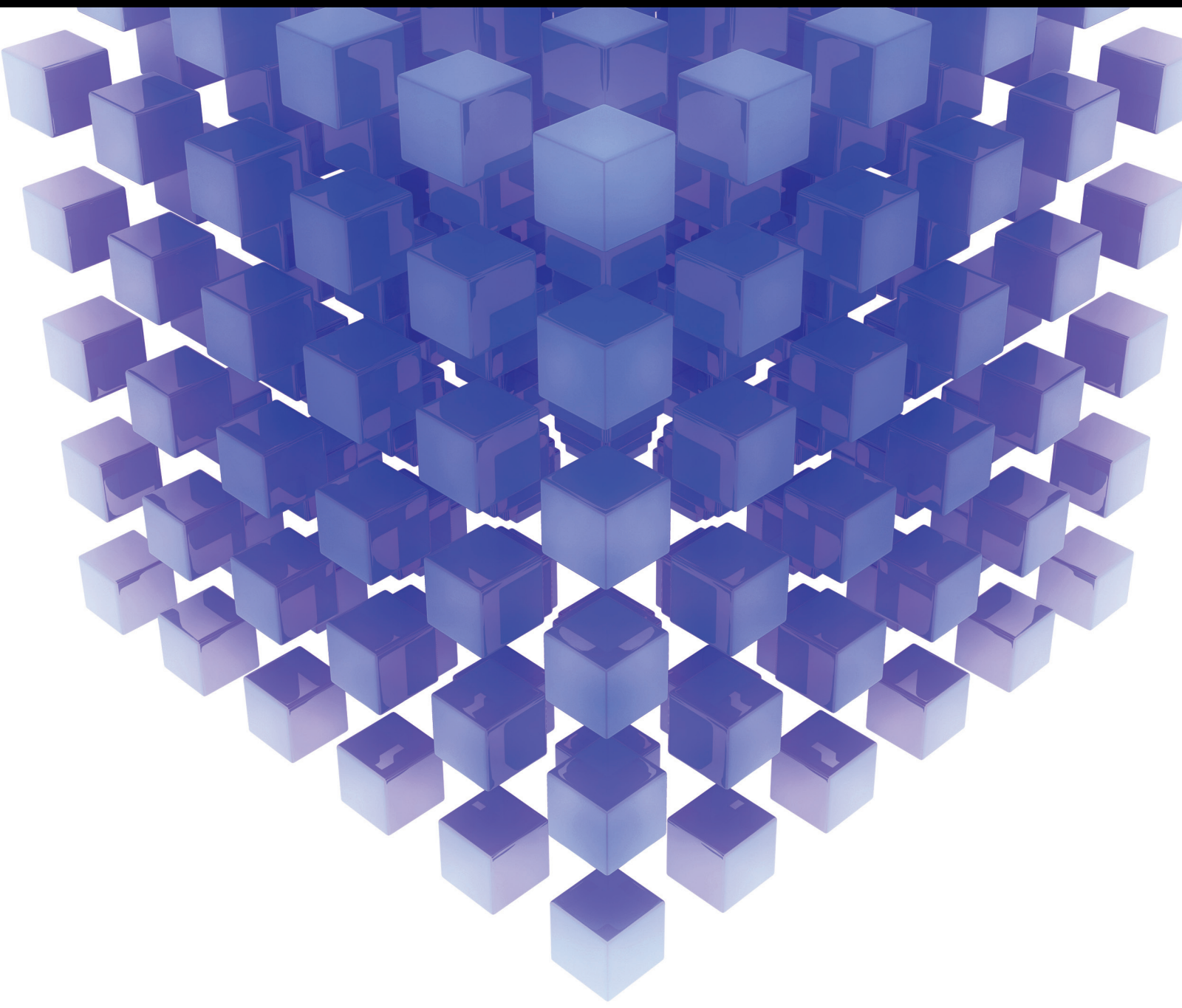
Mathematical Problems in Engineering

Control, Analysis, and Modeling of Vehicular Systems

Guest Editors: Ilse Cervantes, Sheldon S. Williamson, Ali Davoudi, and Luis Alvarez-Icaza



Control, Analysis, and Modeling of Vehicular Systems

Mathematical Problems in Engineering

Control, Analysis, and Modeling of Vehicular Systems

Guest Editors: Ilse Cervantes, Sheldon S. Williamson, Ali Davoudi, and Luis Alvarez-Icaza

Copyright © 2014 Hindawi Publishing Corporation. All rights reserved.

This is a special issue published in "Mathematical Problems in Engineering." 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

Mohamed Abd El Aziz, Egypt
Eihab M. Abdel-Rahman, Canada
Rashid K. Abu Al-Rub, USA
Sarp Adali, South Africa
Salvatore Alfonzetti, Italy
Igor Andrianov, Germany
Sebastian Anita, Romania
W. Assawinchaichote, Thailand
Erwei Bai, USA
Ezzat G. Bakhoum, USA
José Manoel Balthazar, Brazil
R. K. Bera, India
Christophe Bérenguer, France
Jonathan N. Blakely, USA
Stefano Boccaletti, Spain
Stephane P.A. Bordas, USA
Daniela Boso, Italy
M. Boutayeb, France
Michael J. Brennan, UK
Salvatore Caddemi, Italy
Piermarco Cannarsa, Italy
Jose E. Capilla, Spain
Carlo Cattani, Italy
Marcelo Moreira Cavalcanti, Brazil
Diego J. Celentano, Chile
Mohammed Chadli, France
Arindam Chakraborty, USA
Yong-Kui Chang, China
Michael J. Chappell, UK
Kui Fu Chen, China
Xinkai Chen, Japan
Kue-Hong Chen, Taiwan
Jyh-Horng Chou, Taiwan
Slim Choura, Tunisia
Cesar Cruz-Hernandez, Mexico
Swagatam Das, India
Filippo de Monte, Italy
Antonio Desimone, Italy
Yannis Dimakopoulos, Greece
Baocang Ding, China
Joao B. R. Do Val, Brazil
Daoyi Dong, Australia
B. Dubey, India
Horst Ecker, Austria
M. Onder Efe, Turkey
Elmetwally Elabbasy, Egypt
A. Elías-Zúñiga, Mexico
Anders Eriksson, Sweden
Vedat S. Erturk, Turkey
Moez Feki, Tunisia
Ricardo Femat, Mexico
Robertt A. Fontes Valente, Portugal
Claudio R. Fuerte-Esquivel, Mexico
Zoran Gajic, USA
Ugo Galvanetto, Italy
Xin-Lin Gao, USA
Furong Gao, Hong Kong
Behrouz Gatmiri, Iran
Oleg V. Gendelman, Israel
Didier Georges, France
Paulo Batista Gonçalves, Brazil
Oded Gottlieb, Israel
Fabrizio Greco, Italy
Quang Phuc Ha, Australia
M. R. Hajj, USA
Tony Sheu Wen Hann, Taiwan
Thomas Hanne, Switzerland
K. R. (Stevanovic) Hedrih, Serbia
M.I. Herreros, Spain
Wei-Chiang Hong, Taiwan
Jaromir Horacek, Czech Republic
Gordon Huang, Canada
Huabing Huang, China
Chuangxia Huang, China
Yi Feng Hung, Taiwan
Hai-Feng Huo, China
Asier Ibeas, Spain
Anuar Ishak, Malaysia
Reza Jazar, Australia
Zhijian Ji, China
Jun Jiang, China
J. J. Judice, Portugal
Tadeusz Kaczorek, Poland
Tamas Kalmar-Nagy, USA
Tomasz Kapitaniak, Poland
Hamid Reza Karimi, Norway
Metin O. Kaya, Turkey
Nikolaos Kazantzis, USA
Farzad Khani, Iran
Kristian Krabbenhoft, Australia
Ren-Jieh Kuo, Taiwan
Jurgen Kurths, Germany
Claude Lamarque, France
Usik Lee, Korea
Marek Lefik, Poland
Stefano Lenci, Italy
Roman Lewandowski, Poland
Shihua Li, China
Ming Li, China
Jian Li, China
Shanling Li, Canada
Teh-Lu Liao, Taiwan
Panos Liatsis, UK
Jui-Sheng Lin, Taiwan
Yi-Kuei Lin, Taiwan
Shueei M. Lin, Taiwan
Wanquan Liu, Australia
Yuji Liu, China
Bin Liu, Australia
Paolo Lonetti, Italy
Vassilios C. Loukopoulos, Greece
Junguo Lu, China
Chien-Yu Lu, Taiwan
Alexei Mailybaev, Brazil
Manoranjan K. Maiti, India
Oluwole Daniel Makinde, South Africa
Rafael Martinez-Guerra, Mexico
Driss Mehdi, France
Roderick Melnik, Canada
Xinzhu Meng, China
Yuri Vladimirovich Mikhlin, Ukraine
Gradimir Milovanovic, Serbia
Ebrahim Momoniat, South Africa
Trung Nguyen Thoi, Vietnam
Hung Nguyen-Xuan, Vietnam
Ben T. Nohara, Japan
Sotiris K. Ntouyas, Greece
Gerard Olivar, Colombia
Claudio Padra, Argentina
Bijaya Ketan Panigrahi, India
Francesco Pellicano, Italy
Matjaz Perc, Slovenia
Vu Ngoc Phat, Vietnam
Maria do Rosário Pinho, Portugal
Alexander Pogromsky, The Netherlands

Seppo Pohjolainen, Finland
Stanislav Potapenko, Canada
Sergio Preidikman, USA
Carsten Proppe, Germany
Hector Puebla, Mexico
Justo Puerto, Spain
Dane Quinn, USA
K. R. Rajagopal, USA
Gianluca Ranzi, Australia
Sivaguru Ravindran, USA
G. Rega, Italy
Pedro Ribeiro, Portugal
J. Rodellar, Spain
Rosana Rodriguez-Lopez, Spain
Alejandro J. Rodriguez-Luis, Spain
Ignacio Romero, Spain
Hamid Ronagh, Australia
Carla Roque, Portugal
Rubén Ruiz García, Spain
Manouchehr Salehi, Iran
Miguel A. F. Sanjuán, Spain
Ilmar Ferreira Santos, Denmark
Nickolas S. Sapidis, Greece
Evangelos J. Sapountzakis, Greece
Bozidar Sarler, Slovenia
Andrey V. Savkin, Australia
Massimo Scalia, Italy
Mohamed A. Seddeek, Egypt
Alexander P. Seyranian, Russia
Leonid Shaikhet, Ukraine
Cheng Shao, China
Bo Shen, Germany
Zhan Shu, UK
Jian-Jun Shu, Singapore
Dan Simon, USA
Luciano Simoni, Italy
Grigori M. Sisoev, UK
Christos H. Skiadas, Greece
Davide Spinello, Canada
Sri Sridharan, USA
Rolf Stenberg, Finland
Changyin Sun, China
Jitao Sun, China
Xi-Ming Sun, China
Andrzej Swierniak, Poland
Yang Tang, Germany
Allen Tannenbaum, USA
Cristian Toma, Romania
Irina N. Trendafilova, UK
Alberto Trevisani, Italy
Jung-Fa Tsai, Taiwan
Kuppalapalle Vajravelu, USA
Victoria Vampa, Argentina
Josep Vehi, Spain
Stefano Vidoli, Italy
Xiaojun Wang, China
Dan Wang, China
Youqing Wang, China
Cheng C. Wang, Taiwan
Yijing Wang, China
Yongqi Wang, Germany
Moran Wang, China
Gerhard-Wilhelm Weber, Turkey
Jeroen A.S. Witteveen, The Netherlands
Kwok-Wo Wong, Hong Kong
Ligang Wu, China
Zhengguang Wu, China
Gongnan Xie, China
Wang Xing-yuan, China
Xuping Xu, USA
Xi Frank Xu, USA
Jun-Juh Yan, Taiwan
Xing-Gang Yan, UK
Suh-Yuh Yang, Taiwan
Mahmoud T. Yassen, Egypt
Mohammad I. Younis, USA
Bo Yu, China
Huang Yuan, Germany
S.P. Yung, Hong Kong
Ion Zaballa, Spain
Ashraf M. Zenkour, Saudi Arabia
Jianming Zhan, China
Xu Zhang, China
Yingwei Zhang, China
Lu Zhen, China
Liancun Zheng, China
Jian Guo Zhou, UK
Zexuan Zhu, China
Mustapha Zidi, France

Contents

Hindawi Publishing Corporation
Mathematical Problems in Engineering
Volume 2014, Article ID 301595, 3 pages
http://dx.doi.org/10.1155/2014/301595

Editorial

Control, Analysis, and Modeling of Vehicular Systems

Ilse Cervantes,[1] Sheldon S. Williamson,[2] Ali Davoudi,[3] and Luis Alvarez-Icaza[4]

[1] *Institute for Scientific and Technological Research of San Luis Potosi (IPICyT), Camino a la Presa San Jose 2055, Col. Lomas 4ta, 78216 San Luis Potosi, SLP, Mexico*
[2] *Department of Electrical and Computer Engineering, Concordia University, Montreal, QC, Canada H4G 2M1*
[3] *Electrical Engineering Department, University of Texas at Arlington, Arlington, TX 76019, USA*
[4] *Institute of Engineering, Universidad Nacional Autonoma de Mexico, 04510 Mexico City, DF, Mexico*

Correspondence should be addressed to Ilse Cervantes; ilse@ipicyt.edu.mx

Received 29 April 2014; Accepted 29 April 2014; Published 1 July 2014

Copyright © 2014 Ilse Cervantes 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.

1. Introduction

Automation has played a key role in the development of new technologies and in the maturation and improvement of existing ones. Automatic control launched its development with electronic devices since the mid-twentieth century, but many of the key mathematical tools for dynamic analysis and control synthesis were developed much earlier. It was necessary to join the mathematics with the electronics and other related areas such as simulation and modeling to raise this promising, now essential area of knowledge.

The number of automatic processes applied to vehicles and transportation systems has increased dramatically in recent years due to the growing urge of environmental-friendly and autonomous transportation. As technology advances, more complicated mathematical problems have to be solved in order to satisfy the increasing astringent expectations on safety, reliability, and performance of vehicular systems.

Our intention with this special issue was to provide an opportunity for researchers to share their latest theoretical and technological achievements in control, modeling, and analysis of vehicular systems and to become a forum to contribute, from a mathematical point of view, to the solution of some of the challenges that the development of new trends of transportation has arisen.

This special issue focuses on the state-of-the-art research and development in the areas of control and analysis of transportation systems as well as in current trends as, for example, transportation electrification, vehicular communication systems, and driver assistance applications. For this special issue, we received many submissions and after conducting a rigorous review process, we have accepted only 11 high-quality papers.

The articles in this special issue are classified into three topics: those related to stabilization and control of vehicular systems, which concern either with control synthesis and stabilization capabilities of controllers of vehicular systems or dealing with vehicular related applications, for example, a haptic steering by wire system based on high gain GPI observers, the application of a dynamic surface control to lateral vehicle control, or a vehicular wireless communication problem solved from a control point of view.

The second topic in this issue regards the application of analysis or numerical methods to vehicular problems, among these we can cite a sound signal processing algorithm for vehicular-type recognition or an approximation to the vehicle routing problem using an ant colony algorithm. Finally, the third topic regards transportation electrification problems, which concern modeling or control problems related to the electrification of vehicles, for example, the development of control strategies for energy management of an hybrid electric bus.

In the rest of this document we give a short description of the papers in this special issue. We sincerely hope that this special issue serves as a reference for research in the area

of applied mathematics to transportation and provides an incentive to attract talents which contribute to this fascinating area.

2. Stability and Control of Vehicular or Vehicular-Related Systems

One of the topics covered in this special issue is related to the control and stabilization of vehicular systems, including theoretical analysis of stabilization and its use in the solution of vehicle related problems. There are six papers related to this topic. The paper entitled "*Analysis and optimization of resource control in high-speed railway wireless networks*" by S. Xu et al. studies the joint optimal design of admission control and resource allocation for multimedia services delivery in high-speed railway wireless networks. Its main contribution is to propose a stochastic optimization framework for transmission, which focuses on the joint admission control and resource allocation problem under an average power constraint. Using virtual queues, the joint admission control and resource allocation problem, originally formulated as a stochastic optimization problem, is transformed into a queue stability problem.

On the other hand, the transformed problem is decomposed into three subproblems: admission control, utility maximization, and resource allocation. The use of the drift-plus-penalty approach transforms the resource allocation problem into a single variable problem, which is used to develop a distributed dynamic packet loading algorithm with guaranteed global optimality. To the best of our knowledge, it is the first time that the Lyapunov optimization theory is extended into the HSR wireless networks.

Another paper of this topic is "*Stability of gain scheduling control for aircraft with highly nonlinear behavior*" by F. Mendez-Vergara et al. The main goal of this work is to study the stability properties of an aircraft with nonlinear behavior. The problem of body stabilization and attitude control is solved using a piecewise linear output feedback and asymptotical stability of the task coordinates origin and safety of the operation in the entire flight envelope. The paper departs from a general description of the aircraft that uses piecewise nonlinear models. This paper has two main contributions: firstly, to deal with nonlinearities of the description of the aircraft from a unified point of view, in which the effects of the parametric uncertainty, structural uncertainty, and some control failures can be studied with a single description and secondly, to propose relaxed switching conditions for stability that can depend either on time, output, or the entire state and they can be applied using a supervisory structure which evaluates the satisfaction of the stability conditions. The result proposed by the authors is interesting since even if the behavior of the aircraft is highly nonlinear, the controller can be synthesized using a piecewise linear description of the system in a given operation points, and the performance of the system can be set using a trade-off between the size of integral terms and the size of the operating domain.

On the other hand, in "*A vehicle haptic steering by wire system based on high gain GPI observers*" by A. Rodriguez Angeles et al., a vehicle steering by wire (SBW) haptic system based on high gain generalized proportional integral (GPI) observers is introduced. A steer-by-wire system is a modular electric power assisted steering device with an absence of steering column, steering shaft, or mechanical connection between the steering wheel and the road wheel. SBW requires control actions to steering tracking and the observers in this paper are used for the estimation of dynamic perturbations that are present at the tire and steering wheel. The estimated dynamic effects and position tracking errors are feedback yielding a master-slave haptic system with bilateral communication. The main contribution of this paper is to propose a steering by wire (SBW) haptic system that requires the nonlinear system input gain of the system and that uses only encoders to feedback purposes. The tracking error between the desired angle settled by the human operator, and the tire orientation angle can be made arbitrarily small by a proper gain selection. Using experiments, the authors show that the controller is robust to uncertainty on the input gain and that it is able to cancel dynamic perturbation effects such as friction and aligning forces on the tire.

Along the line of developing controllers for advanced driver assistance systems, there is the paper "*Robust switched predictive braking control for rollover prevention in wheeled vehicles*" by M. R. Licea and I. Cervantes. This paper proposes a differential braking rollover mitigation strategy for wheeled vehicles. Using a polytopic description of the vehicle that includes suspension effects and translational and rotational dynamics of the vehicle, the authors propose a velocity-dependent domain partition and a predictive action for preventing rollover. To predict the behavior of the vehicle, the authors use a linear model and a low-order model of the driver behavior. The proposed methodology, which constitutes the main contribution of the paper, is adaptable via switching actions, interfering with driver's acceleration/braking actions only when a risk of rollover is foreseen. The proof of stability of the scheme is performed in two steps; firstly, stability conditions for the hybrid closed-loop system are derived with respect to a switching surface that depends on norm of system states. Secondly, the authors prove that the switching surface can be time varying, leading to a more relaxed switching conditions by the use of prediction.

An interesting paper that performs an extension of known control methodologies is the paper "*Dynamic surface control and its application to lateral vehicle control*" by B. Song et al. This paper extends the design and analysis methodology of dynamic surface control (DSC) for a class of nonlinear systems where the nonlinear functions are included as forcing terms. In many applications such as rotational mechanical systems, sinusoidal functions are included in the equation of motions; however, when the sinusoidal function is used as a forcing term for DSC, the stability analysis is complicated by highly nonlinear functions resulting from the low-pass filter dynamics. With a modification of input variables to the filter dynamics, the burden of mathematical analysis can be reduced and stability conditions in linear matrix inequality are formed to guarantee the quadratic stability. The illustration of this proposed strategy is applied to lateral vehicle control for

forward automated driving and backward parallel parking at a low speed.

The last paper of this topic is "H^{∞} *controller design for asynchronous multirate sampled-data systems*" by X. Sun and W. Mao. Multirate sampled-data systems are the product of different sampling rates of the signals of interest. In the implementation of energy management strategies of hybrid electrical vehicles, a hierarchical controller is designed to perform the power split among the sources, but as the system becomes large and complex, an unavoidable situation of sensors and control inputs evolving at different sampling rates may be present. In these cases, multirate sampled-data systems are better choices for modeling in order to improve the closed-loop performance. This paper proposes an observer-based H^{∞} controller which guarantees that the H^{∞} norm of the closed system is less than a given attenuation level. To improve the performance further, the exogenous signals sampled at different rates are lifted to an appropriate signal rate, while the endogenous signals are not lifted to avoid a causal constraint. The controller is obtained by solving a matrix inequality.

3. Analysis and Numerical Methods Applied to Vehicular Problems

Vehicle-type recognition, which can be used in intrusion detection, transportation, and border monitoring among others, is a demanding application for wireless sensor networks. Traditionally, sensor nodes detect and recognize vehicles from their acoustic or seismic signals using wavelet based or spectral feature extraction methods. However, such methods are demanding in computational power and energy and are difficult to implement on low cost sensor nodes with limited resources. In "*Vehicular type recognition in sensor networks using improved time encoded signal processing algorithm*" by Y. Wang et al., the use of time encoded (acoustic) signal processing algorithm for vehicle recognition is investigated. An improved time encoded signal processing is proposed as the feature extraction method according to the characteristics of the vehicle sound signal. Recognition procedure is accomplished using a support vector machine and a k-nearest neighbor classifier. The authors validate their proposed methodology using experimental results and illustrate that the ITESP features have a better performance compared with the conventional TESP methodologies.

On the other hand, a paper that looks into the path planning of an emergency rescue vehicle in an interesting way is "*An improved ant colony algorithm and its application in vehicle routing problem*" by M. Huang and P. Ding. The vehicle routing problem is defined as a transportation and a distribution problem with the constrained of vehicle load, travel time, or distance. In this paper, an ant colony optimization algorithm is applied to the optimal path planning of an emergency rescue vehicle. With respect to existing methodologies, the proposed algorithm uses a new transition probability function adding the angle factor function and visibility function, while setting penalty function in a new pheromone updating model to improve the accuracy of the route searching. The results presented by the authors show that the proposed method is better than the traditional ant colony algorithm for vehicle routing path planning and that the method can plan a more rational rescue path focused on a real traffic situation.

Robust object detection and tracking is a major technological issue that has to be solved for vehicle automation, in particular for automated steering and acceleration. In "*Invariant Hough random ferns for object detection and tracking*" by Y. Lin et al., a robust object detection methodology under changes in object appearance, scale, partial occlusions, and pose variation is proposed. The methodology incorporates rotation and scale invariance into a local feature description, random ferns classifier training, and Hough voting stages. The contributions of the authors are three: the rotation invariant local binary feature based on polar coordinates is shown to be invariant to image rotation; the combined use of Hough transform with random ferns classifier provides object detection regardless of partial occlusions and nonrigid deformations; and finally, a refined top-down segmentation algorithm based on a clustering scheme is presented. The authors show using experimental results that both object segmentation and long-term object tracking are accurate and robust in a variety of complex scenarios.

4. Transportation Electrification

J. Wang et al. in the paper "*The development and verification of a novel ECMS of hybrid electric bus*" proposed an adaptable equivalent consumption minimization strategy for a hybrid electric bus, which coordinates the relationship between the gear-shifting and motor assistance. The strategy also performs drive cycle recognition to update the charge and discharge coefficients in order to achieve better fuel economy in each drive cycle. Hardware-in-the-loop tests in a series-parallel hybrid electric bus show that the proposed method can achieve significant fuel economy with respect to other existing methodologies and provide new design ideas for the gear-shifting rules. Finally, in "*Sample-data modeling of a zero voltage transition DC-DC converter for on-board battery charger in EV*" by T. R. Granados-Luna et al., a sample-data model for a battery charger is proposed. The authors analyze the idealized waveforms of the converter to perform a dynamical analysis or the system; then a phase control strategy is modelled to obtain a large-signal model of the converter. Using this model, a half-cycle, sample-data linear model is obtained, which helps to provide the small-signal transfer functions of the converter.

Acknowledgments

We are grateful to our reviewers who dedicated their time in reviewing the submitted papers and provided valuable suggestions to the authors.

Ilse Cervantes
Sheldon S. Williamson
Ali Davoudi
Luis Alvarez-Icaza

Hindawi Publishing Corporation
Mathematical Problems in Engineering
Volume 2014, Article ID 981276, 14 pages
http://dx.doi.org/10.1155/2014/981276

Research Article

A Vehicle Haptic Steering by Wire System Based on High Gain GPI Observers

A. Rodriguez-Angeles,[1,2] H. Sira-Ramirez,[1] and J. A. Garcia-Antonio[1]

[1] *Center for Research and Advanced Studies (CINVESTAV-IPN), Electrical Engineering Department, Mechatronics Group, Avenida IPN, No. 2508, Col. San Pedro Zacatenco, 07630 Mexico City, DF, Mexico*
[2] *Dynamics and Control Group, Department of Mechanical Engineering, Eindhoven University of Technology, P.O. Box 513, 5600 MB Eindhoven, The Netherlands*

Correspondence should be addressed to A. Rodriguez-Angeles; aangeles@cinvestav.mx

Received 12 November 2013; Accepted 31 March 2014; Published 30 June 2014

Academic Editor: Ilse Cervantes

Copyright © 2014 A. Rodriguez-Angeles 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 vehicle steering by wire (SBW) haptic system based on high gain generalized proportional integral (GPI) observers is introduced. The observers are considered for the estimation of dynamic perturbations that are present at the tire and steering wheel. To ensure efficient tracking between the commanded steering wheel angle and the tire orientation angle, the estimated perturbations are on line canceled. As to provide a haptic interface with the driver, the estimated dynamic effects at the steering rack are fed back to the steering wheel, yielding a master-slave haptic system with bilateral communication. For implementation purposes few sensors and minimum knowledge of the dynamic model are required, which is a major advantage compared to other approaches. Only position tracking errors are fed back, while all other signals are estimated by the high gain GPI observers. The scheme is robust to uncertainty on the input gain and cancels dynamic perturbation effects such as friction and aligning forces on the tire. Experimental results are presented on a prototype platform.

1. Introduction

Automobile dynamics, safety, and comfort are taking a major role in vehicle design [1]. The steering system, or rather the interface between the driver and the vehicle, is becoming an important part of automobile design tasks [2–4]. The basic design of the automobile steering system has changed little since the steering wheel invention. The driver commanded direction is transmitted by a column or steering shaft, including universal joints and gearboxes, to the front tires. Assisted steering systems were introduced in 1950, with the hydraulic power steering system. Since then, the assistance unit has become standard in modern vehicles. The steering system complements the effort required by the driver to steer the vehicle, by using hydraulic pressure. Meanwhile, electrical assisted steering systems have been introduced recently, which eliminate the requirement of a hydraulic pump or diminish the required effort provided by the hydraulic power, yielding electrohydraulic or electrical assistance to the steering rack [5, 6].

The steering systems can be classified into three main groups: mechanical steering systems [7], hydraulic power and electrohydraulic power systems, and electric power assisted systems (EPAS) [8–10]. Among the EPAS, there is a new technology called steering by wire (SBW), whose main characteristic is to completely do away with as many mechanical components (steering shaft, column, gear reduction mechanism, etc.) as possible [11]; see Figure 1.

In general SBW systems exhibit several advantages, such as the following.

(1) The absence of steering column simplifies the car interior design.

(2) The absence of steering shaft, column and gear reduction mechanism, allows much better space utilization in the engine compartment.

(3) The steering mechanism can be designed and installed as a modular unit.

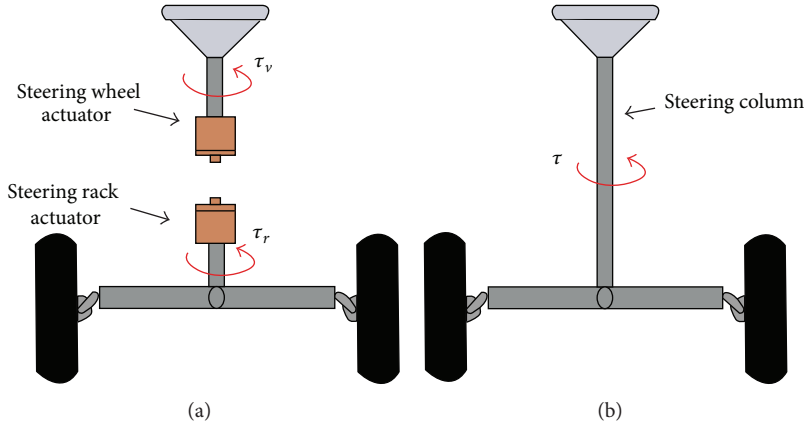


Figure 1: (a) SBW; (b) conventional steering system.

(4) Without mechanical connection between the steering wheel and the road wheel, it is less likely that the impact of a frontal crash will force the steering wheel to invade into the driver's survival space.

(5) SBW system characteristics can easily be adjusted to optimize the steering response and driving feeling [12, 13].

Therefore, SBW systems allow better structures for crash energy absorption [14] and present benefits related to passengers' comfort and driver feeling.

While substituting conventional steering systems by SBW, there is a major issue related to what the driver feels through the steering wheel. The forces, torques, and driving conditions transmitted by the steering wheel to the driver are important for a proper and safe vehicle driving. Therefore, the SBW must provide the driver the opportunity to "feel" what the road driving conditions are, just as the steering column does. As a consequence, the SBW mechanism must also behave as a haptic device [13, 14].

Several dynamic models related to haptic SBW systems are presented in the literature; see [13–16]. In these models, uncertain and perturbation terms appear which affect the dynamics and the behavior of the SBW. Phenomena such as friction, damping, inertia, and aligning forces, among others, are considered in the dynamic model. The influence of those uncertain and perturbation terms is rather important to guarantee a good performance of the SBW and to provide a reliable feedback of the road and driving conditions to the driver. Dynamic phenomena effects on the SBW are difficult to be known with certainty due to the changing conditions on the road and the driving, for example, speed, roadway texture, tire wear, tire air pressure, and rain. For the purpose of estimating the effects of such unknown perturbations, the use of observers has been proposed [11, 17]. The implementation of such observers often requires the use of specialized sensors, such as GPS and INS, measurement of lateral acceleration [12], and current and torque motors measurements [15]. Furthermore, most of the proposed observers highly rely on complicated dynamic models of the phenomena that are present in the steering system. Besides uncertainty on the dynamics phenomena, there exists also the problem of model changes in accordance with the driving conditions and the wide variety of vehicle models [1, 16].

All that was previously described makes it difficult to determine a precise mathematical model and a control strategy for the SBW system. As a solution, model-free observers might be considered for estimating unmodeled dynamics and perturbations effects on SBW systems.

In this work a vehicle SBW system is proposed based on high gain generalized proportional integral (GPI) observers. The use of extended and high gain observers is rather common in active disturbance rejection control (ADRC), [18, 19], and some applications of such approach can be revised in [20–22].

In the proposed approach high gain GPI observer helps in determining the effects of uncertain dynamics phenomena and additive perturbations on the steering rack and tire, such as continuous and discontinuous friction, aligning forces, inertia effects, and damping. The purpose of the high gain GPI observer is twofold: firstly to help a PD controller to actively reject dynamic perturbations and secondly to provide the forces and torques that are fed back to the driver through the steering wheel, therefore closing a haptic loop. The proposed approach is only based on angle position feedback; therefore, it can be implemented with common low cost encoder sensors as in the experimental platform here considered.

The proposed SBW system behaves as a self-contained sensor for tire forces, since such forces directly influence vehicle dynamics. Although all uncertainties and dynamic perturbations are lumped into an estimated linear term, the approach shows good performance and a proper estimation of the combined dynamic effects that are present at the SBW system. Experimental results on a real platform show good agreement with the theoretical results, that allows concluding convergence of the tracking and estimation errors to a small vicinity around zero.

2. SBW Dynamic Model

When considering a vehicle dynamics there are several involved phenomena, such as lateral and longitudinal forces,

tire-road friction [16, 23]. A common approach on vehicle modeling dynamics is the so-called bicycle model [12, 16, 24], which considers the chassis. However for the purpose of this paper and taking into account that the experimental platform only considers one tire (half of the steering rack), only the steering system dynamics is to be taken into account.

As result of the SBW design of the experimental platform, the steering wheel system depicted in Figure 3 and the steering rack system in Figure 5 are very similar. Thus, based on their similarities, the dynamic model of the steering rack subsystem is obtained alike that of the steering wheel. But, at the steering rack system the tire-road dynamic phenomena are included, such as tire-road interaction and aligning forces.

For modeling, control, and implementation purposes the SBW system might be viewed as a master-slave system with bilateral communication. In such approach, the steering wheel acts as a master subsystem, capturing the driver commanded steering angle, and the steering rack is seen as a slave subsystem that has to track the operator commanded steer angle. The bilateral communication implies that the steering wheel sends the commanded angle to the steering rack, while from the steering rack, dynamic effects are estimated and sent to the steering wheel subsystem. Furthermore each subsystem is provided with a PD controller with active disturbance rejection and high gain GPI observers. The overall system thus conforms a bilateral control scheme as in [24].

2.1. Steering Wheel Subsystem. The dynamics of the steering wheel (Figure 2) or master system are affected by physical phenomena that are presented in its three fundamental components: the wheel, the reduction gearbox, and the DC motor (Figure 3). In particular the DC motor is in charge of reflecting to the driver the forces that are estimated by the high gain GPI observer from the dynamical phenomena presented on the steering rack and tire (slave subsystem). Uncertain dynamic effects and perturbation on the wheel subsystem are estimated by another high gain GPI observer and canceled through a PD controller with disturbance rejection. Dynamic effects on the steering wheel include discontinuous Coulomb friction; other works such as [12, 24] considered that Coulomb friction phenomena affect only the steering rack and tire subsystem. However the use of a reduction gearbox induces discontinuous friction phenomena, even dead zones may occur depending on the reduction gear.

Figure 3 shows an equivalent mechanical-electrical diagram of the steering wheel subsystem. It is shown how the dynamics of its three components are interconnected. For the electrical part the relevant parameters are V_v the input voltage, L_v armature inductance, R_v armature resistance, i_v armature current, i_f current through the winding field, e_v electromotive force, and k_2 electromotive force constant gain. Meanwhile, the mechanical relevant parameters are ϕ_v output angle of the motor, J_1 inertia moment at the reduction input, B_1 damping coefficient at the reduction input, F_1 Coulomb friction coefficient at the reduction input, k_1 motor torque constant gain, $\tau_{v'}$ motor torque, F_{Tv} Coulomb friction coefficient at gearbox, B_{Tv} damping coefficient at gearbox, J_2 inertia moment at the reduction output, B_2 damping coefficient at

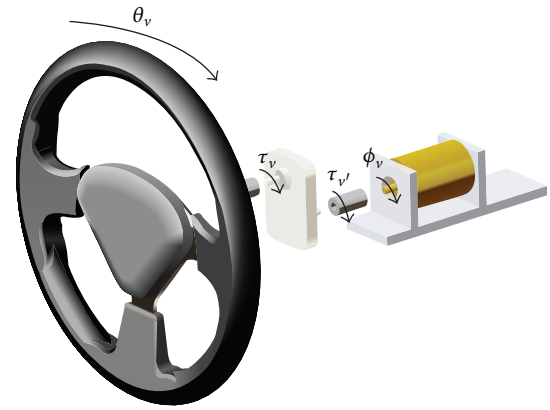


Figure 2: Steering wheel subsystem.

the reduction output, F_2 Coulomb friction coefficient at the reduction output, τ_v torque at the reduction output, C_v wheel Coulomb friction coefficient, B_v wheel damping coefficient, J_v wheel inertia moment, $N_v = N_{v1}/N_{v2}$ gearbox reduction ratio, N_{v1} the number of teeth on the input gear, N_{v2} the number of teeth on the output gear, and θ_v wheel angular position.

For the electrical part of the steering wheel system, the dynamic relation between the torque at the reduction output τ_v and the input voltage V_v is given by

$$\tau_v = \frac{N_v k_1}{R_v} V_v(t) - \frac{N_v^2 k_1 k_2}{R_v} \dot{\theta}_v. \tag{1}$$

Meanwhile, taking into account the reduction gearbox, the dynamic relation between τ_v and the wheel angular position θ_v is represented by

$$\tau_v = J_{Tv} \ddot{\theta}_v + B_{Tv} \dot{\theta}_v + F_{Tv} \operatorname{sign}\left(\dot{\theta}_v\right), \tag{2}$$

where the equivalent coefficients of the reduction gearbox are given by

$$\begin{aligned} J_{Tv} &= \left(\frac{N_{v2}}{N_{v1}}\right)^2 J_1 + J_2, \\ B_{Tv} &= \left(\frac{N_{v2}}{N_{v1}}\right)^2 B_1 + B_2, \\ F_{Tv} &= \left(\frac{N_{v2}}{N_{v1}}\right)^2 F_1 + F_2. \end{aligned} \tag{3}$$

Finally, taking into account (1) and (2) and the dynamic effects of the steering wheel, it is obtained that the dynamic model of the steering wheel subsystem is given by (4), with $J_{ve} = J_v + J_{Tv}$ being the equivalent inertia of the steering wheel system:

$$\begin{aligned} &J_{ve} \ddot{\theta}_v + \left(B_{Tv} + B_v\right) \dot{\theta}_v + \left(F_{Tv} + C_v\right) \operatorname{sign}\left(\dot{\theta}_v\right) \\ &\quad = \frac{N_v k_1}{R_v} V_v(t) - \frac{N_v^2 k_1 k_2}{R_v} \dot{\theta}_v. \end{aligned} \tag{4}$$

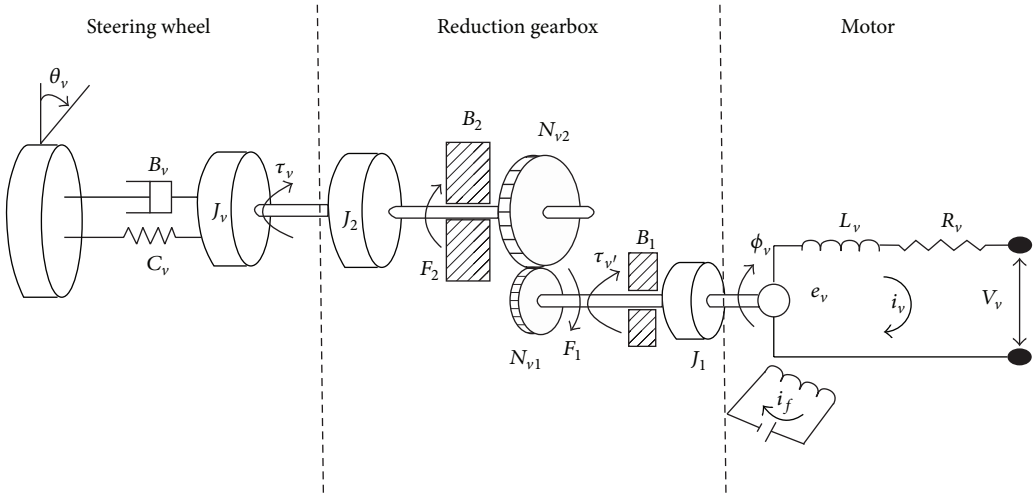


Figure 3: Equivalent mechanical-electrical diagram of the steering wheel subsystem.

2.2. Steering Rack Subsystem and Tire Dynamics. Figure 4 shows the steering rack and tire subsystem (slave subsystem), while a schematic diagram is presented in Figure 5, where the main components (tire, reduction gearbox, and DC motor) are represented by its electrical and mechanical counterparts. Notice the similarities between the steering wheel subsystem depicted in Figure 3 and the steering rack subsystem in Figure 5. Thus, based on their similarities, the dynamic model of the steering rack subsystem is obtained alike that of the steering wheel. However, because the tire-road interaction, there are dynamic effects such as friction and aligning forces that must to be considered [16, 17]. Also longitudinal forces affect the tire dynamics, but at this point, such forces are disregarded and would be further included when vehicle chassis is to be considered.

From Figure 5 and by taking into account the electrical and mechanical parameters, the dynamic model of the steering rack and tire subsystem is obtained as given by (5). The aligning forces at the tire, also referred to as self-aligning forces in the literature, are represented by τ_{SAT}. Meanwhile, the relevant electrical parameters in model (5) are V_r the input voltage, L_r armature inductance, R_r armature resistance, i_r armature current, i_f current through the winding field, e_r electromotive force, and k_2 electromotive force constant gain. For the mechanical part, the parameters are ϕ_r output angle of the motor, J_a inertia moment at the reduction input, B_a damping coefficient at the reduction input, F_a Coulomb friction coefficient at the reduction input, k_1 motor torque constant gain, F_{Tr} Coulomb friction coefficient at gearbox, B_{Tr} damping coefficient at gearbox, $\tau_{r'}$ motor torque, J_b inertia moment at the reduction output, B_b damping coefficient at the reduction output, F_b Coulomb friction coefficient at the reduction output, τ_r rack torque, C_r rack Coulomb friction coefficient, B_r rack damping coefficient, J_r rack inertia moment, $N_r = N_{r1}/N_{r2}$ gearbox reduction ratio, N_{r1} the number of teeth on the input gear, N_{r2} the number of teeth on the output gear, θ_r rack angular position (tire angular position), and $J_{re} = J_r + J_{Tr}$ the equivalent inertia of the steering rack subsystem:

$$J_{re}\ddot{\theta}_r + \left(B_{Tr} + B_r\right)\dot{\theta}_r + \left(F_{Tr} + C_r\right)\operatorname{sign}\left(\dot{\theta}_r\right) + \tau_{\mathrm{SAT}} = \frac{N_r k_1}{R_r} V_r(t) - \frac{N_r^2 k_1 k_2}{R_r}\dot{\theta}_r. \tag{5}$$

The equivalent coefficients of the rack reduction gearbox are given by

$$J_{Tr} = \left(\frac{N_{r2}}{N_{r1}}\right)^2 J_a + J_b,$$
$$B_{Tr} = \left(\frac{N_{r2}}{N_{r1}}\right)^2 B_a + B_b, \tag{6}$$
$$F_{Tr} = \left(\frac{N_{r2}}{N_{r1}}\right)^2 F_a + F_b.$$

2.2.1. Aligning Force at Tire. The aligning torque τ_{SAT} occurs because of the existence of caster and kingpin angles in the steering mechanical structure; this torque induces steering pull. Both linear and nonlinear models for lateral aligning forces have been proposed, mainly based on cornering stiffness, the slip angle, and the so-called bicycle model of a front steering vehicle [16, 24]. While a vehicle is turning, tire forces acting on the steering system tend to resist steering motion away from the straight-ahead position. These forces are due to self-aligning torque, that is a function of the steering geometry, particularly caster angle, tire deformation, and vehicle linear velocity. Since the considered experimental platform is only composed by half of the steering rack (see Figure 8), then only one tire is taken into account.

The aligning torque at the front left wheel is modeled as presented in [23]. Figure 6 depicts the forces and variables related to aligning tire torque. The subindex fl indicates that

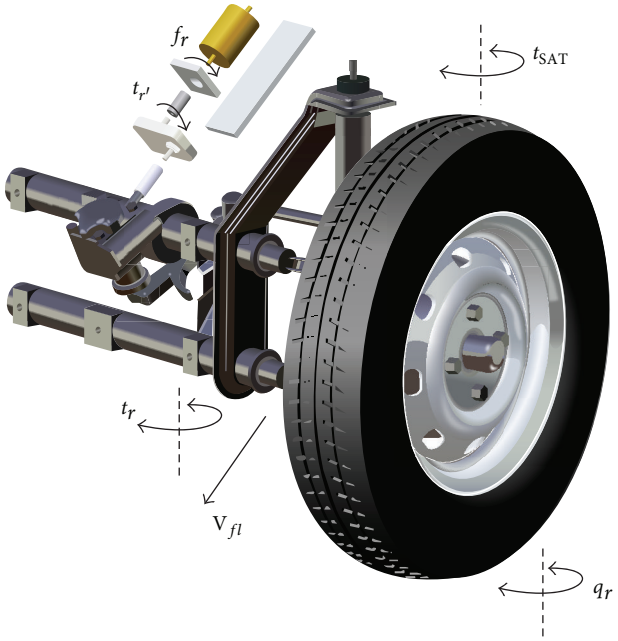


Figure 4: Steering rack and tire subsystem.

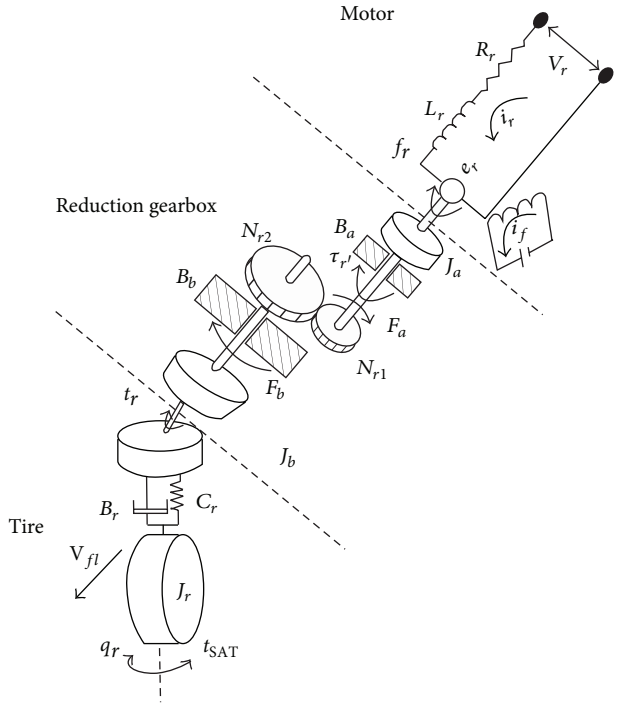


Figure 5: Equivalent mechanical-electrical diagram of the steering rack and tire subsystem.

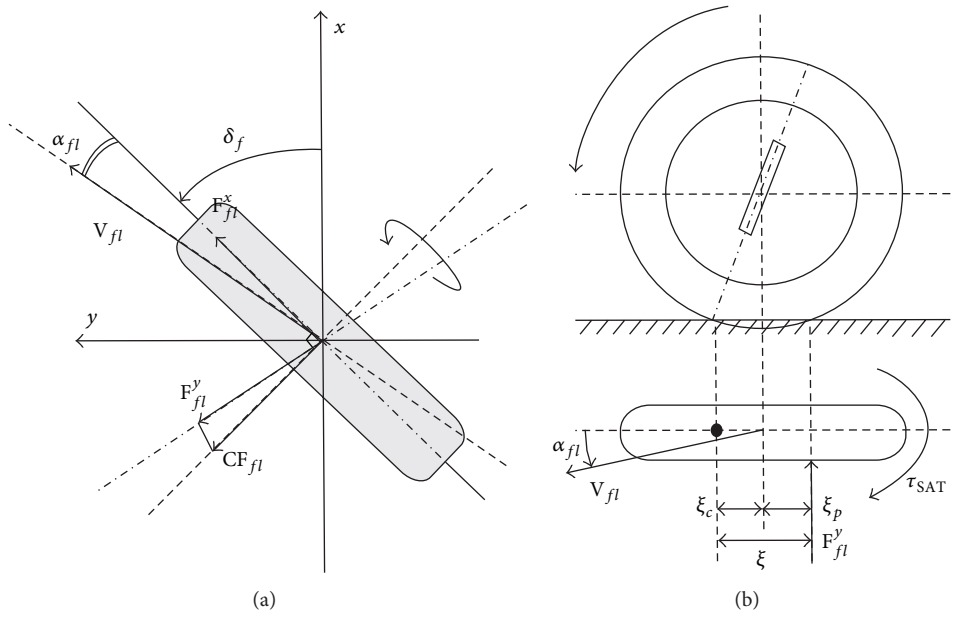


Figure 6: Tire force at front left wheel: (a) tire forces; (b) self-aligning torque [23].

front left tire is considered, while the superindices x and y reflect the axes on which forces act. F^x and F^y are driving and lateral forces, respectively. F^y is generated by sidesliping of the tire. CF stands for cornering force. F^y_{fl} is the front left tire lateral force, α_{fl} is the front tire slip angle, ξ_p is the pneumatic trail, that is, the distance from the center of tire to the application of lateral force, ξ_c is the caster or mechanical trail, that is, the distance from the tire center to the point on the ground about which the tire pivots as a result of the wheel caster angle, and V_{fl} is the velocity of the tire at its center. Then the self-aligning torque is given by

$$\tau_{\text{SAT}} = \left(\xi_c + \xi_p\right) F^y_{fl}\left(\alpha_{fl}\right). \tag{7}$$

According to [23], in the linear region of tire characteristics, the front left tire cornering force CF_{fl} depends on cornering stiffness C_f and the slipping angle α_{fl}; that is,

$$CF_{fl} = -C_f \alpha_{fl}. \tag{8}$$

Then from Figure 6, the relation between F^y_{fl} and CF_{fl} is given by

$$F^y_{fl} = \frac{CF_{fl}}{\cos \alpha_{fl}}. \tag{9}$$

Note that the tire slip angle α_{fl} is not easy to be measured; thus it must be estimated. In [12], it is mentioned that on production cars' sideslip angle is typically derived from integration of inertial sensors (lateral acceleration), but this estimation error is prone to uncertainty and errors due to sensor drift. Furthermore, ξ_c and ξ_p are only approximately known; however, in order to obtain a linear model, ξ_c and ξ_p are assumed to be constant and known. Same consideration applies to the front cornering stiffness C_f.

Remark 1. Notice that the dynamic models here represented by (4) and (5) are not used for the design of the PD controller and the high gain GPI observers. The only parameter that is required for the high gain GPI observer is the input gain. Nevertheless, the dynamic models are introduced to show the dynamic phenomena that affect the SBW system.

3. PD plus High Gain GPI Observer

The design of the controller and observer for both the steering wheel and the steering rack is very similar, but for the sake of clarity, each controller and observer would be presented separately.

3.1. Steering Wheel Controller and Observer. Based on the dynamic model of the steering wheel or master subsystem (4), the nonlinear terms, including perturbations and unmodeled dynamics, are substituted by a time-variant disturbance term $\varphi(t)$. Lumping all uncertainties and unknown dynamics on the time-varying term $\varphi(t)$ might be considered as a rough approximation; nevertheless under certain conditions on boundedness and smoothness such approximation results are valid (see the appendix and [20, 21]). In state space dynamics representation, $x_1 = \theta_v$ and $x_2 = \dot{\theta}_v$; the model (4) is represented as a linear perturbed system depicted in

$$\begin{aligned} \dot{x}_1 &= x_2, \\ \dot{x}_2 &= \frac{N_v k_1}{J_{ve} R_v} V_v(t) + \varphi(t), \end{aligned} \tag{10}$$

where $\varphi(t)$ is by assumption a uniformly, absolutely, bounded, time-varying function, that represents all nonlinear and

uncertain terms, that is, all disturbances affecting the linear simplified model (see (11)):

$$\varphi(t) = -\frac{1}{J_{ve}}\left\{\left(B_{Tv} + B_v\right)\dot{\theta}_v + \left(F_{Tv} + C_v\right)\operatorname{sign}\left(\dot{\theta}_v\right) + \frac{N_v^2 k_1 k_2}{R_v}\dot{\theta}_v\right\}. \tag{11}$$

Note that the equality established by (11) is enforced by the proposed model (4); if other dynamic effects or perturbations are to be considered, these phenomena would be included in the term $\varphi(t)$ as well, such that the approximated linear model (10) remains valid in local basis. The relation (11) is presented for the sake of completeness, since the term $\varphi(t)$ is estimated by a high gain GPI observer.

Based on the state space representation (10), a high gain GPI observer is proposed for the linear perturbed system, given by (12). The number of integrators required at the high gain GPI observer depends on the order and complexity of the nonlinearities represented by the term $\varphi(t)$. This is equivalent to truncating a Taylor or Fourier series expansion, but in this case perfect approximation would imply an infinite integrators chain. Proper determination of the order of the extended observer would imply knowledge of the nonlinearities of the term $\varphi(t)$ as shown in [25]. The order of the observer and the relation between the high gain and the final estimation bound have been analyzed for some particular systems and under certain considerations [18, 19, 25]. Despite the difficulties in determining the order of the extended observer from a theoretical point of view, for some practical situations they can be adjusted by looking after an approximated model and measurements on the systems. In this particular case the number of integrators was adjusted by trial and error based on the value of the estimation error $\widetilde{e}_v = x_1 - \widehat{x}_1$ of a set of experiments. The term $\widehat{\varphi}_1$ corresponds to an estimate of the disturbance term $\varphi(t)$. Notice that the high gain GPI observer (12) only requires measurement of the steering wheel angle $x_1 = \theta_v$ minimizing the amount of sensors on the system:

$$\begin{aligned}
\dot{\widehat{x}}_1 &= \widehat{x}_2 + \lambda_{7v}\left(x_1 - \widehat{x}_1\right),\\
\dot{\widehat{x}}_2 &= \frac{N_v k_1}{J_{ve} R_v} V_v(t) + \widehat{\varphi}_1 + \lambda_{6v}\left(x_1 - \widehat{x}_1\right),\\
\dot{\widehat{\varphi}}_1 &= \widehat{\varphi}_2 + \lambda_{5v}\left(x_1 - \widehat{x}_1\right),\\
\dot{\widehat{\varphi}}_2 &= \widehat{\varphi}_3 + \lambda_{4v}\left(x_1 - \widehat{x}_1\right),\\
\dot{\widehat{\varphi}}_3 &= \widehat{\varphi}_4 + \lambda_{3v}\left(x_1 - \widehat{x}_1\right),\\
\dot{\widehat{\varphi}}_4 &= \widehat{\varphi}_5 + \lambda_{2v}\left(x_1 - \widehat{x}_1\right),\\
\dot{\widehat{\varphi}}_5 &= \widehat{\varphi}_6 + \lambda_{1v}\left(x_1 - \widehat{x}_1\right),\\
\dot{\widehat{\varphi}}_6 &= \lambda_{0v}\left(x_1 - \widehat{x}_1\right).
\end{aligned} \tag{12}$$

The estimation error of the high gain GPI observer, $\widetilde{e}_v$, satisfies the perturbed dynamics given by (13), which corresponds to a nonhomogeneous linear dynamic system, with input given by the sixth time derivative of φ. Similar convergence analyses are done in [18, 19, 25], where there are also considerations on boundedness and smoothness of the perturbations and uncertainties, as those imposed here on the term φ:

$$\widetilde{e}_v^{(8)} + \lambda_{7v}\widetilde{e}_v^{(7)} + \cdots + \lambda_{2v}\ddot{\widetilde{e}}_v + \lambda_{1v}\dot{\widetilde{e}}_v + \lambda_{0v}\widetilde{e}_v = \varphi^{(6)}. \tag{13}$$

Since φ is by assumption a uniformly, absolutely, bounded, and time-varying function, then, to achieve convergence of the error dynamics to a small vicinity around the origin of the estimation error phase coordinates, the gains λ_{iv}, $i = 0, 1, 2, \ldots, 7$ are chosen such that the roots of the associated characteristic polynomial are located far enough into the left half of the complex plane, that usually implies high gain values. From the results reported on ADRC, see [18, 19], it follows that closed loop stability properties can be established by considering slow and fast dynamics, usually of the control and the observer, respectively. This is, the observer dynamics should be faster than those of the controller, and for that a pole location problem can be considered based on (13). Among all possibilities to select the gains λ_{iv}, $i = 0, 1, 2, \ldots, 7$, a simple selection of the gains is done by assigning the associated characteristic polynomial to a suitable polynomial with known adequate roots. Since the order of the high gain observer error dynamics (13) is an even number, then a power of a second-order polynomial might be considered, as shown in (14), with $\xi_{\text{obs},v}$ and $\omega_{n(\text{obs},v)}$ being coefficients related to damping and natural frequency, respectively:

$$s^8 + \lambda_{7v}s^7 + \lambda_{6v}s^6 + \cdots + \lambda_{2v}s^2 + \lambda_{1v}s + \lambda_{0v} = \left(s^2 + 2\xi_{\text{obs},v}\omega_{n(\text{obs},v)}s + \omega_{n(\text{obs},v)}^2\right)^4. \tag{14}$$

Given that the observer (12) explicitly estimates the perturbation term $\varphi(t)$, then it is possible to propose a PD control with disturbance rejection (15). $\dot{\theta}_v^*(t)$ and $\ddot{\theta}_v^*(t)$ are the desired angular velocity and acceleration on the steering wheel, which might be obtained from the corresponding time derivatives of the angular position imposed by the driver $\theta_v^*(t)$, although this might raise problems due to noise:

$$V_v(t) = -\frac{J_{ve}R_v}{k_1 N_v}\left[\widehat{\varphi}_1 + \omega_{n(c,v)}^2\left(\theta_v - \theta_v^*\right) + 2\xi_{c,v}\omega_{n(c,v)}\left(\widehat{x}_2 - \dot{\theta}_v^*\right) - \ddot{\theta}_v^* + K_h\widehat{\sigma}_1\right]. \tag{15}$$

Notice that in (15), $\widehat{x}_2$ corresponds to the estimate of the steering wheel angular velocity $\dot{\theta}_v(t)$, and $\widehat{\varphi}_1$ is the estimated disturbance term acting on the steering wheel subsystem $\varphi(t)$, so that both estimates are simultaneously obtained by the high gain GPI observer (12). Meanwhile, the term $\widehat{\sigma}_1$ corresponds to the estimate of the disturbance signal on the steering rack subsystem and it is also obtained by a high gain GPI observer designed for such particular system (19).

By defining the tracking error at the steering wheel subsystem as $e_v = \theta_v - \theta_v^*$, then the closed loop dynamics formed

by (10) with the controller (15) and the high gain observer (12) are given by the nonhomogeneous linear equation (16). Once more following the results from [18, 19], selection of the control gains should imply that the observer error dynamics (13) must be faster than the controller dynamics:

$$\ddot{e}_v + 2\xi_{c,v}\omega_{n(c,v)}\dot{e}_v + \omega_{n(c,v)}^2 e_v = \left(\varphi - \widehat{\varphi}_1\right) + 2\xi_{c,v}\omega_{n(c,v)}\dot{\widetilde{e}}_v - K_h\widehat{\sigma}_1. \tag{16}$$

The closed loop tracking error dynamics (16) are excited by three terms, from which the terms $(\varphi - \widehat{\varphi}_1)$ and $\dot{\widetilde{e}}_v$ correspond to estimation errors of the high gain GPI observer (12), but as shown before, by a proper selection of the GPI observer gains, these errors might be arbitrarily small around a vicinity of zero. The term $-K_h\widehat{\sigma}_1$, on the other hand, cannot be neglected, since it recreates the feedback reflection from the tire system to the wheel system. This feedback force reflection is recreated by inducing a tracking error between the driver commanded angular position and the wheel angular position settled by the wheel DC motor. This fact establishes a commitment between reflecting forces from the tire system and forcing the tracking error e_v to zero. Asymptotic convergence of the tracking error e_v to a small vicinity around zero can be obtained by a proper selection of the gains $\xi_{c,v}$ and $\omega_{n(c,v)}$.

The feedback reflection of the estimated disturbance term $\widehat{\sigma}_1$ from the steering rack subsystem to the steering wheel subsystem yields the haptic loop in the form of a feedback signal proportional to the lumped perturbation torques present at the rack system. This recreates, to the driver, the forces and dynamical effects of torques affecting the steering rack and tire mechanism. The gain K_h weighs the amount of haptic feedback to the driver and can be related to driving sensation and feeling of the road conditions.

Remark 2. Note that in this proposed approach the driver "feels" proportional forces and disturbances present at the steering rack and tire, with proportional gain K_h, which is the simplest feedback situation. In works such as [12], active modification of the handling and steering system are proposed based on composed functions or adaptive gains. Such active feedback adjustment is common in vehicles where different stiffness is felt at low and high vehicle speed, as well as a modification of the steering angle ratio between steering wheel and tire is induced. Such modifications can be introduced through the proposed controller (15) by defining a feedback function of the estimated variable $\widehat{\sigma}_1$.

3.2. Steering Rack Controller and Observer. From Figures 2, 3, 4, and 5, it is straightforward to find out the similarities between the steering wheel subsystem and the rack and tire subsystem. Due to the similarities between both subsystems, also a high gain observer and PD controller with disturbance rejection similar to (12) and (15) are proposed for the steering rack and tire subsystem. Thus, based on the dynamic model of the steering rack subsystem (5), the nonlinear terms and uncertain dynamics are substituted by a time-variant disturbance term $\sigma(t)$. Then, the state space dynamic model (5) is represented as in (17), with states $y_1 = \theta_r$ and $y_2 = \dot{\theta}_r$:

$$\begin{aligned} \dot{y}_1 &= y_2, \\ \dot{y}_2 &= \frac{N_r k_1}{J_{re}R_r}V_r(t) + \sigma(t) \end{aligned} \tag{17}$$

with

$$\sigma(t) = -\frac{1}{J_{re}}\left\{\left(B_{Tr} + B_r\right)\dot{\theta}_r + \left(F_{Tr} + C_r\right)\operatorname{sign}\left(\dot{\theta}_r\right) + \frac{N_r^2 k_1 k_2}{R_r}\dot{\theta}_r + \tau_{\text{SAT}}\right\}. \tag{18}$$

For the steering rack and tire subsystem, the equality established by (18) is enforced by the proposed model (5), if other dynamic effects or perturbations are to be considered, these phenomena would be included in the term $\sigma(t)$ as well, such that the approximated linear model (17) remains valid in local basis. The term $\sigma(t)$ would be estimated by a high gain GPI observer (19), its estimate is denoted by the variable $\widehat{\sigma}_1$:

$$\begin{aligned} \dot{\widehat{y}}_1 &= \widehat{y}_2 + \lambda_{7r}\left(y_1 - \widehat{y}_1\right), \\ \dot{\widehat{y}}_2 &= \frac{N_r k_1}{J_{re}R_r}V_r(t) + \widehat{\sigma}_1 + \lambda_{6r}\left(y_1 - \widehat{y}_1\right), \\ \dot{\widehat{\sigma}}_1 &= \widehat{\sigma}_2 + \lambda_{5r}\left(y_1 - \widehat{y}_1\right), \\ \dot{\widehat{\sigma}}_2 &= \widehat{\sigma}_3 + \lambda_{4r}\left(y_1 - \widehat{y}_1\right), \\ \dot{\widehat{\sigma}}_3 &= \widehat{\sigma}_4 + \lambda_{3r}\left(y_1 - \widehat{y}_1\right), \\ \dot{\widehat{\sigma}}_4 &= \widehat{\sigma}_5 + \lambda_{2r}\left(y_1 - \widehat{y}_1\right), \\ \dot{\widehat{\sigma}}_5 &= \widehat{\sigma}_6 + \lambda_{1r}\left(y_1 - \widehat{y}_1\right), \\ \dot{\widehat{\sigma}}_6 &= \lambda_{0r}\left(y_1 - \widehat{y}_1\right). \end{aligned} \tag{19}$$

The observer estimation error $\widetilde{e}_r = y_1 - \widehat{y}_1$ satisfies the dynamics given by

$$\widetilde{e}_r^{(8)} + \lambda_{7r}\widetilde{e}_r^{(7)} + \cdots + \lambda_{2r}\ddot{\widetilde{e}}_r + \lambda_{1r}\dot{\widetilde{e}}_r + \lambda_{0r}\widetilde{e}_r = \sigma^{(6)}. \tag{20}$$

The convergence properties of the estimation error at the steering rack $\widetilde{e}_r$ to a small vicinity of the origin can be established in a similar manner as for the steering wheel subsystem. This is assigning the associated characteristic polynomial to a suitable polynomial with known adequate roots, as in (14), ensuring faster observer dynamics than the controller dynamics.

Based on the estimated angular velocity of the steering rack subsystem $\widehat{y}_2$ along with the perturbation estimate $\widehat{\sigma}_1$, an active disturbance rejection control is synthesized in the form of a PD controller with a perturbation term rejection. This controller is proposed in (21). Notice that the reference

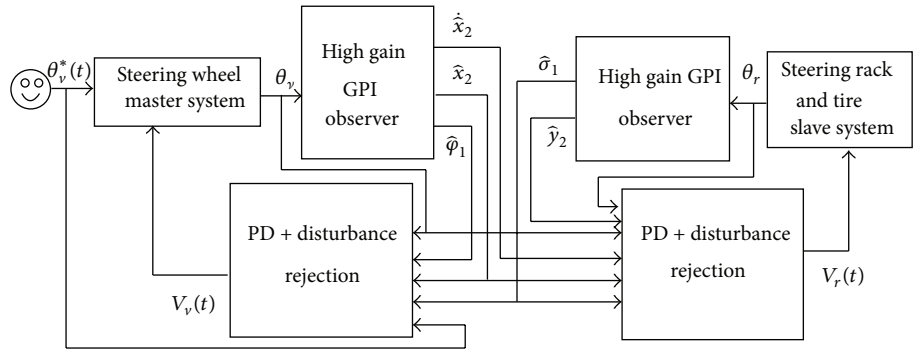


Figure 7: SBW schematic interconnections.

trajectory for the PD steering rack controller corresponds to the steering wheel angle $\theta_v(t)$ and the angular velocity and acceleration estimates $\hat{x}_2$ and $\dot{\hat{x}}_2$, which are obtained by the high gain GPI observer (12). Therefore, the closed loop system can be seen as a master-slave system with a haptic loop and bilateral communication. The haptic loop is due to feedback reflecting the estimate of the steering rack perturbation term $\hat{\sigma}_1$ to the steering wheel PD control, given by (15). Thus, through the steering wheel the human operator is able to feel the dynamic phenomena that affect the steering rack and tire:

$$V_r(t) = -\frac{J_{re}R_r}{k_1 N_r}\left[\hat{\sigma}_1 + \omega_{n(c,r)}^2\left(\theta_r - \theta_v(t)\right) + 2\xi_{c,r}\omega_{n(c,r)}\left(\hat{y}_2 - \hat{x}_2\right) - \dot{\hat{x}}_2\right]. \tag{21}$$

By defining the tracking error at the steering rack and tire subsystem as $e_r = \theta_r - \theta_v$, then the closed loop dynamics formed by (17) with the controller (21) and the high gain observer (19) are given by the nonhomogeneous linear equation

$$\ddot{e}_r + 2\xi_{c,r}\omega_{n(c,r)}\dot{e}_r + \omega_{n(c,r)}^2 e_r = \left(\sigma - \hat{\sigma}_1\right) + 2\xi_{c,r}\omega_{n(c,r)}\left(\dot{\tilde{e}}_v - \dot{\tilde{e}}_r\right) - \ddot{\tilde{e}}_v. \tag{22}$$

The closed loop tracking error dynamics (22) are excited by estimation errors of the high gain GPI observers (12) and (19); nevertheless, by a proper selection of the GPI observer gains, these errors might be arbitrarily small around a vicinity of zero. Thus asymptotic convergence of the tracking error e_r to a small vicinity around zero can be obtained by a proper selection of the gains $\xi_{c,r}$ and $\omega_{n(c,r)}$.

Notice that the full control system given by the controllers (15) and (21) together with the high gain GPI observers (12) and (19) only require measurement of the steering wheel angle θ_v and the tire orientation angle θ_r. It is also important to point out that the proposed approach is based on the input gain of the systems; thus minimum knowledge of the dynamic models is required. Figure 7 shows a schematic representation of the proposed SBW system and the interconnections between the steering wheel and steering rack and tire subsystem.

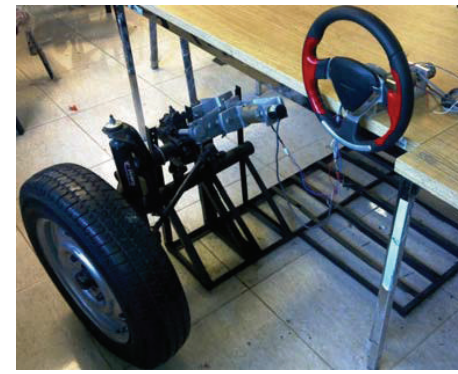


Figure 8: SBW experimental platform.

4. Experimental Results

For experimental purposes, a low cost platform has been built (see Figure 8). The platform is provided with a steering wheel, and half of the steering rack of a beetle VW vehicle, including suspension system and tire. This type of steering system is still available in economic commercial cars. The steering column was modified as shown in Figures 2 and 4, including 2 DC motors and encoders. The communication and control are programmed in a PC with Windows as operative system.

The whole control approach of the SBW haptic system, PD, and high gain GPI observer (12) and (15) for the steering wheel and (19) and (21) for the steering rack are programmed using general blocks of Simulink available in MATLAB. The electronics and mechanical components of the experimental platform are listed as follows. The angular positions on the steering wheel and on the rack are measured by incremental encoders OMRON, model E6B2-CWZ1X, 2000PPR, 0.5 M. The PC is equipped with a data acquisition card Sensoray model 626, which is programmed to work with a sampling period of 0.0005 seconds. Two operational amplifiers, model STK4050 II, are used to condition the control voltages $V_v(t)$ and $V_r(t)$ that are sent to the DC motors Nisca Motor model NC5475B. Each motor is connected to a gearbox, such that for the steering wheel a reduction ratio of $N_v = 16$ is obtained, and for the steering rack the reduction ratio is $N_r = 48$. Figure 9 shows a schematic connection of the SBW haptic system.

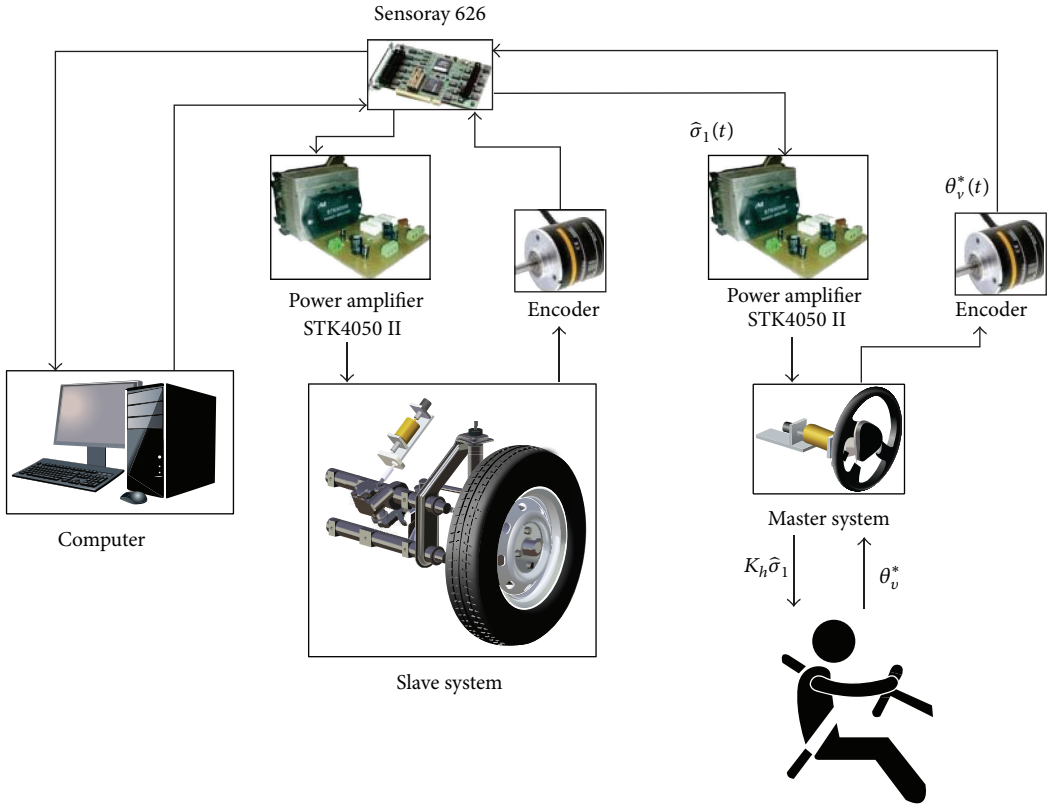


Figure 9: Schematic diagram of the experimental SBW.

Following the results from [18, 19], selection of the control and observer gains should imply that the observer error dynamics must be faster than the controller dynamics. All gains were tuned by using (14), by considering pure real roots and taking into account the performance of the tracking errors e_v and e_r. For the master subsystem the control gains involved in (15) are $\xi_{c,v} = 1.01$, $\omega_{n(c,v)} = 5.92$, and for its GPI observer (12) $\xi_{\text{obs},v} = 2.2$, $\omega_{n(\text{obs},v)} = 15.2$ meanwhile, for the steering rack (slave subsystem) the control gains, see (21), they are $\xi_{c,r} = 2.53$, $\omega_{n(c,r)} = 5.92$, and for the GPI observer (19) the gains are $\xi_{\text{obs},r} = 10$, $\omega_{n(\text{obs},r)} = 25$.

Two cases are compared: first the SBW is tested without traction, and second tire traction at a speed of 9.4 m/s (33.84 km/hr) is supplied. This comparison is to look after how tire velocity affects the aligning forces and thus the reaction forces presented at tire and steering rack subsystem. The same platform as sketched in Figure 9 is used for both cases, with and without tire traction; only the motor in charge of traction is turned on or off; this emulates steering the vehicle when being parked and in motion. It is also important to emphasize that the controller and observers are the same in both cases; that is, nothing is changed in their structure. This fact implies a realistic and robust application of the proposed approach in a vehicle, such that the same controller and observers work for all possible vehicle scenarios. The control gains that are used in both cases are those described at previous paragraph, which were tuned for the case without tire traction. For the sake of comparison first the results without traction are presented; then the results with traction are shown.

4.1. SBW System without Tire Traction. This case emulates steering the vehicle when being parked, such that greater effort is required to steer the tire than when being in motion. In some commercial vehicles, a larger steering ratio, between steering wheel and tire, is set when the vehicle moves at low speed or is standstill. This produces a less stiff or softer steering wheel sensation. In this experiment and for comparison purposes the same steering ratio and feedback gain for the haptic loop (K_h) are kept for traction and without traction cases.

The convergence between the steering rack angle (tire angle orientation) θ_r and the human desired angular position given by the desired steering wheel angle θ_v^* allows concluding good tracking and performance of the SBW system; see Figure 10. This is further supported by the tracking error $e = \theta_r - \theta_v^*$ that converges to a small vicinity around zero; see Figure 11. Thus, it can be concluded that the high gain GPI observer (19) properly estimates the disturbance signal at the steering rack and tire $\widehat{\sigma}_1(t)$; see Figure 12. Therefore the PD plus disturbance compensation (21) is able to reject the perturbation signal effects. On the other hand, feedback of the dynamic perturbation phenomena $\widehat{\sigma}_1(t)$ from the steering rack to the steering wheel generates a haptic loop such that the operator "feels" the dynamic perturbations that affect the

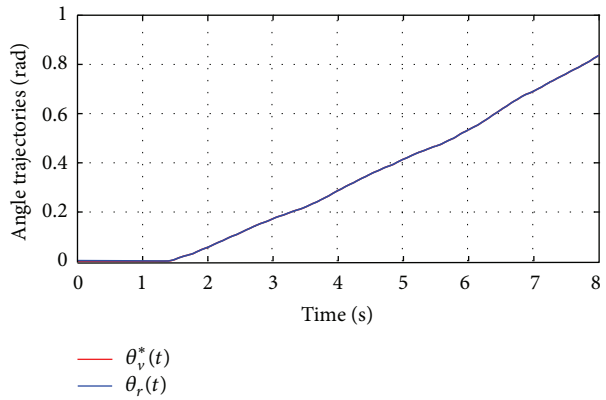


FIGURE 10: Rack angular position and its desired human reference $\theta_v(t)^*$ without tire traction.

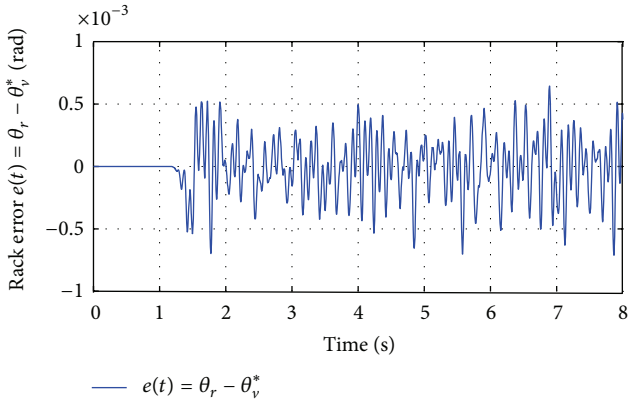


FIGURE 11: Rack tracking error $e_r(t)$ without tire traction.

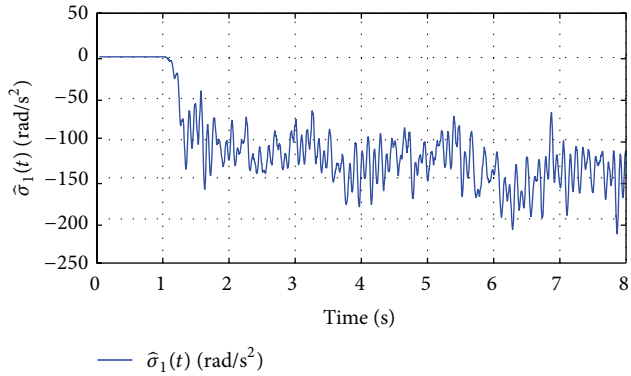


FIGURE 12: Disturbance steering rack signal $\widehat{\sigma}_1(t)$ without tire traction.

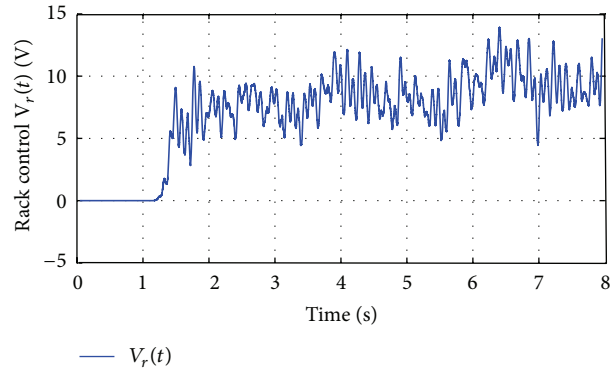


FIGURE 13: Voltage at the steering rack DC motor $V_r(t)$ without tire traction.

steering rack and tire. The DC motor voltage on the steering rack subsystem $V_r(t)$ required to properly track the rack angular reference is shown in Figure 13.

4.2. SBW System with Tire Traction. This case emulates steering the vehicle when being in motion, such that less effort is required to steer the tire than when being parked. In some commercial vehicles a smaller steering ratio, between steering wheel and tire, is set as higher the vehicle speed, as well as a stiffer steering wheel feeling, this is the so called "sport" mode. Although such adaptive response of the SBW can be considered in the proposed controller (15), in this experiment and for comparison purposes the same steering ratio and feedback gain for the haptic loop (K_h) are kept for traction and without traction cases.

When tire traction is considered (9.4 m/s), convergence between the steering rack angle θ_r and the human desired angular position θ_v^* allows concluding good tracking and performance of the SBW system; see Figure 14. This is further supported by the tracking error $e_r = \theta_r - \theta_v^*$ that converges to a small vicinity around zero; see Figure 15. Thus, the high gain GPI observer (19) properly estimates the perturbation signal at the steering rack and tire $\widehat{\sigma}_1(t)$; see Figure 16. Thus, the PD plus disturbance compensation (21) is able to reject the perturbation signal effects. On the other hand, feedback of the disturbance phenomena $\widehat{\sigma}_1(t)$ from the steering rack to the steering wheel generates a haptic loop such that the operator "feels" the dynamic perturbation that affect the steering rack and tire. The DC motor voltage on the steering rack subsystem $V_r(t)$ required to properly track the rack angular reference is shown in Figure 17.

4.3. Comparison Results and Discussion. It is important to point out that the commanded steering wheel reference was provided by a human operator steering the wheel. Therefore, although the same desired trajectory was intended, there are differences on such reference; see Figures 10 and 14. Nevertheless, some conclusions can be drawn from the comparison study.

In both cases, with and without tire traction, the performance of the SBW is rather acceptable; that is, the tracking error $e = \theta_r - \theta_v^*$ converges to a small vicinity around zero, as predicted by the closed loop error dynamics (22); that is, $\theta_r \to \theta_v \to \theta_v^*$; see Figures 11 and 15. It is important to note that the control and observer gains were tuned for the without tire traction case; nevertheless, the magnitude and performance of the tracking error $e = \theta_r - \theta_v^*$ is very similar

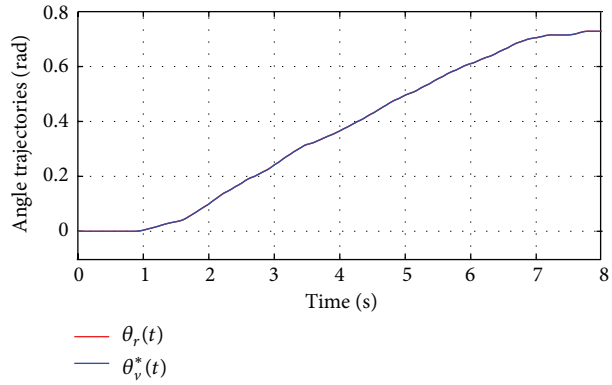


FIGURE 14: Rack angular position and its desired human reference $\theta_v(t)^*$ with tire traction.

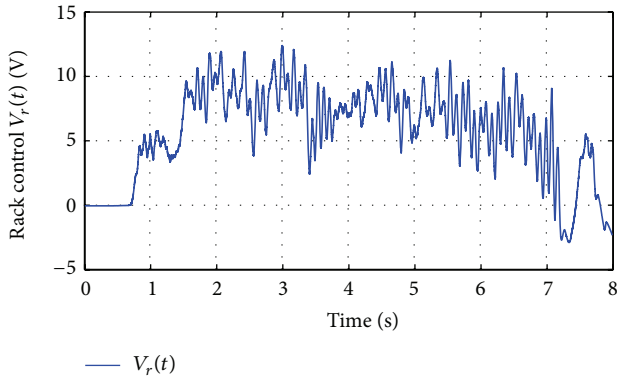


FIGURE 15: Rack tracking error $e_r(t)$ with tire traction.

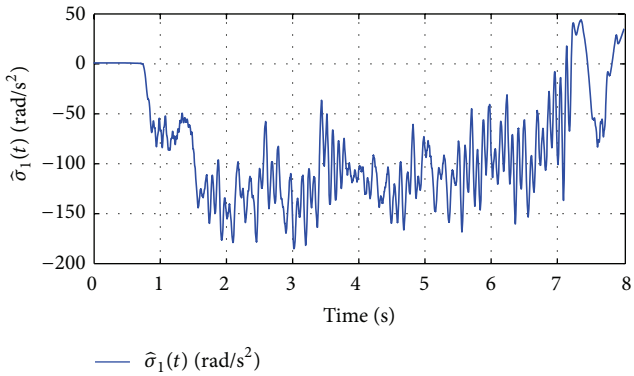


FIGURE 16: Disturbance steering rack signal $\widehat{\sigma}_1(t)$ with tire traction.

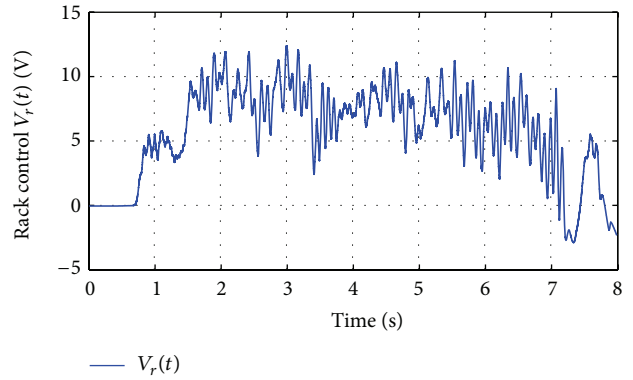


FIGURE 17: Voltage at the steering rack DC motor $V_r(t)$ with tire traction.

for both cases. Therefore, it can be concluded that uncertainty and dynamic perturbations are properly estimated and compensated by the high gain GPI observers in both cases, with and without traction.

Even though similar tracking error convergence $e = \theta_r - \theta_v^*$ is obtained for both cases, there are relevant differences on magnitude and shape of the estimated perturbation term $\widehat{\sigma}_1(t)$ (Figures 12 and 16), as well as for the control action $V_r(t)$ (Figures 13 and 17). These differences are due to the effects of tire traction, particularly, aligning torques and road-tire friction. The estimated perturbation signals $\widehat{\sigma}_1(t)$ show that for the case with traction, the perturbation torque is smaller in magnitude that for the without tire traction. This fact translates in the driver feeling less effort to steer the wheel when there is traction, than without traction. This "feeling" is common when driving a vehicle at high speed, when less effort is required to steer, compared to the steering effort when the vehicle speed is very low or even being parked.

Notice that some commercial vehicles actively adapt the steering ratio, between steering wheel and tire, and stiffness steering wheel feeling, as function of the vehicle speed, for example, the parking and sport mode. Although such adaptive response of the SBW can be considered in the proposed controller (15), in the comparison study here carried out the same steering ratio and feedback gain for the haptic loop (K_h) are kept for traction and without traction cases. Adaptive behavior of the proposed SBW haptic system can be considered by modifying the feedback of the disturbance and dynamic effects of the steering rack and tire.

5. Conclusions and Perspectives

This paper has presented a haptic steering by wire system based on high gain GPI observers. The high gain GPI technique requires a minimum amount of information from the dynamic model; only the nonlinear system input gain is required; as for sensors only encoders are employed, which is an advantage compared to other techniques of SBW systems that required specialized sensors. The tracking error between the desired angle settled by the human operator, the steering wheel angle, and the tire orientation angle can be made arbitrarily small by a proper gain selection.

The technique is capable of estimating the perturbation inputs and the uncertain terms that affect the SBW dynamics. In particular, the SBW system represents a twofold application of the GPI observer: on the one hand it helps in rejection or compensating the perturbation in each subsystem; on the other hand the estimated perturbations on the steering rack are fed back to the steering wheel to yield a haptic system, including tire self-aligning torque.

It is worth to notice that the high gain GPI observer together with the PD control relies only on measurements of an encoder, thus highly diminishing the use of sensors. Then the proposed SBW system behaves as a self-contained sensor for steering rack and tire forces. To validate this last conclusion, as future work, the installation of force, acce leration, and other sensors are being considered. This will allow establishing clear comparisons between the estimated perturbation and measurements on the experimental platform.

Appendix

Ultra Models

Considering the single-input single-output, n-dimensional perturbed nonlinear system is given by (A.1). The initial conditions at t_0 are given by the set $\mathscr{Y}_0 = \{y_0, \dot{y}_0, \ldots, y_0^{(n-1)}\}$. Assume that the input $u = u(t)$ and the gain input function $\phi(t, y)$ are perfectly known. The nonlinear input gain function $\phi(t, y(t))$ is assumed to be bounded and uniformly bounded away from zero:

$$y^{(n)} = \psi\left(t, y, \dot{y}, \ldots, y^{(n-1)}\right) + \phi\left(t, y\right) u. \tag{A.1}$$

Definition A.1. Let $\xi(t)$ be a time function, which represents the additive nonlinear term $\psi(t, y, \dot{y}, \ldots y^{(n-1)})$ in (A.1), particularized for $y = y(t)$, where the function $y(t)$ is the solution of (A.1) for a given input function $u(t)$ and for a set of initial conditions $\mathscr{Y}_0$. That is,

$$\xi\left(t\right) = \psi\left(t, y\left(t\right), \dot{y}\left(t\right), \ldots, y^{(n-1)}\left(t\right)\right). \tag{A.2}$$

Defining a global *ultra model* of the system (A.1), the system represented by (A.3), subject to the corresponding set of initial conditions $\mathscr{Z}_0$ satisfies $\mathscr{Z}_0 = \mathscr{Y}_0$:

$$z^{(n)} = \xi\left(t\right) + \phi\left(t, z\left(t\right)\right) u\left(t\right). \tag{A.3}$$

The term $\xi(t)$ in (A.3) is obviously *a priori* unknown; nevertheless, it is algebraically observable in the sense of Diop and Fliess [26]. The uncertainty in $\xi(t)$ is basically due to parameters and unknown nonlinearities in the term $\psi(\cdot)$ in (A.1). It may also represent exogenous perturbations which are unmodeled. As a time function, $\xi(t)$, and a finite number of its time derivatives (e.g., p derivatives) are assumed to be uniformly absolutely bounded. According to Gliklikh [27], without this hypothesis, there is not any solutions to the original system. In open loop, the nominal control input $u^*(t)$, corresponds, due to differential flatness of the system [28], to an smooth output reference trajectory $y^*(t)$.

Notice that for any bounded input function $u(t)$, and a given initial condition set $\mathscr{Y}_0$, the solution of system (A.1) $y(t)$ satisfies, as a time function, the identity (A.4) [29]:

$$y^{(n)} = \psi\left(t, y, \dot{y}\right), \ldots, y^{(n-1)} + \phi\left(t, y\right) u\left(t\right). \tag{A.4}$$

Theorem A.2. *The system* (A.3) *is trajectory equivalent to the nonlinear system* (A.1)*. That is, both systems present exactly the same solution trajectories $y(t) = z(t)$ for all t, for a particular common initial condition set $\mathscr{Y}_0$.*

Proof. Define $\varepsilon = y(t) - z(t)$; then (A.3) and (A.1) imply that ε satisfies $\varepsilon^{(n)}(t) = [\phi(t, y(t)) - \phi(t, z(t))]u(t)$, with zero initial conditions, $\varepsilon^{(k)}(t_0) = 0$, $k = 0, 1, \ldots, n-1$. Meanwhile, $\varepsilon^{(n)}(t_0) = 0$ and $\varepsilon^{(n+j)}(t_0) = 0$, for all j. Therefore, it follows that $\varepsilon(t) = 0$ for all t; thus systems (A.3) and (A.1) are trajectory equivalent.

The consequence of the previously stated fact is that the nonlinear systems (A.1) can be seen as the linear perturbed system (A.3), including the exogenous presence of time-variant term: $\xi(t) = \psi(t, y(t), \ldots, y^{(n-1)}(t))$. These two systems (A.1) and (A.3) are identical on the precise sense that their trajectories are the same for any time interval. Thus, distinction between $z(t)$ and $y(t)$ becomes irrelevant. Any system of the form (A.1) can be examined through the equivalent trajectory of system (A.3), which can be seen, without any ambiguity, as in

$$y^{(n)} = \xi\left(t\right) + \phi\left(t, y\right) u. \tag{A.5}$$

The system (A.5) which is trajectory equivalent to the original nonlinear system (A.1) lacks the state time-dependent additive nonlinear term. In practice, the value of the time-dependent nonlinearities, lumped in the term $\xi(t)$, might be unknown and thus it can be seen as a perturbation to the system. The use of (A.5) greatly simplifies the simultaneous observation problem of the nonlinear perturbation, which is state dependent, and of the state of the system (A.1) itself, based on the input u and output y signals. □

Conflict of Interests

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

Acknowledgment

The authors acknowledge support from CONACyT, Mexico, through Project 133527.

References

[1] R. Bosch, *Bosch Automotive Handbook*, John Wiley & Sons, Chicester, UK, 7th edition, 2008.

[2] J. Reimpell, H. Stoll, and J. Betzler, *The Automotive Chassis: Engineering Principles*, Butterworth-Heinemann, London, UK, 2001.

[3] J. Y. Wong, *Theory of Ground Vehicles*, John Wiley & Sons, New York, NY, USA, 1978.

[4] T. K. Garrett, K. Newton, and W. Steeds, *The Motor Vehicle*, Butterworth-Heinemann, London, UK, 2001.

[5] S. Amberkar, M. Kushion, K. Eschtruth, and F. Bolourchi, "Diagnostic development for an electric power steering system," SAE Technical Paper 2000-01-0819, 2000.

[6] D. Peter and R. Gerhard, "Electric power steering—the first step on the way to steer by wire," SAE Technical Paper 1999-01-0401, 1999.

[7] J. W. Post, *Modeling, simulation and testing of automobile power steering system for the evaluation of on-center handling [Ph.D.*

thesis], Graduate School Clemson University, Clemson, SC, USA, 1995.

[8] J. Kim and J. Song, "Control logic for an electric power steering system using assist motor," *Mechatronics*, vol. 12, no. 3, pp. 447–459, 2002.

[9] A. Zaremba and R. I. Davis, "Dynamic analysis and stability of a power assist steering system," in *Proceedings of the American Control Conference*, pp. 4253–4257, June 1995.

[10] A. T. Zaremba, M. K. Liubakka, and R. M. Stuntz, "Control and steering feel issues in the design of an electric power steering system," in *Proceedings of the American Control Conference (ACC '98)*, pp. 36–40, June 1998.

[11] A. E. Cetin, M. A. Adli, D. E. Barkana, and H. Kucuk, "Implementation and development of an adaptive steering-control system," *IEEE Transactions on Vehicular Technology*, vol. 59, no. 1, pp. 75–83, 2010.

[12] P. Yih and J. C. Gerdes, "Modification of vehicle handling characteristics via steer-by-wire," *IEEE Transactions on Control Systems Technology*, vol. 13, no. 6, pp. 965–976, 2005.

[13] A. Baviskar, J. R. Wagner, D. M. Dawson, D. Braganza, and P. Setlur, "An adjustable steer-by-wire haptic-interface tracking controller for ground vehicles," *IEEE Transactions on Vehicular Technology*, vol. 58, no. 2, pp. 546–554, 2009.

[14] A. Bertacchini, L. Tamagnini, and P. Pavan, "Force feedback in steer-by-wire systems: architecture and experimental results," in *Proceedings of the International Symposium on Industrial Electronics 2006 (ISIE '06)*, pp. 3050–3055, IEEE, Montreal, Canada, July 2006.

[15] B.-H. Nguyen and J.-H. Ryu, "Direct current measurement based steer-by-wire systems for realistic driving feeling," in *Proceedings of the International Symposium on Industrial Electronics (ISIE '09)*, pp. 1023–1028, Seoul, Republic of Korea, July 2009.

[16] L. Li, F. Wang, and Q. Zhou, "Integrated longitudinal and lateral tire/road friction modeling and monitoring for vehicle motion control," *IEEE Transactions on Intelligent Transportation Systems*, vol. 7, no. 1, pp. 1–19, 2006.

[17] C. Kim, K. Hong, W. Yoo, and Y. Park, "Tire-road friction estimation for a wheel-driven field robot," in *Proceedings of the Society of Instrument and Control Engineers (SICE '08)*, pp. 782–787, Tokyo, Japan, August 2008.

[18] B. Guo and Z. Zhao, "On convergence of the nonlinear active disturbance rejection control for MIMO systems," *SIAM Journal on Control and Optimization*, vol. 51, no. 2, pp. 1727–1757, 2013.

[19] W. Zhou, S. Shao, and Z. Gao, "A stability study of the active disturbance rejection control problem by a singular perturbation approach," *Applied Mathematical Sciences*, vol. 3, no. 10, pp. 491–508, 2009.

[20] H. Sira-Ramírez, "On the generalized PI sliding mode control of DC-to-DC power converters: a tutorial," *International Journal of Control*, vol. 76, no. 9-10, pp. 1018–1033, 2003.

[21] H. Sira-Ramirez, F. Gonzalez-Montanez, J. A. Cortes-Romero, and A. Luviano-Juarez, "A robust linear field-oriented voltage control for the induction motor: experimental results," *IEEE Transactions on Industrial Electronics*, vol. 60, no. 8, pp. 3025–3033, 2013.

[22] R. E. Bou Serhal and H. K. Khalil, "Application of the extended high gain observer to underactuated mechanical systems," in *Proceedings of the American Control Conference (ACC '12)*, pp. 4727–4732, June 2012.

[23] Y. Yamaguchi and T. Murakami, "Adaptive control for virtual steering characteristics on electric vehicle using steer-by-wire system," *IEEE Transactions on Industrial Electronics*, vol. 56, no. 5, pp. 1585–1594, 2009.

[24] J. Im, F. Ozaki, N. Matsunaga, and S. Kawaji, "Control of steering-by-wire system using bilateral control scheme with passivity approach," in *Proceedings of the International Conference on Control, Automation and Systems (ICCAS '07)*, pp. 1488–1493, October 2007.

[25] H. Khalil and L. Praly, "High-gain observers in nonlinear feedback control," *International Journal of Robust and Nonlinear Control*, vol. 24, no. 6, pp. 993–1015, 2014.

[26] S. Diop and M. Fliess, "Nonlinear observability, identifiability, and persistent trajectories," in *Proceedings of the 30th IEEE Conference on Decision and Control*, pp. 714–719, Brighton, UK, December 1991.

[27] Y. E. Gliklikh, "Necessary and sufficient conditions for global-in-time existence of solutions of ordinary, stochastic, and parabolic differential equations," *Abstract and Applied Analysis*, vol. 2006, Article ID 39786, 17 pages, 2006.

[28] H. Sira-Ramirez and S. Agrawal, *Differentially Flat Systems*, Marcel Dekker, New York, NY, USA, 2004.

[29] L. S. Pontryagin, *Differential Equations*, Addison-Wesley, Reading, Mass, USA, 1962.

Hindawi Publishing Corporation
Mathematical Problems in Engineering
Volume 2014, Article ID 712360, 15 pages
http://dx.doi.org/10.1155/2014/712360

Research Article

Sample-Data Modeling of a Zero Voltage Transition DC-DC Converter for On-Board Battery Charger in EV

Teresa R. Granados-Luna,[1] Ismael Araujo-Vargas,[2] and Francisco J. Perez-Pinal[3]

[1] *Coacalcos Institute of Tecnological Studies, 16 de Septiembre Avenue No. 54, Col. Cabecera Municipal, 55700 Coacalco de Berriozabal, MEX, Mexico*
[2] *School of Mechanical and Electrical Engineering, National Polytechnic Institute of Mexico, ESIME Cul., Santa Ana Avenue No. 1000, Col. San Francisco Culhuacan, 04430 Coyoacan, DF, Mexico*
[3] *Automotive Mechanical Engineering Department, Polytechnic University of Pachuca, Ex Hacienda de Santa Barbara, Carretera Pachuca Cd. Sahagún, Km. 20, 43830 Zempoala, HGO, Mexico*

Correspondence should be addressed to Ismael Araujo-Vargas; iaraujo@ipn.mx

Received 30 November 2013; Accepted 5 February 2014; Published 2 June 2014

Academic Editor: Sheldon S. Williamson

Copyright © 2014 Teresa R. Granados-Luna 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.

Battery charger is a key device in electric and hybrid electric vehicles. On-board and off-board topologies are available in the market. Lightweight, small, high performance, and simple control are desired characteristics for on-board chargers. Moreover, isolated single-phase topologies are the most common system in Level 1 battery charger topologies. Following this trend, this paper proposes a sampled-data modelling strategy of a zero voltage transition (ZVT) DC-DC converter for an on-board battery charger. A piece-wise linear analysis of the converter is the basis of the technique presented such that a large-signal model and, therefore, a small-signal model of the converter are derived. Numerical and simulation results of a 250 W test rig validate the model.

1. Introduction

Advanced vehicular systems are based on the more electric systems (MES) concept. MES is the intensive application of power electronic converters (PEC), electric machines (EM), and advanced embedded control systems to aeronautical, automotive, and maritime systems. MES was initially applied to aeronautical systems toward the reduction and/or substitution of mechanical, pneumatic, and hydraulic systems, that is, the more electric aircraft (mea) and totally integrated more electric systems (TIMES), [1]. MES are more efficient compared to their counterparts due to (i) small utilization of electric energy, (ii) high energy efficiency, (iii) reduced weight, and (iv) low maintenance [2]. After that, MES was implemented in automotive sector resulting in the more electric vehicle (MEV). MEV includes electric vehicles (EV), hybrid electric vehicles (HEV), and plug-in hybrid electric vehicles (PHEV) [3]. In particular, MES applied to vehicular systems has become popular due to the market introduction of the HEV Toyota Prius in 1997 [4]. HEV are being developed by companies like BMW, Chrysler, Daimler AG, General Motors, PSA Peugeot Citroen, Suzuki Motor Corp, Toyota, and Volkswagen. Motivations to develop EV, HEV, and PHEV are based on economic, environmental, and energetic facts. Regardless of these kinds of configurations, at least two different sources of energy are needed to achieve the same performance compared to an internal combustion engine (ICE). Indeed, at least one EM and PEC are needed in the propulsion stage at any EV, HEV, and PHEV configuration. Series, parallel, series/parallel, and integrated starter alternator (ISA) with its optional plug-in capability are typical configurations available in the market.

PHEV uses an off-board or on-board charger similar to EV. The standard SAEJ1772 is used in North America and comprises three charge methods: AC level 1 (supply voltage varies from 120VAC 1-phase), AC level 2 (208V to 240VAC and 600V DC maximum; with a maximum current (amps-continuous) from 12A, 32A and 400A), and DC charging.

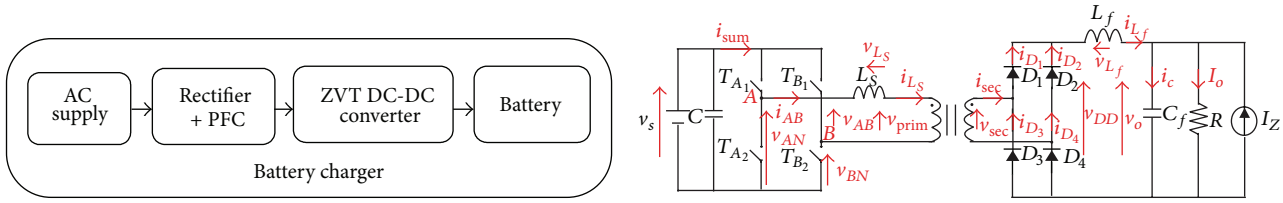


Figure 1: System configuration (a) block diagram and (b) phase-shift controlled ZVT DC-DC converter.

Additionally, SAEJ1772 provides a guide to the AC level-3 vehicle, an on-board charger capable of accepting energy from an AC supply source at a nominal voltage of 208V and 240VAC and a maximum current of 400A. In addition, SAEJ1772 provides information about the coupler requirements, general electric vehicle supply equipment (EVSE) requirements, control and data, and general conductive charging system description [5].

Single- and three-phase, isolated and nonisolated, and unidirectional or bidirectional configurations have been proposed in literature as battery chargers, such as reported in [6]. Methods to improve their performance are using one or several combinations of the following techniques: power factor correction (PFC); interleaved, multicell, and resonant configurations; soft/hard switching; zero voltage switching (ZVS); and zero current switching (ZCS). Moreover, the control algorithms include proportional integral (PI), proportional-integral-derivative (PID), sliding modes, fuzzy logic, and adaptive neural network. Following this trend, this paper proposed a sample-data modeling strategy of a DC-DC ZVT to understand its dynamic characteristics as an on-board battery charger. In this topology, the switches are turned on during zero voltage reducing the switching losses; as a result, a compact, lightweight system with high switching frequency can be designed. A typical peak current method is used in this work for control purposes resulting in a simple and inexpensive control law.

This paper is organized as follows. The principle of operation of the converter is described in Section 2 using idealized waveforms. Then, a mathematical analysis based on a piece-wise linear analysis is provided in Section 3, where a phase control strategy is modeled to obtain a large-signal model of the converter. Using this model, a half-cycle, sample-data linear model is obtained, which helps provide the final small-signal transfer functions of the converter. Numerical and simulation results of a 250 W prototype are presented to validate the model obtained. Final conclusions are summarized in Section 5.

2. Principle of Operation of the Converter

2.1. Circuit Description. A typical system configuration for an EV battery charger is shown in Figure 1(a), which normally consists of a boost power factor corrected (PFC) rectifier connected to an AC supply and a high frequency (HF) DC-DC converter to regulate the load of the batteries. The topology of the DC-DC ZVT converter is shown in Figure 1(b), which has a full bridge inverter supplied with a DC voltage source; a HF transformer with a turns ratio of $1:N$ to generate a quasisquare, phase-controlled wave; a stray inductance connected in series to the inverter output, mainly formed by the leakage inductance of the transformer; a full-bridge rectifier connected to the transformer secondary side; and, then, a LC filter to smooth the pulsating rectified voltage waveform of the output of the rectifier. The model also considers a disturbance current source in parallel with the load.

The left-hand leg of the inverter, denoted by leg A, is used as the reference to describe the converter operation. The switches of leg A operate complementarily with fixed duty ratios of 50% at high frequency. The switches of the right-hand inverter leg, leg B, also operate complementarily with fixed duty ratios of 50%, but the operation of leg B is delayed by $\delta T/2$ respective to leg A, where T is the switching period and δ is the phase control variable, which ranges from 0 to 1.

The steady state operation of the circuit of Figure 1 may be explained with the steady state, voltage, and current waveforms of Figure 2.

The first four waveforms shown in Figure 2 are the states of the switches of the inverter leg A. Then, the third and fourth waves show the states of the leg B switches, which are delayed by $\delta T/2$ respective to the first and second waves of Figure 2. The fifth and sixth waveforms of Figure 2 are the inverter output voltage, $v_{AB} = v_{A_2} - v_{B_2}$, and output current, i_{AB} (also i_{L_S}). The next two waveforms of this figure are the rectifier output voltage, v_{DD}, and the filter inductor current, i_{L_f}. i_{L_f} is a continuous wave with a small ripple component, which is also present in i_{L_S}, but amplified by the turns ratio N and reverted during the negative semicycle of v_{AB}. The last three waveforms of Figure 2 are the current of diodes D_1 to D_4, i_{D_1} to i_{D_4}, and the supply current waveform i_{sum}. When $v_{AB} = V_{\text{in}}$, the current i_{L_S} is positive and flows through D_1 and D_4, whereas when $v_{AB} = -V_{\text{in}}$, the current i_{L_S} is negative and flows through D_2 and D_3 since these diodes are positively biased. When the v_{AB} waveform changes from zero to $\pm V_{\text{in}}$, the diodes D_1 and D_4 are naturally commutated, short-circuiting the transformer secondary winding due to the overlapped operation of the diodes. The duration of the diodes overlap, T_{OL}, causes a fast reversal of the inverter primary current i_{L_S}, being limited by the inductance L_S, which prevents a short circuit of the inverter output.

The production of T_{OL} may be described using the waves i_{D_1} to i_{D_4} of Figure 2. When v_{AB} changes from zero to V_{in}, the currents i_{D_1} and i_{D_4} rise from zero to the output current level

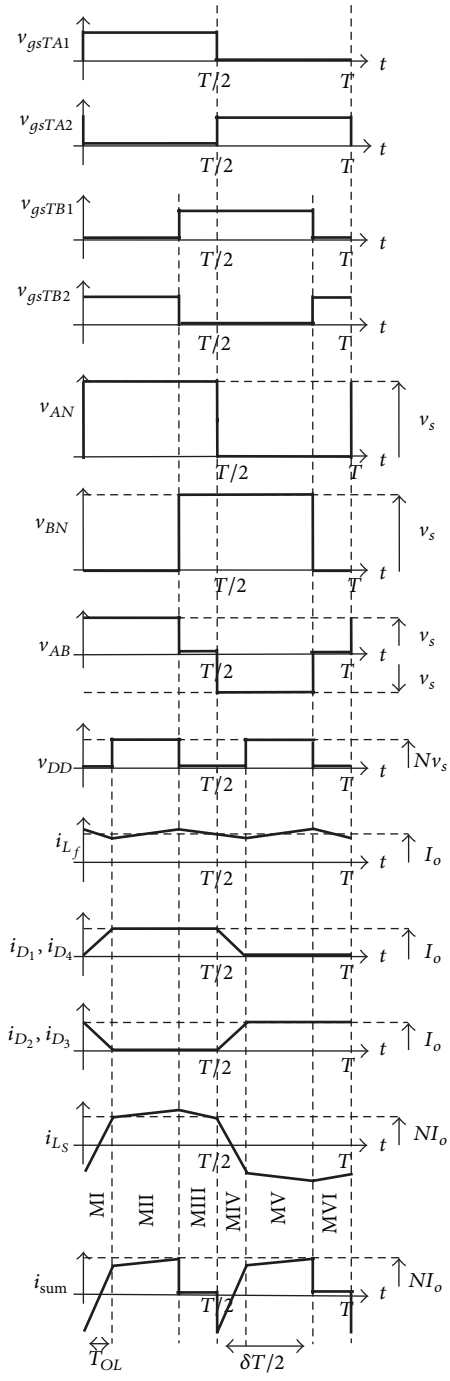


Figure 2: Ideal waveforms of the converter.

and i_{D_2} and i_{D_3} fall to zero, whereas when v_{AB} changes from zero to $-V_{\text{in}}$, i_{D_2} and i_{D_3} rise from zero to the output current level and i_{D_1} and i_{D_4} fall to zero. During the T_{OL} period, a gradual current transfer is effectuated from one diode pair to the other, in such a way that i_{L_f} continues the slight current slope which feeds the load. The steady state value of T_{OL} may be calculated assuming that the rate of change of the current reversal of i_{AB}, di_{AB}/dt, only depends on the supply voltage and the amplitude of the DC output filter current, I_O, [3], which may be expressed as

$$T_{OL} = \frac{2NL_S I_O}{V_s}. \tag{1}$$

2.2. Piece-Wise Analysis of the ZVT Converter. From Figure 2, i_{LS} presents six different behaviour intervals, which may be termed operating modes I to VI.

For each mode of operation a different circuit configuration may be obtained, which is shown in Figure 3. These equivalent circuits may be described using the state-space equation (2) and the state-space output expression (3):

$$\dot{X} = A_n X + B_n U, \tag{2}$$

$$Y = F_n X + G_n U, \tag{3}$$

where $X = \begin{bmatrix} i_{L_S} & i_{L_f} & v_o \end{bmatrix}^T$ is the state vector, $U = \begin{bmatrix} V_s & I_Z \end{bmatrix}^T$ is the input vector, I_Z is the output current disturbance, $Y = \begin{bmatrix} i_{\text{sec}} & i_{D_1} & i_{D_2} & i_{\text{sum}} \end{bmatrix}^T$ is the output vector, i_{sum} is the supply current, and A_n, B_n, F_n, and G_n are the state matrixes of the six operating modes, being $n = 1, 2, \ldots, 6$.

Mode I is formed when T_{A_1} and T_{B_2} are in the on state, T_{A_2} and T_{B_1} are in the off state, and D_1 to D_4 are conducting due to the overlap rectifier phenomena. The equivalent circuit of Mode I is shown in Figure 3(a), and the equations that describe this mode are shown in (2) and (3) with $n = 1$. The matrixes A_1, B_1, F_1, and G_1 are listed in Table 5.

In Mode II, the state of the inverter switches is exactly as that of Mode I, but D_1 and D_4 are conducting and D_2 and D_3 are off. The equivalent circuit of Mode II is shown in Figure 3(b), and again (2) and (3) describe Mode II, but with $n = 2$. The matrixes A_2, B_2, F_2, and G_2 are shown in Table 5.

In Mode III, T_{A_1} and T_{B_1} are in the on state, T_{A_2} and T_{B_2} are in the off state, D_1 and D_4 are conducting, and D_2 and D_3 are off. The equivalent circuit of Mode III is shown in Figure 3(c), being (2) and (3) with $n = 3$ the mathematical model of this mode. Again, the matrixes A_3, B_3, F_3, and G_3 are shown in Table 5.

Mode IV is a mirror of Mode I, but with T_{A_1} and T_{B_2} in the off state and T_{A_2} and T_{B_1} in the on state, whilst Modes V and VI are mirrors of Modes II and III, respectively, since the state of the switches and diodes is complementary to that of Modes II and III. Again, (2) and (3) describe Modes IV, V, and VI but with $n = 4$, 5, and 6, respectively. The corresponding matrixes to these operating modes are shown in Table 5.

2.3. Current Control Loop Description. The circuit shown in Figure 4 is a DC-DC ZVT converter with peak current control loop, which has a current transducer with gain R_s that senses i_{sum}, one *SR* flip-flop and two *D* flip-flops, a clock signal, v_{CLK}, a sawtooth generator, v_{SAW}, and the reference current level, $v_{i\text{REF}}$, which is provided by an outer voltage loop.

The operation of the circuit of Figure 4 may be explained with the state voltage and current waveforms of Figure 5. The first waveform shown in Figure 5 is the clock signal of system, v_{CLK}. The second waveform is i_{sum} plotted together with the

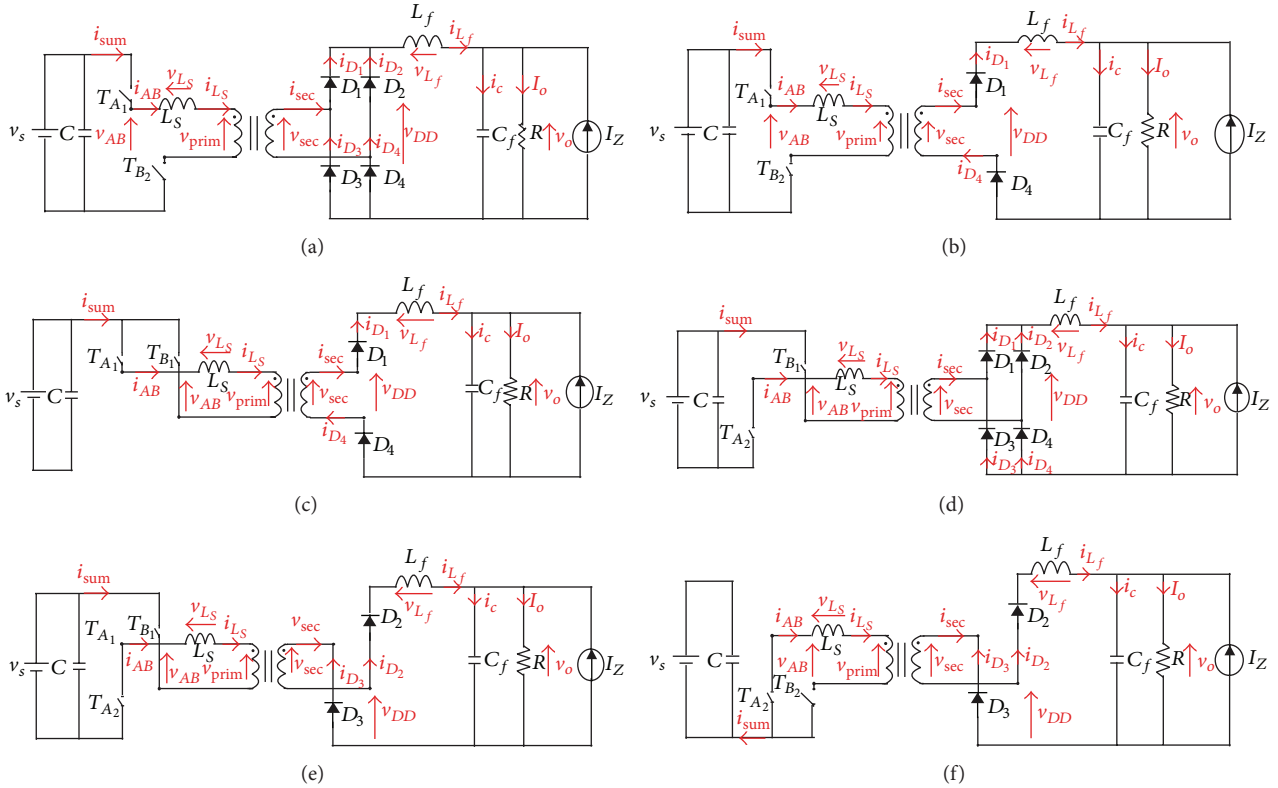


Figure 3: Equivalent circuits formed from the operation ZVT DC-DC converter. (a) Mode I, (b) Mode II, (c) Mode III, (d) Mode IV, (e) Mode V, and (f) Mode VI.

deference of v_{iREF} with v_{SAW}, where v_{SAW} is a negative slope synchronized with v_{CLK}, while v_{iREF} is the current reference that regulates the peak level of i_{sum}. v_{COMP} is the state of the output comparator, which is the third waveform of this figure. The fourth and fifth waveforms are the *SR* flip-flop outputs Q_A and Q_B, with Q_A set to the on state by v_{CLK} and to the off state when v_{COMP} switches to the on state. The fifth and sixth waveforms are the outputs of the first flip-flop, $v_{g_s T_{A1}}$ and $v_{g_s T_{A2}}$, which are controlled by the rising edge of Q_A, whereas $v_{g_s T_{B1}}$ and $v_{g_s T_{B2}}$, the outputs of the second *D* flip-flop, which are the last waves of this figure, are controlled by the rising edge of Q_B.

2.4. Numerical Estimation of the T_{OL} and δ Periods. T_{OL} defines the duration of Modes I and IV and may be numerically estimated by determining the instant when either i_{D_2} or i_{D_1} reaches zero. T_{OL} may also be calculated using the Newton-Raphson method, [4]. This numerical method may be implemented using

$$T_{OL}(n+1) = T_{OL}(n) - \frac{f(T_{OL})}{f'(T_{OL})}, \tag{4}$$

where

$$f(T_{OL}) = F_{D_1}(\varphi_1(T_{OL})X(t_1)) + \left(F_{D_1}A_1^{-1}[\varphi_1(T_{OL}) - I_d]B_1 + G_{D_1}\right)U,$$

$$f'(T_{OL}) = F_{D_1}A_1 e^{A_1 T_{OL}} X(t_1) + F_{D_1}\left[\sum_j^{\infty} \frac{(j+1)A_1^j T_{OL}^j}{(j+1)!}\right] B_1 U. \tag{5}$$

Equations of (5) use the output equation that includes i_{D_1}, which determine the duration of Mode I, whereas the duration of Mode IV is determined by i_{D_2}, such that (6) may be rewritten as follows:

$$f(T_{OL}) = F_{D_2}\left(\varphi_4(T_{OL})X\left(\frac{T}{2}\right)\right) + \left(F_{D_2}A_4^{-1}[\varphi_4(T_{OL}) - I_d]B_4 + G_{D_4}\right),$$

$$f'(T_{OL}) = F_{D_2}A_4 e^{A_4 T_{OL}} X\left(\frac{T}{2}\right) + F_{D_2}\left[\sum_j^{\infty} \frac{(j+1)A_4^j T_{OL}^j}{(j+1)!}\right] B_4 U. \tag{6}$$

δ defines the duration of Modes II and V and may be numerically estimated by determining the instant when the equation $v_{iREF} - v_{SAW}$ is equal to $R_s i_{sum}$. δ may also

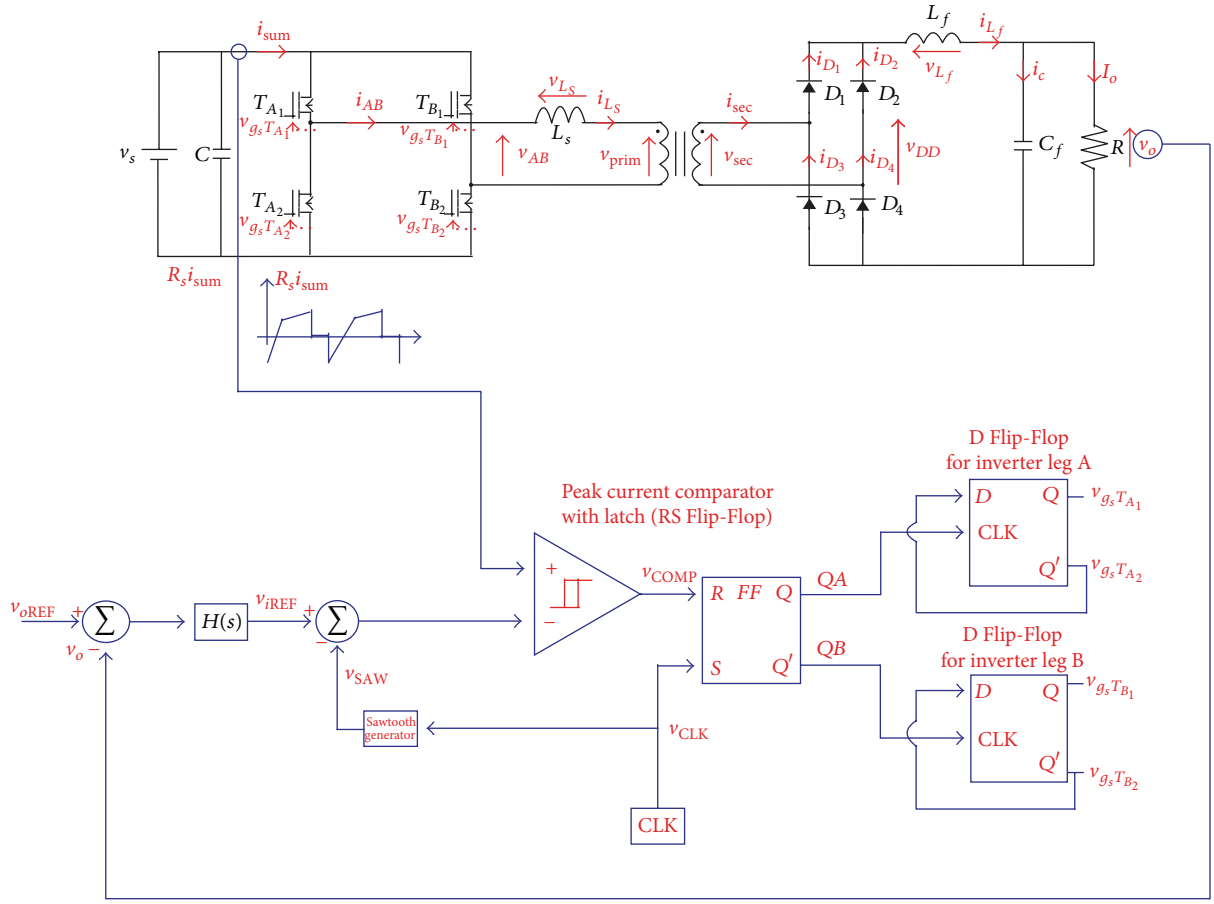


Figure 4: DC-DC ZVT converter with current control loop.

be calculated using the Newton-Raphson method, [4]. This numerical method may be implemented using

$$\delta(n+1) = \delta(n) - \frac{i_{2,\text{sum}}(\delta)}{i'_{2,\text{sum}}(\delta)}, \tag{7}$$

where

$$\begin{aligned} i_{2,\text{sum}}(\delta) &= -\left(\frac{1}{(R_s)}\right)\left[\left(v_v - v_p\right)\left(\frac{\delta T/2}{T/2}\right) + v_p\right], \\ f'^{(\delta)} &= \left(v_v - v_p\right)\delta \\ &\quad - F_{2,\text{sum}}\left(A_2\frac{T}{2}\right) \\ &\quad \times \left(\varphi_2\left(\frac{\delta T}{2} - T_{OL}\right)X\left(t_1 + T_{OL}\right)\right) \\ &\quad + F_{2,\text{sum}}\frac{T}{2}\left[\sum_j^{\infty}\frac{(j+1)\,A_2^j((\delta T/2) - T_{OL})^j}{(j+1)!}\right]U. \end{aligned} \tag{8}$$

Equations of (8) use the output equation that includes $i_{2,\text{sum}}$ and the expression $v_{iREF} - v_{SAW} = R_s i_{sum}$, which determines the duration of Mode II, whereas the duration of Mode V is determined by $i_{5,\text{sum}}$ and the expression $v_{iREF} - v_{SAW} = R_s i_{sum}$, such that (9) may be rewritten as follows:

$$\begin{aligned} i_{5,\text{sum}(\delta)} &= -\frac{\left(v_v - v_p\right)((\delta T/2)/(T/2)) + v_p}{(R_s)} \\ i'_{5,\text{sum}}(\delta) &= \left(v_v - v_p\right)\delta \\ &\quad - F_{5,\text{sum}}\left(A_5\frac{T}{2}\right) \\ &\quad \times \left(\varphi_5\left(\frac{\delta T}{2} - T_{OL}\right)X\left(t_1 + T_{OL}\right)\right) \\ &\quad + F_{5,\text{sum}}\frac{T}{2}\left[\sum_j^{\infty}\frac{(j+1)\,A_5^j((\delta T/2) - T_{OL})^j}{(j+1)!}\right]U. \end{aligned} \tag{9}$$

3. Modeling of the ZVT Converter with Current Control Loop

3.1. Piece-Wise Linear Model. Equation (2) may be used to develop a piece-wise linear model of the converter of Figure 1

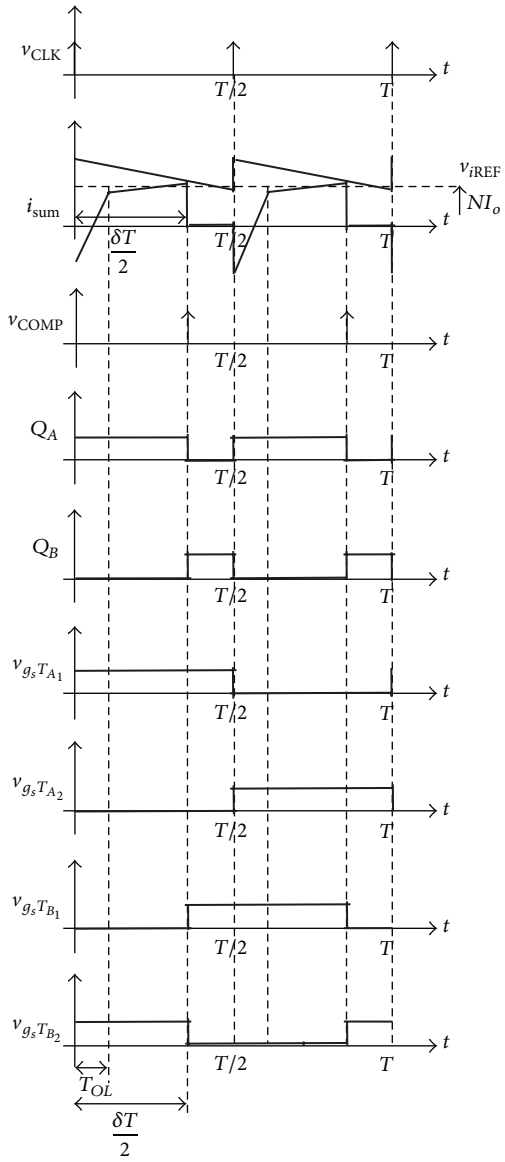


Figure 5: Ideal waveform DC-DC ZVT converter with current control loop.

throughout all the modes of operation. The solution of (2) may be expressed as

$$X\left(t_{n-1}+t_n\right)=\varphi_n\left(t_n\right)X\left(t_1\right)+A_n^{-1}\left[\varphi_n\left(t_n\right)-I_d\right]B_nU, \tag{10}$$

where

$$\varphi_n\left(t_n\right)=e^{A_nt_n}, \tag{11}$$

which defines in the first term of (10) the natural response of the system along the period of time $t_n - t_{n-1}$ with the initial condition $X(t_{n-1})$ at Mode n. The second term of (10) is the steady state response, which is obtained by using the convolution integral. Therefore, using (10) for the time interval of $t_1 \le t < t_1 + T_{OL}$, together with the matrixes of Mode I, the solution of the state vector for Mode I is obtained as follows:

$$X\left(t_1+T_{OL}\right)=\varphi_1\left(T_{OL}\right)X\left(t_1\right)+A_1^{-1}\left[\varphi_1\left(T_{OL}\right)-I_d\right]B_1U. \tag{12}$$

In a similar way, the solutions for Modes II to VI at the time intervals $t_1 + T_{OL} \le t < t_1 + \delta T/2$, $t_1 + \delta T/2 \le t < t_1 + T/2$, $t_1 + T/2 \le t < t_1 + T/2 + T_{OL}$, $t_1 + T/2 + T_{OL} \le t < t_1 + (1+\delta)T/2$, and $t_1 + (1+\delta)T/2 \le t < t_1 + T$ become

$$X\left(t_1+\frac{\delta T}{2}\right)=\varphi_2\left(\frac{\delta T}{2}-T_{OL}\right)X\left(t_1+T_{OL}\right)+A_2^{-1}\left[\varphi_2\left(\frac{\delta T}{2}-T_{OL}\right)-I_d\right]B_2U, \tag{13}$$

$$X\left(t_1+\frac{T}{2}\right)=\varphi_3\left(\frac{T}{2}-\frac{\delta T}{2}\right)X\left(t_1+\frac{\delta T}{2}\right)+A_3^{-1}\left[\varphi_3\left(\frac{T}{2}-\frac{\delta T}{2}\right)-I_d\right]B_3U, \tag{14}$$

$$X\left(t_1+\frac{T}{2}+T_{OL}\right)=\varphi_4\left(T_{OL}\right)X\left(t_1+\frac{T}{2}\right)+A_4^{-1}\left[\varphi_4\left(T_{OL}\right)-I_d\right]B_4U, \tag{15}$$

$$X\left(t_1+\frac{(1+\delta)T}{2}\right)=\varphi_5\left(\frac{\delta T}{2}-T_{OL}\right)X\left(t_1+\frac{T}{2}+T_{OL}\right)+A_5^{-1}\left[\varphi_5\left(\frac{\delta T}{2}-T_{OL}\right)-I_d\right]B_5U, \tag{16}$$

$$X\left(t_1+T\right)=\varphi_6\left(\frac{(1-\delta)T}{2}\right)X\left(t_1+\frac{(1+\delta)T}{2}\right)+A_6^{-1}\left[\varphi_6\left(\frac{(1-\delta)T}{2}\right)-I_d\right]B_6U. \tag{17}$$

3.2. Large-Signal Model. The large-signal model of the ZVT converter may be obtained by substituting (12) in (13), (13) in (14), (14) in (15), (15) in (16), and (16) in (17), in such a way that a single expression is obtained as shown in

$$\begin{aligned}X\left(t_1+T\right)&=\varphi_6\left(\frac{(1-\delta)T}{2}\right)\varphi_5\left(\frac{\delta T}{2}-T_{OL}\right)\varphi_4\left(T_{OL}\right)\varphi_3\\&\times\left(\frac{(1-\delta)T}{2}\right)\varphi_2\left(\frac{\delta T}{2}-T_{OL}\right)\varphi_1\left(T_{OL}\right)X\left(t_1\right)\end{aligned}$$

$$
\begin{aligned}
&+\left[\varphi_6\left(\frac{(1-\delta)T}{2}\right)\right.\\
&\quad\times\left[\varphi_5\left(\frac{\delta T}{2}-T_{OL}\right)\right.\\
&\qquad\times\left[\varphi_4\left(T_{OL}\right)\left[\varphi_3\left(\frac{(1-\delta)T}{2}\right)\right.\right.\\
&\qquad\qquad\times\varphi_2\left(\frac{\delta T}{2}-T_{OL}\right)\\
&\qquad\qquad\times A_1^{-1}\left[\varphi_1\left(T_{OL}\right)-I_d\right]B_1\\
&\qquad\qquad+A_2^{-1}\left[\varphi_2\left(\frac{\delta T}{2}-T_{OL}\right)\right.\\
&\qquad\qquad\qquad\left.\left.-I_d\right]B_2\right]\\
&\qquad+A_3^{-1}\left[\varphi_3\left(\frac{(1-\delta)T}{2}\right)-I_d\right]B_3\Bigg]\\
&\quad+A_4^{-1}\left[\varphi_4\left(T_{OL}\right)-I_d\right]B_4\Bigg]\\
&+A_5^{-1}\left[\varphi_5\left(\frac{\delta T}{2}-T_{OL}\right)-I_d\right]B_5\Bigg]\\
&+A_6^{-1}\left[\varphi_6\left(\frac{(1-\delta)T}{2}\right)-I_d\right]B_6U.
\end{aligned}
\tag{18}
$$

3.3. Half-Cycle Model. The waveform i_{L_S} of Figure 2 shows that operation modes IV, V, and VI are mirrors of Modes I, II, and III, respectively; therefore, the first three operation modes are sufficient to describe the function of the converter. A particular matrix titled as W may relate the modes I, II, and III with the modes IV, V, and VI, which satisfies the condition $WW = I_d$, the identity matrix. Therefore, the half-cycle model of the ZVT converter may be obtained replacing the terms A_4, A_5, A_6, B_4, B_5, and B_6 by expressions WA_1, WA_2, WA_3, WB_1, WB_2, and WB_3, respectively, in (18), such that the half-cycle model is

$$
\begin{aligned}
&X\left(t_1+\frac{T}{2}\right)\\
&\quad=\varphi_3\left(\frac{(1-\delta)T}{2}\right)\varphi_2\left(\frac{\delta T}{2}-T_{OL}\right)\varphi_1\left(T_{OL}\right)X\left(t_1\right)\\
&\qquad+\left[\varphi_3\left(\frac{(1-\delta)T}{2}\right)\varphi_2\,\varphi_3\left(\frac{(1-\delta)T}{2}\right)\varphi_2\right.\\
&\qquad\quad\times\left(\frac{\delta T}{2}-T_{OL}\right)A_1^{-1}\left[\varphi_1\left(T_{OL}\right)-I_d\right]B_1\\
&\qquad\quad+\varphi_3\left(\frac{(1-\delta)T}{2}\right)A_2^{-1}\left[\varphi_2\left(\frac{\delta T}{2}-T_{OL}\right)-I_d\right]B_2\\
&\qquad\quad\left.+A_3^{-1}\left[\varphi_3\left(\frac{(1-\delta)T}{2}\right)-I_d\right]B_3\right]U\\
&\qquad+W\left(\varphi_3\left(\frac{(1-\delta)T}{2}\right)\left(\varphi_2\left(\frac{\delta T}{2}-T_{OL}\right)\right.\right.\\
&\qquad\qquad\times W\varphi_1\left(T_{OL}\right)X\left(t_1\right)\\
&\qquad\qquad+\varphi_3\left(\frac{(1-\delta)T}{2}\right)\\
&\qquad\qquad\times\varphi_2\left(\frac{\delta T}{2}-T_{OL}\right)A_1^{-1}\\
&\qquad\qquad\left.\times\left[\varphi_1\left(T_{OL}\right)-I_d\right]B_1\right)\\
&\qquad+\varphi_3\left(\frac{(1-\delta)T}{2}\right)A_2^{-1}\\
&\qquad\times\left[\varphi_2\left(\frac{\delta T}{2}-T_{OL}\right)-I_d\right]B_2\\
&\qquad+A_3^{-1}\left[\varphi_3\left(\frac{(1-\delta)T}{2}\right)-I_d\right]WB_3\Bigg)U\left(t_1\right).
\end{aligned}
\tag{19}
$$

The previous equation may be rewritten as $X(t_1 + T/2) = A_{\mathrm{MC}}X(t_1) + B_{\mathrm{MC}}U(t_1)$ for practical purposes.

3.4. Sample-Data, Small-Signal Linear Model of the Converter in Open-Loop Conditions. The equation $X(t_1 + T/2) = A_{\mathrm{MC}}X(t_1) + B_{\mathrm{MC}}U(t_1)$ may be used as a half-cycle, discrete model of the ZVT converter, which may be written as

$$
X'_{K+1} = A_{\mathrm{MC}}X_K + B_{\mathrm{MC}}U_K, \tag{20}
$$

where $X'_{K+1} = X(t_1 + T/2)$, $X_K = X(t_1)$, and $U_K = U(t_1)$

A sample-data, small-signal model may be obtained by using the Taylor series, (21), and using small-signal perturbations as δx_K, δU_K, δT_{OL_k} and $\delta\delta_K$. One has

$$
\mathbb{f}\left(x\right) = \mathbb{f}\left(x_0\right) + \mathcal{J}\left(x - x_0\right). \tag{21}
$$

Using this equation, the sample-data, small-signal linear model becomes

$$
\begin{aligned}
\delta x_{k+1} \approx{}&\frac{\partial x_{k+1}}{\partial x_k}\delta x_k + \frac{\partial x_{k+1}}{\partial U_k}\delta U_k\\
&+\frac{\partial x_{k+1}}{\partial \delta}\delta\delta_k + \frac{\partial x_{k+1}}{\partial T_{OL}}\delta T_{OL_k}.
\end{aligned}
\tag{22}
$$

The solution of the partial derivatives is $\partial x_{k+1}/\partial x_k = A_{\mathrm{MC}}$, $\partial x_{k+1}/\partial U_k = B_{\mathrm{MC}}$, $\partial x_{k+1}/\partial \delta = C_\delta$, and $\partial x_{k+1}/\partial T_{OL} = D_{TOL}$. δT_{OLk} may be determined utilizing the restriction equation of T_{OL}, whereas $\delta\delta_K$ is obtained using the restriction equation of δ.

3.5. Restriction Equations of the Control Loop. The restriction equations for T_{OL} may be obtained analyzing the waveforms i_{LS}, i_{D_1,D_3}, and i_{sum}, which are shown in Figure 2 during the Mode I. The slope of i_{sum} during Mode I is named g_1 and is determined by the rate of change of i_{L_S}; that is, $g_1 = V_{s_k}/L_S$,

while the slopes of i_{D_1,D_3} during the Mode I are named g_3, which are contrary and of lower amplitude than g_1; that is, $g_3 = -g_1/2N$. The restriction equation of T_{OL} may be determined by integrating the waveforms $i_{D1,D3}$ during Mode I:

$$\frac{i_{LS_k}}{2N} + \frac{i_{Lf_k}}{2} + \frac{V_{s_k}}{L_S} T_{OL_k} = 0. \tag{23}$$

The previous equation is not linear; therefore, it is necessary to use the Taylor series to obtain a linearized model:

$$\delta T_{OL_k} = \frac{2NL_S}{V_{S_{CD}}} \begin{bmatrix} -\frac{1}{2N} & -\frac{1}{2} & -\frac{T_{OL_{CD}}}{2NL_S} \end{bmatrix} \begin{bmatrix} \delta i_{LS_k} \\ \delta i_{Lf_k} \\ \delta V_{s_k} \end{bmatrix}. \tag{24}$$

The restriction equation of δ may be obtained analyzing the waveforms i_{L_S} and i_{sum} with $v_{i\text{REF}} - v_{\text{SAW}}$ during Mode II (Figures 2 and 5). v_{SAW} is determined by $M(\delta_k T/2)$ and the slope of i_{sum} during Mode II, named g_2, and may be obtained by integrating again i_{sum} when $R_s i_{\text{sum}} = v_{i\text{REF}} - v_{\text{SAW}}$:

$$R_s \left(i_{LS_k} + g_1 T_{OL_k} + g_2 \left(\frac{\delta_k T}{2} - T_{OL_k} \right) \right) = V_{i\text{REF}} - M \frac{\delta_k T}{2}. \tag{25}$$

The Taylor series is used to linearize the previous equation, and therefore the restriction equation of δ becomes

$$\delta\delta_k = \frac{1}{\Delta\delta} \begin{bmatrix} a_{\delta 1} & a_{\delta 2} & a_{\delta 3} & a_{\delta 4} & a_{\delta 5} \end{bmatrix} \begin{bmatrix} \delta i_{LS_k} \\ \delta i_{Lf_k} \\ \delta V_{s_k} \\ \delta V_{i\text{REF}_k} \\ \delta V_{o_k} \end{bmatrix}, \tag{26}$$

where $\Delta_\delta = -(2/T)/(MT/2R_S + 1/(L_S + (1/N^2)L_f))(v_{s_{CD}} - (1/N)v_{0_{CD}})$, $a_{\delta 1} = \partial i_{LS_k}/\partial i_{LS_k}$, $a_{\delta 2} = \partial i_{LS_k}/\partial i_{Lf_k}$, $a_{\delta 3} = \partial i_{LS_k}/\partial V_{S_k}$, $a_{\delta 4} = \partial i_{LS_k}/\partial V_{\text{iREF}_k}$, and $a_{\delta 5} = \partial i_{LS_k}/\partial V_{o_k}$.

3.6. Sample-Data, Small-Signal Linear Model of the Converter in Closed-Loop Conditions. The sample-data, small-signal linear model, may be obtained by substituting $\delta_{T_{OLk}}$ and $\delta\delta_K$, (24) and (26), respectively, in (22):

$$\begin{aligned} \delta x_{k+1} &= A_{\text{MC}} \begin{bmatrix} \delta i_{LS_k} \\ \delta i_{Lf_k} \\ \delta V_{s_k} \end{bmatrix} + B_{\text{MC}} \begin{bmatrix} \delta V_{s_k} \\ \delta I_{Z_k} \end{bmatrix} \\ &+ C_\delta \left(\frac{1}{\Delta\delta} \begin{bmatrix} a_{\delta 1} & a_{\delta 2} & a_{\delta 3} & a_{\delta 4} & a_{\delta 5} \end{bmatrix} \begin{bmatrix} \delta i_{LS_k} \\ \delta i_{Lf_k} \\ \delta V_{s_k} \\ \delta V_{i\text{REF}_k} \\ \delta V_{o_k} \end{bmatrix} \right) \\ &+ D_{T_{OL}} \left(\frac{2NL_S}{V_{S_{CD}}} \begin{bmatrix} -\frac{1}{2N} & -\frac{1}{2} & -\frac{T_{OL_{CD}}}{2NL_S} \end{bmatrix} \begin{bmatrix} \delta i_{LS_k} \\ \delta i_{Lf_k} \\ \delta V_{s_k} \end{bmatrix} \right), \end{aligned} \tag{27}$$

Table 1: Operating parameters.

Supply voltage	200 V ± 20%
Output voltage	48 V
Maximum power	250 W
Minimum power	50 W
Output voltage ripple	25 mV
Output current ripple	300 mA
Maximum phase	50%
Switching frequency	50 kHz

such that the small-signal model becomes

$$\delta x_{k+1} = A_{cl} \delta x_k + \omega_{cl} \delta w_{cl_k}, \tag{28}$$

where $\delta x_{k+1} = \begin{bmatrix} \delta i'_{LS_k} & \delta i'_{Lf_k} & \delta V'_{s_k} \end{bmatrix}^T$, $\delta x_k = \begin{bmatrix} \delta i_{LS_k} & \delta i_{Lf_k} & \delta V_{s_k} \end{bmatrix}'$, and $\delta w_{cl_k} = \begin{bmatrix} \delta V_{s_k} & \delta V_{i\text{REF}_k} & \delta I_{Z_k} \end{bmatrix}^T$.

Equation (28) may be solved by using the Z transform, such that δx_K becomes

$$\delta x_k = \left(I_d - A_{cl} \right)^{-1} Z \omega_{cl} \delta w_{cl_k}. \tag{29}$$

3.7. Transfer Function. To verify the dynamic characteristics of the converter is necessary to analyze the transfer functions that relate v_o with V_s, $v_{i\text{REF}}$, and I_z, which are the throughput input-to-output DC voltage transfer function, $H_v(z) = v_o(z)/v_s(z)$, the control-to-output transfer function, $T_{OL}(z) = v_o(z)/v_{i\text{REF}}(z)$, and the output impedance transfer function $Z_o(z) = v_o(z)/I_z(z)$, respectively.

The magnitude and phase of each transfer function may be obtained using a Bode diagram, whereas the root locus technique may be employed to describe the behaviour of the poles and zeros of (30). One has

$$\begin{aligned} H_v(z) &= \frac{z + a_1}{z^2 + b_1 z + c_1}, \\ T_{OL}(z) &= \frac{a_2 z^2 + b_2 z + c_2}{z^2 + b_1 z + c_1}, \\ Z_o(z) &= \frac{a_3 z^2 + b_3 z - c_3}{z^2 + b_1 z + c_1}. \end{aligned} \tag{30}$$

4. Verification of the Proposed Model

4.1. Prototype Operating Parameters. A 250 W ZVT DC-DC prototype converter was designed under the analysis described in [7] to verify the large-signal model of (14). Table 1 shows the operating parameters of the converter.

The steady-state output voltage, V_O, may be calculated as the average of the rectified voltage v_{DD}, such that at full load

$$V_o = \delta_{\max} N V_{\text{in}} - \frac{4N^2 I_{o(\max)} L_S}{T}. \tag{31}$$

Taking the assumption shown in (31), the converter component values must comply with the zero-voltage switching

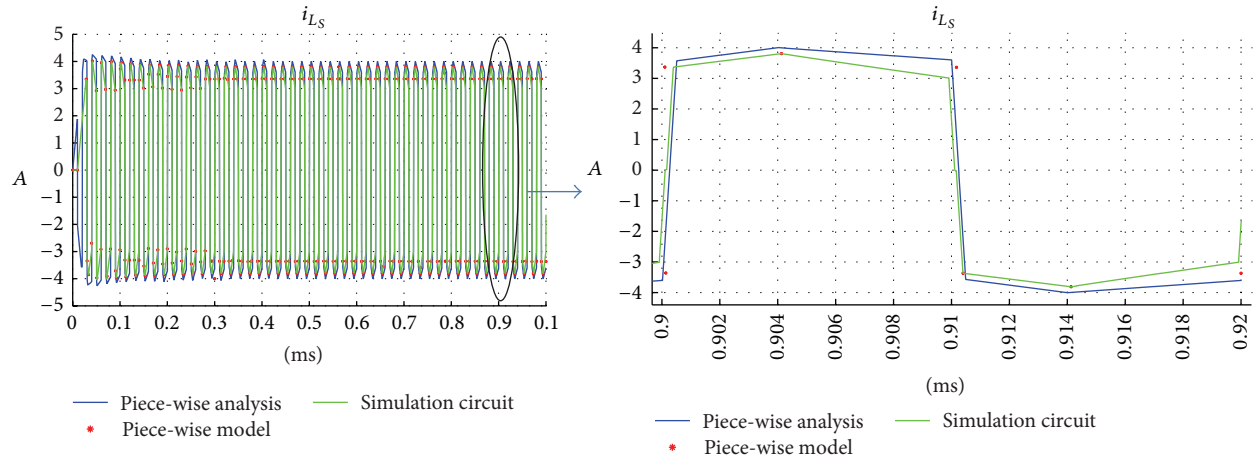


FIGURE 6: i_{L_S} current waveform obtained with the piece-wise linear model and a Micro-Cap simulation.

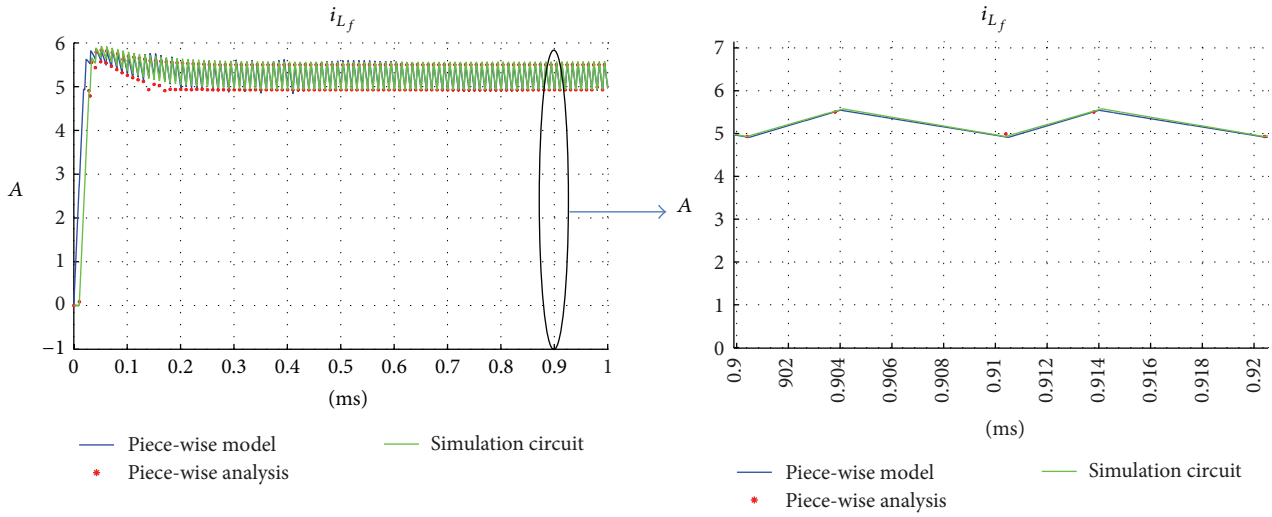


FIGURE 7: i_{L_f} current waveform obtained with the piece-wise linear model and a Micro-Cap simulation.

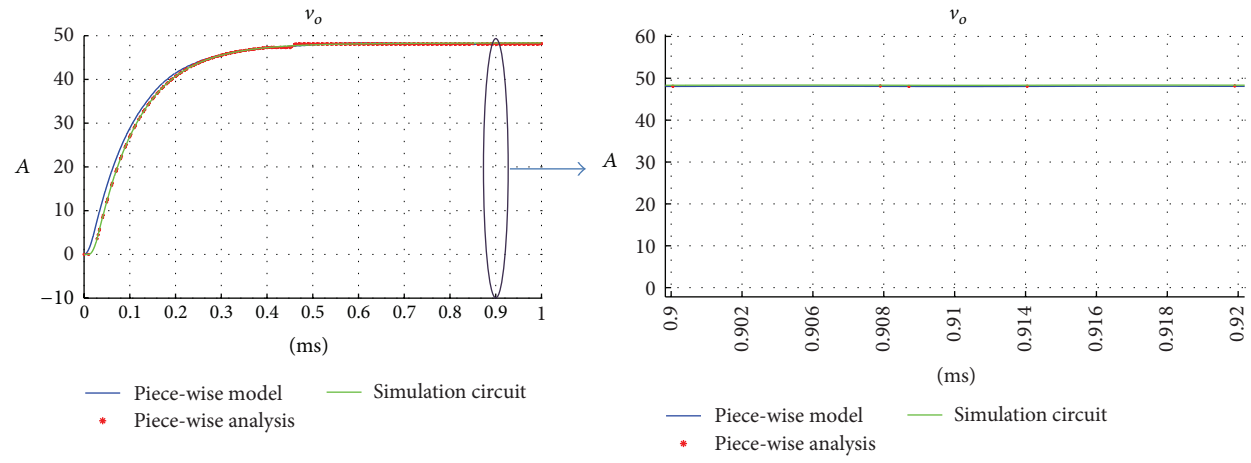


FIGURE 8: v_o voltage waveform obtained with the piece-wise linear model and a Micro-Cap simulation.

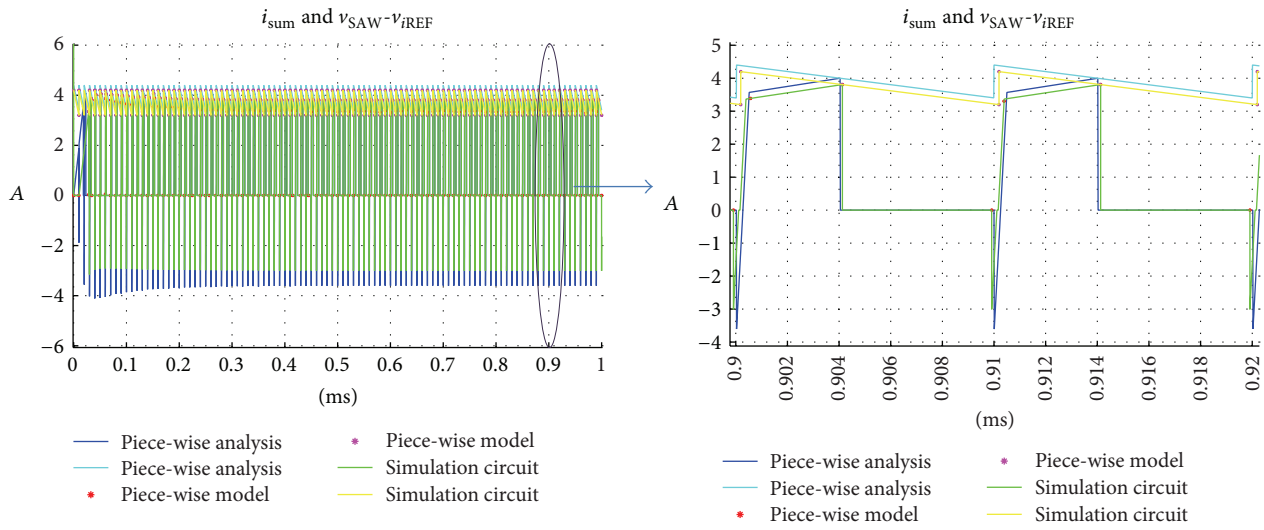


FIGURE 9: Waveforms of i_{sum} and $v_{SAW} - v_{iREF}$ obtained with the piece-wise linear model and a Micro-Cap simulation.

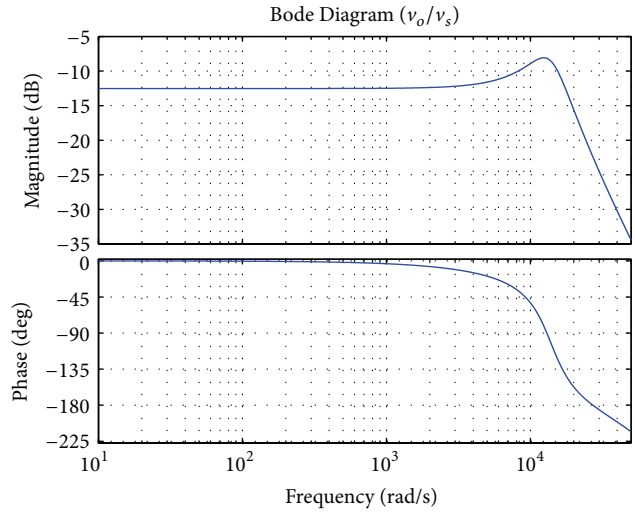


FIGURE 10: Bode plot of $H_v(z)$.

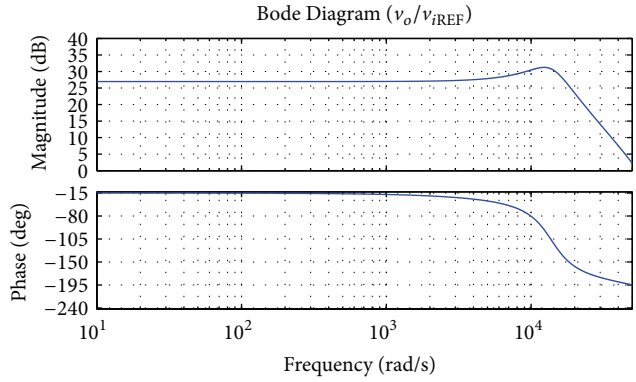


FIGURE 11: Bode plot of $T_{OL}(z)$.

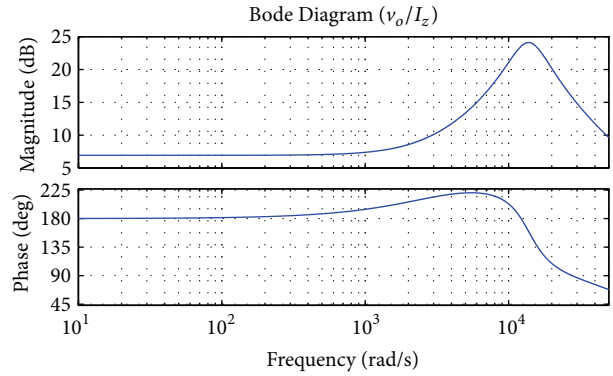


Figure 12: Bode plot of $Z_o(z)$.

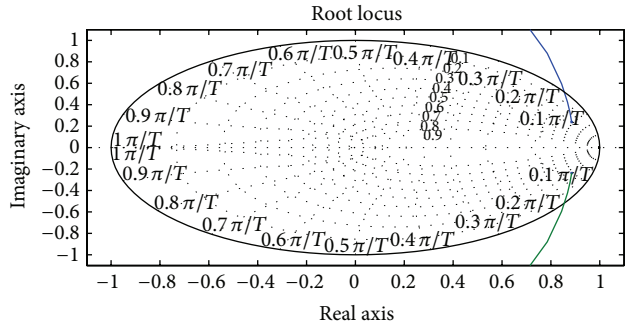


Figure 13: Root locus of $H_v(z)$.

Table 2: Component values used in the 250 W prototype.

Devices	Value
Stray inductance (L_S)	16.35 uH
Filter inductor (L_f)	528 uH
Filter capacitor (C_f)	12.18 uF
Resistive load	9.216 Ω
Transformer turns ratio N	0.683

Table 3: Verification of the large-signal model.

	$X(t_1)$ large-signal model	$X(t_1 + T)$ large-signal model	$X(t_1 + T)$ verification	% error
i_{L_S}	−3.5292	−3.529	−3.5304	0.04%
i_{L_f}	5.0171	5.017	5.0063	0.2%
V_o	48.035	48.019	48.002	0.4%

Table 4: Verification of the half-cycle model.

	$X(t_1)$ half-cycle model	$X(t_1 + T/2)$ half-cycle model	$X(t_1 + T/2)$ verification	% error
i_{L_S}	−3.415	3.47	3.4153	0.016%
i_{L_f}	4.954	4.957	4.8431	0.023%
V_o	48.53	48.53	47.1245	0.029%

phenomena to reduce the transistor switching losses under a certain load range. For instance, L_S should be large enough to keep the converter operation with ZVT under a low load condition, whilst N should be small to maintain regulated output voltage for a maximum input voltage. The list of parameters shown in Table 1, together with (31), defines the component values that may be used to keep the DC-DC converter operating with the ZVT effect within a load range of 50 W to 250 W, and the values shown in Table 2 were decided to be appropriate for the converter design. The output filter components of the rectifier were determined with the output voltage ripple, ΔV_O, and the filter inductor current ripple, ΔI_O, which may be calculated by using

$$\Delta I_o = \frac{2V_o}{\omega L_f}\left\{\frac{\pi}{2}\sin\left(\cos^{-1}\left(\frac{2}{\pi}\right)\right) - \cos^{-1}\left(\frac{2}{\pi}\right)\right\},$$
$$\Delta V_o = \frac{T\Delta I_o}{8C}. \tag{32}$$

4.2. Simulation and Results. The piece-wise linear model, the large-signal model, and the half-cycle model of the converter were verified by iterative program developed in MatLab. The piece-wise model was solved using the Runge-Kutta numerical method and using a small simulation step time, together with T_{OL} and δ, to calculate the duration of each operating mode, whereas the large-signal model and half-cycle model

Table 5: Definition of matrix for each mode of operation.

Mode I	A_1	$\begin{bmatrix} i'_{L_S} \\ i'_{L_f} \\ V'_o \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & -\frac{1}{L_f} \\ 0 & \frac{1}{C_f} & -\frac{1}{RC_f} \end{bmatrix} \begin{bmatrix} i_{L_S} \\ i_{L_f} \\ V_o \end{bmatrix} + \begin{bmatrix} \frac{1}{L_S} & 0 \\ 0 & 0 \\ 0 & \frac{1}{C_f} \end{bmatrix} \begin{bmatrix} V_s \\ I_Z \end{bmatrix}$
	B_1	$\begin{bmatrix} i_{\text{sec}} \\ i_{D_1} \\ i_{D_2} \\ i_{\text{sum}} \end{bmatrix} = \begin{bmatrix} \frac{1}{N} & 0 & 0 \\ \frac{1}{2N} & \frac{1}{2} & 0 \\ -\frac{1}{2N} & \frac{1}{2} & 0 \\ 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} i_{L_S} \\ i_{L_f} \\ V_o \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} V_s \\ I_Z \end{bmatrix}$
Mode II	A_2	$\begin{bmatrix} i'_{L_S} \\ i'_{L_f} \\ V'_o \end{bmatrix} = \begin{bmatrix} 0 & 0 & -\frac{N}{\left(N^2L_S + L_f\right)} \\ 0 & 0 & -\frac{1}{\left(N^2L_S + L_f\right)} \\ 0 & \frac{1}{C_f} & -\frac{1}{RC_f} \end{bmatrix} \begin{bmatrix} i_{L_S} \\ i_{L_f} \\ V_o \end{bmatrix} + \begin{bmatrix} \frac{N^2}{\left(N^2L_S + L_f\right)} & 0 \\ \frac{N}{\left(N^2L_S + L_f\right)} & 0 \\ 0 & \frac{1}{C_f} \end{bmatrix} \begin{bmatrix} V_s \\ I_Z \end{bmatrix}$
	B_2	$\begin{bmatrix} i_{\text{sec}} \\ i_{D_1} \\ i_{D_2} \\ i_{\text{sum}} \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \\ 1 & N & 0 \end{bmatrix} \begin{bmatrix} i_{L_S} \\ i_{L_f} \\ V_o \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} V_s \\ I_Z \end{bmatrix}$
Mode III	A_3	$\begin{bmatrix} i'_{L_S} \\ i'_{L_f} \\ V'_o \end{bmatrix} = \begin{bmatrix} 0 & 0 & -\frac{N}{\left(N^2L_S + L_f\right)} \\ 0 & 0 & -\frac{1}{\left(N^2L_S + L_f\right)} \\ 0 & \frac{1}{C_f} & -\frac{1}{RC_f} \end{bmatrix} \begin{bmatrix} i_{L_S} \\ i_{L_f} \\ V_o \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ 0 & \frac{1}{C_f} \end{bmatrix} \begin{bmatrix} V_s \\ I_Z \end{bmatrix}$
	B_3	$\begin{bmatrix} i_{\text{sec}} \\ i_{D_1} \\ i_{D_2} \\ i_{\text{sum}} \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} i_{L_S} \\ i_{L_f} \\ V_o \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} V_s \\ I_Z \end{bmatrix}$
Mode IV	A_4	$\begin{bmatrix} i'_{L_S} \\ i'_{L_f} \\ V'_o \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & -\frac{1}{L_f} \\ 0 & \frac{1}{C_f} & -\frac{1}{RC_f} \end{bmatrix} \begin{bmatrix} i_{L_S} \\ i_{L_f} \\ V_o \end{bmatrix} + \begin{bmatrix} -\frac{1}{L_S} & 0 \\ 0 & 0 \\ 0 & \frac{1}{C_f} \end{bmatrix} \begin{bmatrix} V_s \\ I_Z \end{bmatrix}$
	B_4	$\begin{bmatrix} i_{\text{sec}} \\ i_{D_1} \\ i_{D_2} \\ i_{\text{sum}} \end{bmatrix} = \begin{bmatrix} \frac{1}{N} & 0 & 0 \\ \frac{1}{2N} & \frac{1}{2} & 0 \\ -\frac{1}{2N} & \frac{1}{2} & 0 \\ -1 & 0 & 0 \end{bmatrix} \begin{bmatrix} i_{L_S} \\ i_{L_f} \\ V_o \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} V_s \\ I_Z \end{bmatrix}$
Mode V	A_5	$\begin{bmatrix} i'_{L_S} \\ i'_{L_f} \\ V'_o \end{bmatrix} = \begin{bmatrix} 0 & 0 & \frac{N}{\left(N^2L_S + L_f\right)} \\ 0 & 0 & -\frac{1}{\left(N^2L_S + L_f\right)} \\ 0 & \frac{1}{C_f} & -\frac{1}{RC_f} \end{bmatrix} \begin{bmatrix} i_{L_S} \\ i_{L_f} \\ V_o \end{bmatrix} + \begin{bmatrix} -\frac{N^2}{\left(N^2L_S + L_f\right)} & 0 \\ \frac{N}{\left(N^2L_S + L_f\right)} & 0 \\ 0 & \frac{1}{C_f} \end{bmatrix} \begin{bmatrix} V_s \\ I_Z \end{bmatrix}$
	B_5	$\begin{bmatrix} i_{\text{sec}} \\ i_{D_1} \\ i_{D_2} \\ i_{\text{sum}} \end{bmatrix} = \begin{bmatrix} 0 & -1 & 0 \\ 0 & 0 & 0 \\ 0 & 1 & 0 \\ -1 & N & 0 \end{bmatrix} \begin{bmatrix} i_{L_S} \\ i_{L_f} \\ V_o \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} V_s \\ I_Z \end{bmatrix}$

Table 5: Continued.

Mode VI	A_6	$\begin{bmatrix} i'_{L_S} \\ i'_{L_f} \\ V'_o \end{bmatrix} = \begin{bmatrix} 0 & 0 & \frac{N}{\left(N^2 L_S + L_f\right)} \\ 0 & 0 & -\frac{1}{\left(N^2 L_S + L_f\right)} \\ 0 & \frac{1}{C_f} & -\frac{1}{RC_f} \end{bmatrix} \begin{bmatrix} i_{L_S} \\ i_{L_f} \\ V_o \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ 0 & \frac{1}{C_f} \end{bmatrix} \begin{bmatrix} V_s \\ I_Z \end{bmatrix}$
	B_6	$\begin{bmatrix} i_{\text{sec}} \\ i_{D_1} \\ i_{D_2} \\ i_{\text{sum}} \end{bmatrix} = \begin{bmatrix} 0 & -1 & 0 \\ 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} i_{L_S} \\ i_{L_f} \\ V_o \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} V_s \\ I_Z \end{bmatrix}$

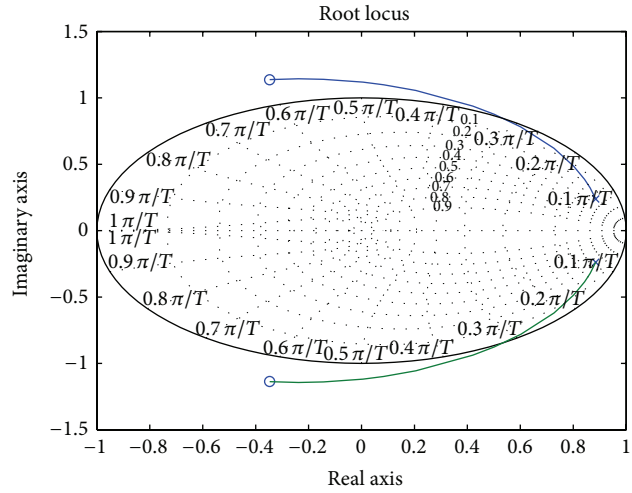


Figure 14: Root locus of $T_{OL}(z)$.

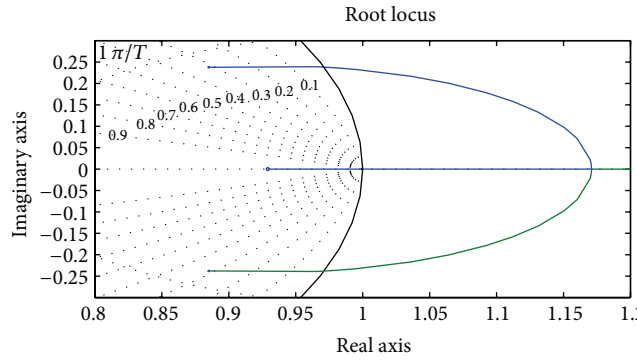


Figure 15: Root locus of $Z(z)$.

were solved by calculating T_{OL} and δ with the Newton-Raphson method. Figures 6 to 9 show a comparison of the results obtained with the piece-wise model and simulation results obtained with Micro-Cap. The waveforms plotted on these figures are the transformer primary-side current i_{L_S}, Figure 6, the filter inductor current i_{L_f}, Figure 7, the output voltage v_O, Figure 8, and the supply current i_{sum} together with $v_{\text{SAW}} - v_{i\text{REF}}$, Figure 9. Tables 3 and 4 show a comparison of the instantaneous values of the state vector obtained at the end of a full cycle in steady state conditions, which verifies the exactitude of the large-signal model, whereas Table 4 shows those of the half-cycle model. Both tables list results together with instantaneous results obtained with Micro-Cap. The value of T_{OL} calculated for Modes I and IV is 0.727 μs while δ is 0.3998.

Figures 10, 11, and 12 show the bode diagram of the transfer functions $H_v(z)$, $T_{OL}(z)$, and $Z_o(z)$ calculated with the symbolic equation tool of MatLab, whilst Figures 13, 14, and 15 show the roots locus of $H_v(z)$, $T_{OL}(z)$, and $Z_o(z)$.

The magnitude of $H_v(z)$ at low frequency converges to the steady-state DC of $v_{o-\text{DC}}/V_s$, while the magnitude of $T_{OL}(z)$ reveals the gain of the control-to-output under dynamic

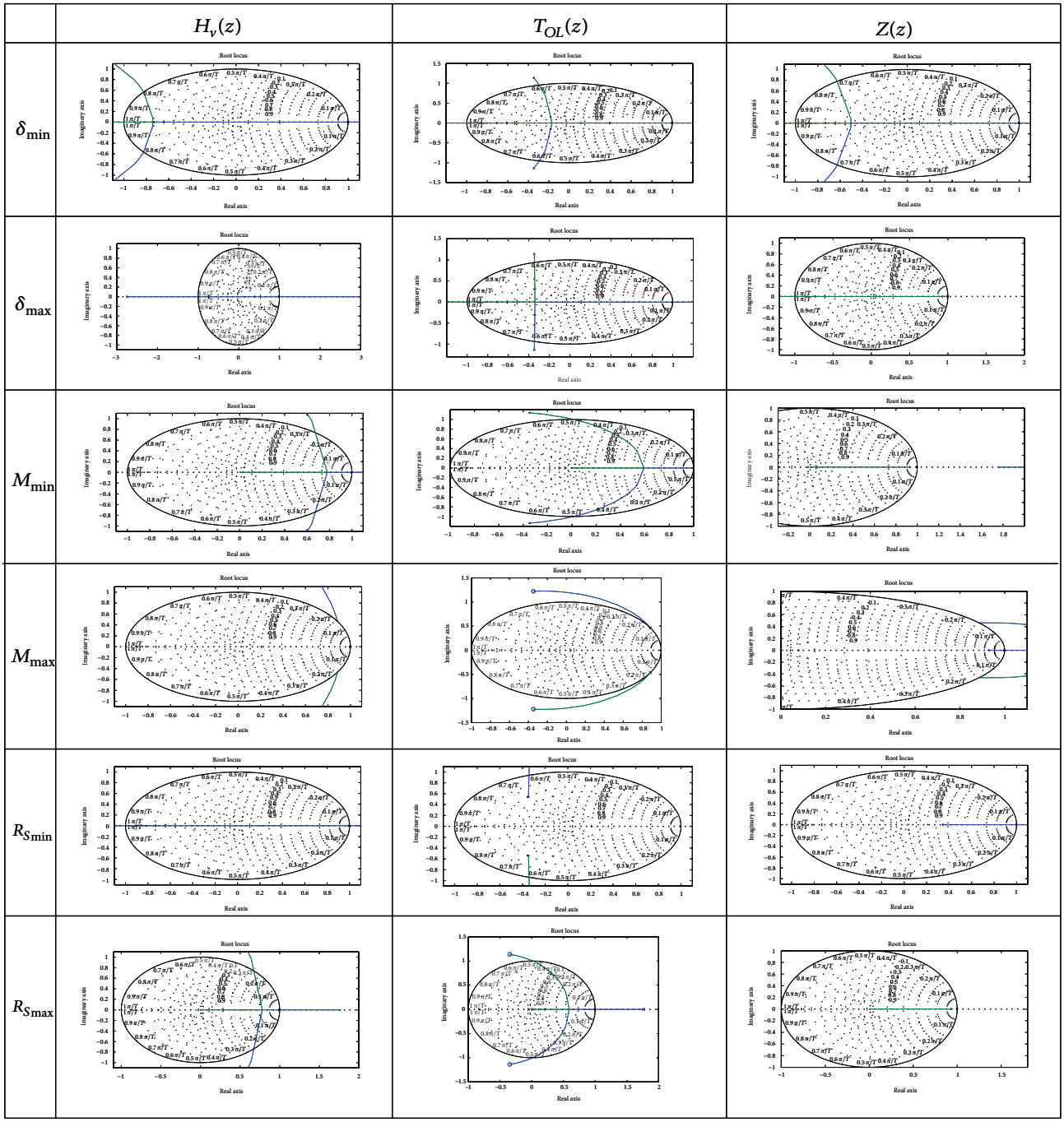


Figure 16: Root locus plots of $H_v(z)$, $T_{OL}(z)$, and $Z_O(z)$ obtained by ranging of δ, M, and R_s.

conditions, and the magnitude of $Z(z)$ converges to small value below and above the resonant frequency determined by $1/\sqrt{L_f + N^2 L_S/C_f}$.

The poles of $H_V(z)$ are conjugates complex roots, whereas there is a single zero located at -2.75. $T_{OL}(z)$ is steady with a gain lower than 0.596 dB, and its poles are conjugates complex roots with two steady zeros. $Z_o(z)$ is steady with a gain lower than 0.0682 dB, and its poles are conjugates complex roots, and there are two zeros located at -3.93 and 0.93.

To design a controller is necessary to determine the behaviour of the poles and zeros of the transfer functions with respect to variations of δ, M, and R_s. Figure 16 shows diverse root locus plots of $H_V(z)$, $T_{OL}(z)$, and $Z_o(z)$ by ranging the values of δ, M, and R_s.

5. Conclusion

This paper presented the mathematical derivation of a sample-data, small-signal model for a ZVT DC-DC converter. The method used a piece-wise linear analysis to obtain a large-signal model which was verified with numerical predictions that depend on the integration step size to obtain high accuracy. A sample-data, half-cycle linear model was derived using the large-signal model, such that a dynamic model of the converter was obtained by using a linear approximation. A comparison of the instantaneous values listed in Tables 3 and 4 showed that there is close correspondence between the derived models and the circuit simulation, especially with the half-cycle model. Also, Bode diagrams and a root locus analysis showed that the control system may be steady if δ is defined within an interval of 0.3 to 0.6, R_s within 0.9 V/A to 1.2 V/A, and M within 0.044 V/s to 0.1 V/s. Therefore, the half-cycle model is more accurate than the large-signal model, and it is useful to determine a small-signal, sample-data model of the DC-DC converter for dynamic studies. The presented method may help to power electronic practitioners to derive discrete transfer functions of soft switched DC-DC converter and understand the dynamic behaviour of power electronic systems, which serves as a basic principle to design a controller for the converter outer loop.

Conflict of Interests

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

Acknowledgments

The authors are grateful to the National Council of Science and Technology of Mexico (CONACyT), the Postgraduate and Research Department (SEPI) of the School of Mechanical and Electrical Engineering, Campus Culhuacan, of the National Polytechnic Institute (IPN) of Mexico, and the Technological Studies Superiors of Coacalco for their encouragement and the realization of the prototype.

References

[1] F. J. Perez-Pinal, *The Electric Vehicle: Design Stages Considerations*, Editorial Academica Española, Madrid, Spain, 2011.

[2] F. J. Perez-Pinal, C. Nunez, R. Alvarez, and M. Gallegos, "Step by step design of the power stage of a light electric vehicle," *International Review of Electrical Engineering*, vol. 3, no. 1, pp. 100–108, 2008.

[3] F. J. Perez-Pinal, J. C. Kota-Renteria, J. C. Nuñez-Perez, and N. Al-Mutawaly, "Hybrid conversion kit applied to public transportation: a taxi case solution," *International Review on Modelling and Simulations*, vol. 6, no. 2, pp. 554–559, 2013.

[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] F. J. Perez-Pinal, N. Al-Mutawaly, and J. C. Nuñez-Perez, "Impact of plug-in hybrid electric vehicle in distributed generation and smart grid: a brief review," *International Review on Modelling and Simulations*, vol. 6, no. 3, pp. 795–805, 2013.

[6] 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.

[7] A. Tapia-Hernández, I. Araujo-Vargas, M. Ponce-Silva, and M. Ponce-Flores, "Design of a ZVT DC-DC converter with stray components integration for a public-transport electric vehicle," in *Proceedings of the International Symposium on Power Electronics, Electrical Drives, Automation and Motion (SPEEDAM '10)*, pp. 478–483, June 2010.

Hindawi Publishing Corporation
Mathematical Problems in Engineering
Volume 2014, Article ID 906367, 12 pages
http://dx.doi.org/10.1155/2014/906367

Research Article

Stability of Gain Scheduling Control for Aircraft with Highly Nonlinear Behavior

Fany Mendez-Vergara, Ilse Cervantes, and Angelica Mendoza-Torres

Hybrid Systems Laboratory, Institute for Scientific and Technological Research of San Luis Potosi (IPICyT), Camino a la Presa San Jose 2055, Colonia Lomas 4ta, 78216 San Luis Potosi, SLP, Mexico

Correspondence should be addressed to Ilse Cervantes; ilse@ipicyt.edu.mx

Received 30 November 2013; Accepted 21 April 2014; Published 21 May 2014

Academic Editor: Sheldon S. Williamson

Copyright © 2014 Fany Mendez-Vergara 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 main goal of this work is to study the stability properties of an aircraft with nonlinear behavior, controlled using a gain scheduled approach. An output feedback is proposed which is able to guarantee asymptotical stability of the task-coordinates origin and safety of the operation in the entire flight envelope. The results are derived using theory of hybrid and singular perturbed systems. It is demonstrated that both body velocity and orientation asymptotic tracking can be obtained in spite of nonlinearities and uncertainty. The results are illustrated using numerical simulations in F16 jet.

1. Introduction

Gain scheduled control (GSC) is widely employed in flight control applications. The principle of operation of such a controller relies on the use of a piece-wise linear approximation of aircraft dynamics to conclude stability via direct or indirect arguments relating the linear approximation and the original nonlinear dynamics [1]. Usually, stability of every linear system is analyzed, and the design of an interpolation controller to guarantee stable switching is performed as a separated control design task [2–4].

Since the flight dynamics have nonlinear nature, the linear approximation of the system leads to controllers with stability and performance limitations. To overcome such limitations there exist in the literature a variety of control methodologies which provide important advances in GSC and hybrid and nonlinear approaches [5–8]; see also GSC surveys [1, 9–19].

Robustness of controllers is crucial not only due to the high nonlinear behavior of the system, but also because aeronautical systems are subjected to external perturbations and they must have some fault tolerant capabilities. There exist in the literature a variety of works leading the control of these systems from a robust or adaptive point of view. In [14] a multi-input, multioutput nonlinear model predictive controller is proposed that makes use of a neurofuzzy predictor model. The authors analyze the robustness of the proposed control in presence of uncertainties in the physical parameters. In [15] attitude stabilization with external disturbance, unknown inertia parameters, and actuator uncertainties is analyzed. An adaptive control is proposed and input-to-state stability is derived. Reference [16] studies the spacecraft attitude tracking problem; the conditions that guarantee that the attitude and body-rate error are ultimately bounded are derived and bounds for the steady-state tracking errors are also provided. Reference [17] proposes an observer-controller control scheme to solve the output feedback attitude control of a rigid body with bounded input.

In [10] a robust sum-of-squares (SOS) LMI (linear matrix inequality) method to design a nonlinear controller for longitudinal dynamics with parametric uncertainties is proposed. The control design problem is formulated as a robust SOS problem and then an LMI representation is derived using an algorithm that is solved via a stochastic ellipsoid method. In [11] an output feedback to stabilize a planar nonlinear model of a vertical take-off and landing (PVTOL) aircraft is proposed. To address the lack of full state measurement, a finite time convergent observer is

designed. Reference [12] addresses the nonlinear tracking problem of a roll-rate-sensorless vertical take-off and landing aircraft in the presence of roll-angle measurement delay. The authors develop an extended attitude observer and a tracking controller based on the observer by using the first-order Pade approximation signal of the delayed roll angle, two global coordinate transformations, and an extension of the backstepping technique.

In many of the works above, it is not clear if the controller can deal with structural uncertainty that may arise from a failure, or from the nonlinearities of the system. Moreover, some of them make use of observers to estimate the entire state, which usually leads to high-order controllers. Some of these drawbacks have made researches investigate improved GSC strategies that can ensure the stability of the system in spite of nonlinearities but still with simple controllers. Many of these works make use of recent advances in switching and hybrid systems.

Recently, switched LPV (linear parameter varying) systems have gained the attention of researchers due to the fact that stabilization can also be derived using a multiple Lyapunov function approach. In [20] it is studied the GSC using a switched polytopic approach. In such work, it was proved that the switched polytopic system is input-to-state bounded if the subsystems share a common Lyapunov function, and average dwell time is no smaller than a given constant. On the other hand, [21] proposed a hybrid tracking control for aircraft described by discrete-time stochastic linear hybrid systems; the flight logic is described by discrete dynamics governed by Markov chains. Conditions for convergence of a model adaptive estimation algorithm are proposed, and a new residual-mean interacting multiple model algorithm is proposed.

Reference [18] uses a piecewise linear description of the system and variations in operation points modeled as switches between constituent linearized system dynamics to analyze the control tracking problem. An adaptive control scheme is proposed that can achieve asymptotic tracking performance if the reference input is sufficiently rich and the switches are sufficiently slow. Reference [19] develops a hybrid control strategy that provides autonomous transition between hovered and leveled flights to a model-scale fixed-wing aircraft. The controller, together with an appropriate choice of reference maneuvers, ensures that a transition from hovered flight to level flight, or vice versa, is achieved. In [22], a sensor-based backstepping method is implemented on a fixed-wing aircraft with its focus on handling structural changes caused by damages. The controller combines nonlinear dynamic inversion, which is highly dependent on the knowledge of the system structure, with a sensor-based backstepping control approach.

In this work, we are focused on the velocity regulation and orientation of nonlinear aircraft. It is known (see [23]) that the more general description of the flight dynamics is piecewise nonlinear since forces and torques involved may vary dramatically with altitude, angle of attack, side-slip angle, and mach number, among others. That is, in the best situation, multiple nonlinear description of the aircraft constitutes the more reliable description that can be obtained. Under this premise, we depart from a six degree of freedom, nonlinear piece-wise description of the aircraft to study the stability of the system under a simple linear dynamic output feedback (as in GSC). Using theory of switched hybrid systems and singular perturbation systems, we demonstrate that the origin of the closed-loop piecewise nonlinear description of the aircraft is asymptotically stable, even in presence of uncertainties that can be either structural or parametric. This result constitutes the main contribution of this paper.

The tracking control strategy proposed here achieves a desired linear velocity and attitude, while ensuring body rate stabilization in spite of nonlinearities and model uncertainty. As in GSC, the proposed controller is able to adapt its operation based on the suitable definition of the switching criterion. Stability conditions for every operation mode and the overall switching system are derived; such piece-wise nonlinear analysis has not been performed before, to the authors' knowledge.

Based on the aircraft movement restrictions, a cascade control structure is proposed to solve the underactuated task of solving simultaneously body rate stabilization and attitude regulation. The attitude problem is formulated using the kinematic restrictions relating angular body velocity and quaternions, avoiding in this way the drawback of multiple singularities in a given operation condition. It is shown that the proposed controller is able to robustly stabilize the origin of the controlled nonlinear dynamics of the aircraft in spite of parametric or structural uncertainties, as long as they satisfy being bounded and continuous Lipschitz functions of their arguments in the operation domain. Numerical simulations of an F16 aircraft are used to illustrate advantages of the proposed control.

The paper begins with the nonlinear piecewise description of the aircraft and the control problem formulation. Section 3 introduces some preliminary results to establish the stability and the equilibrium uniqueness of some auxiliary systems, while Section 4 is devoted to deriving the main result of the paper. Illustrations of the results are given in Section 5 using a fixed-wing aircraft. Finally, Section 6 summarizes the main contribution of the proposed structure and presents some conclusions.

2. Problem Statement and Dynamic Description of the Aircraft

In this work, we depart from a nonlinear description of the aircraft using wind axis. As it is known, the nonlinear behavior of the aircraft depends crucially on the movement task performed. Typically the model parameters are obtained about certain operation conditions. Due to this fact, the nonlinear behavior of the aircraft is better described using piecewise continuous models. In [23], it is stated that the nonlinear description of the aircraft depends on variables such as dynamic pressure, altitude, angle of attack, and sideslip angle among the most important. Let the nonlinear

model j be valid in the domain Ω_j. Departing form this premise, the piecewise nonlinear description of the system can be written without loss of generality as constituted by a linear function and a nonlinear function ($^j\overline{\Phi}(x)$) clustering high order terms and uncertainty, depending exclusively on system's states. In this work, we assume that such description is valid on a domain Ω_j. In this way, the airplane movement description and orientation is given by

$$\begin{aligned} \dot{x} &= {}^jA\,x(t) + {}^jB\,u(t) + {}^j\overline{\Phi}(x), \\ \dot{\vartheta} &= T(\vartheta)\,\omega_w, \end{aligned} \tag{1}$$

where $x = [V_T, \beta, \alpha, P, Q, R]^T$ are deviation variables with respect to an operation point, $\vartheta = [\eta, \epsilon]^T$, $\epsilon = [\epsilon_1, \epsilon_2, \epsilon_3]^T$, and $\omega_w = [P, Q, R]^T$. Moreover,

$$T(\vartheta) \triangleq \begin{bmatrix} -\frac{1}{2}\epsilon^T \\ \frac{1}{2}\left[\eta I - S(\epsilon)\right] \end{bmatrix}, \tag{2}$$

where

$$S(\epsilon) = \begin{bmatrix} 0 & -\epsilon_3 & \epsilon_2 \\ \epsilon_3 & 0 & -\epsilon_1 \\ -\epsilon_2 & \epsilon_1 & 0 \end{bmatrix}, \tag{3}$$

with ^jA, ^jB, and $^j\overline{\Phi}(x)$ given by

$$^jA = m\begin{bmatrix} X_V + X_{T_V}\cos(\alpha_e + \alpha_T) & 0 & X_\alpha & 0 & 0 & 0 \\ 0 & Y_\beta & 0 & Y_p & 0 & Y_r - V_{T_e} \\ Z_V - X_{T_V}(\alpha_e + \alpha_T) & 0 & Z_\alpha & 0 & V_{T_e} + Z_q & 0 \\ 0 & \mu L_\beta + \sigma_1 N_\beta & 0 & \mu L_p + \sigma_1 N_p & 0 & \mu L_r + \sigma_1 N_r \\ M_V + M_{T_V} & 0 & M_\alpha + M_{T\alpha} & 0 & M_q & 0 \\ 0 & \mu N_\beta + \sigma_2 L_\beta & 0 & \mu N_p + \sigma_2 L_p & 0 & \mu N_r + \sigma_2 L_r \end{bmatrix}, \tag{4}$$

$$^jB = \begin{bmatrix} \frac{1}{\mathrm{m}}\frac{\partial D_a}{\partial \delta_T} & \frac{1}{\mathrm{m}}\frac{\partial D_a}{\partial \delta_e} & \frac{1}{\mathrm{m}}\frac{\partial D_a}{\partial \delta_a} & \frac{1}{\mathrm{m}}\frac{\partial D_a}{\partial \delta_r} \\ \frac{1}{\mathrm{m}V_T}\frac{\partial C_a}{\partial \delta_T} & \frac{1}{\mathrm{m}V_T}\frac{\partial C_a}{\partial \delta_e} & \frac{1}{\mathrm{m}V_T}\frac{\partial C_a}{\partial \delta_a} & \frac{1}{\mathrm{m}V_T}\frac{\partial C_a}{\partial \delta_r} \\ \frac{1}{\mathrm{m}V_T\cos\beta}\frac{\partial L}{\partial \delta_T} & \frac{1}{\mathrm{m}V_T\cos\beta}\frac{\partial L}{\partial \delta_e} & \frac{1}{\mathrm{m}V_T\cos\beta}\frac{\partial L}{\partial \delta_a} & \frac{1}{\mathrm{m}V_T\cos\beta}\frac{\partial L}{\partial \delta_r} \\ J_w^{-1}\frac{\partial l_w}{\partial \delta_T} & J_w^{-1}\frac{1}{I'_x}\frac{\partial l_w}{\partial \delta_e} & J_w^{-1}\frac{1}{I'_x}\frac{\partial l_w}{\partial \delta_a} & J_w^{-1}\frac{1}{I'_x}\frac{\partial l_w}{\partial \delta_r} \\ J_w^{-1}\frac{\partial m_w}{\partial \delta_T} & J_w^{-1}\frac{1}{I'_x}\frac{\partial m_w}{\partial \delta_e} & J_w^{-1}\frac{1}{I'_x}\frac{\partial m_w}{\partial \delta_a} & J_w^{-1}\frac{1}{I'_x}\frac{\partial m_w}{\partial \delta_r} \\ J_w^{-1}\frac{\partial n_w}{\partial \delta_T} & J_w^{-1}\frac{1}{I'_x}\frac{\partial n_w}{\partial \delta_e} & J_w^{-1}\frac{1}{I'_x}\frac{\partial n_w}{\partial \delta_a} & J_w^{-1}\frac{1}{I'_x}\frac{\partial n_w}{\partial \delta_r} \end{bmatrix}, \tag{4}$$

$$^j\overline{\Phi}(\mathbf{x}) = \begin{bmatrix} \frac{1}{\mathrm{m}}\left\{D_{a_e} + \sum D_a(V_T, \beta, \alpha, P, Q, R) - g\cos\alpha\cos\beta\sin\theta + g\sin\beta\sin\phi\cos\theta + g\sin\alpha\cos\phi\cos\theta\right\} + \overline{\varphi_1} \\ \left\{\frac{1}{\mathrm{m}V_T}\right\}\left\{\mathrm{m}V_T R + C_{a_e} + \sum C_a(V_T, \beta, \alpha, P, Q, R) + g\cos\alpha\sin\beta\sin\theta \right. \\ \left. + g\cos\beta\sin\phi\cos\theta - g\sin\alpha\sin\beta\cos\phi\cos\theta\right\} + \overline{\varphi_2} \\ \left\{\frac{1}{\mathrm{m}V_T\cos\beta}\right\}\left\{\mathrm{m}V_T\{Q\cos\beta - P\sin\beta\} + L_e + \sum L(V_T, \beta, \alpha, P, Q, R) + g\sin\alpha\sin\theta \right. \\ \left. + g\cos\alpha\cos\phi\cos\theta + \frac{\partial L}{\partial \delta_e}\delta_e + \frac{\partial L}{\partial \delta_r}\delta_r\right\} + \overline{\varphi_3} \\ J_w^{-1}\left\{PQI'_{xz} + QRI'_y + RQI'_z + l_{w_e} + \sum l_w(V_T, \beta, \alpha, P, Q, R)\right\} + \overline{\varphi_4} \\ J_w^{-1}\left\{-P\left\{RI'_z + PI'_{xz}\right\} - R\left\{PI'_{xz} + PI'_z\right\} + m_{w_e} + \sum m_w(V_T, \beta, \alpha, P, Q, R)\right\} + \overline{\varphi_5} \\ J_w^{-1}\left\{PQI'_z - QPI'_y + RQI'_{xz} + n_{w_e} + \sum n_w(V_T, \beta, \alpha, P, Q, R)\right\} + \overline{\varphi_6} \end{bmatrix}, \tag{5}$$

where $\overline{\varphi_i}$ $i = 1, \ldots, 6$ account for state-dependent uncertainties, and $\sum D_a(V_T, \beta, \alpha, P, Q, R)$, $\sum C_a(V_T, \beta, \alpha, P, Q, R)$, and so forth, are terms accounting for higher order nonlinear terms as shown in the following example:

$$\begin{aligned}
&\sum D_a\left(V_T, \beta, \alpha, P, Q, R\right) \\
&\quad = \left\{\frac{\partial^2 D_a}{\partial V_T^2}\frac{V_T^2}{2!} + \frac{\partial^3 V_T}{\partial V_T^3}\frac{V_T^3}{3!} + \frac{\partial^4 V_T}{\partial V_T^4}\frac{V_T^4}{4!} + \cdots\right\} \\
&\qquad + \left\{\frac{\partial^2 D_a}{\partial \alpha^2}\frac{\alpha^2}{2!} + \frac{\partial^3 D_a}{\partial \alpha^3}\frac{\alpha^3}{3!} + \frac{\partial^4 D_a}{\partial \alpha^4}\frac{\alpha^4}{4!} + \cdots\right\} \\
&\qquad + \left\{\frac{\partial^2 D_a}{\partial \beta^2}\frac{\beta^2}{2!} + \frac{\partial^3 D_a}{\partial \beta^3}\frac{\beta^3}{3!} + \frac{\partial^4 D_a}{\partial \beta^4}\frac{\beta^4}{4!} + \cdots\right\} \\
&\qquad + \left\{\frac{\partial^2 D_a}{\partial P^2}\frac{P^2}{2!} + \frac{\partial^3 D_a}{\partial P^3}\frac{P^3}{3!} + \frac{\partial^4 D_a}{\partial P^4}\frac{P^4}{4!} + \cdots\right\} \\
&\qquad + \left\{\frac{\partial^2 D_a}{\partial Q^2}\frac{Q^2}{2!} + \frac{\partial^3 D_a}{\partial Q^3}\frac{Q^3}{3!} + \frac{\partial^4 D_a}{\partial Q^4}\frac{Q^4}{4!} + \cdots\right\} \\
&\qquad + \left\{\frac{\partial^2 D_a}{\partial R^2}\frac{R^2}{2!} + \frac{\partial^3 D_a}{\partial R^3}\frac{R^3}{3!} + \frac{\partial^4 D_a}{\partial R^4}\frac{R^4}{4!} + \cdots\right\}.
\end{aligned} \tag{6}$$

Moreover the output and input of the system are $y = [V_T, P, Q, R]^T$ $u^T = [\delta_T \ \ \delta_e \ \ \delta_a \ \ \delta_r]$ where δ_T is the throttle position and δ_e, δ_a, and δ_r are the control surface positions. The following is assumed:

(A.1) ${}^j\overline{\Phi}$ is a bounded, continuous differentiable function of x in Ω_j.

The assumption (A.1) implies that ${}^j\overline{\Phi}$ is Lipschitz with respect to x in Ω_j. Notice that the Euler angles in (5) are related to the quaternions (2)-(3).

2.1. Problem Formulation. Let the airplane model be given by (1)–(6); find an output dependent piecewise linear control such that

$$\lim_{t\to\infty}\begin{bmatrix} V_T \\ P \\ Q \\ R \\ \eta \\ \epsilon \end{bmatrix} = \begin{bmatrix} V_{T_{\text{ref}}} \\ 0 \\ 0 \\ 0 \\ \eta_{\text{ref}} \\ \epsilon_{\text{ref}} \end{bmatrix}, \tag{7}$$

for every $x \in D$, where $D = \cup\Omega_j$ and $V_{T_{\text{ref}}}$, η_{ref}, and ϵ_{ref} are the desired values of the aircraft velocity magnitude and quaternion orientation.

2.2. Control Structure. To accomplish both the movement and orientation control defined in (7), it is necessary to define the regulation error as follows:

$$e_y = y - y_{\text{ref}}, \tag{8}$$

$$e_\vartheta = \vartheta - \vartheta_{\text{ref}}. \tag{9}$$

Notice that (8) leads to the following implicit definition of the state error ${}^jC\, e_x = e_y$. By substituting (8), (9) in (1) the regulation error dynamics is obtained as follows:

$$\begin{aligned}
\dot{e}_x &= {}^jA\, e_x + {}^jB\, u + {}^j\Phi_1\left(e_x\right), \\
e_y &= {}^jC\, e_x, \\
\dot{e}_\vartheta &= T\left(e_\vartheta\right)\boldsymbol{\omega}_{\mathbf{e}},
\end{aligned} \tag{10}$$

with

$$\boldsymbol{\omega}_{\mathbf{e}} \triangleq \omega_w - \omega_{\text{ref}}, \tag{11}$$

where $\omega_w = [P, Q, R]^T$ and

$${}^j\Phi_1\left(e_x\right) \triangleq {}^j\overline{\Phi}\left(e_x + x_{\text{ref}}\right) + {}^jA\, x_{\text{ref}}. \tag{12}$$

To achieve (7) two feedback loops can be used: one that depends on the velocity regulation error (body stabilization task) and another that performs the orientation regulation. In this work the following gain scheduling control structure is studied:

$$u = u_1 + u_2, \tag{13}$$

where

$$u_1 = -{}^j\overline{K}_{p,1}e_y - {}^j\varepsilon_1\,{}^jK_{I,1}\int e_y, \tag{14}$$

$$u_2 = -{}^j\overline{K}_{p,2}e_\vartheta - {}^j\varepsilon_2\,{}^jK_{I,2}\int e_\vartheta, \tag{15}$$

with $j = 1, \ldots, s$ where s is the number of modes defined by the design.

${}^j\overline{K}_{p,1}$, ${}^jK_{I,1} \in \mathbf{R}^{4\times 4}$ are the control gains for the body stabilization loop for every mode j.

${}^j\overline{K}_{p,2}$, ${}^jK_{I,2} \in \mathbf{R}^{4\times 4}$ are the control gains for the orientation control loop for every mode j.

${}^j\varepsilon_1$ and ${}^j\varepsilon_2 \in \mathbf{R}$ are constant.

The approach proposed in this work is to use the integral parameters epsilon 1 and epsilon 2 as perturbation parameters in order to establish stability conditions for the nonlinear closed-loop system. The schematic diagram of the controller (13)–(15) is given in Figure 1. To solve the regulation problem, consider the following additional assumptions:

(A.2) the triplet $({}^jA, {}^jB, {}^jC)$ is controllable and observable for every operation mode;

(A.3) the operation domain or flight envelope D satisfies $D = \cup\Omega_j$ with $\text{int}\,\Omega_j \bigcap \text{int}\,\Omega_i = \emptyset$ for $i \neq j$ and $\Omega_j \neq \emptyset$.

3. Preliminaries

In this section, important results to deduce the main contribution of this paper are introduced. To this end, let us analyze

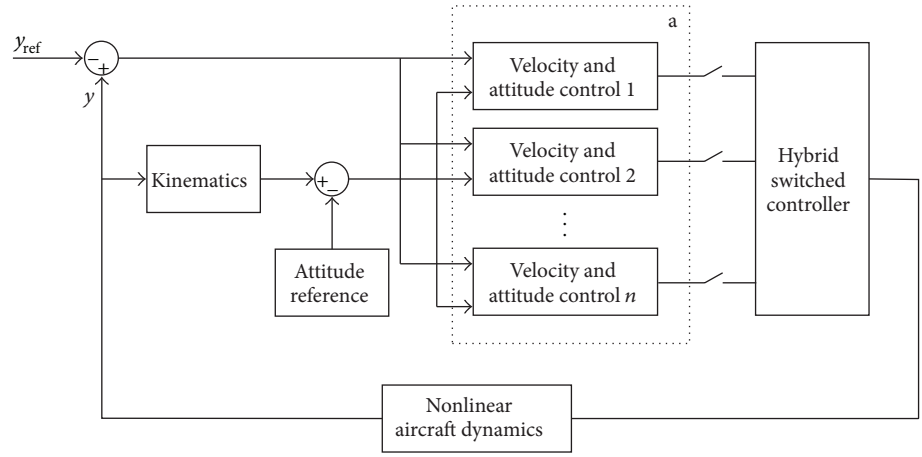


Figure 1: Controller structure.

the aircraft dynamics (10) under the action of the control law (14), which is given by

$$\dot{e} = \begin{bmatrix} {}^{j}A_1 & I \\ -{}^{j}\varepsilon_1\, {}^{j}B\, {}^{j}K_{I,1}\, {}^{j}C & 0_{6\times 6} \end{bmatrix} e + \begin{bmatrix} {}^{j}\Phi_1(e_x) \\ 0 \end{bmatrix}, \tag{16}$$

where e_x, e_s are the state and quaternion tracking error, respectively, $e = [e_x, e_s]^T \in \mathbf{R}^{12}$ is a virtual state of the system associated with the integral action, and

$${}^{j}A_1 \triangleq {}^{j}A - {}^{j}B\, {}^{j}\overline{K}_{p,1}\, {}^{j}C \in \mathbf{R}^{6\times 6}. \tag{17}$$

The equilibrium point of (16) is given by

$$e_{x,\mathrm{eq}} = -{}^{j}A_1^{-1}\left[b_I + {}^{j}\Phi_1\left(e_{x,\mathrm{eq}}\right)\right] \triangleq g\left(b_I, e_{x,\mathrm{eq}}\right), \tag{18}$$

where $\dot{e}_s = 0 \Rightarrow e_s = b_I$ and b_I is a constant vector. Regarding the properties of the system at equilibrium the following propositions are stated.

Proposition 1. *There exists an explicit function $e_{x,eq} = h(b_I)$ derived from* (18) *if*

$$\left\| {}^{j}A_1^{-1} \right\| < \frac{1}{{}^{j}L_{d1}}, \tag{19}$$

where ${}^{j}L_{d1}$ is the Lipschitz constant of ${}^{j}\Phi_1(e_x)$.

Proof. Notice that

$$\begin{aligned} \left\| -{}^{j}A_1^{-1} \nabla_{e_{x,\mathrm{eq}}}\, {}^{j}\Phi_1\left(e_{x,\mathrm{eq}}\right) \right\| &< \left\| {}^{j}A_1^{-1} \right\| \left\| \nabla_{e_{x,\mathrm{eq}}}\, {}^{j}\Phi_1\left(e_{x,\mathrm{eq}}\right) \right\| \\ &< \left\| {}^{j}A_1^{-1} \right\| {}^{j}L_{d1}. \end{aligned} \tag{20}$$

From the hypothesis, we have that $\| {}^{j}A_1^{-1} \|\, {}^{j}L_{d1} < 1$; therefore $\| - {}^{j}A_1^{-1} \nabla_{e_{x,\mathrm{eq}}}\, {}^{j}\Phi_1(e_{x,\mathrm{eq}}) \| < 1$. Furthermore since ${}^{j}\Phi_1(e_{x,\mathrm{eq}})$ is differentiable, from (18) we have

$$\left\| \nabla_{e_{x,\mathrm{eq}}} g\left(e_{x,\mathrm{eq}}\right) \right\| = \left\| -{}^{j}A_1^{-1} \nabla_{e_{x,\mathrm{eq}}}\, {}^{j}\Phi_1\left(e_{x,\mathrm{eq}}\right) \right\| < 1. \tag{21}$$

Using the contraction mapping theorem [24], there exists a unique solution of (18); hence there exists an explicit map $e_{x,\mathrm{eq}} = h(b_I)$.

In other words, if ${}^{j}\overline{K}_{p,1} \triangleq {}^{j}\gamma\, {}^{j}K_{p,1}$, where ${}^{j}\gamma > 0$, hence (20) can be rewritten as

$$\begin{aligned} \left\| {}^{j}\gamma \left(\frac{{}^{j}A}{{}^{j}\gamma} - {}^{j}B\, {}^{j}K_{p,1}\, {}^{j}C \right) \right\| &\leq \left| {}^{j}\gamma \right| \left\| \left(\frac{{}^{j}A}{{}^{j}\gamma} - {}^{j}B\, {}^{j}K_{p,1}\, {}^{j}C \right) \right\|, \\ {}^{j}L_{d1} &< {}^{j}\gamma \left\| \frac{{}^{j}A}{{}^{j}\gamma} - {}^{j}B\, {}^{j}K_{p,1}\, {}^{j}C \right\|. \end{aligned} \tag{22}$$

From (22) it can be seen that if ${}^{j}\gamma \to \infty$, then ${}^{j}\gamma \,\| ({}^{j}A / {}^{j}\gamma) - {}^{j}B\, {}^{j}K_{p,1}\, {}^{j}C \| \to \infty$. Such results imply the existence of a sufficiently large ${}^{j}\gamma$ such that condition of this proposition is satisfied. □

Proposition 2. *Considering the map*

$$e_s = -{}^{j}A_1 e_x - {}^{j}\Phi_1\left(e_x\right), \tag{23}$$

there exist real negative eigenvalues of ${}^{j}A_1$, such that $[\nabla_{e_x} e_s]^{-1}$ is positive definite.

Proof. Computing the e_x-derivative of (23), we have

$$\nabla_{e_x} e_s = -{}^{j}A_1 - \nabla_{e_x}\, {}^{j}\Phi_1\left(e_x\right). \tag{24}$$

According to Theorem 3 in [25] and Assumption (A.2), the eigenvalues of (17) can be arbitrarily assigned (by virtue of gain ${}^{j}\overline{K}_{p,1}$). Therefore, they are chosen for ${}^{j}A_1$ to be negative definite. Notice that in view of this fact ${}^{j}A_1$ is Hurwitz and $-{}^{j}A_1$ is positive definite. Moreover, given (17), if ${}^{j}\overline{K}_{p,1} = {}^{j}\gamma\, {}^{j}K_{p,1}$, there exists a sufficiently large value of ${}^{j}\gamma$ such that the term ${}^{j}A_1$ prevails over $\nabla_{e_x}\, {}^{j}\Phi_1(e_x)$, which by virtue of Assumption (A.1) is bounded in Ω_j. These arguments lead

us to deduce the existence of controller gains ${}^{j}\overline{K}_{p,1}$ such that $\nabla_{e_x} e_s$ is positive definite, which implies that $[\nabla_{e_x} e_s]^{-1}$ is positive definite. □

4. Main Results

In this section, the main contributions of this work are stated in Theorems 3 and 4. These theorems contain results regarding the stability of every operation mode (i.e., stability of a two-loop controller acting in a given operation region) and the stability of the entire operation region or flight envelope conditions.

4.1. Stability of Every Operation Mode. This stability analysis is performed in two parts; firstly the stability of the body stabilization loop is analyzed and afterwards the orientation stabilization loop. The stability analysis makes use of singularly perturbed systems [26], since the idea is to use integral control gains as perturbation parameters. In other words, since the integral actions constitute slow actions of the controller, the closed-loop system can be described naturally using two systems: a fast system given by the aircraft dynamics under a pure proportional action and a slow system given by the integral dynamics of the regulation error. The singular perturbed system arises as the integral gains approach zero. The interested reader is referenced to [26, 27] for a more detailed explanation of singularly perturbed systems.

Theorem 3. *Let the dynamics of the aircraft be described by* (10)–(15) *in the set* Ω_j, *let* ${}^{j}K_{I,i} \in \mathbf{R}^{4\times 4}$, $i \in \{1, 2\}$, $j \in \{1, \ldots, s\}$, *and let* ${}^{j}\overline{K}_{p,1}$ *be as in Propositions 1 and 2; then there exist constants* $0 <^{j} \varepsilon_1 < \varepsilon_{\max}$ *and* $0 <^{j} \varepsilon_2 < \varepsilon_{\max}$, *such that the origin of* (10) *under the action of control* (13) *is asymptotically stable.*

Proof. As a first step, consider the aircraft under the body stabilization loop. From Proposition 1, a unique isolated equilibrium exists; let us denote such equilibrium values as $e_{x,\mathrm{eq}}$ and $e_{s,\mathrm{eq}}$. In view of this fact, the following coordinates are introduced to shift such equilibrium to the origin

$$\begin{aligned} z &\triangleq e_x - e_{x,\mathrm{eq}}, \\ w &\triangleq e_s - e_{s,\mathrm{eq}}. \end{aligned} \tag{25}$$

At this point, let us use variable ${}^{j}\varepsilon_1 > 0$ as a perturbation parameter; hence the closed-loop system (10)–(13) can be rewritten as a standard singularly perturbed system given by

$$\begin{bmatrix} {}^{j}\varepsilon_1 z' \\ w' \end{bmatrix} = \begin{bmatrix} {}^{j}A_1 & I \\ -{}^{j}B\,{}^{j}K_{I,1}\,{}^{j}C & 0_{4\times 4} \end{bmatrix} \begin{bmatrix} z \\ w \end{bmatrix} + \begin{bmatrix} {}^{j}\Phi_2(z) \\ 0 \end{bmatrix}, \tag{26}$$

where ${}^{j}\Phi_2(z) \triangleq^{j} \Phi_1(z + e_{x,\mathrm{eq}})$, $t' = {}^{j}\epsilon_1 t$, $z \triangleq z(t')$, $z' \triangleq dz/dt'$, $w \triangleq w(t')$, and $w' \triangleq dw/dt'$.

The boundary-layer and reduced systems associated with (26) are given, respectively, by

$$\dot{z} = {}^{j}A_1 z + {}^{j}\Phi_2(z), \tag{27}$$

$$0 = {}^{j}A_1 z + w + {}^{j}\Phi_2(z), \tag{28}$$

$$w' = -{}^{j}B\,{}^{j}K_{I,1}\,{}^{j}C\,z. \tag{29}$$

Observe that (27) can be seen as the aircraft system under pure proportional action. On the other hand, notice that (28) is a steady-state relationship between z and w that satisfies $z = -{}^{j}A_1^{-1}[w + {}^{j}\Phi_2(z)]$, which has the same structure of (18); therefore Proposition 1 can be used; hence, for a suitable proportional gain (i.e., sufficiently large ${}^{j}\gamma$), there exists an explicit function $z \triangleq h(w)$ such that

$$w' = -{}^{j}B\,{}^{j}K_{I,1}\,{}^{j}C\,h(w). \tag{30}$$

Firstly, we will show that the linear output feedback (14) is able to track asymptotically the velocity of the nonlinear aircraft. Observe that (30) can be seen as the aircraft system under pure integral action. To derive the stability properties of (26), we can use systems (27) and (30).

We start by showing that, for a sufficiently large ${}^{j}\gamma$, the origin of (27) is exponentially stable. To this end, let us consider the following scalar function:

$${}^{j}V_1 = z^T\,{}^{j}P_1 z, \tag{31}$$

where ${}^{j}P_1 \in \mathbf{R}^{6\times 6}$ is positive definite. Since ${}^{j}A_1$ is Hurwitz, the time derivative of (31) along (27) is given by

$${}^{j}\dot{V}_1 = -z^T\,{}^{j}Q_1 z + 2z^T\,{}^{j}P_1\,{}^{j}\Phi_2(z), \tag{32}$$

with ${}^{j}Q_1$ positive definite, since ${}^{j}\Phi_2(z)$ is Lipschitz; hence

$$\begin{aligned} {}^{j}\dot{V}_1 &\leq -\|z\|^2 \left\|{}^{j}Q_1\right\| + 2\|z\|^2 \left\|{}^{j}P_1\right\| L_1, \\ {}^{j}\dot{V}_1 &\leq \left(-\left\|{}^{j}Q_1\right\| + 2L_1\left\|{}^{j}P_1\right\|\right)\|z\|^2, \end{aligned} \tag{33}$$

where L_1 is the Lipschitz constant of ${}^{j}\Phi_2(z)$. To ensure that $(-\|{}^{j}Q_1\| + 2L_1\|{}^{j}P_1\|)$ is negative definite, it is required that $2L_1\|{}^{j}P_1\| \leq \|{}^{j}Q_1\|$, since $\|{}^{j}Q_1\| = \|{}^{j}A_1^T\,{}^{j}P_1 + {}^{j}P_1\,{}^{j}A_1\|$ and ${}^{j}\overline{K}_{p,1} = {}^{j}\gamma\,{}^{j}K_{p,1}$ with ${}^{j}\gamma > 0$; hence

$$\begin{aligned} \left\|{}^{j}Q_1\right\| &= \left\|{}^{j}A_1^T\,{}^{j}P_1 + {}^{j}P_1\,{}^{j}A_1\right\| \\ &= \left\| {}^{j}\gamma \left[\frac{{}^{j}A}{{}^{j}\gamma} - {}^{j}B\,{}^{j}K_{p,1}\,{}^{j}C\right]^T {}^{j}P_1 \right. \\ &\qquad \left. + {}^{j}\gamma\,{}^{j}P_1 \left[\frac{{}^{j}A}{{}^{j}\gamma} - {}^{j}B\,{}^{j}K_{p,1}\,{}^{j}C\right] \right\| \\ &= \left\|{}^{j}\gamma\right\| \left\| \left[\frac{{}^{j}A}{{}^{j}\gamma} - {}^{j}B\,{}^{j}K_{p,1}\,{}^{j}C\right]^T {}^{j}P_1 \right. \\ &\qquad \left. + {}^{j}P_1 \left[\frac{{}^{j}A}{{}^{j}\gamma} - {}^{j}B\,{}^{j}K_{p,1}\,{}^{j}C\right] \right\|. \end{aligned} \tag{34}$$

From (34) it can be observed that $\lim_{{}^{j}\gamma \to \infty} \| {}^{j}Q_1 \| = \infty$, which implies that, for a sufficiently large ${}^{j}\gamma$, the origin of (27) is exponentially stable.

On the other hand, to study stability of the reduced system (30), consider the scalar function ${}^{j}V_2 = (1/2) w^T w$. Let $E \triangleq {}^{j}B\, {}^{j}K_{I,1}\, {}^{j}C$; hence the time derivative of ${}^{j}V_2$ along (27) is ${}^{j}\dot{V}_2 = -w^T E h(w)$.

Using the result in Proposition 2, $\nabla_w h(w) = [\nabla_{e_x} e_s]^{-1}$ is positive definite, which in turn implies that ${}^{j}\dot{V}_2$ is positive semidefinite. Using the Lasalle invariance principle and provided that ${}^{j}\dot{V}_2 = 0$ only in $z = 0$, $w = 0$, it is possible to conclude that the origin of (27) is asymptotically stable.

At this point, asymptotic stability of (16) can be concluded using Theorem 2.1 in [26]. The stability properties of the reduced and boundary layer system imply the existence of a parameter $\varepsilon_{\max} > 0$ such that, for every $0 < {}^{j}\varepsilon_1 < \varepsilon_{\max}$, the origin of (16) is asymptotically stable, with a scalar function

$$ {}^{j}V_3 = \frac{\left({}^{j}V_1 + {}^{j}V_2 \right)}{{}^{j}\varepsilon_1}. \tag{35} $$

So far, it has been shown that the linear output feedback (14) is able to asymptotically stabilize the velocity of the nonlinear aircraft.

To show that the second linear loop (15) is able to regulate the orientation, consider the following: let us compute the closed-loop system (10)–(15) under steady-state conditions:

$$ 0 = f\left(\sigma_{\text{eq}}\right) - \Lambda\left(e_{\vartheta,\text{eq}}; {}^{j}K_{p,2}\right) + \overline{B}\left(\rho_{\text{eq}}\right), \tag{36} $$

$$ 0 = T\left(e_{\vartheta,\text{eq}}\right)[P, Q, R]^T, \tag{37} $$

$$ 0 = {}^{j}\varepsilon_2\, {}^{j}K_{I,2} e_{\vartheta,\text{eq}}, \tag{38} $$

where the subindex eq is the value of the variable under equilibrium conditions, $\sigma = [z, w]^T$, $f(\sigma)$ is the right-hand side of (26), ρ is the virtual state associated with the integral action, $\overline{B}(\rho) \triangleq [{}^{j}B\, \rho, 0]^T$, and

$$ \Lambda\left(e; {}^{j}K_{p,2}\right) \triangleq \begin{bmatrix} {}^{j}B\, {}^{j}K_{p,2} e_\vartheta \\ 0 \end{bmatrix}. \tag{39} $$

From (38) it is clear that $e_{\vartheta,\text{eq}} = 0$, since ${}^{j}K_{I,2}$ is full rank. Observe that $T(0) = 0$, $\Lambda(0; {}^{j}K_{p,2}) = 0$, and $\overline{B}(\rho_{\text{eq}})$ is a constant vector, so (36) can be rewritten as

$$ 0 = f\left(\sigma_{\text{eq}}\right) + \overline{B}\left(\rho_{\text{eq}}\right). \tag{40} $$

Observe that (40) has the same structure as (18), except for the translation term $\overline{B}(\rho_{\text{eq}})$. Hence Proposition 1 can be used to guarantee the existence and uniqueness of the equilibrium point. That implies that in steady-state conditions for every value of the state one and only one integral action corresponds. As before, let us consider system (26) under the action of controller (15) as a singularly perturbed system, with the gain of the integral action of the second loop as a new perturbation parameter. In this way, such closed-loop system can be rewritten as:

$$ \begin{aligned} {}^{j}\varepsilon_2 \widetilde{\sigma}' &= f\left(\widetilde{\sigma}'\right) - \Lambda\left(\widetilde{e}'_\vartheta; {}^{j}K_{p,2}\right) + \overline{B}\left(\widetilde{\rho}'\right), \\ {}^{j}\varepsilon_2 \widetilde{e}'_\vartheta &= T\left(\widetilde{e}'_\vartheta\right)[P, Q, R]^T, \\ \widetilde{\rho}' &= -{}^{j}K_{I,2} \widetilde{e}'_\vartheta, \end{aligned} \tag{41} $$

with $\widetilde{\sigma} = [\widetilde{z}, \widetilde{w}] \triangleq \sigma - \sigma_{\text{eq}} = [z - z_{\text{eq}}, w - w_{\text{eq}}]$, $\widetilde{\rho} \triangleq \rho - \rho_{\text{eq}}$, $\tau' = {}^{j}\varepsilon_2 t$ is the new time scale, and $\widetilde{e}_\vartheta \triangleq e_\vartheta - e_{\vartheta,\text{eq}}$, $\widetilde{\sigma}' \triangleq d\widetilde{\sigma}/d\tau'$, $\widetilde{\rho}' \triangleq d\widetilde{\rho}/d\tau'$, and $\widetilde{e}'_\vartheta \triangleq d\widetilde{e}_\vartheta/d\tau'$. Observe that (41) has an equilibrium point in $(\widetilde{\sigma}, \widetilde{e}_\vartheta, \widetilde{\rho}) = (0, 0, 0)$. Moreover, observe that $(\widetilde{\sigma}, \widetilde{e}_\vartheta, \widetilde{\rho}) \to (0, 0, 0)$, as $\tau' \to \infty$ implies that $(\widetilde{\sigma}, \widetilde{e}_\vartheta, \widetilde{\rho}) \to (0, 0, 0)$ as $t \to \infty$.

At this point, let us compute the reduced and boundary layer systems associated with (41), which are given, respectively, by

$$ \widetilde{\rho}' = -{}^{j}K_{I,2}\, {}^{j}K_{p,2}^{-1} \widetilde{\rho}, \tag{42} $$

$$ \dot{\widetilde{\sigma}} = f(\widetilde{\sigma}). \tag{43} $$

Notice that (43) has been studied before; the origin of such system is asymptotically stable for $0 < {}^{j}\varepsilon_1 < \varepsilon_{\max}$. On the other hand, the stability of the reduced system (42) can be readily established by suitable gains $-{}^{j}K_{I,2}^{j} K_{p,2}$ that make the matrix Hurwitz, hence, the existence of a Lyapunov function ${}^{j}V_4 = \widetilde{\rho}^T\, {}^{j}P_4 \widetilde{\rho}$ associated with (42) can be derived such that ${}^{j}\dot{V}_4 = -\widetilde{\rho}^T\, {}^{j}Q_4 \widetilde{\rho} < 0$.

Given the arguments above and the Lipschitz nature of ${}^{j}\Phi_2(\widetilde{z})$, Theorem 2.1 in [26] implies the existence of a parameter $\varepsilon_{\max} > 0$ such that for all $0 < {}^{j}\varepsilon_2 < \varepsilon_{\max}$ the origin of (41) is asymptotically stable. □

So far, it has been proven that a linear control is able to stabilize both velocity and orientation of the nonlinear aircraft; however, such result has been derived from the hypothesis that the system description is continuous. To relax such assumption, in the following we will study the stability of a family of piecewise continuous functions, which describes the dynamic of the aircraft at a family of desired operation points.

4.2. Stability of the Switching Structure. Conditions in Theorem 3 guarantee the stability in a local domain Ω_j. However, they do not constitute sufficient evidence to guarantee the stability of the entire flight envelope. To obtain such evidence, the piecewise continuous system under the action of discrete event switching has to be studied.

Since the origin of every continuous description of the system can be defined using different operation points (reference), the switching among modes can be described using discrete event switching depending on time, system's state, or both.

In view of this fact, it is necessary to study the requirements on the discrete event actions that ensure asymptotic stability of the closed-loop hybrid system. The following theorem establishes such conditions.

Theorem 4. *Assume that conditions of Theorem 3 are satisfied and let*

$$^{j}V_{tot}(\pi) = z^{T}\,{}^{j}P_{1}z + \frac{1}{2}w^{T}w + \tilde{\rho}^{T}\,{}^{j}P_{4}\tilde{\rho} \triangleq \pi^{T}\,{}^{j}\hat{P}\,\pi, \quad (44)$$

with $\pi^{T} = [z, w, \tilde{\rho}]$. *The origin of the hybrid system* (10)–(13) *is asymptotically stable for the entire flight envelope using switching criteria guaranteeing an average dwell time for the jth mode,* $T_{AD} > 0$, *such that*

$$^{j}V_{tot}\left(\pi\left({}^{j+}t_{l}\right)\right) \geq {}^{j}V_{tot}\left(\pi\left({}^{j-}t_{l+1}\right)\right), \quad (45)$$

with ${}^{j+}t_{l}$, ${}^{j-}t_{l}$ *are the initial time and final time when the jth mode becomes active for the lth time, respectively.*

Proof. The proof will be performed using Theorem 7 in [28]. According to the result in [28], the existence of a differentiable and radially unbounded scalar function $^{j}V_{\text{tot}} > 0$ must be shown such that $^{j}\dot{V}_{\text{tot}} \leq 0$ along trajectories of (41). Let us consider the following quadratic positive definite scalar function:

$$^{j}V_{\text{tot}}(\pi) = z^{T}\,{}^{j}P_{1}z + \frac{1}{2}w^{T}w + \tilde{\rho}^{T}\,{}^{j}P_{4}\tilde{\rho}. \quad (46)$$

From Theorem 3,

$$\alpha_{1}(\|\pi\|) \leq {}^{j}V_{\text{tot}}(\pi) \leq \alpha_{2}(\|\pi\|), \quad (47)$$

with $\alpha_{2}(\|\pi\|) = 3\max_{j}\{\lambda_{\max}({}^{j}P_{1})/{}^{j}\varepsilon_{1}\,{}^{j}\varepsilon_{2}, 1/{}^{j}\varepsilon_{1}\,{}^{j}\varepsilon_{2}, \lambda_{\max}({}^{j}P_{4})/{}^{j}\varepsilon_{2}\}\|\pi\|^{2}$, and $\alpha_{1}(\|\pi\|) = 3\min_{j}\{\lambda_{\min}({}^{j}P_{1})/{}^{j}\varepsilon_{1}\,{}^{j}\varepsilon_{2}, 1/{}^{j}\varepsilon_{1}\,{}^{j}\varepsilon_{2}, \lambda_{\min}({}^{j}P_{4})/{}^{j}\varepsilon_{2}\}\|\pi\|^{2}$. On the other hand, the time derivative of (44) along (10)–(13) is given by

$$\begin{aligned} &{}^{j}\dot{V}_{\text{tot}}(\pi) \\ &\leq -\left\{\left\{-\left\|{}^{j}Q_{1}\right\| + 2L_{1}\left\|{}^{j}P_{1}\right\|\right\}\|z\|^{2} + w^{T}Eh(w) + \tilde{\rho}^{T}\,{}^{j}Q_{4}\tilde{\rho}\right\} \\ &\triangleq -{}^{j}W_{\text{tot}}(\pi). \end{aligned} \quad (48)$$

From Theorem 3 in this work, the origin of every operation mode j is asymptotically stable. Using this property and Remark 6 in [28] it is possible to conclude the satisfaction of Conditions 1 and 2 of Theorem 7 in [28], that is, the existence of a class $\mathcal{K}_{\infty}$ function (κ) such that $|\pi(t)| \leq \kappa(\|{}^{j}W_{\text{tot}}(\pi)\|_{[t,t+\tau]})$.

Finally, Remark 7 and Lemma 8 in [28] let us conclude that if a minimum average dwell time $T_{\text{AD}} > 0$,

$$^{j}V_{\text{tot}}\left(\pi\left({}^{j+}t_{l}\right)\right) \geq {}^{j}V_{\text{tot}}\left(\pi\left({}^{j-}t_{l+1}\right)\right), \quad (49)$$

with ${}^{j-}t_{l}$, ${}^{j+}t_{l}$ are the initial time and final time when the jth mode becomes active for the lth time, respectively, the origin of the hybrid system is asymptotically stable in D, where $D = \cup\,\Omega_{j}$. □

Remark 5. Theorem 4 provide us with guidelines to design switching laws for the hybrid system. According to such theorem, the gain scheduling control must use a supervisory action that decides whether a given mode is active using switching laws that can be either time or state dependent, as long as a minimum average dwell time is satisfied for every mode. This average time can be arbitrary small as long as $T_{\text{AD}} > 0$. The stability result in Theorem 4 actually establishes that the system will be stable if chattering does not occur. The phenomenon of chattering occurs when infinite switching actions occur in finite time and it may lead to finite time destabilization.

Remark 6. The results discussed in Theorems 3 and 4 are in fact quite general, and can be used to other systems that admit the same description, including aircrafts with some actuator failures. That is, the same result can be applied to other highly nonlinear systems as long as Assumptions (A.1)–(A.3) are satisfied and the piecewise description of the system fits the model in (10)–(13).

5. Illustrative Results

The aim of this section is to illustrate the theoretical results above. To this end, the nonlinear model of F16 (see [29]) is simulated using Simulink. In order to illustrate the effect of the uncertainty, the simulations are conducted as follows: firstly, a set of linear approximations are proposed at different operation points, each of these approximations is associated with a domain Ω_{j} and the linear approximation is used only to perform the controller tuning at the corresponding modes. Secondly, the gains are used to control a piecewise nonlinear model of the system whose modes coincide with the modes of the linear counterpart. For illustration purposes only two flight modes are used, namely, steady-state fight at different cruise velocities.

Firstly, the effect of the integral action is evaluated tracking a sinusoidal function. The switching criterion is chosen to be velocity dependent and at the surface $\nu = 325$ ft/s, so to perform the task the aircraft must switch consecutively between Mode 1 and Mode 2. The results are displayed in Figures 2, 3, and 4.

Figure 2 shows the time evolution of the cruise velocity for three different integral gains, which, according to Theorem 3, are parameterized using $^{j}\varepsilon_{1}$. Figure 2 shows that for simplicity ${}^{1}\varepsilon_{1} = {}^{2}\varepsilon_{1} = \varepsilon_{1}$ have been chosen and the three integral parameters are denoted as $\varepsilon_{1,a}$, $\varepsilon_{1,b}$, $\varepsilon_{1,c}$. Theorem 3 establishes the existence of a maximum integral gain such that the system is asymptotically stable. For $\varepsilon_{1,a} < \varepsilon_{1,b}$, the velocity reference can be followed; however, for $\varepsilon_{1,c} \gg \varepsilon_{1,b}$, the system leads to instability.

In Figure 3 it can be observed that for $\varepsilon_{1,a}$ the body stabilization task is able to avoid rotation, still with a short transient period. Such transient response is reduced with $\varepsilon_{1,b}$ and finally with $\varepsilon_{1,c}$ the aircraft is unstable. That is, moderate values of the integral action result in a successful velocity and orientation tracking and good performance. However since the maximum value of the integral gain that can be

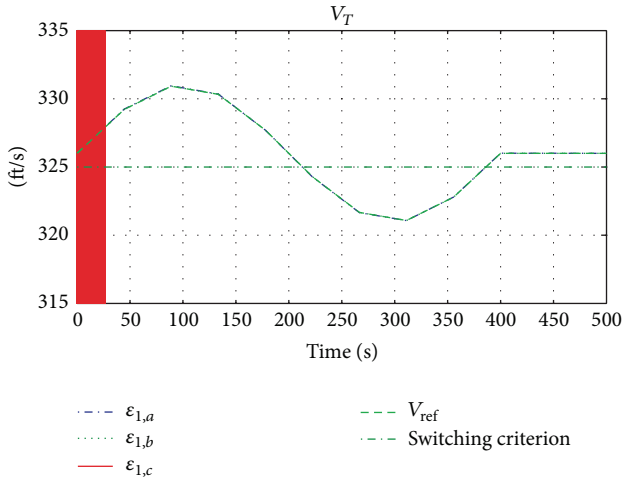


FIGURE 2: Time evolution of aircraft cruise velocity for three different integral gains ($\varepsilon_{1,a} > \varepsilon_{1,b} > \varepsilon_{1,c}$).

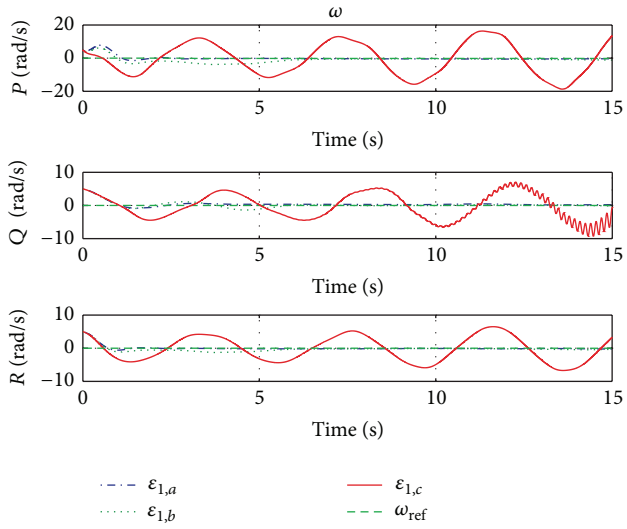


FIGURE 3: Time evolution of the aircraft attitude for three different integral gains ($\varepsilon_{1,a} > \varepsilon_{1,b} > \varepsilon_{1,c}$).

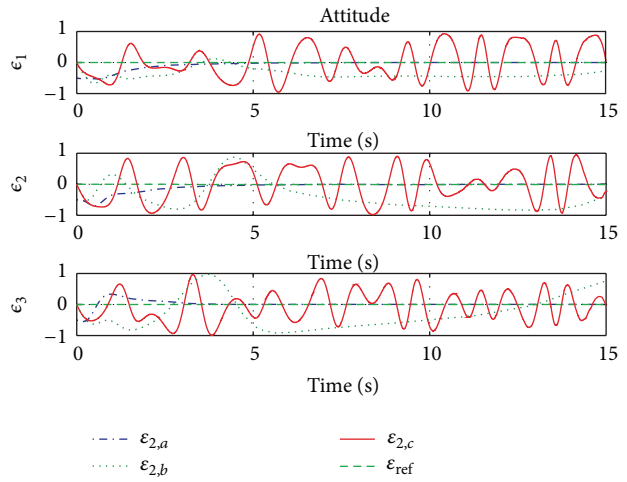


FIGURE 4: Time evolution of the angular velocity for three different integral gains ($\varepsilon_{2,a} > \varepsilon_{2,b} > \varepsilon_{2,c}$).

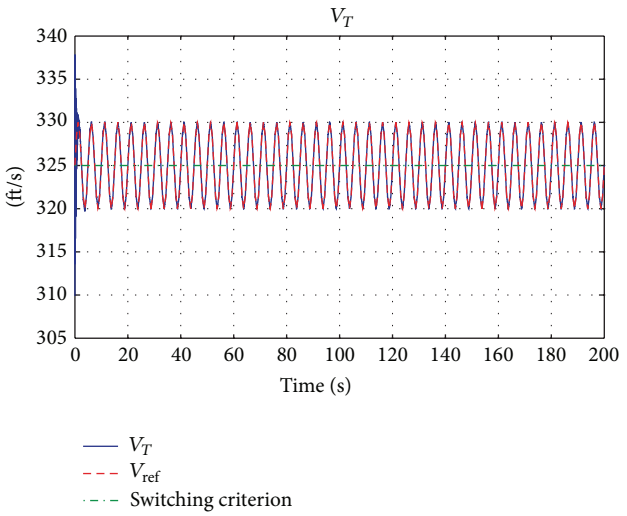


FIGURE 5: Time evolution of aircraft cruise velocity under fast switching conditions.

used is upper bounded, a performance criterion to choose the integral gains can be used to define properly the operation domains Ω_j.

Similar observations can be performed for the orientation loop (see Figure 4), which integral gains can be parameterized using ${}^j\varepsilon_2$. For simplicity, the integral parameters are chosen such that ${}^1\varepsilon_2 = {}^2\varepsilon_2 = \varepsilon_2$. Three different gains are evaluated $\varepsilon_{2,a} > \varepsilon_{2,b} > \varepsilon_{2,c}$. For small integral gains, the orientation can be successfully tracked ($\varepsilon_{2,a}$); however, higher gains lead firstly to a performance degradation ($\varepsilon_{2,b}$) and secondly to instability ($\varepsilon_{2,c}$).

Finally, to illustrate the result in Theorem 4, fast switching actions are induced by following a high frequency sinusoidal velocity reference (see Figures 5 and 6). The switching conditions are velocity dependent and set at a cruise velocity of 35 ft/sec. In Figure 6 it is possible to observe the corresponding behavior of the angular velocity and the aircraft attitude. Notice that even if the model is switching continuously, the evolution of the system remains stable and, for the controller gains in the simulations, they are able to track correctly the desired reference. Such result agrees with theoretical results in Theorem 4.

6. Conclusions

In this work, the stability properties of an aircraft with nonlinear behavior controlled with a gain scheduled approach are studied. It is shown that the inclusion of a linear integral action to an output feedback linear controller makes it possible to guarantee asymptotical tracking of velocity and orientation in spite of nonlinearities and uncertainties. This result is interesting since even if the behavior of the aircraft

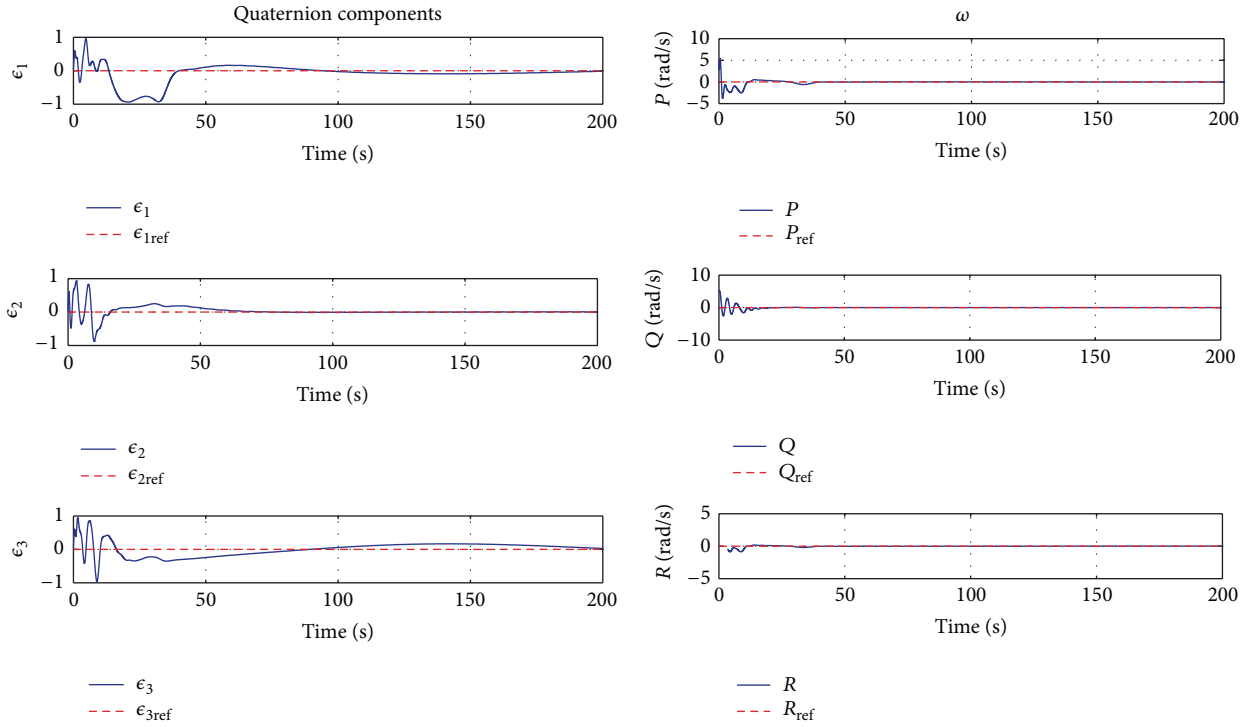


FIGURE 6: Time evolution of (a) aircraft attitude and (b) angular velocity under fast switching conditions.

system is highly nonlinear, the controller gains can be tuned using linear approximations of the system in given operation points and the performance of the system can be set using a trade-off between the size of the integral gains and the size of the operating domain Ω_j. The switching criteria can be either time or state dependent as long as a minimum residence time ($T_{AD} > 0$) is used in every mode and it can be designed using a supervisory structure which evaluates the satisfaction of condition (45). Due to this fact the control scheme is flexible and physical or user-defined restrictions can be taken into account.

Nomenclature

α:	Angle of attack
α_e:	Angle of attach at the operation point
α_T:	Drift angle
β:	Side slip angle
${}^j\gamma$:	Scalar associated with the proportional gain of the first loop
γ_e:	Flight path angle
Γ:	$J'_Z J'_X - J'^2_{XZ}$
δ_T:	Throttle position
δ_e, δ_a, and δ_r:	Control surface positions (elevator, aileron, and rudder)
${}^j\varepsilon_1$:	Scalar associated with the integral gain of the first loop
${}^j\varepsilon_2$:	Scalar associated with the integral gain of the second loop
$\eta, \epsilon_1, \epsilon_2, \epsilon_3$:	Quaternion components
μ:	$J'_Z J'_X / \Gamma$
σ_1:	$J'_Z J'_{XZ} / \Gamma$
σ_2:	$J'_X J'_{XZ} / \Gamma$
ϕ, θ, ψ:	Euler angles
${}^j\overline{\Phi}(x)$:	Nonlinearities vector
Ω_j:	Mode domain
${}^jA$:	State matrix for mode j
${}^jB$:	Control matrix for mode j
C:	Output matrix
D:	Flight envelope domain
D_a, L, C_a:	Drag, lift, and crosswind forces
D_{ae}, L_e, C_{ae}:	Drag, lift, and crosswind forces at the operation point
F_T:	Traction force
g:	Gravity acceleration
I:	Identity matrix
i:	Loop index
$I'_x, I'_y, I'_z, I'_{xz}$:	Inertia matrix components (body frame)
j:	Waypoint or mode index
$J_w, J'_X, J'_Y, J'_Z, J'_{XZ}$:	Inertia matrix components (wind frame)
l_w, m_w, n_w:	Rolling, pitching, and yawing moment
$l_{w_e}, m_{w_e}, n_{w_e}$:	Rolling, pitching, and yawing moment at the operation point
L_β:	$(1/J'_X)(\partial l_w / \partial \beta)$

L_p: $(1/J'_X)(\partial l_w/\partial P)$
L_r: $(1/J'_X)(\partial l_w/\partial R)$
m: Aircraft mass
M_α: $(1/J'_Y)(\partial m_w/\partial \alpha)$
M_q: $(1/J'_Y)(\partial m_w/\partial Q)$
m_w: Trust moment (body frame)
M_{T_α}: $(1/J'_Y)(\partial m_T/\partial \alpha)$
M_V: $(1/J'_Y)(\partial m_w/\partial V_T)$
M_{T_V}: $(1/J'_Y)(\partial m_T/\partial V_T)$
N_β: $(1/J'_Z)(\partial n_w/\partial \beta)$
N_p: $(1/J'_Z)(\partial n_w/\partial P)$
N_r: $(1/J'_Z)(\partial n_w/\partial R)$
P, Q, R: Angular velocity components (roll, pitch, and yaw)
u: Controlled input vector
V_T: Magnitude of airplane linear velocity
V_{T_e}: Airplane linear velocity at the operation point
x: State vector
X_α: $(1/\mathrm{m})[L - (\partial D_a/\partial \alpha)]$
X_V: $(-1/\mathrm{m})(\partial D_a/\partial V_T)$
X_{T_V}: $(1/\mathrm{m})(\partial F_T/\partial V_T)$
y: Output vector
Y_β: $(1/\mathrm{m})[-D_a + (\partial C_a/\partial \beta)]$
Y_p: $(1/\mathrm{m})(\partial C_a/\partial P)$
Y_r: $(1/\mathrm{m})(\partial C_a/\partial R)$
Z_α: $(-1/\mathrm{m})[D_a + (\partial D_a/\partial \alpha)]$
Z_q: $(-1/\mathrm{m})(\partial L/\partial Q)$
Z_V: $(-1/\mathrm{m})(\partial L/\partial V_T)$.

Conflict of Interests

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

Acknowledgments

Fany Mendez-Vergara and Angelica Mendoza-Torres acknowledge financial support from CONACYT. Ilse Cervantes and Angelica Mendoza-Torres acknowledge financial support from CONACYT-FORDECYT Grant no. 190966. The authors acknowledge technical support from Irwin Allen Diaz-Diaz.

References

[1] D. J. Leith and W. E. Leithead, "Survey of gain-scheduling analysis and design," *International Journal of Control*, vol. 73, no. 11, pp. 1001–1025, 2000.

[2] D. J. Stilwell, "State-space interpolation for a gain-scheduled autopilot," *Journal of Guidance, Control, and Dynamics*, vol. 24, no. 3, pp. 460–465, 2001.

[3] B. P. Rasmussen and Y. J. Chang, "Stable controller interpolation and controller switching for LPV systems," *Journal of Dynamic Systems, Measurement and Control*, vol. 132, no. 1, Article ID 011007, 12 pages, 2010.

[4] G. Cai, C. Hu, B. Yin, H. He, and X. Han, "Gain-scheduled H_2 controller synthesis for continuous-time polytopic LPV systems," *Mathematical Problems in Engineering*, vol. 2014, Article ID 972624, 14 pages, 2014.

[5] D. J. Leith and W. E. Leithead, "Gain-scheduled control: relaxing slow variation requirements by velocity-based design," *Journal of Guidance, Control, and Dynamics*, vol. 23, no. 6, pp. 988–1000, 2000.

[6] D. Enns, D. Bugajskia, R. Hendricka, and G. Stein, "Dynamic inversion: an envolving methodology for flight control design," *International Journal of Control*, vol. 59, no. 1, pp. 71–91, 1994.

[7] A. Packard, "Gain scheduling via linear fractional transformations," *Systems & Control Letters*, vol. 22, no. 2, pp. 79–92, 1994.

[8] P. Apkarian, P. Gahinet, and G. Becker, "Self-scheduled H_∞ control of linear parameter-varying systems: a design example," *Automatica*, vol. 31, no. 9, pp. 1251–1261, 1995.

[9] W. J. Rugh and J. S. Shamma, "Research on gain scheduling," *Automatica*, vol. 36, no. 10, pp. 1401–1425, 2000.

[10] A. Ataei and Q. Wang, "Non-linear control of an uncertain hypersonic aircraft model using robust sum-of-squares method," *IET Control Theory & Applications*, vol. 6, no. 2, pp. 203–215, 2012.

[11] M. T. Frye, S. Ding, C. Qian, and S. Li, "Fast convergent observer design for output feedback stabilisation of a planar vertical takeoff and landing aircraft," *IET Control Theory & Applications*, vol. 4, no. 4, pp. 690–700, 2010.

[12] B. Zhu, X. Wang, and K.-Y. Cai, "Tracking control for angular-rate-sensorless vertical take-off and landing aircraft in the presence of angular-position measurement delay," *IET Control Theory & Applications*, vol. 4, no. 6, pp. 957–969, 2010.

[13] R. Zhang, Q. Quan, and K.-Y. Cai, "Attitude control of a quadrotor aircraft subject to a class of time-varying disturbances," *IET Control Theory & Applications*, vol. 5, no. 9, pp. 1140–1146, 2011.

[14] R. Kamyar and E. Taheri, "Aircraft optimal terrain/threat-based trajectory planning and control," *Journal of Guidance, Control, and Dynamics*, vol. 37, no. 2, pp. 466–483, 2014.

[15] Q. Hu, B. Xiao, D. Wang, and E. K. Poh, "Attitude control of spacecraft with actuator uncertainty," *Journal of Guidance, Control, and Dynamics*, vol. 36, no. 6, pp. 1771–1776, 2013.

[16] A. H. J. de Ruiter, "Spacecraft attitude tracking with guaranteed performance bounds," *Journal of Guidance, Control, and Dynamics*, vol. 36, no. 4, pp. 1214–1221, 2013.

[17] J. Hu and H. Zhang, "Bounded output feedback of rigid-body attitude via angular velocity observers," *Journal of Guidance, Control, and Dynamics*, vol. 36, no. 4, pp. 1240–1248, 2013.

[18] Q. Sang and G. Tao, "Multivariable adaptive piecewise linear control design for NASA generic transport model," *Journal of Guidance, Control, and Dynamics*, vol. 5, no. 35, pp. 1559–1567, 2012.

[19] P. Casau, D. Cabecinhas, and C. Silvestre, "Hybrid control strategy for the autonomous transition flight of a fixed-wing aircraft," *IEEE Transactions on Control Systems Technology*, vol. 6, no. 21, pp. 2194–2211, 2013.

[20] 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.

[21] I. Hwang, H. Balakrishnan, and C. Tomlin, "State estimation for hybrid systems: applications to aircraft tracking," *IEE Proceedings. Control Theory & Applications*, vol. 153, no. 5, pp. 556–566, 2006.

[22] L. G. Sun, C. C. de Visser, Q. P. Chu, and Q. P. Falkena, "Hybrid sensor-based backstepping control approach with its application to fault-tolerant flight control," *Journal of Guidance, Control, and Dynamics*, vol. 1, no. 37, pp. 59–71, 2014.

[23] J. Roskam, *Airplane Flight Dynamics and Automatic Flight Controls. Part I*, John Wiley & Sons, 3rd edition, 1995.

[24] A. N. Kolmogorov and S. V. Formin, *Elements of the Theory of the Functional Analysis-I: Metric and Normed Sapaces*, Courier Dover Publications, Mineola, NY, USA, 1957.

[25] H. Kimura, "Pole assignment by gain output feedback," *IEEE Transactions on Automatic Control*, vol. 20, no. 4, pp. 509–516, 1975.

[26] P. Kokotović, H. K. Khalil, and J. O'Reilly, *Singular Perturbation Methods in Control: Analysis and Design*, vol. 25, Prentice Hall, Englewood Cliffs, NJ, USA, 3rd edition, 1999.

[27] F. C. Hoppensteadt, "Singular perturbations on the infinite interval," *Transactions of the American Mathematical Society*, vol. 123, no. 2, pp. 521–535, 1966.

[28] J. P. Hespanha, D. Liberzon, D. Angeli, and E. D. Sontag, "Nonlinear norm-observability notions and stability of switched systems," *IEEE Transactions on Automatic Control*, vol. 50, no. 2, pp. 154–168, 2005.

[29] B. L. Stevens and F. L. Lewis, *Aircraft Control and Simualtion*, John Wiley & Sons, New York, NY, USA, 2nd edition, 2003.

Hindawi Publishing Corporation
Mathematical Problems in Engineering
Volume 2014, Article ID 693607, 10 pages
http://dx.doi.org/10.1155/2014/693607

Research Article

Dynamic Surface Control and Its Application to Lateral Vehicle Control

Bongsob Song,[1] J. Karl Hedrick,[2] and Yeonsik Kang[3]

[1] *Department of Mechanical Engineering, Ajou University, Suwon 443-749, Republic of Korea*
[2] *Department of Mechanical Engineering, University of California, Berkeley, CA 94720, USA*
[3] *Department of Automotive Engineering, Kookmin University, Seoul 136-702, Republic of Korea*

Correspondence should be addressed to Bongsob Song; bsong@ajou.ac.kr

Received 4 December 2013; Accepted 31 March 2014; Published 29 April 2014

Academic Editor: Ilse Cervantes Camacho

Copyright © 2014 Bongsob Song 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 extends the design and analysis methodology of dynamic surface control (DSC) in Song and Hedrick, 2011, for a more general class of nonlinear systems. When rotational mechanical systems such as lateral vehicle control and robot control are considered for applications, sinusoidal functions are easily included in the equation of motions. If such a sinusoidal function is used as a forcing term for DSC, the stability analysis faces the difficulty due to highly nonlinear functions resulting from the low-pass filter dynamics. With modification of input variables to the filter dynamics, the burden of mathematical analysis can be reduced and stability conditions in linear matrix inequality form to guarantee the quadratic stability via DSC are derived for the given class of nonlinear systems. Finally, the proposed design and analysis approach are applied to lateral vehicle control for forward automated driving and backward parallel parking at a low speed as well as an illustrative example.

1. Introduction

The dynamic surface control (DSC), one of robust nonlinear control techniques, has been developed with a wide spectrum of applications including throttle/brake control on automated vehicles [1], underactuated ship control [2], and robot control [3]. This control technique is a dynamic extension of multiple sliding surface control with a series of first order low-pass filters to avoid an "explosion of terms" [4]. The existence of DSC gains and filter time constants for semiglobal stability was theoretically proved in [4]. Recently, a noble analysis method in the framework of convex optimization has been introduced to allow us to find a quadratic Lyapunov function numerically for a class of nonlinear systems called "strict-feedback" form as follows [3]:

$$\begin{aligned} \dot{x}_i &= x_{i+1} + f_1(x_1, \ldots, x_i) \quad \text{for } i = 1, \ldots, n-1, \\ \dot{x}_n &= u + f_n(x_1, \ldots, x_n). \end{aligned} \tag{1}$$

Furthermore, if x_{i+1} in (1) is replaced by $g_{i+1}(x_{i+1})$ where g_{i+1} and $[\partial g_{i+1}/\partial x]$ are continuous and invertible, the design procedure proposed by Swaroop et al. [4] and Gerdes and Hedrick [5] can be still applied for the given system.

However, this replacement induces another highly nonlinear function resulting from the low-pass filter error dynamics when stability analysis is performed. The following example illustrates the design approach of DSC as well as the difficulty that this paper seeks to solve:

$$\dot{x}_1 = \tan x_2 + f_1(x_1), \tag{2}$$

$$\dot{x}_2 = u, \tag{3}$$

where f_1 and $[\partial f_1/\partial x_1]$ are continuous on $\mathscr{D}$; for example, $\mathscr{D} = \{x \in \mathfrak{R}^2 \mid |x_1| \leq 1, |x_2| \leq \pi/4\}$; thus both $\tan x_2$ and f_1 are bounded on $\mathscr{D}$. The control objective is to stabilize the system; that is, $x_1 \to 0$. First, define the first error surface as $S_1 = x_1$. After taking its derivative along the trajectory of (2)

$$\dot{S}_1 = \tan x_2 + f_1. \tag{4}$$

Then, the synthetic input, which is forced to drive $S_1 \to 0$, is derived as

$$\tan \overline{x}_2 = -f_1 - KS_1 \Longrightarrow \overline{x}_2 = \tan^{-1}(-f_1 - KS_1), \tag{5}$$

where K is a controller gain. We now define the second sliding surface $S_2 = x_2 - x_{2d}$, where x_{2d} equals $\overline{x}_2$ passed through a first order low-pass filter; that is,

$$\tau \dot{x}_{2d} + x_{2d} = \overline{x}_2, \qquad x_{2d}(0) = \overline{x}_2(0), \tag{6}$$

where τ is the filter time constant. Finally, the control input is derived as

$$u = \dot{x}_{2d} - KS_2 = \frac{\overline{x}_2 - x_{2d}}{\tau} - KS_2. \tag{7}$$

Next, the stability analysis is investigated based on the closed-loop dynamics as suggested in [3]. If both $\tan \overline{x}_2$ and $\tan x_{2d}$ are added and subtracted in (2) and u in (7) is put in (3), the closed-loop dynamics is written as

$$\begin{aligned} \dot{x}_1 &= (\tan x_2 - \tan x_{2d}) + (\tan x_{2d} - \tan \overline{x}_2) + \tan \overline{x}_2 + f_1, \\ \dot{x}_2 &= \dot{x}_{2d} - KS_2. \end{aligned} \tag{8}$$

By use of (5) and definitions of S_1 and S_2, (8) is rewritten as

$$\begin{aligned} \dot{S}_1 &= (\tan x_2 - \tan x_{2d}) + (\tan x_{2d} - \tan \overline{x}_2) - KS_1, \\ \dot{S}_2 &= -KS_2. \end{aligned} \tag{9}$$

Since the first order low-pass filter in (6) is added, the filter dynamics should be included in the closed-loop dynamics for stability analysis. After defining the filter error, $\xi = x_{2d} - \overline{x}_2$, the augmented closed-loop dynamics is summarized as

$$\begin{aligned} \dot{S}_1 &= -KS_1 + (\tan x_2 - \tan x_{2d}) + (\tan x_{2d} - \tan \overline{x}_2) \\ &:= -KS_1 + w_1 + w_2, \\ \dot{S}_2 &= -KS_2, \\ \dot{\xi} &= -\frac{\xi}{\tau} - \frac{d}{dt}\left\{\tan^{-1}(-f_1 - KS_1)\right\} \\ &= -\frac{\xi}{\tau} + \frac{1}{1 + (f_1 + KS_1)^2}\left(\dot{f}_1 + K\dot{S}_1\right) \\ &:= -\frac{\xi}{\tau} + \eta(S_1)\left(\dot{f}_1 + K\dot{S}_1\right). \end{aligned} \tag{10}$$

Since the function, $\tan x$, is locally Lipschitz, there exists $\gamma > 0$ such that

$$\begin{aligned} |\tan x_2 - \tan x_{2d}| \le \gamma |x_2 - x_{2d}| = \gamma |S_2|, |\tan x_{2d} - \tan \overline{x}_2| \\ \le \gamma |x_{2d} - \overline{x}_2| = \gamma |\xi|, \end{aligned} \tag{11}$$

where γ is a Lipschitz constant on $\mathscr{D}$. Using the continuity of f_1 and $[\partial f_1/\partial x_1]$ in (2), it is also shown that the last term of the third row in (10) is bounded on $\mathscr{D}$. Therefore, the existence of the controller gain K and filter time constant τ for semiglobal stability can be shown as suggested in [5].

However, this fact does not tell us whether the closed-loop system is stable for the given K and τ. To answer the question, we may need to find a Lyapunov function candidate explicitly and one of the possible analysis approaches is based on linear matrix inequality. To apply this approach to (10), it is necessary to write it in matrix form as

$$\begin{aligned} &\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -K\eta & 0 & 1 \end{bmatrix} \begin{bmatrix} \dot{S}_1 \\ \dot{S}_2 \\ \dot{\xi} \end{bmatrix} \\ &\quad = \begin{bmatrix} -K & 0 & 0 \\ 0 & -K & 0 \\ 0 & 0 & -\dfrac{1}{\tau} \end{bmatrix} \begin{bmatrix} S_1 \\ S_2 \\ \xi \end{bmatrix} + \begin{bmatrix} 1 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & \eta \end{bmatrix} \begin{bmatrix} w_1 \\ w_2 \\ \dot{f} \end{bmatrix}. \end{aligned} \tag{12}$$

While the next procedure is to investigate whether (12) is in a class of linear differential inclusions classified in [6], the inclusion of the nonlinear function $\eta(S_1)$ in (12) results in the mathematical difficulty of stability analysis.

The contribution of this paper is to extend a design and analysis methodology of DSC for a more general class of nonlinear systems as shown in (2) and (3). The consideration of this class of nonlinear systems is motivated when rotational mechanical systems are considered for applications; that is, sinusoidal functions are in general included in the equation of motions. As one of the applications, the proposed control approach is applied to lateral vehicle control for forward automated driving and backward parallel parking at a low speed. Finally, its performance will be validated via simulations.

2. Problem Statement

Consider the class of nonlinear systems

$$\begin{aligned} \dot{x}_1 &= g_2(x_2) + f_1(x_1), \\ \dot{x}_2 &= g_3(x_3) + f_2(x_1, x_2), \\ &\vdots \\ \dot{x}_{n-1} &= g_n(x_n) + f_{n-1}(x_1, \ldots, x_{n-1}), \\ \dot{x}_n &= u + f_n(x_1, \ldots, x_n), \end{aligned} \tag{13}$$

where f_i and $[\partial f_i/\partial x]$ are continuous on $\mathscr{D}_i \subset \mathscr{D} \subset \mathfrak{R}^n$ and $f_i : \mathscr{D}_i \to \mathfrak{R}$ is in *strict-feedback* form in the sense that the f_i depends only on $x_1, \ldots, x_i$. It is implied that f_i is locally Lipschitz and $[\partial f_i(x)/\partial x]$ is bounded on $\mathscr{D}_i$ [7]. Therefore, there exists a constant $\gamma_i > 0$ such that

$$\left\| \frac{\partial f_i}{\partial x} \right\| = \left\| \begin{bmatrix} \dfrac{\partial f_i}{\partial x_1} & \cdots & \dfrac{\partial f_i}{\partial x_i} \end{bmatrix} \right\| := \|J_i\| \le \gamma_i \tag{14}$$

for all x on $\mathscr{D}_i$.

The nonlinear function g_i is also locally Lipschitz; that is, there exists a constant $\lambda_i > 0$ such that

$$|g_i(a) - g_i(b)| = \lambda_i |a - b|, \quad i = 2, \ldots, n. \tag{15}$$

In addition, there exist differentiable functions $q_i : \mathscr{E}_i \to \mathfrak{R}$, where $\mathscr{E}_i = \{y \in \mathfrak{R} \mid y = g_i(x) \text{ for all } x \in \mathscr{D}_i\}$, which are inverses of the g_i in the sense that

$$q_i(g_i(c)) = c, \quad i = 2, \ldots, n, \tag{16}$$

and $[\partial q_i/\partial g_i]$ is bounded on $\mathcal{D}$; that is, there exists a constant $\delta > 0$ such that

$$\left\| \left[\frac{\partial q_2}{\partial g_2} \quad \cdots \quad \frac{\partial q_n}{\partial g_n} \right] \right\| \leq \delta. \tag{17}$$

3. Analysis and Design of DSC

3.1. Design Procedure. Although the proposed design procedure is quite similar to the standard one described in [4], an outline of the design procedure is as follows. Define the first error surface as $S_1 := x_1 - x_{1d}$, where x_{1d} is the desired value as the control objective. After taking the time derivative of S_1 along the trajectory of (13),

$$\dot{S}_1 = g_2(x_2) + f_1(x_1) - \dot{x}_{1d}. \tag{18}$$

The surface error S_1 will converge to zero if $S_1\dot{S}_1 < 0$; however there is no direct control over the surface dynamics. If g_2 is considered as the forcing term for the surface dynamics, then the sliding condition outside some boundary layer is satisfied if $g_2 = \overline{g}_2$, where

$$\overline{g}_2(\overline{x}_2) = \dot{x}_{1d} - f_1(x_1) - K_1 S_1, \qquad \overline{x}_2 = q_2(\overline{g}_2), \tag{19}$$

where q_2 is the inverse of g_2.

The next step is to force $x_2 \to \overline{x}_2$, so define $S_2 := x_2 - x_{2d}$, where $x_{2d} = q_2(g_{2d})$ and g_{2d} is obtained after passing through a first order low-pass filter; that is,

$$\tau_2 \dot{g}_{2d} + g_{2d} = \overline{g}_2, \qquad g_{2d}(0) := \overline{g}_2(0). \tag{20}$$

It is noted that this procedure is different from the one explained in the introduction. That is, $\overline{g}_2$ instead of $\overline{x}_2$ passes through the filter and the inverse function of the filtered signal is used to define x_{2d}. After taking a derivative of S_2 along the trajectory of (13), the resulting synthesis term, $\overline{g}_3$, is derived as

$$\overline{g}_3(\overline{x}_3) = \dot{x}_{2d} - f_2(x_1, x_2) - K_2 S_2, \qquad \overline{x}_3 = q_3(\overline{g}_3), \tag{21}$$

where

$$\dot{x}_{2d} = \frac{\partial q_2}{\partial g_{2d}} \dot{g}_{2d} = \frac{\partial q_2}{\partial g_{2d}} \frac{\overline{g}_2 - g_{2d}}{\tau_2} \tag{22}$$

and the last equality comes from (20).

Similarly, continuing this process for each consecutive state, define the ith error surface as $S_i = x_i - x_{id}$ where $x_{id} = q_i(g_{id})$ and $\overline{g}_{i+1}$ is

$$\begin{aligned} \overline{g}_{i+1}(\overline{x}_{i+1}) &= \dot{x}_{id} - f_i(x_1, \ldots, x_i) - K_i S_i, \\ \overline{x}_{i+1} &= q_{i+1}(\overline{g}_{i+1}), \end{aligned} \tag{23}$$

where

$$\dot{x}_{id} = \frac{\partial q_i}{\partial g_{id}} \dot{g}_{id} = \frac{\partial q_i}{\partial g_{id}} \frac{\overline{g}_i - g_{id}}{\tau_i}. \tag{24}$$

Then, $g_{(i+1)d}$ is obtained by filtering $\overline{g}_{i+1}$; that is,

$$\tau_{i+1} \dot{g}_{(i+1)d} + g_{(i+1)d} = \overline{g}_{i+1}, \qquad g_{(i+1)d}(0) := \overline{g}_{i+1}(0). \tag{25}$$

After continuing this procedure for $1 \leq i \leq n-1$, define $S_n := x_n - x_{nd}$, where $x_{nd} = q_n(g_{nd})$. Finally, the control input is derived as

$$u = \dot{x}_{nd} - f_n(x_1, \ldots, x_n) - K_n S_n, \tag{26}$$

where

$$\dot{x}_{nd} = \frac{\partial q_n}{\partial g_{nd}} \dot{g}_{nd} = \frac{\partial q_n}{\partial g_{nd}} \frac{\overline{g}_n - g_{nd}}{\tau_n}. \tag{27}$$

3.2. Augmented Error Dynamics. The closed-loop error dynamics will be derived for stability analysis in this section. After subtracting and adding $\overline{g}_{i+1}$ and $g_{(i+1)d}$ and using (26) in u, the closed-loop dynamics of (13) can be written as

$$\begin{aligned} \dot{x}_i &= \overline{g}_{i+1} + [g_{i+1} - g_{(i+1)d}] + [g_{(i+1)d} - \overline{g}_{i+1}] + f_i \\ &\qquad \text{for } i = 1, \ldots, n-1, \\ \dot{x}_n &= \dot{x}_{nd} - K_n S_n. \end{aligned} \tag{28}$$

By use of (23) and the definition of error surfaces, the above equations can be described in terms of the error surfaces of DSC as follows:

$$\begin{aligned} \dot{S}_i &= -K_i S_i + h_{i+1} + [g_{(i+1)d} - \overline{g}_{i+1}] \quad \text{for } i = 1, \ldots, n-1, \\ \dot{S}_n &= -K_n S_n, \end{aligned} \tag{29}$$

where $h_{i+1} = g_{i+1} - g_{(i+1)d}$.

In addition, we need to consider the augmented error dynamics due to inclusion of a set of the first order low-pass filters. Let us define the filter error as $\xi_i := g_{id} - \overline{g}_i$ for $2 \leq i \leq n$. Then, the filter dynamics are

$$\dot{\xi}_i = \dot{g}_{id} - \frac{d\overline{g}_i}{dt} = -\frac{\xi_i}{\tau_i} - \frac{d\overline{g}_i}{dt}, \tag{30}$$

where the last equality comes from (25). By taking a derivative of (23), we can write $d\overline{g}_i/dt$ as

$$\begin{aligned} \frac{d\overline{g}_2}{dt} &= -\dot{f}_1 + \ddot{x}_{1d} - K_1 \dot{S}_1, \\ \frac{d\overline{g}_i}{dt} &= -\dot{f}_{i-1} + \ddot{x}_{(i-1)d} - K_{i-1}\dot{S}_{i-1} \quad \text{for } i = 3, \ldots, n. \end{aligned} \tag{31}$$

Combining (30) with (31), we have the filter error dynamics,

$$\begin{aligned} \dot{\xi}_2 - K_1 \dot{S}_1 &= -\frac{\xi_2}{\tau_2} + \dot{f}_1 - \ddot{x}_{1d}, \\ \dot{\xi}_i - K_{i-1}\dot{S}_{i-1} &= -\frac{\xi_i}{\tau_i} + \dot{f}_{i-1} - \ddot{x}_{(i-1)d} \quad \text{for } i = 3, \ldots, n. \end{aligned} \tag{32}$$

Therefore, the overall error dynamics, (29) and (32), can be given as

$$\begin{aligned} \dot{S}_i &= -K_i S_i + \xi_{i+1} + h_{i+1} \quad \text{for } i = 1, \ldots, n-1, \\ \dot{S}_n &= -K_n S_n, \\ \dot{\xi}_{j+1} - K_j \dot{S}_j &= -\frac{\xi_{j+1}}{\tau_{j+1}} + \dot{f}_j - \ddot{x}_{jd} \quad \text{for } j = 1, \ldots, n-1. \end{aligned} \tag{33}$$

Furthermore, (33) can be written in matrix form as follows:

$$\begin{bmatrix} \mathbf{I}_n & \mathbf{0} \\ -\mathbf{K}_0 & \mathbf{I}_{n-1} \end{bmatrix} \begin{bmatrix} \dot{S} \\ \dot{\xi} \end{bmatrix} = \begin{bmatrix} -\mathbf{K} & \mathbf{I}_{n\times(n-1)} \\ \mathbf{0} & -\mathbf{\Gamma} \end{bmatrix} \begin{bmatrix} S \\ \xi \end{bmatrix} + \begin{bmatrix} \mathbf{I}_{n\times(n-1)} & \mathbf{0} \\ \mathbf{0} & \mathbf{I}_{n-1} \end{bmatrix} \begin{bmatrix} h \\ \dot{f} \end{bmatrix} + \begin{bmatrix} \mathbf{0} \\ -\mathbf{I}_{n-1} \end{bmatrix} \ddot{x}_d, \tag{34}$$

where the vectors are defined as

$$\begin{aligned} S &= \begin{bmatrix} S_1 & \cdots & S_n \end{bmatrix}^T \in \mathfrak{R}^n, \qquad \xi = \begin{bmatrix} \xi_2 & \cdots & \xi_n \end{bmatrix}^T \in \mathfrak{R}^{n-1}, \\ h &= \begin{bmatrix} h_2 & \cdots & h_n \end{bmatrix}^T \in \mathfrak{R}^{n-1}, \\ \dot{f} &= \begin{bmatrix} \dot{f}_1 & \dot{f}_2 & \cdots & \dot{f}_{n-1} \end{bmatrix}^T \in \mathfrak{R}^{n-1}, \\ \ddot{x}_d &= \begin{bmatrix} \ddot{x}_{1d} & \cdots & \ddot{x}_{(n-1)d} \end{bmatrix}^T \in \mathfrak{R}^{n-1} \end{aligned} \tag{35}$$

and the submatrices are

$$\begin{aligned} \mathbf{K} &= \operatorname{diag}\left(K_1, \ldots, K_n\right), \\ \mathbf{K}_0 &= \left[\operatorname{diag}\left(K_1, \ldots, K_{n-1}\right) 0_{n-1}\right] \in \mathfrak{R}^{(n-1)\times n}, \\ \mathbf{\Gamma} &= \operatorname{diag}\left(\frac{1}{\tau_2}, \ldots, \frac{1}{\tau_n}\right) \in \mathfrak{R}^{(n-1)\times(n-1)}. \end{aligned} \tag{36}$$

Since the first block matrix in (34) is invertible such that

$$\begin{bmatrix} \mathbf{I}_n & \mathbf{0} \\ -\mathbf{K}_0 & \mathbf{I}_{n-1} \end{bmatrix}^{-1} = \begin{bmatrix} \mathbf{I}_n & \mathbf{0} \\ \mathbf{K}_0 & \mathbf{I}_{n-1} \end{bmatrix}, \tag{37}$$

after multiplying the inverse matrix to both sides in (34), the augmented closed-loop error dynamics are rewritten as

$$\begin{aligned} \frac{d}{dt} \begin{bmatrix} S \\ \xi \end{bmatrix} &= \begin{bmatrix} -\mathbf{K} & \mathbf{I}_{n\times(n-1)} \\ \mathbf{K}_0\mathbf{K} & \mathbf{\Lambda} \end{bmatrix} \begin{bmatrix} S \\ \xi \end{bmatrix} \\ &\quad + \begin{bmatrix} \mathbf{I}_{n\times(n-1)} & \mathbf{0} \\ \mathbf{K}_{1:(n-1)} & \mathbf{I}_{n-1} \end{bmatrix} \begin{bmatrix} h \\ m \end{bmatrix} + \begin{bmatrix} \mathbf{0} \\ -\mathbf{I}_{n-1} \end{bmatrix} \ddot{x}_d, \end{aligned} \tag{38}$$

where

$$\mathbf{K}_{1:(n-1)} = \operatorname{diag}\left(K_1, \ldots, K_{n-1}\right), \qquad \mathbf{\Lambda} = \mathbf{K}_{1:(n-1)} - \mathbf{\Gamma}. \tag{39}$$

Since $\dot{x}_j$ is written in a function of S, ξ, h, and $\dot{x}_1$ based on (29)

$$\begin{aligned} \dot{x}_1 &= -K_1 S_1 + \xi_2 + h_2 + \dot{x}_{1d}, \\ \dot{x}_i &= -K_i S_i + \xi_{i+1} + h_{i+1} + \dot{x}_{id} = -K_i S_i + \xi_{i+1} \\ &\quad + h_{i+1} - \frac{\partial q_i}{\partial g_{id}} \frac{\xi_i}{\tau_i} \quad \text{for } i = 2, \ldots, n-1, \\ \dot{x}_n &= -K_n S_n + \dot{x}_{nd} = -K_n S_n - \frac{\partial q_n}{\partial g_{nd}} \frac{\xi_n}{\tau_n}, \end{aligned} \tag{40}$$

the time derivative of f_i can be decomposed into three parts as follows:

$$\begin{aligned} \dot{f}_i &= \sum_{j=1}^{i} \frac{\partial f_i}{\partial x_j} \dot{x}_j \\ &= \begin{bmatrix} \dfrac{\partial f_i}{\partial x_1} & \cdots & \dfrac{\partial f_i}{\partial x_i} \end{bmatrix} \\ &\quad \times \left(\begin{bmatrix} -K_1 & 0 & \cdots & 0 & 1 & 0 & \cdots & 0 \\ 0 & -K_2 & \cdots & 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & -K_i & 0 & 0 & \cdots & 1 \end{bmatrix} z + \begin{bmatrix} h_2 \\ h_3 \\ \vdots \\ h_{i+1} \end{bmatrix} \right) \\ &\quad + \begin{bmatrix} \dfrac{\partial f_i}{\partial x_1} & \cdots & \dfrac{\partial f_i}{\partial x_i} \end{bmatrix} \begin{bmatrix} 0 \\ -\dfrac{\partial q_2}{\partial g_{2d}} \dfrac{\xi_2}{\tau_2} \\ \vdots \\ -\dfrac{\partial q_i}{\partial g_{id}} \dfrac{\xi_i}{\tau_i} \end{bmatrix} + \frac{\partial f_i}{\partial x_1} \dot{x}_{1d} \\ &:= p_{1i} + p_{2i} + p_{3i}. \end{aligned} \tag{41}$$

Therefore, (38) is rewritten as

$$\begin{aligned} \frac{d}{dt} \begin{bmatrix} S \\ \xi \end{bmatrix} &= \begin{bmatrix} -\mathbf{K} & \mathbf{I}_{n\times(n-1)} \\ -\mathbf{K}_0\mathbf{K} & \mathbf{\Lambda} \end{bmatrix} \begin{bmatrix} S \\ \xi \end{bmatrix} \\ &\quad + \begin{bmatrix} \mathbf{I}_{n\times(n-1)} & \mathbf{0} & \mathbf{0} \\ \mathbf{K}_{1:(n-1)} & \mathbf{I}_{n-1} & \mathbf{I}_{n-1} \end{bmatrix} \begin{bmatrix} h \\ p_1 \\ p_2 \end{bmatrix} \\ &\quad + \begin{bmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{I}_{n-1} & -\mathbf{I}_{n-1} \end{bmatrix} \begin{bmatrix} p_3 \\ \ddot{x}_d \end{bmatrix} \\ &\Longrightarrow \dot{z} = A_{cl} z + B_w w + B_r r, \end{aligned} \tag{42}$$

where the error state $z = \begin{bmatrix} S^T & \xi^T \end{bmatrix}^T \in \mathfrak{R}^{2n-1} := \mathfrak{R}^{n_z}$, $w = \begin{bmatrix} h^T & p_1^T & p_2^T \end{bmatrix}^T \in \mathfrak{R}^{3n-3} := \mathfrak{R}^{n_w}$, and $r = \begin{bmatrix} p_3^T & \ddot{x}_{1d} \end{bmatrix}^T \in \mathfrak{R}^n$.

Finally, we need to determine the upper bound of w in (42). Using the assumptions (15) in Section 2, the upper bound of w_i for $i = 1, \ldots, n-1$ is

$$\begin{aligned} \left|w_i\right| = \left|h_{i+1}\right| &= \left|g_{i+1} - g_{(i+1)d}\right| \le \lambda_{i+1} \left|x_{i+1} - x_{(i+1)d}\right| \\ &= \lambda_{i+1} \left|S_{i+1}\right| := \left|C_{zi} z\right|. \end{aligned} \tag{43}$$

Using (14), the upper bound of w_i for $i = n, \ldots, 2n-2$ is

$$\begin{aligned} \left|w_i\right| = \left|p_{1j}\right| &= \left\|J_j \left(C_z z + D_w w\right)\right\| \le \gamma_j \left\|\widetilde{C}_{zj} z + \widetilde{D}_{wj} w\right\| \\ &:= \left\|C_{zi} z + D_{wi} w\right\|, \quad j = 1, \ldots, n-1, \end{aligned} \tag{44}$$

where $J_j = [\partial f_j/\partial x_1 \cdots \partial f_j/\partial x_j]$, $C_{zi} = \gamma_j \widetilde{C}_{zj}$, and $D_{wi} = \gamma_j \widetilde{D}_{wj}$. Similarly, the upper bound of w_i for $i = 2n-1, \ldots, n_w$ is obtained using (17)

$$\begin{aligned} |w_i| &= |p_{2j}| \\ &= \left\| - \begin{bmatrix} \dfrac{\partial f_j}{\partial x_2}\dfrac{\partial q_2}{\partial g_{2d}} & \cdots & \dfrac{\partial f_j}{\partial x_j}\dfrac{\partial q_j}{\partial g_{jd}} \end{bmatrix} \begin{bmatrix} \dfrac{\xi_2}{\tau_2} & \cdots & \dfrac{\xi_i}{\tau_i} \end{bmatrix}^T \right\| \\ &\le \delta\gamma_j \left\| \overline{C}_{zj} z \right\| := \left\| C_{zi} z \right\|, \quad j = 1, \ldots, n-1. \end{aligned} \tag{45}$$

Combined with (43), (44), and (45), (42) can be written in diagonal norm-bounded LDI form as follows [6]:

$$\begin{aligned} \dot{z} &= A_{cl} z + B_w w + B_r r, \qquad t = C_z z + D_w w, \\ |w_i| &\le \|t_i\|, \quad i = 1, \ldots, n_w. \end{aligned} \tag{46}$$

3.3. LMI Approach for Stability Analysis. If either stabilization or regulation problem is considered, the first element $\ddot{x}_{1d}$ of $\ddot{x}_d$ in (46) is zero. Furthermore, if $\ddot{x}_d = 0$ for special cases among nonlinear systems in (13), quadratic stability of the resulting closed-loop error dynamics is defined as follows [3].

Definition 1. Let $z = 0$ be an exponentially stable equilibrium point of the closed-loop error dynamics in (46) where $r = 0$ and A_{cl} is Hurwitz for the given set of controller gains; $\Theta = \{K_1, \ldots, K_n, \tau_2, \ldots, \tau_n\}$. Then, a nonlinear system in (13) is *quadratically stabilizable* via DSC if there exists a positive definite matrix P such that

$$\begin{aligned} \frac{d}{dt}V(z) = \frac{d}{dt}\left(z^T P z\right) &= \left(A_{cl} z + B_w w\right)^T P z \\ &+ z^T P \left(A_{cl} z + B_w w\right) < 0. \end{aligned} \tag{47}$$

Furthermore, the quadratic stability under the DSC is guaranteed by the following theorem.

Theorem 2. *Suppose that the closed-loop error dynamics in* (46) *are given for the given set of controller gains,*

$$\Theta = \{K_1, \ldots, K_n, \tau_2, \ldots, \tau_n\}, \tag{48}$$

for all x in a domain $\mathscr{D}$. If there exist $P > 0$ and $\Sigma_B = \mathrm{diag}(\sigma_1, \ldots, \sigma_{n-1}, \sigma_n I, \ldots, \sigma_{n_w} I) \ge 0$ such that

$$\begin{bmatrix} A_{cl}^T P + P A_{cl} + C_z^T \Sigma_B C_z & P B_w + C_z^T \Sigma_B D_w \\ B_w^T P + D_w^T \Sigma_B C_z & D_w^T \Sigma_B D_w - \Sigma_B \end{bmatrix} < 0, \tag{49}$$

where $C_z = \begin{bmatrix} C_{z1}^T & \cdots & C_{zn_w}^T \end{bmatrix}^T$ and $D_w = \begin{bmatrix} D_{w1}^T & \cdots & D_{zn_w}^T \end{bmatrix}^T$, the origin in (46) *is then exponentially stable in $\mathscr{D}$. Thus the nonlinear system* (13) *is quadratically stabilizable via DSC with the given Θ on $\mathscr{D}$.*

For details of the proof of the theorem, readers are referred to Boyd et al. [6].

If a tracking problem is considered ($\ddot{x}_{1d} \neq 0$) and it is assumed that $\ddot{x}_d$ is bounded, the ultimate and quadratic boundedness is defined as follows [3].

Definition 3. The error dynamics in (46) is *quadratically bounded with Lyapunov matrix P* if there exists $P > 0$ such that

$$\begin{aligned} z^T P z > 1 \quad \text{implies } &\left(A_{cl} z + B_w w + B_r r\right)^T P z \\ &+ z^T P \left(A_{cl} z + B_w w + B_r r\right) < 0 \end{aligned} \tag{50}$$

for all nonzero $z \in \mathscr{E}_P = \{z \in \mathfrak{R}^{n_z} \mid z^T P z \ge 1\}$.

It is noted that the assumption that $\ddot{x}_d$ is bounded is feasible because the time derivative of the filtered signal x_d is bounded.

Suppose r in (46) is norm-bounded such that $\|r\| \le r_0$. After defining $\widetilde{r} := r/r_0$ and $\widetilde{B}_r := r_0 B_r$, the error dynamics in (46) is written as

$$\begin{aligned} \dot{z} &= A_{cl} z + B_w w + \widetilde{B}_r \widetilde{r}, \qquad t = C_z z + D_w w, \\ |w_i| &\le \|t_i\|, \qquad \|\widetilde{r}\| \le 1. \end{aligned} \tag{51}$$

Without loss of generality, it can be considered that $\widetilde{r}$ is a unit-peak input. Then, the following theorem describes the condition for guaranteeing quadratic tracking as well as the computation of the matrix P for a given set of controller gains.

Theorem 4. *For the given set of controller gains, Θ, suppose that the closed-loop error dynamics in* (51) *is given on the domain $\mathscr{D}$ and x_{1d} is a feasible output trajectory. The closed-loop error dynamics is quadratically bounded with Lyapunov matrix P if there exist $P > 0$, $\Sigma_B = \mathrm{diag}(\sigma_1, \ldots, \sigma_{n-1}, \sigma_n I, \ldots, \sigma_{n_w} I) \ge 0$, and $\alpha \ge 0$ such that*

$$\begin{bmatrix} A_{cl}^T P + P A_{cl} + \alpha P + C_z^T \Sigma_B C_z & P B_w + C_z^T \Sigma_B D_w & P \widetilde{B}_r \\ B_w^T P + D_w^T \Sigma_B C_z & D_w^T \Sigma_B D_w - \Sigma_B & 0 \\ \widetilde{B}_r^T P & 0 & -\alpha I \end{bmatrix} < 0, \tag{52}$$

where $\widetilde{B}_r = r_0 B_r$, $C_z = \begin{bmatrix} C_{z1}^T & \cdots & C_{zn_w}^T \end{bmatrix}^T$, and $D_w = \begin{bmatrix} D_{w1}^T & \cdots & D_{zn_w}^T \end{bmatrix}^T$.

Readers are referred to Boyd et al. [6] for the proof and the definition of the feasible output trajectory is explained in [3, 4].

3.4. Illustrative Example. Consider the example in (2) and (3) where $f_1 = -x_1^2$ and the domain $\mathscr{D}$ is defined as

$$\mathscr{D} = \left\{ x \in \mathfrak{R}^2 \mid |x_1| \le 1, |x_2| \le \frac{\pi}{4} \right\}. \tag{53}$$

Then, (5) becomes

$$\overline{g}_2\left(\overline{x}_2\right) = \tan \overline{x}_2 = x_1^2 - K S_1, \tag{54}$$

and a first order low-pass filter is defined as

$$\tau \dot{g}_{2d} + g_{2d} = \overline{g}_2, \qquad g_{2d}(0) = \overline{g}_2(0). \tag{55}$$

It is noted that $\overline{g}_2$ instead of $\overline{x}_2$ is passed through the low-pass filter. After defining the second sliding surface $S_2 = x_2 - x_{2d}$,

where $x_{2d} = \tan^{-1}(g_{2d})$, and taking its derivative, the control input is obtained by

$$u = \dot{x}_{2d} - KS_2, \tag{56}$$

where

$$\dot{x}_{2d} = \frac{d}{dt}\tan^{-1}(g_{2d}) = \frac{1}{1+g_{2d}^2}\dot{g}_{2d} = \frac{1}{1+g_{2d}^2}\frac{\overline{g}_2 - g_{2d}}{\tau} \tag{57}$$

and the last equality comes from (55).

If a new filter error, $\xi_2 = g_{2d} - \overline{g}_2$, is defined, the augmented closed-loop dynamics can be written as

$$\begin{aligned}
\dot{S}_1 &= -KS_1 + \xi_2 + (\tan x_2 - \tan x_{2d}) = -KS_1 + \xi_2 + h_2,\\
\dot{S}_2 &= -KS_2,\\
\dot{\xi}_2 &= \frac{d}{dt}(g_{2d} - \overline{g}_2) = -\frac{\xi_2}{\tau} + \frac{d}{dt}\{f_1 + KS_1\}\\
&= -\frac{\xi_2}{\tau} + \dot{f}_1 + K\dot{S}_1,
\end{aligned} \tag{58}$$

where $h_2 = \tan x_2 - \tan x_{2d}$ and $\dot{g}_{2d} = -\xi_2/\tau$ from (55). Therefore, (58) is written in matrix form as

$$\begin{aligned}
&\begin{bmatrix} 1 & 0 & 0\\ 0 & 1 & 0\\ -K & 0 & 1 \end{bmatrix}\begin{bmatrix} \dot{S}_1\\ \dot{S}_2\\ \dot{\xi}_2 \end{bmatrix}\\
&\quad = \begin{bmatrix} -K & 0 & 1\\ 0 & -K & 0\\ 0 & 0 & -\frac{1}{\tau} \end{bmatrix}\begin{bmatrix} S_1\\ S_2\\ \xi_2 \end{bmatrix} + \begin{bmatrix} 1 & 0\\ 0 & 0\\ 0 & 1 \end{bmatrix}\begin{bmatrix} h_2\\ \dot{f}_1 \end{bmatrix},
\end{aligned} \tag{59}$$

$$\begin{bmatrix} \dot{S}_1\\ \dot{S}_2\\ \dot{\xi}_2 \end{bmatrix} = \begin{bmatrix} -K & 0 & 1\\ 0 & -K & 0\\ -K^2 & 0 & K - \frac{1}{\tau} \end{bmatrix}\begin{bmatrix} S_1\\ S_2\\ \xi_2 \end{bmatrix} + \begin{bmatrix} 1 & 0\\ 0 & 0\\ K & 1 \end{bmatrix}\begin{bmatrix} h_2\\ \dot{f}_1 \end{bmatrix}.$$

The upper bound of $\begin{bmatrix} h_2 & \dot{f}_1 \end{bmatrix}^T$ can be determined as

$$\begin{aligned}
|h_2| &= |\tan x_2 - \tan x_{2d}| \le |x_2 - x_{2d}|\\
&= |S_2| = |[0 \;\; 1 \;\; 0]\, z| := |c_{z1} z|,\\
|\dot{f}_1| &= \left|\frac{\partial f_1}{\partial x_1}\dot{x}_1\right| = |-2x_1\dot{x}_1| = |-2x_1(-KS_1 + \xi_2 + h_2)|\\
&\le 2|-KS_1 + \xi_2 + h_2| = 2|[-K \;\; 0 \;\; 1]\, z + [1 \;\; 0]\, w|\\
&:= |c_{z2} z + d_{w2} w|,
\end{aligned} \tag{60}$$

where the first inequality comes from a Lipschitz condition such that $|\tan x - \tan y| \le |x - y|$ for all $x, y \in \mathcal{D}$ and the second inequality comes from a fact that $\partial f_1/\partial x_1$ is bounded on $\mathcal{D}$.

Finally, the augmented error dynamics can be written in diagonal norm-bounded LDI form as follows:

$$\begin{aligned}
\dot{z} &= A_{cl} z + B_w w, \qquad t = C_z z + D_w w,\\
|w_i| &\le |t_i|, \quad i = 1, 2,
\end{aligned} \tag{61}$$

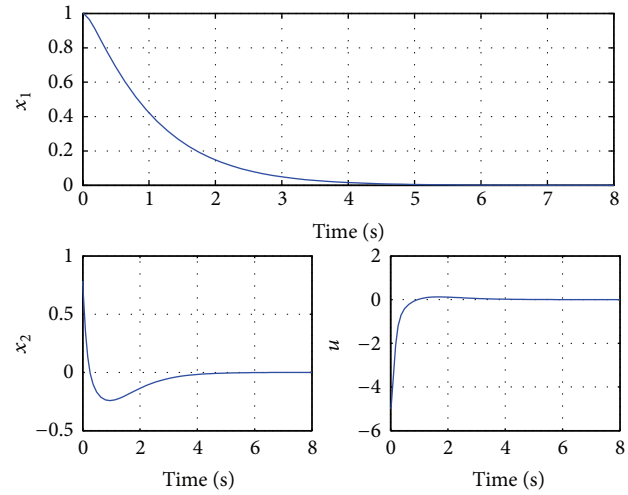


Figure 1: Time responses of x and u for the given initial condition, $[x_1(0) \;\; x_2(0)] = [1 \;\; \pi/4]$.

where

$$C_z = \begin{bmatrix} c_{z1}\\ c_{z2} \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0\\ -2K & 0 & 2 \end{bmatrix}, \qquad D_w = \begin{bmatrix} 0_2^T\\ d_{w2} \end{bmatrix} = \begin{bmatrix} 0 & 0\\ 2 & 0 \end{bmatrix}. \tag{62}$$

When the controller gains are given as $K = 1$ and $\tau = 0.1$, LMI (49) is solved numerically in the framework of convex optimization using CVX [8]. It is shown that the closed-loop system is quadratically stable by finding the feasible solution of LMI (49). For the given controller gains and initial condition, the time responses of x and u are shown in Figure 1 and $x(t) \to 0$ as $t \to \infty$. Thus, it is validated that the result of quadratic stability analysis based on an LMI approach is equivalent to simulation results.

4. Application to Lateral Vehicle Control

The proposed control approach is applied to design the robust lateral control algorithm for autonomous valet parking (AVP). It is assumed that the position and heading angle information is provided via either infrastructure sensors and vehicle to infrastructure (V2I) communication [9] or an in-vehicle sensor such as DGPS [10]. The objective of the lateral controller is to perform two different maneuvers for AVP, that is, forward driving and backward parallel parking. Therefore, it is necessary for the lateral controller to be robust enough to track desired trajectories for different driving maneuvers.

4.1. Vehicle Model. While the bicycle model has been used widely for design of a lateral controller for high speed driving on highway [11, 12], a vehicle is driving at low speed for AVP and thus slip angle can be neglected in this study. Therefore,

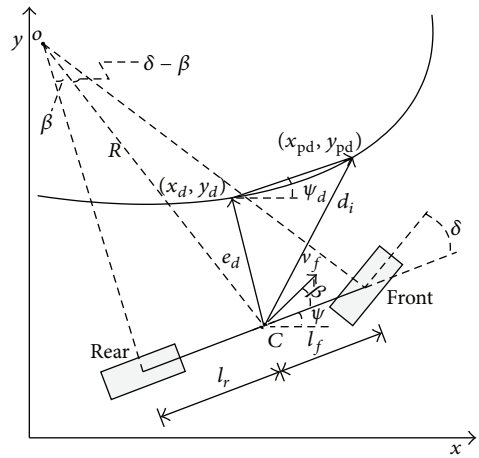


FIGURE 2: Schematic of kinematic model and error definition.

the following kinematic model is used for both forward driving and backward parallel parking (refer to Figure 2) [13]:

$$\begin{aligned}\dot{x} &= v_i \cos(\psi + \beta),\\ \dot{y} &= v_i \sin(\psi + \beta),\\ \dot{\psi} &= \frac{v_i}{l}\cos\beta\tan\delta = \frac{v_i}{l_j}\sin\beta := \frac{v_i}{l_j} g_2(\beta),\end{aligned} \tag{63}$$

where the subscript i represents the driving maneuver; that is, $i = f$ for forward driving and $i = r$ for backward parking,

$$\beta = \tan^{-1}\left(\frac{l_i}{l}\tan\delta\right), \qquad v_i = \begin{cases} v_f & \text{for } i = f,\\ -v_r & \text{for } i = r,\end{cases}$$
$$j = \begin{cases} r & \text{for } i = f,\\ f & \text{for } i = r.\end{cases} \tag{64}$$

If dynamics of a steering actuator from steering wheel angle command to steering angle of the vehicle is considered, the following equation of motion may be added:

$$\tau_s\dot{\theta} + \theta = u, \qquad |\delta| = |R_s\theta| \le \delta_{\max}, \tag{65}$$

where $\delta_{\max}$ is the maximum steering angle.

4.2. Controller Design. With consideration of operating conditions such as low speed and small slip angle, the nonlinear kinematic model with actuator dynamics in (63) and (65) is used for design of the lateral controller. Then, the proposed control approach based on DSC is applied to the kinematic model as follows. First, the first error surface is defined using the idea of preview control suggested in [14] (see in Figure 2):

$$S_{1i} = e_{bi} + d_i e_\psi = e_b + d_i(\psi_d - \psi), \tag{66}$$

where the lateral error e_b is defined as

$$\begin{aligned} e_{bi} &= \begin{cases} 0 & \text{if } c = 0,\\ -\operatorname{sign}(c)\cdot\|e_d\| & \text{otherwise,}\end{cases}\\ e_d &= \begin{bmatrix} x_d - x\\ y_d - y\end{bmatrix} := \begin{bmatrix} e_x\\ e_y\end{bmatrix},\\ a &= \begin{bmatrix} x_{pd} - x_d\\ y_{pd} - y_d\end{bmatrix} \in \Re^2, \qquad b = \begin{bmatrix} x - x_d\\ y - y_d\end{bmatrix} \in \Re^2,\\ c &= \det\left(\begin{bmatrix} a^T\\ b^T\end{bmatrix}\right),\end{aligned} \tag{67}$$

the point (x_d, y_d) is the closest point with respect to current position, and the preview distance d_i and desired heading angle ψ_d are

$$d_i = \left\|\begin{bmatrix} x_{pd} - x\\ y_{pd} - y\end{bmatrix}\right\|, \qquad \psi_d = \tan^{-1}\left(\frac{x_{pd} - x_d}{y_{pd} - y_d}\right). \tag{68}$$

For instance, sign(c) is negative for the given scenario in Figure 2 because the rotational direction from vector a to b defined in (67) is clockwise. Thus, the resulting positive lateral error implies a steering wheel angle command in the counterclockwise direction. If it is assumed that the preview distance is a constant, the point (x_{pd}, y_{pd}) in Figure 2 can be calculated with respect to the given desired trajectory and the desired heading angle is then determined.

After taking a derivative of S_{1i} along the trajectory of (63), the derivative of S_{1i} is

$$\dot{S}_{1i} = \dot{e}_{bi} + d_i\dot{\psi}_d - \frac{d_i v_i}{l_j}\sin\beta. \tag{69}$$

It is remarked that d_i can be chosen as a variable if necessary. For the simplicity of derivation, it is assumed to be constant. To make S_{1i} go to zero, let $\dot{S}_{1i} = -K_{1i}S_{1i}$, where K_{1i} is a controller gain. Then the desired steering angle is obtained as

$$\overline{g}_2(\overline{\beta}) = \sin\overline{\beta} = \frac{l_j}{d_i v_i}\left(\dot{e}_{bi} + d_i\dot{\psi}_d + K_{1i}S_{1i}\right), \tag{70}$$

where

$$\begin{aligned}\dot{e}_{bi} &= \operatorname{sign}(c)\cdot\dot{e}_d,\\ \dot{e}_d &= \frac{e_x + e_y}{e_d}\left(\dot{e}_x + \dot{e}_y\right)\\ &= \frac{e_x + e_y}{e_d}\left\{\left(\dot{x}_d - v_i\cos(\psi + \beta)\right)\right.\\ &\qquad \left. + \left(\dot{y}_d - v_i\sin(\psi + \beta)\right)\right\},\end{aligned}$$

$$\dot{\psi}_d = \frac{d}{dt}\tan^{-1}\left(\frac{x_{pd}-x_d}{y_{pd}-y_d}\right) = \frac{1}{1+\left(\left(x_{pd}-x_d\right)/\left(y_{pd}-y_d\right)\right)^2} \times \frac{\left(\dot{x}_{pd}-\dot{x}_d\right)\left(y_{pd}-y_d\right)-\left(x_{pd}-x_d\right)\left(\dot{y}_{pd}-\dot{y}_d\right)}{\left(y_{pd}-y_d\right)^2}. \tag{71}$$

It is noted that all of $\dot{x}_d$, $\dot{x}_{pd}$, $\dot{y}_d$, and $\dot{y}_{pd}$ are known for the given desired trajectory.

Then, the second error surface is defined as $S_{2i} = \beta - \beta_{\text{des}}$, where $\beta_{\text{des}} = q_2(g_{2d}) = \sin^{-1}(g_{2d})$ and g_{2d} is calculated after passing through a first order low-pass filter as follows:

$$\tau_2\dot{g}_{2d} + g_{2d} = \overline{g}_2. \tag{72}$$

After differentiating S_{2i} and using (65), the resulting equation is

$$\begin{aligned}\dot{S}_{2i} &= \dot{\beta} - \dot{\beta}_{\text{des}} \\ &= \frac{l_j\left(1+\tan^2\delta\right)}{l\left(1+\left(l_r^2/l^2\right)\tan^2\delta\right)}R_s\dot{\theta} - \dot{\beta}_{\text{des}} \\ &= \frac{R_s l_j\left(1+\tan^2\delta\right)}{l\left(1+\left(l_j^2/l^2\right)\tan^2\delta\right)}\frac{u-\theta}{\tau_s} - \dot{\beta}_{\text{des}},\end{aligned} \tag{73}$$

where

$$\dot{\beta} = \frac{d}{dt}\tan^{-1}\left(\frac{l_j\tan\delta}{l}\right) = \frac{1}{1+\left(l_j\tan\delta/l\right)^2}\frac{l_j}{l}\sec^2\delta\cdot\dot{\delta}. \tag{74}$$

Let $\dot{S}_{2i} = -K_{2i}S_{2i}$, where K_{2i} is a controller gain. Then the desired steering wheel angle is obtained as

$$u = \theta + \frac{\tau_s\left(l^2 + l_j^2\tan^2\delta\right)}{R_s\cdot l\cdot l_j\left(1+\tan^2\delta\right)}\left(\dot{\beta}_{\text{des}} - K_{2i}S_{2i}\right), \tag{75}$$

where

$$\dot{\beta}_{\text{des}} = \frac{\partial q_2}{\partial g_{2d}}\dot{g}_{2d} = \frac{1}{\sqrt{1-g_{2d}^2}}\frac{\overline{g}_2 - g_{2d}}{\tau_2}, \tag{76}$$

and the last equality comes from (72).

4.3. Stability Analysis. If $\sin\beta_{\text{des}}$ and $\sin\overline{\beta}$ are added and subtracted in (69) and u in (75) is put in (73), the augmented closed-loop error dynamics is

$$\begin{aligned}\dot{S}_{1i} &= \dot{e}_{bi} + d_i\dot{\psi}_d - \frac{d_i v_i}{l_j}\left(\sin\beta - \sin\beta_{\text{des}}\right) \\ &\quad - \frac{d_i v_i}{l_j}\left(\sin\beta_{\text{des}} - \sin\overline{\beta}\right) - \frac{d_i v_i}{l_j}\sin\overline{\beta},\end{aligned}$$

$$\begin{aligned}\dot{S}_{2i} &= -K_{2i}S_{2i}, \\ \dot{\xi}_{2i} &= \dot{g}_{2d} - \dot{\overline{g}}_2 = -\frac{\xi_2}{\tau_2} - \frac{l_j}{d_i v_i}\left(\ddot{e}_{yi} + d_i\ddot{\psi}_d + K_{1i}\dot{S}_{1i}\right),\end{aligned} \tag{77}$$

where it is assumed that v_i is constant with respect to time. Using (70), (77) is written as

$$\begin{aligned}\dot{S}_{1i} &= -K_{1i}S_{1i} - \frac{d_i v_i}{l_j}\xi_{2i} - \frac{d_i v_i}{l_j}h_2, \\ \dot{S}_{2i} &= -K_{2i}S_{2i}, \\ \frac{l_j}{d_i v_i}K_{1i}\dot{S}_{1i} + \dot{\xi}_{2i} &= -\frac{\xi_2}{\tau_2} - \frac{l_j}{d_i v_i}\left(\ddot{e}_{bi} + d_i\ddot{\psi}_d\right),\end{aligned} \tag{78}$$

where $h_2 = \sin\beta - \sin\beta_{\text{des}}$. As done in the example of Section 3.4, the augmented error dynamics is written in matrix form as follows:

$$\begin{aligned}\begin{bmatrix}1 & 0 & 0\\ 0 & 1 & 0\\ \dfrac{l_j K_{1i}}{(d_i v_i)} & 0 & 1\end{bmatrix}\begin{bmatrix}\dot{S}_{1i}\\ \dot{S}_{2i}\\ \dot{\xi}_{2i}\end{bmatrix} &= \begin{bmatrix}-K_{1i} & 0 & -\dfrac{d_i v_i}{l_j}\\ 0 & -K_{2i} & 0\\ 0 & 0 & -\dfrac{1}{\tau_2}\end{bmatrix}\begin{bmatrix}S_{1i}\\ S_{2i}\\ \xi_{2i}\end{bmatrix} \\ &\quad + \begin{bmatrix}-\dfrac{d_i v_i}{l_j}\\ 0\\ 0\end{bmatrix}h_2 + \begin{bmatrix}0 & 0\\ 0 & 0\\ -\dfrac{l_j}{(d_i v_i)} & -\dfrac{l_j}{v_i}\end{bmatrix}\begin{bmatrix}\ddot{e}_{bi}\\ \ddot{\psi}_d\end{bmatrix}, \\ \begin{bmatrix}\dot{S}_{1i}\\ \dot{S}_{2i}\\ \dot{\xi}_{2i}\end{bmatrix} &= \begin{bmatrix}-K_{1i} & 0 & -\dfrac{d_i v_i}{l_j}\\ 0 & -K_{2i} & 0\\ \dfrac{l_j K_{1i}^2}{(d_i v_i)} & 0 & K_{1i} - \dfrac{1}{\tau_2}\end{bmatrix}\begin{bmatrix}S_{1i}\\ S_{2i}\\ \xi_{2i}\end{bmatrix} \\ &\quad + \begin{bmatrix}-\dfrac{d_i v_i}{l_j}\\ 0\\ K_{1i}\end{bmatrix}h_2 - \frac{l_j}{v_i}\begin{bmatrix}0 & 0\\ 0 & 0\\ \dfrac{1}{d_i} & 1\end{bmatrix}\begin{bmatrix}\ddot{e}_{bi}\\ \ddot{\psi}_d\end{bmatrix}.\end{aligned} \tag{79}$$

Furthermore, the upper bound of h_2 is obtained as $|h_2| = |\sin\beta - \sin\beta_{\text{des}}| \le |\beta - \beta_{\text{des}}| = |S_2|$ because it is Lipschitz, and both $\dot{e}_{bi}$ and $\dot{\psi}_d$ in (71) are differentiable on $\mathscr{D}$; $\begin{bmatrix}\ddot{e}_{bi} & \ddot{\psi}_d\end{bmatrix}^T$ is bounded with respect to the desired trajectory. Thus, without loss of generality, it is assumed that $\begin{bmatrix}\ddot{e}_{bi} & \ddot{\psi}_d\end{bmatrix}^T$ is norm-bounded such that $\left\|\begin{bmatrix}\ddot{e}_{bi} & \ddot{\psi}_d\end{bmatrix}^T\right\| \le r_0$. Therefore, the

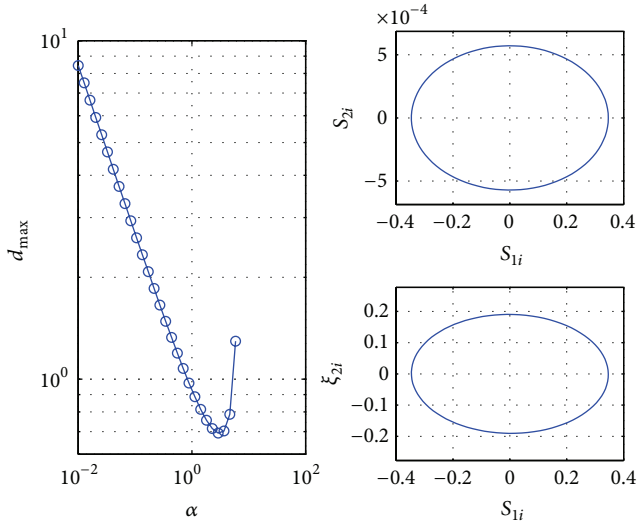


FIGURE 3: Minimum of the maximum diameter of ellipsoid along line search of α.

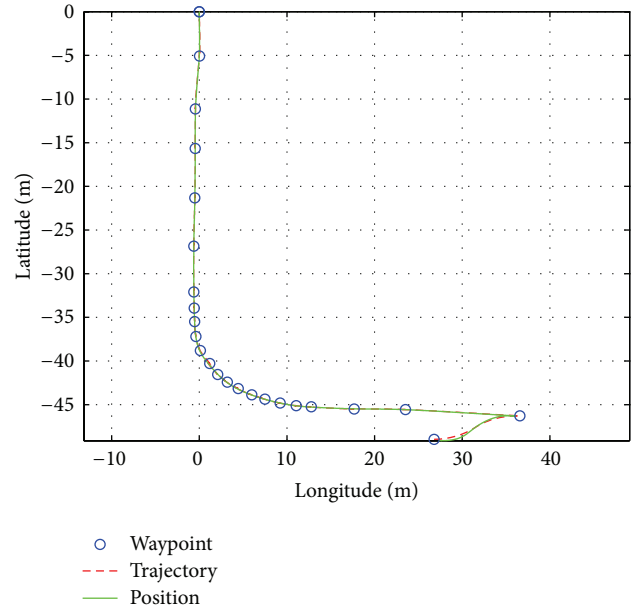


FIGURE 4: Desired trajectory and position of a vehicle for autonomous valet parking.

augmented error dynamics can be written in LDI form of (51) as

$$\begin{aligned} \dot{z}_i &= A_i z_i + B_w w + \widetilde{B}_r \widetilde{r}_i, \qquad t = C_z z, \quad i = f, r, \\ |w| &\le |t|, \quad \|\widetilde{r}\| \le 1, \end{aligned} \tag{80}$$

where $z_i = \begin{bmatrix} S_{1i} & S_{2i} & \xi_{2i} \end{bmatrix}^T \in \Re^3$, $w = h_2$, and $C_z = \begin{bmatrix} 0 & 1 & 0 \end{bmatrix}$.

Suppose $v_i = 3$ (m/s), $l_j = 1.35$ (m) in (79), and the control parameters are assigned as $K_{1i} = K_{2i} = 3$, $\tau_2 = 0.02$, and $d_f = 10$ and $d_r = 1$. When r_0 is assumed to be 5, LMI (52) is solved iteratively for a fixed α by minimizing the largest semiaxis (i.e., maximizing the smallest eigenvalue of P) [3]. That is, after the 40 logarithmically equally spaced points between 10^{-2} and 10^2 are generated for α's, the minimum of the maximum diameter, which is $d_{\max} = 2/\sqrt{\lambda_{\min}(P)}$, is obtained when $\alpha = 2.8943$ (in the left plot of Figure 3). Then the 20 linearly equally spaced points between 0.3455 and 0.5541 are generated and the iterative computation of LMI (52) is performed for each α. Finally, for $\alpha = 0.5212$, the corresponding maximum diameter of the ellipsoid, $d_{\max}$, is 0.6925 which is the semiaxis in the S_{2i} axis. It is remarked that the size of the ultimate and quadratic error bound is roughly proportional to the magnitude of r_0 and thus more accurate estimation of the error bound relies on better estimate of r_0 resulting from the desired trajectory. In consequence, it is expected that S_1 is bounded for the given set of control parameters. Furthermore, the relative degree of the error dynamics is one and it is shown that its internal dynamics is input-state stable [14]. Thus, this implies that both e_{bi} and e_ψ are bounded if S_1 is bounded.

4.4. Simulation Results. Suppose 24 waypoints are given a priori as shown in Figure 4. Forward driving maneuver is assigned from the first waypoint to 23rd waypoint and backward parallel parking maneuver is requested from 23rd to 24th waypoint. Moreover, both straight and curved road geometry is considered in this driving scenario; that is, from the first to 7th waypoints are for the straight road, from 7th to 20th waypoint for the curved road with about 12 (m) radius of curvature, and from 20th to 23rd waypoint again for the straight road as shown in Figure 4.

Let the additional system parameters in (63) and (65) be $\tau_s = 0.1$, $\delta_{\max} = 36$ (degree), and $R_s = 1/17.12$ for simulations. When the control parameters used for stability analysis above are applied to the proposed lateral controller, time responses of steering and heading angle are shown in Figure 5 and it is shown that the corresponding lateral error, e_b, is less than 0.3 (m). That is, it is shown that the lateral error is bounded as expected above. It is validated that the proposed controller enables the vehicle to perform two different maneuvers with only different value of d_i.

5. Conclusions

This paper developed the analysis and design method of DSC for a class of nonlinear systems where the nonlinear functions are included as forcing terms of DSC. The proposed control approach was applied to lateral vehicle control for forward driving and parallel parking maneuvers at low speed. With modification of input variables to the filter dynamics, it was shown that most of results in [3] could be used for the new class of nonlinear systems. Thus, the stability conditions in linear matrix inequality form were presented to guarantee the quadratic stability and boundedness via DSC for the given class of nonlinear systems. Furthermore, the quadratic Lyapunov functions were calculated numerically in the framework of convex optimization for a lateral vehicle control problem as well as an illustrative example. It was validated that the analysis results agreed with ones of simulation.

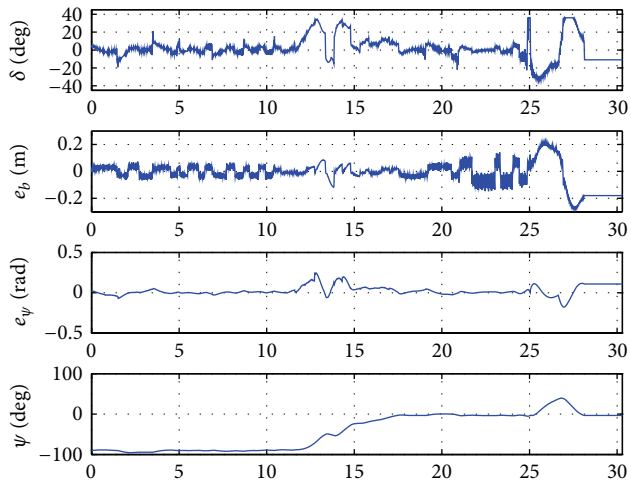


FIGURE 5: Time responses of errors, steering, and heading angle.

Nomenclature

x: Vehicle position in longitudinal direction
y: Vehicle position in lateral direction
v: Vehicle velocity
l_f: Length from a center of mass to front wheel axle
l_r: Length from a center of mass to rear wheel axle
l: Length of wheelbase, that is, $l = l_f + l_r$
δ: Front steering angle
θ: Steering wheel angle
u: Steering wheel angle command
R_s: Steering gear ratio, that is, $R_s = \delta/\theta$
e_b: Lateral position error
ψ: Heading angle
e_ψ: Heading angle error
d_i: Preview distance
τ_s: Time constant for steering actuator
β: Slip angle.

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 Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education, Science and Technology (no. 2009-0075110). It was also supported in part by the research project funded by LS Mtron.

References

[1] J. K. Hedrick and P. P. Yip, "Multiple sliding surface control: Theory and application," *Journal of Dynamic Systems, Measurement and Control*, vol. 122, no. 4, pp. 586–593, 2000.

[2] D. Chwa, "Global tracking control of underactuated ships with input and velocity constraints using dynamic surface control method," *IEEE Transactions on Control Systems Technology*, vol. 19, no. 6, pp. 1357–1370, 2011.

[3] B. Song and J. K. Hedrick, *Dynamic Surface Control of Uncertain Nonlinear Systems: An LMI Approach*, Springer, New York, NY, USA, 2011.

[4] 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.

[5] J. C. Gerdes and J. K. Hedrick, "'Loop-at-a-time' design of dynamic surface controllers for nonlinear systems," *Journal of Dynamic Systems, Measurement and Control*, vol. 124, no. 1, pp. 104–110, 2002.

[6] S. Boyd, L. El Ghaoui, E. Feron, and V. Balakrishnan, *Linear Matrix Inequalities in System and Control Theory*, vol. 15 of *SIAM Studies in Applied Mathematics*, SIAM, Philadelphia, Pa, USA, 1994.

[7] H. K. Khalil, *Nonlinear Systems*, Prentice Hall, New York, NY, USA, 3rd edition, 2002.

[8] M. Grant and S. Boyd, "CVX: matlab software for disciplined convex programming," version 1.21, 2011, http://cvxr.com/cvx/.

[9] B. Song, "Copperative lateral control for autonomous valet parking," *International Journal of Autonomotive Technology*, vol. 14, no. 4, pp. 633–640, 2013.

[10] M. Omae, H. Shimizu, and T. Fujioka, "GPS-based automatic driving control in local area with course of large curvature and parking space," *Vehicle System Dynamics*, vol. 42, no. 1-2, pp. 59–73, 2004.

[11] H. Peng and M. Tomizuka, "Preview control for vehicle lateral guidance in highway automation," *Journal of Dynamic Systems, Measurement and Control*, vol. 115, no. 4, pp. 679–686, 1993.

[12] R. Rajamani, H.-S. Tan, B. K. Law, and W.-B. Zhang, "Demonstration of integrated longitudinal and lateral control for the operation of automated vehicles in platoons," *IEEE Transactions on Control Systems Technology*, vol. 8, no. 4, pp. 695–708, 2000.

[13] R. Rajamani, *Vehicle Dynamics and Control*, Springer, New York, NY, USA, 2006.

[14] R. Rajamani, C. Zhu, and L. Alexander, "Lateral control of a backward driven front-steering vehicle," *Control Engineering Practice*, vol. 11, no. 5, pp. 531–540, 2003.

Hindawi Publishing Corporation
Mathematical Problems in Engineering
Volume 2014, Article ID 356250, 12 pages
http://dx.doi.org/10.1155/2014/356250

Research Article

Robust Switched Predictive Braking Control for Rollover Prevention in Wheeled Vehicles

Martín Antonio Rodríguez Licea and Ilse Cervantes

Hybrid Systems Laboratory, Applied Mathematics Division, Institute for Scientific and Technological Research of San Luis Potosi (IPICyT), Camino a la Presa San Jose 2055, Col. Lomas 4ta Seccion, 78216 San Luis Potosi, SLP, Mexico

Correspondence should be addressed to Martín Antonio Rodríguez Licea; martin_antonio_r@hotmail.com

Received 29 December 2013; Accepted 24 March 2014; Published 24 April 2014

Academic Editor: Sheldon S. Williamson

Copyright © 2014 M. A. R. Licea and I. Cervantes. 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 aim of this paper is to propose a differential braking rollover mitigation strategy for wheeled vehicles. The strategy makes use of a polytopic (piecewise linear) description of the vehicle and includes translational and rotational dynamics, as well as suspension effects. The braking controller is robust and the system states are predicted to estimate the rollover risk up to a given time horizon. In contrast to existing works, the switched predictive nature of the control allows it to be applied only when risk of rollover is foreseen, interfering a minimum with driver's actions. The stability of the strategy is analyzed and its robustness is illustrated via numerical simulations using CarSim for a variety of vehicles.

1. Introduction

Vehicle rollover is a serious problem that commonly involves sport utility vehicles (SUVs), trailers, trucks, and, in general, vehicles with a high center of gravity (CoG). Many quantifications of rollover risk exist in the literature [1–5]. The most commonly used is the load transfer ratio (LTR), also known as rollover index (RI) proposed by National Highway Traffic Safety Administration (NAFTA) [1], which is an index that measures the balance of vertical forces at the tire-road contact points:

$$\mathrm{RI} = \mathrm{LTR} = \frac{\sum F_{\mathrm{ZL}} - \sum F_{\mathrm{ZR}}}{\sum F_{\mathrm{ZL}} + \sum F_{\mathrm{ZR}}}, \tag{1}$$

where F_{ZL}, F_{ZR} are the sum of left and right side tire-road vertical forces, respectively. RI is dimensionless and constitutes a balance between right and left vertical forces, with RI = ±1 indicating imminent rollover (the whole weight of the vehicle is on its left or right side) and RI = 0 indicating a perfect balanced vehicle. The RI is used to describe rollover risk in any four-wheeled vehicle since it is related to a balance of vertical forces. |RI| values below 0.6 are considered safe [1]. Since RI is a force balance, it is directly related to the car lateral inclination and it can be measured as a function of chassis roll angle (φ) and chassis roll angle velocity ($\dot{\varphi}$). Although $\varphi, \dot{\varphi}$ can be measured it is not the case of the tire-road contact point forces. The estimation of these forces is complex to obtain since it depends nonlinearly on the vehicle's roll angle and speed and also on the type of terrain, tires, suspension, driver behavior, type of vehicle, and so forth; accordingly, there are several works in the literature dealing with the estimation of these forces [5–8].

Unfortunately, the RI estimation and a detonation of a warning of rollover cannot be used alone to mitigate the rollover phenomenon since human reaction delay and driver experience are limiting factors. Rollover can be avoided only by using effective timely actions on either active suspension, active steering, active differential, or differential braking [9–13]. Among the aforementioned actuators, perhaps the differential braking strategy is the easiest to implement since only antilock braking system (ABS) is necessary, contrasting to the other actuators which require costly and/or nonstandard devices. Since the aim of this work is to propose a differential braking control strategy, the literature review will be focused on these works only.

There exist in the literature works dealing with robust, switched, and/or predictive differential braking controllers. In particular, robustness of a control is a desired property

since nonlinearities of the dynamics, particularly those of tires and suspension, are hard to describe. Furthermore, exhaustive nonlinear models cannot be used for control design purposes since the problem becomes hardly tractable. Instead, approximated (linear) models are used to describe the vehicle dynamics and to simplify control synthesis. Since a single linear model cannot reproduce the system's complexity, polytopic or piecewise linear models are better options that guarantee tractability of the control design and of the stability analysis problem [14–16].

Robust differential braking controllers are proposed in [17–19]. In [17], a parametric uncertainty robust control strategy is proposed using a linear description of the vehicle including the suspension dynamics, leading to control actions that are applied continuously at all times. In [18], a robust control design based on linear matrix inequalities (LMIs) is proposed. The control is based on a vehicle description that uses the speed as uncertain parameter. The controller is shown to be robust to parameter uncertainty, but it is not clear if it can deal with the uncertainty in the whole operation domain and with the nonlinearities of the system. An adaptive switching controller is proposed in [19]. Least squares and Kalman filtering techniques are employed to estimate the height of the center of gravity in face of speed variations.

Since timely control actions depend on driver's decisions as well as on system and actuator inner dynamics, more recent works use the prediction of system states to compensate for such dynamics. In [20] a predictive RI strategy based on the present values of RI and steering angle is proposed. The resulting controller displays good performance avoiding rollover; however, the control action is applied at all times, interfering adversely with the driver's commands, even when no rollover risk is foreseen. Other works focusing on predictive and robust control actions are [21, 22]. In [21] a nonlinear inverted pendulum-like vehicle model is used to design a model predictive controller that is able to be applied in real-time. Even if the controller is not shown to be robust, the contribution of the authors is to take into account some of the nonlinearities of the system. In [22] a back-stepping observer is designed to estimate a lateral sliding and rollover indicator online. A maximum velocity tracking is obtained using predictive functional control.

Although these works constitute important advances, either they interfere with the drivers commands expending control actions even when no imminent risk of rollover is detected, or the robustness of the control in face of nonlinearities is not clear.

In this work we are focused on developing a noninvasive braking control that allows the driver to manipulate freely the vehicle until a rollover risk is foreseen. The controller is predictive and robust and guarantees stability in face of uncertainties using a polytopic linear model that takes into account different behaviors at a variety of operation conditions. The control is switched and its on-off control actions are performed using simple time/state dependent criteria based on the prediction. These three features: robustness, prediction, and switching make the controller versatile and its structure and proof of stability constitute the main contribution of this paper.

The prediction of the states is obtained using a set of linear models and an estimation of the future steering input. The steering prediction is obtained using a zero-order model of the driver. The predicted vehicle dynamics are used to estimate the RI behavior up to the time horizon in order to determine whether or not a rollover risk exists. If such possibility is present, a switched controller is used to mitigate the rollover phenomenon. The controller has a supervisory structure and it is comprised of a set of time- and state-dependent switching rules and a feedback braking control action.

The effect of the prediction horizon on system performance, as measured with the RI, is studied and compared against another nonswitched scheme. The proposed strategy has the feature of being simple for implementation purposes and it uses ABS braking technology. Moreover, it does not make use of GPS for driving prediction.

This document is organized as follows: in Section 2 the vehicle model is derived. Sections 3 and 4 introduce, respectively, the problem statement and the main contribution of this paper. The performance of the proposed controller is illustrated using numerical simulations in Section 5. Finally, in Section 6 some conclusions are presented.

2. Vehicle Model

Consider the nonlinear model of the vehicle given by

$$\begin{aligned}
\dot{v}_{\mathrm{CoG}} &= \frac{\sin(\beta)}{m}\left(\sum F_{\mathrm{Sij}} + \left(F_{\mathrm{LFL}} + F_{\mathrm{LFR}}\right)\delta + F_{\mathrm{YT}}\right) \\
&\quad + \frac{\cos(\beta)}{m}\left(\sum F_{\mathrm{Lij}} - \left(F_{\mathrm{SFL}} + F_{\mathrm{SFR}}\right)\delta + F_{\mathrm{XT}}\right), \\
\dot{\beta} &= \frac{\cos(\beta)}{m v_{\mathrm{CoG}}}\left(\sum F_{\mathrm{Sij}} + \left(F_{\mathrm{LFL}} + F_{\mathrm{LFR}}\right)\delta + F_{\mathrm{YT}}\right) \\
&\quad - \frac{\sin(\beta)}{m v_{\mathrm{CoG}}}\left(\sum F_{\mathrm{Lij}} - \left(F_{\mathrm{SFL}} + F_{\mathrm{SFR}}\right)\delta + F_{\mathrm{XT}}\right), \\
\ddot{\psi} &= \left(F_{\mathrm{SFR}} + F_{\mathrm{SFL}} + \left(F_{\mathrm{LFL}} + F_{\mathrm{LFR}}\right)\delta\right)\frac{L_v}{J_Z} - \left(F_{\mathrm{SRR}} + F_{\mathrm{SRL}}\right)\frac{L_h}{J_Z} \\
&\quad + \left(F_{\mathrm{LRR}} + F_{\mathrm{LFR}} - F_{\mathrm{LFL}} - F_{\mathrm{LRL}} + \left(F_{\mathrm{SFL}} - F_{\mathrm{SFR}}\right)\delta\right)\frac{b}{2J_Z}, \\
\ddot{\varphi} &= \frac{1}{J_x}\left(h_{\mathrm{CoG}}\left(\sum F_{\mathrm{Sij}} + \left(F_{\mathrm{LFL}} + F_{\mathrm{LFR}}\right)\delta + F_{\mathrm{YT}}\right) - k\varphi - c\dot{\varphi}\right),
\end{aligned} \tag{2}$$

where F_{Sij} are the lateral tire forces, F_{Lij} are the braking/acceleration forces with $i = \{\mathrm{Front}, \mathrm{Rear}\}$, $j = \{\mathrm{Left}, \mathrm{Right}\}$, δ is the front wheels angle, and F_{XT}, F_{YT} are terms that account for the gravity, rolling resistance, and other forces. L_v, L_h, J_z, and h are suitable constants. See Nomenclature section.

Forces F_{Sij} are highly nonlinear and dependent on the front and rear tire side slip angles (α_v, α_h), as well as on road and tire parameters [23, 24]. Let

$$\begin{aligned}
\alpha_v &= \delta - \beta - \frac{L_v \dot{\psi}}{v_{\text{CoG}}}, \\
\alpha_h &= -\beta + \frac{L_h \dot{\psi}}{v_{\text{CoG}}}, \\
F_{\text{LFL}} + F_{\text{LRL}} &= \begin{cases} u, & u < 0, \\ 0, & u \geq 0, \end{cases} \\
F_{\text{LFR}} + F_{\text{LRR}} &= \begin{cases} -u, & u > 0, \\ 0, & u \leq 0. \end{cases}
\end{aligned} \tag{3}$$

Let $x' = [\beta \ \dot{\psi} \ \dot{\varphi} \ \varphi]$ (where x' denotes the transpose of x); the linearized vehicle dynamics about a speed dependent operation point becomes

$$\dot{x} = Ax + B_\delta \delta(t) + B_u u, \tag{4}$$

with

$$A = \begin{bmatrix} -\dfrac{\sigma J_{x_{\text{eq}}}}{m\widehat{v}_{\text{CoG}} J_x} & \dfrac{\rho J_{x_{\text{eq}}}}{m\widehat{v}_{\text{CoG}}^2 J_x} - 1 & -\dfrac{hc}{J_x \widehat{v}_{\text{CoG}}} & \dfrac{h(mgh-k)}{J_x \widehat{v}_{\text{CoG}}} \\ \dfrac{\rho}{J_z} & -\dfrac{\kappa}{J_z \widehat{v}_{\text{CoG}}} & 0 & 0 \\ -\dfrac{h\sigma}{J_x} & \dfrac{h\rho}{\widehat{v}_{\text{CoG}} J_x} & -\dfrac{c}{J_x} & \dfrac{mgh-k}{J_x} \\ 0 & 0 & 1 & 0 \end{bmatrix},$$

$$B_\delta = \begin{bmatrix} \dfrac{c_v J_{x_{\text{eq}}}}{m J_x \widehat{v}_{\text{CoG}}} & \dfrac{c_v L_v}{J_z} & \dfrac{hc_v}{J_x} & 0 \end{bmatrix}',$$

$$B_u = \begin{bmatrix} 0 & -\dfrac{b}{2J_z} & 0 & 0 \end{bmatrix}', \tag{5}$$

$$\begin{aligned}
\sigma &\triangleq c_v + c_h, \\
\rho &\triangleq c_h L_h - c_v L_v, \\
\kappa &\triangleq c_h L_h^2 + c_v L_v^2, \\
J_{x_{\text{eq}}} &\triangleq J_x + mh^2,
\end{aligned} \tag{6}$$

where $\widehat{v}_{\text{CoG}}$ is the nominal velocity value and u represents the differential braking force. Using this model, the RI is given by (see [25])

$$\text{RI}(t) = -\frac{2}{bmg}\left(c\dot{\varphi} + k\varphi\right). \tag{7}$$

2.1. Nonlinearities and Parametric Uncertainty. The linear model (4)-(5) has been derived as an approximation of the nonlinear dynamics (2), about a given operation point. In the following, we will derive a polytopic linear model of the system such that not only small deviations of the operation point are allowed to describe the system dynamics, but also large ones, as long as they occur in a closed operation domain. To this end, assume that the operation domain $\mathscr{D}$ is divided into a finite number of sets Ξ_p for $p = 1, 2, \ldots, \text{nd}$, each one linked to a unique linear model expression (see Figure 1). Sets Ξ_p satisfy

(1) $\bigcup_{p=1}^{\text{nd}} \Xi_p = \mathscr{D}$,

(2) $\text{Int}(\Xi_p) \bigcap \text{Int}(\Xi_l) = \emptyset$ for $p \neq l$,

(3) $\Xi_p \neq \emptyset$,

where Int stands for the interior of the set. If the description of the system inside $\mathscr{D}$ can be performed using a convex combination of the parameters of every linear model associated with Ξ_p then a robust continuous control can be derived to guarantee the stability in $\mathscr{D}$ [26]. The convex parameter combination leads to a so-called polytopic model. In order to describe such a model, let us introduce the following definitions.

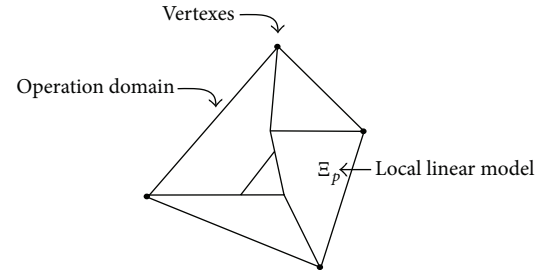


Figure 1: Schematic diagram of the operation domain $\mathscr{D}$. The domain is divided into finite sets Ξ_p and, associated with each one, there is a linear approximation of the system. The description of the system inside $\mathscr{D}$ can be performed using a convex combination of the parameters of the models.

Definition 1. The set

$$\Theta = \left\{\theta \in R \mid \sum_{i=1}^{\text{nv}} \theta_i = 1, \theta_i \geq 0\right\} \tag{8}$$

is called simplex and nv is called the number of vertexes.

Definition 2. The matrix A is called polytopic if

$$A = \left\{A(\theta) \mid \sum_{i=1}^{\text{nv}} A_i \theta_i = A, \theta_i \geq 0\right\}. \tag{9}$$

Using the notations above, the model of the vehicle can be rewritten as

$$\dot{x} = A(\theta)\, x + B_\delta(\theta)\, \delta + B_u(\theta)\, u, \tag{10}$$

where $A(\theta)$, $B_\delta(\theta)$, and $B_u(\theta)$ are polytopic matrices of suitable dimensions and $\delta(t) : \mathbb{R}^+ \to \mathbb{R}$, the front wheels angle generated by the steering input, is a piecewise continuous

function of time. According to Definitions 1 and 2, the parameters in the polytopic matrices vary within closed and bounded ranges, where each extremum is characterized by the simplex vertexes $\theta_i = 1$. As a result, it is possible to divide the domain $\mathscr{D}$ in an arbitrary finite number of subsets Ξ_p (see Figure 1). In other words, model (10) is actually a multilinear model that can be constructed using a finite number of linear approximations (4) at different operating conditions (e.g., vehicle speed, tire pressure, type of terrain, etc.). In Section 5 we illustrate the derivation of the multilinear model for a Cherokee Jeep 2000.

3. Controller Structure

In the literature, the differential braking control problem has been stated as follows.

Problem 3. Consider the vehicle model (10) and assume that the state x is available for measurement; design a control law $u(t)$ such that

$$|\mathrm{RI}(t)| \leq \mathrm{RI}_{\mathrm{ref}}, \tag{11}$$

for all time t and all parameters in the simplex Θ, where $\mathrm{RI}_{\mathrm{ref}}$ stands for a reasonable value of RI such that the vehicle has no rollover risk.

The control objective (11) can be attained using robust control theory if the pair $(A(\theta), B_u(\theta))$ is completely controllable and if it is possible to solve linear matrix inequalities (LMIs) associated with a Lyapunov equation [26]. In this case, a suitable vector K can be found, such that the control law

$$u(t) = Kx, \tag{12}$$

makes system (10) L_∞ stable for all $|\delta| \leq \delta_s$, where δ_s stands for the maximum magnitude of the front wheels angle to avoid violating the restriction $|\mathrm{RI}(t)| \leq \mathrm{RI}_{\mathrm{ref}}$. By solving the LMI it is ensured that $A_2(\theta) = A(\theta) + B_u(\theta)K$ is Hurwitz in the simplex Θ.

This approach has been discussed in the past, for a simpler model in [17, 25]. However, the said approach has the disadvantage of being continuously applied in the vehicle, consuming braking force and interfering constantly with the velocity path set by the driver, even when no risk of rollover exists. To avoid such disadvantages the problem addressed in this paper can be formulated as follows.

Problem 4. Consider the vehicle model (10) with x available for measurement; design a differential braking control that guarantees

$$|\mathrm{RI}(t)| \leq \mathrm{RI}_{\mathrm{ref}}, \tag{13}$$

for all time t and all parameters in the simplex Θ, such that the controller is applied only when a predicted value of RI satisfies $\mathrm{RI} > \mathrm{RI}_{\mathrm{ref}}$.

To solve Problem 4, a two-mode switched, noninvasive approach is proposed in this work. The first mode (Mode 1) coincides with the vehicle dynamics as conducted by the driver (open-loop), while the other one (Mode 2) applies the robust control (12) when necessary to attain the control objective (13).

If the Mode 2 is activated once the boundary $|\mathrm{RI}(t)| = \mathrm{RI}_{\mathrm{ref}}$ is reached, there is no guarantee of the accomplishment of (13). An alternate solution would be to set a lower bound of $\mathrm{RI}(t)$ to compensate for system dynamics. In this work, we use a prediction of the state vector up to a time horizon (T) to switch to Mode 2. If a rollover risk is detected within the time interval $[t, t+T]$, then Mode 2 is activated (i.e., $\mathrm{RI}(t_T) \leq \mathrm{RI}_{\mathrm{ref}}$ with $t_T \in [t, t+T]$).

The rollover risk prediction is a tool that allows us to take corrective actions before the risk is imminent, and as the prediction horizon T becomes larger, it is expected that the control becomes more invasive. The operation modes of the proposed methodology are described below with the help of a switching law:

$$\dot{x} = A_i(\theta)\, x + B_\delta(\theta)\, \delta,$$
$$i = \begin{cases} 1, & \text{if } \left|\mathrm{RI}\left(t_T\right)\right| \leq \mathrm{RI}_{\mathrm{ref}} \wedge t_r > 0 \text{ (Mode 1)}, \\ 2, & \text{if } \left|\mathrm{RI}\left(t_T\right)\right| > \mathrm{RI}_{\mathrm{ref}} \text{ (Mode 2)}, \end{cases} \tag{14}$$
$$t_T \in [t, t+T], \quad T > 0,$$

where $A_1(\theta) = A(\theta)$, $A_2(\theta) = A(\theta) + B_u(\theta)K$, and t_r is a minimum residence time of Mode 2. The block diagram of the proposed controller and the switching algorithm are shown in Figures 2 and 3, respectively.

To predict the behavior of the vehicle using the model (14), the prediction of the driver behavior is needed. Although modeling driving behavior constitutes a relevant and actual topic of research, it is out of the scope of this work. In this paper, we will assume that such model is available. For convenience, the approach in [27] will be used, with a zero order dynamics since it is simple and still adaptive to changing environments.

The proposed control works as follows. The driver provides the front wheels angle and the braking force through the steering wheel and the brake pedal, respectively. At every instant t, RI is compared with $\mathrm{RI}_{\mathrm{ref}}$ along the prediction interval fixed by the prediction horizon. The RI estimation is performed using the vehicle model (14), (7), and the model of the driver behavior. If the threshold $\mathrm{RI}_{\mathrm{ref}}$ is exceeded at some instant along the interval $[t, t+T]$, the Mode 2 is activated (i.e., a future rollover risk is detected); otherwise the Mode 1 is preserved. Once Mode 2 is activated, it continues being active until the rollover risk for the entire prediction horizon has disappeared and a minimum dwell time t_r is elapsed. This dwell time prevents fast switching actions (chattering) improving the performance. The stability of the proposed control is studied in the following section, where stability properties of the switching actions are given.

4. Main Result

Let us denote the solution of (14) at time t with initial condition $x_0 = x(t_0)$, uncertain parameter within the simplex

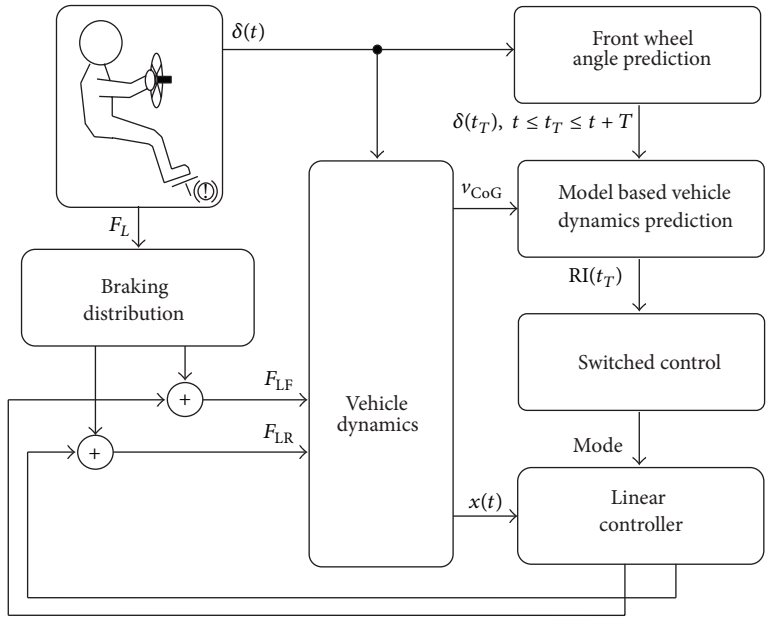


Figure 2: Block diagram of control strategy.

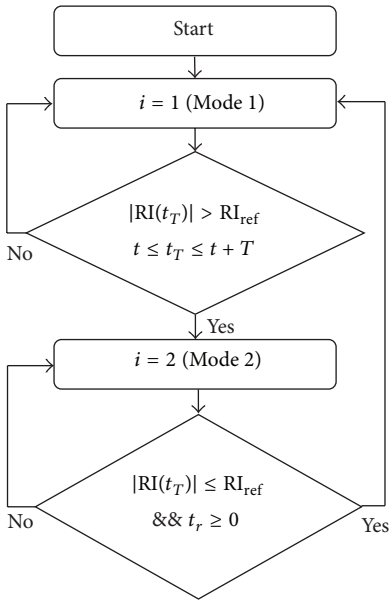


Figure 3: Algorithm of the switching criterion.

Θ, and initial time t_0 as $x(t_0, t, x_0, \theta)$. $B_\xi(z)$ is the ball centered at z with radius ξ (i.e., $B_\xi(z) = \{x \mid \|x - z\| \le \xi\}$).

For $i = 1$, the unperturbed system (14) (i.e., $\delta = 0$) has an open-loop equilibrium in $x' = [\beta_{\text{eq}}, \dot{\psi}_{\text{eq}}, \dot{\varphi}_{\text{eq}}, \varphi_{\text{eq}}] = 0$. By vehicle design, $A_1(\theta)$ is Hurwitz for all $\theta \in \Theta$; otherwise, the unperturbed vehicle would spin or rollover by following a straight line. On the other hand, the existence of a robust controller (12) such that $A_2(\theta)$ is also a Hurwitz matrix is guaranteed by the controllability of the pair $(A(\theta), B_u(\theta))$.

The control objective (13) can be expressed from the stability point of view, as the property of positive invariance in some suitable set (Ω). The trajectories $x(t_0, t, x_0, \theta)$ are positively invariant in Ω; that is, $x(t_0, t, x_0, \theta) \in \Omega$, for all $t \ge 0$.

Consider system (14) with initial conditions within $B_\varepsilon(0)$ with $\varepsilon > 0$. Since for each mode A_i is Hurwitz, then for each $\|x(t_0)\| \le \varepsilon$ there exists a maximum value of $|\widehat{\delta}|_{\max}$ such that the system trajectories will remain within $B_\rho(0)$; that is, the system is input-output stable. The first step in deriving our main contribution is the quantification of such perturbation $|\widehat{\delta}|_{\max}$ under system commutation.

Problem 5. Consider the system

$$\dot{x} = A_i(\theta)\, x + B_\delta(\theta)\, \delta, \tag{15}$$

with $i = 1, 2$ and let

$$\Omega \triangleq \{x(t_0, t, x_0, \theta) \mid |\text{RI}(t)| \le \text{RI}_{\text{ref}}\}. \tag{16}$$

Establish switching conditions for the system (15) such that $x(t_0, t, x_0, \theta) \in \Omega$ for all $t \ge 0$.

Let $i = 2$ and $t_0 = 0$ be the switching instant from Mode 1 to Mode 2. Then, the solution of (15) while evolving in Mode 2 is

$$x(0, t, x_0, \theta) = e^{A_2(\theta)t} x_0 + \int_0^t e^{A_2(\theta)(t-\tau)} B_\delta(\theta)\, \delta(\tau)\, d\tau, \tag{17}$$

$$\text{RI}(t) = C(\theta)\, e^{A_2(\theta)t} x_0 + C(\theta) + \int_0^t e^{A_2(\theta)(t-\tau)} B_\delta(\theta)\, \delta(\tau)\, d\tau, \tag{18}$$

with

$$C(\theta) = \left[0 \quad 0 \quad -\frac{2c(\theta)}{b(\theta)\, m(\theta)\, g} \quad -\frac{2k(\theta)}{b(\theta)\, m(\theta)\, g}\right]. \tag{19}$$

Computing the norm of (18) we have

$$\begin{aligned}\|\mathrm{RI}(t)\| &\le \|C(\theta)\| \left\|e^{A_2(\theta)t}\right\| \|x_0\| \\ &\quad + \|C(\theta)\| \left\|\left[I - e^{A_2(\theta)t}\right]\right\| \left\|A_2^{-1}(\theta) B_\delta(\theta)\right\| |\delta| \\ &\le \|C(\theta)\| \beta_2(\theta) e^{r_2(\theta)t} \varepsilon \\ &\quad + \|C(\theta)\| \left[1 + \beta_2(\theta) e^{r_2(\theta)t}\right] \left\|A_2^{-1}(\theta) B_\delta(\theta)\right\| |\delta| \\ &\le \|C(\theta)\| \beta_2(\theta) \varepsilon \\ &\quad + \|C(\theta)\| \left[1 + \beta_2(\theta)\right] \left\|A_2^{-1}(\theta) B_\delta(\theta)\right\| |\delta| \le \mathrm{RI}_{\mathrm{ref}},\end{aligned} \tag{20}$$

where Euclidean norm is denoted as $\|\cdot\|$, $\|x_0\| \le \varepsilon$, and $\|e^{A_2(\theta)t}\| \le \beta_2(\theta) e^{r_2(\theta)t}$ with $r_2(\theta) < 0$.

From (20) the maximum perturbation that allows $x(0, t, x_0, \theta) \in \Omega$ is

$$|\delta|_{\max} < \pi(\theta, \varepsilon) \triangleq \frac{\mathrm{RI}_{\mathrm{ref}} - \|C(\theta)\| \beta_2(\theta) \varepsilon}{\|C(\theta)\| \left[1 + \beta_2(\theta)\right] \left\|A_2^{-1}(\theta) B_\delta(\theta)\right\|}. \tag{21}$$

Let

$$\theta_{r2} = \arg\min \left(\frac{\mathrm{RI}_{\mathrm{ref}} - \|C(\theta)\| \beta_2(\theta) \varepsilon}{\|C(\theta)\| \left[1 + \beta_2(\theta)\right] \left\|A_2^{-1}(\theta) B_\delta(\theta)\right\|} \right), \tag{22}$$

then

$$|\delta|_{\max}(\varepsilon) = \pi\left(\theta_{r2}, \varepsilon\right). \tag{23}$$

Notice that there exists a maximum bound for the initial conditions, ε, such that $|\delta|_{\max} \ge 0$ and it is given by

$$\frac{\mathrm{RI}_{\mathrm{ref}}}{\left\|C\left(\theta_{r2}\right)\right\| \beta_2\left(\theta_{r2}\right)} = \varepsilon_{\max}. \tag{24}$$

If the initial conditions are at the origin, the maximum perturbation $|\widehat{\delta}|_{\max}$ is

$$\begin{aligned}\left|\widehat{\delta}\right|_{\max} &\le \pi\left(\theta_{r2}, 0\right) \\ &= \frac{\mathrm{RI}_{\mathrm{ref}}}{\left\|C\left(\theta_{r2}\right)\right\| \left[1 + \beta_2\left(\theta_{r2}\right)\right] \left\|A_2^{-1}\left(\theta_{r2}\right) B_\delta\left(\theta_{r2}\right)\right\|}.\end{aligned} \tag{25}$$

From expression (23) the maximum bound of perturbation δ depends on the initial conditions, being maximum at the origin (i.e., $|\widehat{\delta}|_{\max}$). Using this fact one can define a switching criterion using the boundary $\|x_0\| \le \varepsilon^* \le \varepsilon$ as a switching surface, which will be stable for perturbations satisfying $0 < |\delta^*| < |\widehat{\delta}|_{\max}$. That is

$$\varepsilon^* = \frac{\mathrm{RI}_{\mathrm{ref}} - \left\|C\left(\theta_{r2}\right)\right\| \left[1 + \beta_2\left(\theta_{r2}\right)\right] \left\|A_2^{-1}\left(\theta_{r2}\right) B_\delta(\theta)_{r2}\right\| |\delta^*|}{\left\|C\left(\theta_{r2}\right)\right\| \beta_2\left(\theta_{r2}\right)}, \tag{26}$$

where $\delta^* < |\widehat{\delta}|_{\max}$ is a design value that is chosen to compute ε^*. In other words, the computation of (26) only requires the (nominal) parameters of vertexes θ_{r2}.

Let the switching criterion from Mode 1 to Mode 2 be as follows:

$$\text{if } \|x\| \ge \varepsilon^* \text{ then Mode 2 is active,} \tag{27}$$

with ε^* given by (26); then $|\mathrm{RI}(t)| \le \mathrm{RI}_{\mathrm{ref}}$ for $|\delta| \le |\widehat{\delta}|_{\max}$.

Until now we have proved that the switching criterion from Mode 1 to Mode 2 keeps the trajectories of the system within Ω. However, to obtain a noninvasive property of the controller, the control must switch back to open-loop conditions as soon as the risk has passed. Since $A_2(\theta)$ is Hurwitz and $|\delta| \le |\widehat{\delta}|_{\max}$, the trajectories of Mode 2 as $t \to \infty$ will tend to

$$\lim_{t \to \infty} \left|\mathrm{RI}(t)_{\mathrm{Mode\,2}}\right| \le \left\|C(\theta) A_2^{-1}(\theta) B_\delta(\theta)\right\| \left|\widehat{\delta}\right|_{\max} = \Lambda_2, \tag{28}$$

while in Mode 1 the norm of the RI tends to

$$\lim_{t \to \infty} \left|\mathrm{RI}(t)_{\mathrm{Mode\,1}}\right| \le \left\|C(\theta) A_1^{-1}(\theta) B_\delta(\theta)\right\| \left|\widehat{\delta}\right|_{\max} = \Lambda_1. \tag{29}$$

From the controllability property of $(A(\theta), B(\theta))$ there exists a sufficiently large value of K, such that $\Lambda_1 > \Lambda_2$ for all θ in the simplex. Let $\Omega_{\Lambda_1} \triangleq \{x(t_0, t, x_0, \theta) \mid |\mathrm{RI}(t)| < \Lambda_1\}$ and $\Omega_{\Lambda_2} \triangleq \{x(t_0, t, x_0, \theta) \mid |RI(t)| < \Lambda_2\}$; then it is satisfied that $\Omega_{\Lambda_1} \supset \Omega \supset \Omega_{\Lambda_2}$. That $\Omega \supset \Omega_{\Lambda_2}$ is a consequence of (25). On the other hand, the maximum perturbation $|\widehat{\delta}|_{\max,\mathrm{la}}$ such that trajectories remain in Ω at open-loop conditions is

$$|\delta|_{\mathrm{la}} < \pi_{\mathrm{la}}(\theta, \varepsilon) = \frac{\mathrm{RI}_{\mathrm{ref}} - \|C(\theta)\| \beta_1(\theta) \varepsilon}{\|C(\theta)\| \left[1 + \beta_1(\theta)\right] \left\|A_1^{-1}(\theta) B_\delta(\theta)\right\|} \tag{30}$$

with $\|e^{A_1(\theta)t}\| \le \beta_1(\theta) e^{r_1(\theta)t}$, $r_1 < 0$. Let

$$\theta_{r1} = \arg\min \left\{ \frac{\mathrm{RI}_{\mathrm{ref}} - \|C(\theta)\| \beta_1(\theta) \varepsilon}{\|C(\theta)\| \left[1 + \beta_1(\theta)\right] \left\|A_1^{-1}(\theta) B_\delta(\theta)\right\|} \right\}, \tag{31}$$

then

$$|\delta|_{\max,\mathrm{la}}(\varepsilon) < \pi_{\mathrm{la}}\left(\theta_{r1}, \varepsilon\right), \tag{32}$$

which means that the larger value of the perturbation that can be tolerated is at $\varepsilon = 0$; that is,

$$\begin{aligned}\left|\widehat{\delta}\right|_{\max,\mathrm{la}} &< \pi_{\mathrm{la}}\left(\theta_{r1}, 0\right) \\ &= \frac{\mathrm{RI}_{\mathrm{ref}}}{\left\|C\left(\theta_{r1}\right)\right\| \left[1 + \beta_1\left(\theta_{r1}\right)\right] \left\|A_1^{-1}\left(\theta_{r1}\right) B_\delta(\theta)_{r1}\right\|}.\end{aligned} \tag{33}$$

From (33) and (25) we have that $|\widehat{\delta}|_{\max,\mathrm{la}} < |\widehat{\delta}|_{\max}$, for a sufficient large value of K. Therefore if $|\widehat{\delta}|_{\max}$ is applied at Mode 1 the trajectories can be outside Ω. Now, since $A_i(\theta)$ are

Hurwitz, the set Ω_{Λ_i} is attractive for the Mode i; therefore, any switching surface $\|x\| < \varepsilon^*$ with $x \in \Omega$ serves as a switching surface from Mode 2 to Mode 1 to ensure the control objective (13); that is,

$$\text{if } \|x\| < \varepsilon^* \text{ then Mode 1 is active.} \tag{34}$$

Since the condition to switch from Mode 2 to Mode 1 (34) is arbitrarily smaller than that to switch from Mode 1 to Mode 2 (27), fast switching actions may be present. To avoid this phenomenon a minimum residence time ($t_r > 0$) at Mode 2 can be used to avoid performance degradation. That is, if $|\delta| \leq |\widehat{\delta}|_{\max}$ and the switching criterion is defined as follows:

$$\begin{aligned}&\text{if } \|x\| \geq \varepsilon^* \text{ then Mode 2 is active}\\ &\text{if } \|x\| < \varepsilon^* \text{ and Mode 2 has been active at least a time}\\ &t_r > 0 \text{ then Mode 1 is active,}\end{aligned} \tag{35}$$

where ε^* is given by expression (26), then $|\mathrm{RI}(t)| \leq \mathrm{RI}_{\mathrm{ref}}$, for all time $t > 0$.

4.1. Predictive Switching Criterion. The result in the section above establishes sufficient conditions for stability under switching. In the following we will establish a criterion to define the switching surface ε^*, which depends on the prediction of the rollover risk, which constitutes a time varying switching surface. Without loss of generality let the present time $t = 0$; now let us compute the required initial conditions such that $|\mathrm{RI}(T)| = \mathrm{RI}_{\mathrm{ref}}$; since the rollover prediction risk must be performed in open-loop conditions only Mode 1 is analyzed (see Figure 3):

$$\begin{aligned}\left\|\mathrm{RI}_{\mathrm{ref}}\right\| = {} & \left\|C\left(\theta_{r1}\right)\right\| \beta_1\left(\theta_{r1}\right) e^{r_1(\theta_{r1})T} \widetilde{\varepsilon}\\ & + \left\|C\left(\theta_{r1}\right)\right\| \left[1 + \beta_1\left(\theta_{r1}\right) e^{r_1(\theta_{r1})T}\right]\\ & \times \left\|A_1^{-1}\left(\theta_{r1}\right) B_\delta\left(\theta_{r1}\right)\right\| |\delta|,\end{aligned} \tag{36}$$

with $|\delta| \in [0, |\widehat{\delta}|_{\max}]$ hence $\widetilde{\varepsilon}$ is the value of the switching surface at present time to avoid the imminent risk $\|\mathrm{RI}(T)\| = \|\mathrm{RI}_{\mathrm{ref}}\|$. Consider

$$\begin{aligned}\widetilde{\varepsilon} = {} & \mathrm{RI}_{\mathrm{ref}} - \left\|C\left(\theta_{r1}\right)\right\| \left[1 + \beta_1\left(\theta_{r1}\right) e^{r_1(\theta_{r1})T}\right]\\ & \times \left\|A_1(\theta_{r1})^{-1} B_\delta\left(\theta_{r1}\right)\right\| |\delta|\\ & \times \left(\left\|C\left(\theta_{r1}\right)\right\| \beta_1\left(\theta_{r1}\right) e^{r_1(\theta_{r1})T}\right)^{-1}.\end{aligned} \tag{37}$$

Let $|\delta| = |\widehat{\delta}|_{\max}$; hence substituting (25) in (37)

$$\widetilde{\varepsilon} = \frac{\mathrm{RI}_{\mathrm{ref}}\left[1 - Q\left(\theta_{r1}, \theta_{r2}\right)\right]}{\left\|C\left(\theta_{r1}\right)\right\| \beta_1\left(\theta_{r1}\right) e^{r_1(\theta_{r1})T}}, \tag{38}$$

with

$$Q\left(\theta_{r1}, \theta_{r1}\right) = \frac{\left[1 + \beta_1\left(\theta_{r1}\right) e^{r_1(\theta_{r1})T}\right] \left\|A_1(\theta_{r1})^{-1} B_\delta\left(\theta_{r1}\right)\right\|}{\left[1 + \beta_2\left(\theta_{r2}\right)\right] \left\|A_2(\theta_{r2})^{-1} B_\delta\left(\theta_{r2}\right)\right\|}. \tag{39}$$

Table 1: 2000 Cherokee nominal parameters.

c	4000 N·m·s/rad
c_v	90240 N/rad
c_h	180000 N/rad
h	0.375 m
J_z	1280 kg·m^2
J_x	362.6 kg·m^2
k	36075 n·m/rad
L_v	1.102 m
L_h	1.25 m
m	1224 kg
b	1.51 m

Notice that $1 - Q(\theta_{r1}, \theta_{r2})$ can be either positive or negative. Negative values mean that no predicted risk of rollover exists and it is not necessary to switch to Mode 2 at present time. However, as $1 - Q(\theta_{r1}, \theta_{r2})$ becomes positive, the switching is possible with

$$\widetilde{\varepsilon} \geq \varepsilon_{\max}, \tag{40}$$

with $\varepsilon_{\max}$ given by (24). From this point, stability can be derived as in the section above using $\varepsilon^* = \widetilde{\varepsilon}$.

5. Numerical Simulations

In this section, numerical evaluations of the proposed control strategy are used to illustrate the controller performance. Firstly, a numerical sensitivity analysis is performed to determine the vertexes of the polytope corresponding to (10). Afterward, the evaluation of the proposed controller using CarSim and three different vehicles is performed.

5.1. Numerical Sensitivity Analysis and Vertexes Computation. In order to keep to a minimum the number of the vertexes of the polytopic matrix, it is convenient to determine which parameters are more significant in the system description. To this end, a sensitivity study is performed. One of the simplest methods to determine parameter sensitivity is to compute the output response for a variety of maneuvers, in face of parameter variations. The parameters are perturbed using fixed steps of 5%, from a lower to an upper bound. The sensitivity index (SI) is defined as [28]

$$\mathrm{SI} = \frac{\widehat{y}_{\max} - \widehat{y}_{\min}}{\widehat{y}_{\max}} \cdot 100, \tag{41}$$

where y is the measured variable, $\widehat{y}_{\max} = \max\{y_{\max}(\varrho_{\mathrm{lower}}), y_{\max}(\varrho_{\mathrm{lower}} + 0.05\varrho_{\mathrm{lower}}), y_{\max}(\varrho_{\mathrm{lower}} + 0.1\varrho_{\mathrm{lower}}), \ldots, y(\varrho_{\mathrm{upper}})\}$, where ϱ_{lower} and ϱ_{upper} are the lower and upper bound of the parameters. On the other hand, $\widehat{y}_{\min} = \min\{y_{\max}(\varrho_{\mathrm{lower}}), y_{\max}(\varrho_{\mathrm{lower}} + 0.05\varrho_{\mathrm{lower}}), y_{\max}(\varrho_{\mathrm{lower}} + 0.1\varrho_{\mathrm{lower}}), \ldots, y(\varrho_{\mathrm{upper}})\}$. In this paper, y is the rollover index.

Nominal parameters are displayed in Table 1, while Table 2 shows the maximum SI values of the each parameter. It can be observed that the vehicle dynamics are very sensitive to the design dependent values c_v (related with the tire),

Table 2: Sensitivity index.

Parameter	SI
c_v	11.65%
h	10.88%
L_h	9.01%
L_v	6.84%
v_{CoG}	6.76%
m	5.37%
c_h	5.06%
c	1.83%
J_x	1.12%
k	0.62%
J_z	0.2%

h (distance of the CoG from the road), L_v and L_h (distance from CoG to the front, rear axles resp.), and the vehicle speed v_{CoG}. Let us fix L_h and L_v at their nominal values since their variation is usually small; then the vertexes of the matrices are defined with respect to the parameters c_v, h, and v_{CoG} for simplicity. $A(\theta)$ and $B_\delta(\theta)$ can be expressed as a convex function of the time-varying parameter vector:

$$\theta = [\theta_1, \theta_2, \theta_3, \theta_4, \theta_5, \theta_6, \theta_7, \theta_8], \tag{42}$$

as described as follows:

$$\begin{aligned} A(\theta) &= \sum_{i=1}^{8} \theta_i A_i, \\ B_\delta(\theta) &= \sum_{i=1}^{8} \theta_i B_{\delta_i}, \end{aligned} \tag{43}$$

where

$$\begin{aligned} A_1 &= A\left(\underline{c_v}, \underline{h}, \underline{v_{\mathrm{CoG}}}\right), & A_2 &= A\left(\underline{c_v}, \underline{h}, \overline{v_{\mathrm{CoG}}}\right), \\ A_3 &= A\left(\underline{c_v}, \overline{h}, \underline{v_{\mathrm{CoG}}}\right), & A_4 &= A\left(\underline{c_v}, \overline{h}, \overline{v_{\mathrm{CoG}}}\right), \\ A_5 &= A\left(\overline{c_v}, \underline{h}, \underline{v_{\mathrm{CoG}}}\right), & A_6 &= A\left(\overline{c_v}, \underline{h}, \overline{v_{\mathrm{CoG}}}\right), \\ A_7 &= A\left(\overline{c_v}, \overline{h}, \underline{v_{\mathrm{CoG}}}\right), & A_8 &= A\left(\overline{c_v}, \overline{h}, \overline{v_{\mathrm{CoG}}}\right), \\ B_{\delta_1} &= B_\delta\left(\underline{c_v}, \underline{h}, \underline{v_{\mathrm{CoG}}}\right), & B_{\delta_2} &= B_\delta\left(\underline{c_v}, \underline{h}, \overline{v_{\mathrm{CoG}}}\right), \\ B_{\delta_3} &= B_\delta\left(\underline{c_v}, \overline{h}, \underline{v_{\mathrm{CoG}}}\right), & B_{\delta_4} &= B_\delta\left(\underline{c_v}, \overline{h}, \overline{v_{\mathrm{CoG}}}\right), \\ B_{\delta_5} &= B_\delta\left(\overline{c_v}, \underline{h}, \underline{v_{\mathrm{CoG}}}\right), & B_{\delta_6} &= B_\delta\left(\overline{c_v}, \underline{h}, \overline{v_{\mathrm{CoG}}}\right), \\ B_{\delta_7} &= B_\delta\left(\overline{c_v}, \overline{h}, \underline{v_{\mathrm{CoG}}}\right), & B_{\delta_8} &= B_\delta\left(\overline{c_v}, \overline{h}, \overline{v_{\mathrm{CoG}}}\right). \end{aligned} \tag{44}$$

In this paper $\underline{v_{\mathrm{CoG}}} = 25\,\mathrm{m/s} = 90\,\mathrm{Km/h}$, and $\overline{v_{\mathrm{CoG}}} = 40\,\mathrm{m/s} = 144\,\mathrm{Km/h}$. A controller gain vector that guarantees stability for all vertexes is $K = \mathrm{mg}\,[-7.1287 \quad 0.9842 \quad 0.3271 \quad -0.0944]$.

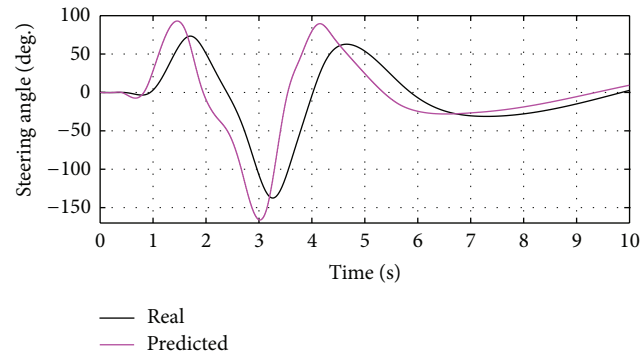


Figure 4: Steering wheel angle real versus predicted.

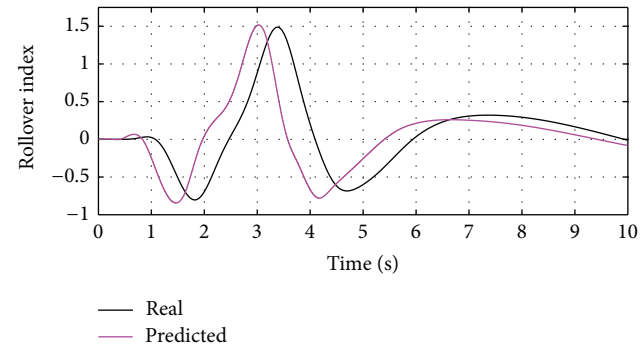


Figure 5: Rollover index real versus predicted (open loop).

5.2. The Predictive Controller. The first step to implement the predictive controller in Figures 2 and 3 is to predict the front wheels angle (δ) given by the driver's behavior. In this work a zero-order model is used to perform the prediction as in [27]. Afterward, the RI at open-loop conditions is computed along a given prediction horizon. The results are displayed in Figures 4 and 5 for a double lane-change maneuver with rollover risk.

Secondly the predicted RI is used to perform the switching, using the algorithm in Figure 3. Two double lane-change maneuvers are chosen to illustrate the performance of the switched controller: (i) a first one with rollover risk along the [3, 3.8] seconds interval in open loop (see Figure 6) and (ii) a second one without rollover risk (see Figure 7).

The comparison of the proposed strategy with respect to a robust nonswitched linear action and open loop dynamics is performed. The proposed controller is able to prevent rollover and to maintain the RI within the specified values. Moreover, the control action is reduced significantly with respect to the nonswitched alternative and it is used only when necessary to avoid rollover (in Figure 7 the braking control actions are not performed).

The effect of the prediction horizon is evaluated in the proposed scheme for a given maximum value of $|\delta| < |\widehat{\delta}|_{\max}$; the results are displayed in Figure 8. As the prediction horizon is increased the controller becomes more invasive; however the actual values of RI are smaller. In every case, RI is kept within the required values.

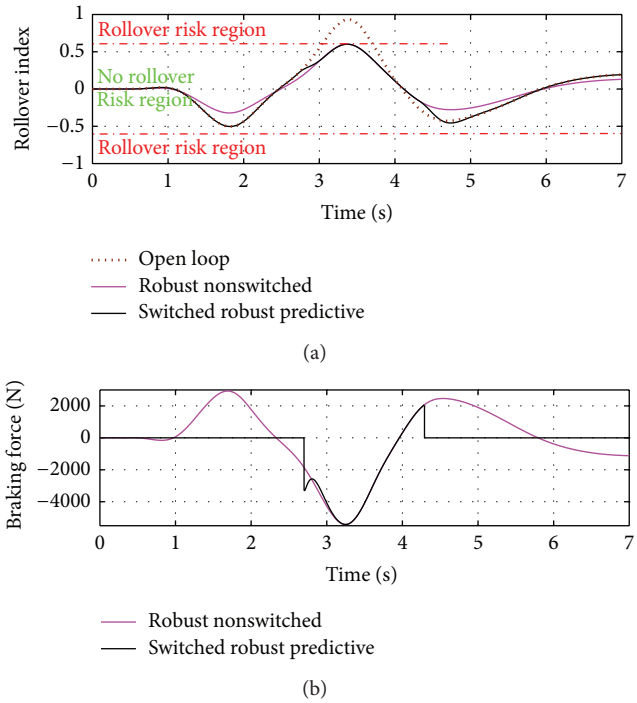


FIGURE 6: Rollover index and braking force evolution for a double lane-change maneuver with rollover risk.

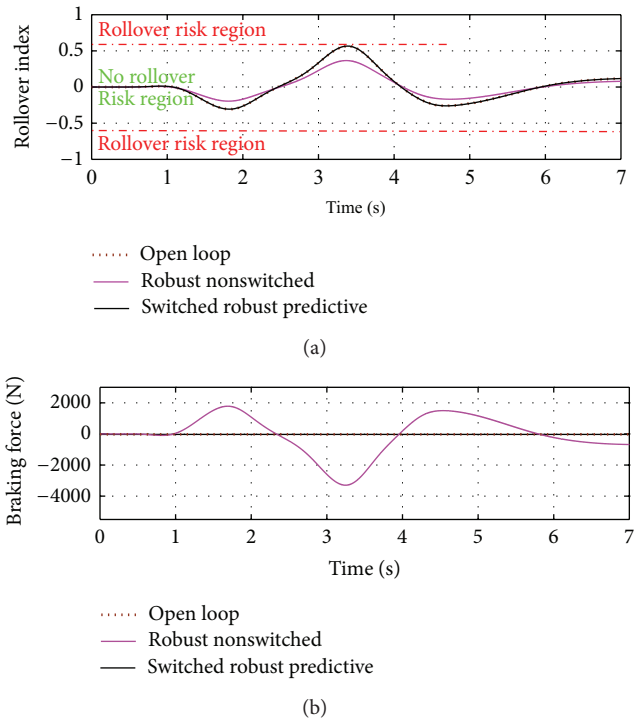


FIGURE 7: Rollover index and braking force evolution for a double lane-change maneuver without rollover risk.

To illustrate the robustness of the proposed scheme, the controller is evaluated using models from CarSim of three vehicles: a 2000 Porsche 911 sport car, a 2002 Mitsubishi Mini U62T utility truck, and a Jeep Cherokee 2000 SUV. The nominal parameters for these vehicles are shown in Table 3.

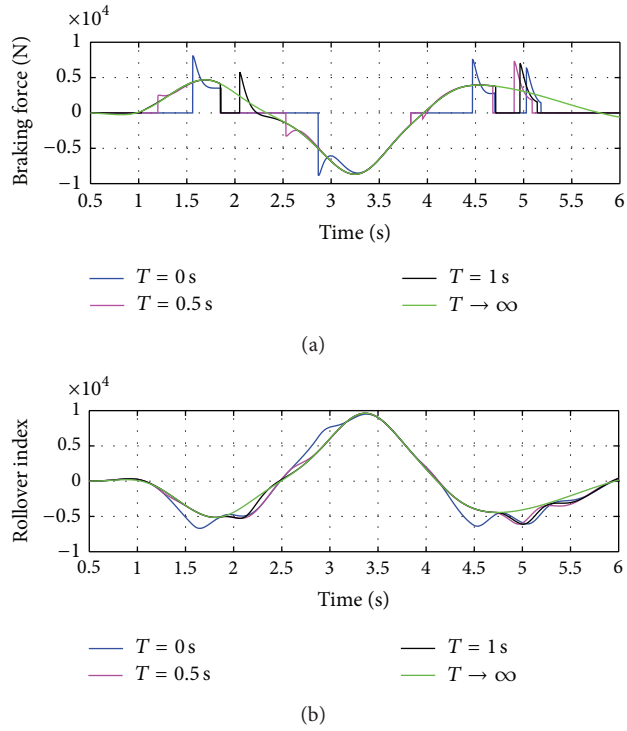


FIGURE 8: Effect of prediction horizon on system performance.

TABLE 3: Nominal parameters.

Value	Cherokee	Porsche	U62T	Units
h	0.375	0.010	0.200	m
J_z	1280	1270	686	kg·m^2
J_x	362.6	614	384	kg·m^2
L_v	1.102	1.525	0.550	m
L_h	1.25	0.825	1.373	m
m	1224	1278	600	kg
b	1.51	1.72	1.26	m

An aggressive double lane-change is performed at 120 Km/h and the open loop behaviors of the three vehicles are shown in Figure 9. Since the Porsche has a smaller h (among other differences), the maneuver does not lead to rollover but to lateral slip, while for the Cherokee and the U62T the rollover is present. In Figure 10 the proposed controller is applied to the three cars. In this case, the controller that uses the model of the Jeep Cherokee is used also for the Porsche and the utility truck. The controller is able to prevent rollover in the Jeep Cherokee, and surprisingly the same controller is able to prevent rollover in the case of the truck, even if considerable differences exist. As can be observed the controller leads to a further decrement on the velocity to compensate for the higher CoG, among other differences. On the other hand, it is also surprising that the skid and 180° spin of the sport vehicle can be prevented.

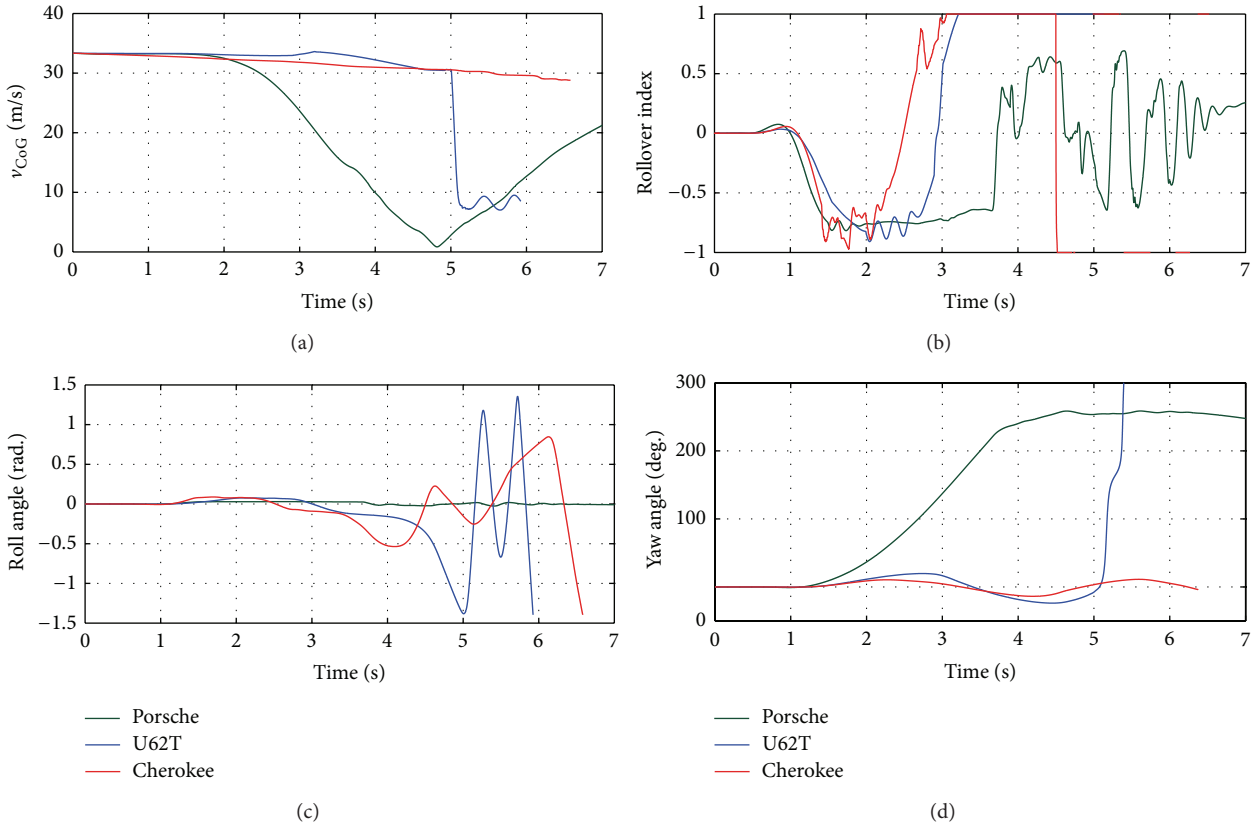


FIGURE 9: Dynamic response comparative in open loop for an aggressive double lane-change maneuver.

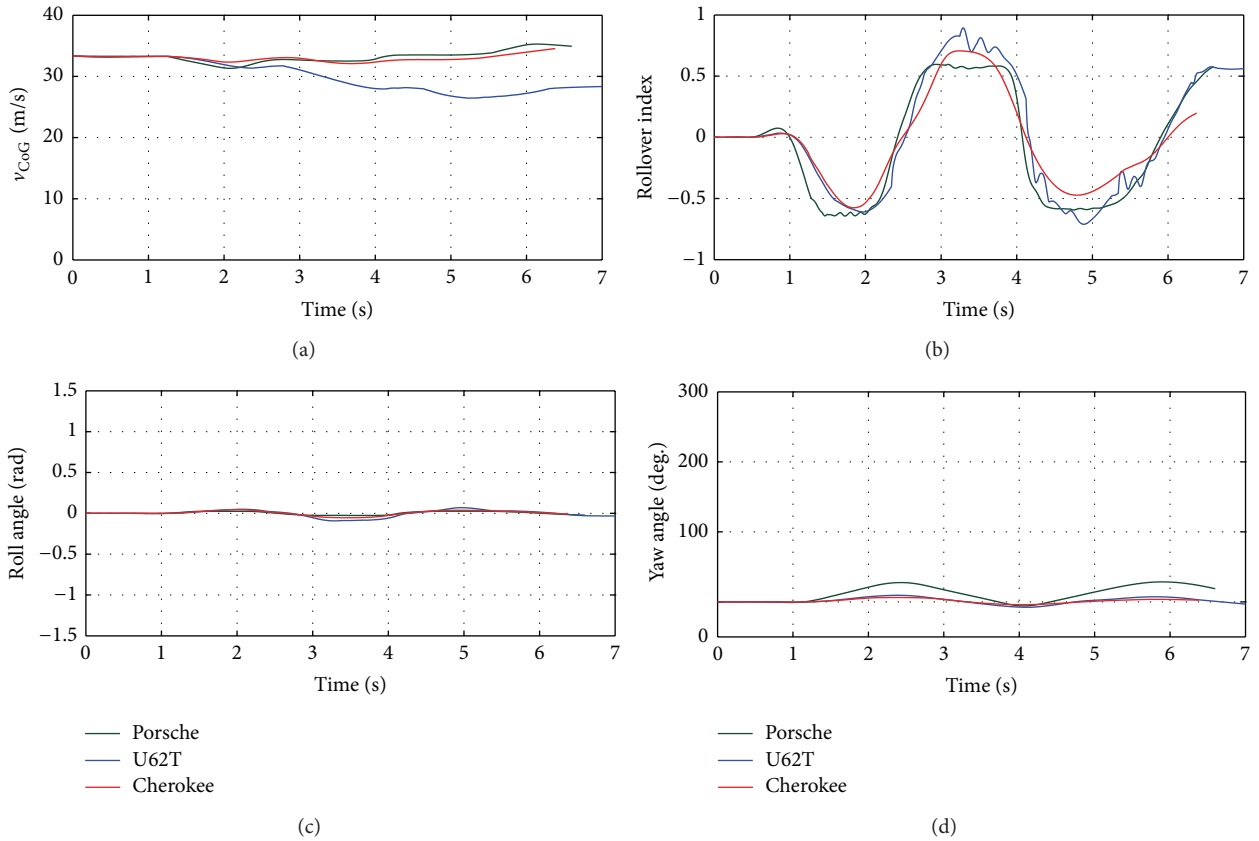


FIGURE 10: Dynamic response comparative with the proposed controller active for an aggressive double lane-change maneuver.

6. Conclusions

In this paper a robust switched controller with a commutation dependent on the RI prediction is proposed. The stability of the controller is derived in two steps. Firstly, by noticing that a state-dependent switching criterion ensures the invariance of a suitable set Ω for a given $|\widehat{\delta}|_{\max}$, and secondly using RI prediction to define a time-varying state-dependent switching criterion. The results are illustrated with simulations and the robustness of the proposed scheme is verified for different vehicle models. The proposed controller is robust and it has displayed an unexpected capacity of preventing skidding; this last feature must be further analyzed.

Nomenclature

α_v, α_h: Front, rear tire side slip angle
β: Vehicle side slip angle
δ: Front wheels angle
ψ: Vehicle yaw angle
φ: Chassis roll angle
b: Track width
c: Damping coefficient
c_v, c_h: Front, rear tire stiffness coefficient
F_{ZL}, F_{ZR}: Left, right contact point vertical forces
g: Gravity
h: Distance of CoG from the road
J_z, J_x: Moment of inertia about the vertical, longitudinal axis
k: Spring stiffness coefficient of the suspension
L_v, L_h: Distance from CoG to the front, rear axle
m: Vehicle mass
RI: Rollover index
T: Prediction horizon
u: Differential braking force
$u_{\max}$: Maximum differential braking force
v_{CoG}: Vehicle speed of the CoG relative to the inertial coordinate system as seen in the vehicle.

Conflict of Interests

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

Acknowledgments

The authors acknowledge financial support from CONACYT GRANTS: SENER 152485 and FORDECYT 190966. The authors also acknowledge technical support from Irwin A. Díaz-Diaz.

References

[1] North American Free Trade Agreement, Land Transportation Standards Subcommittee, "Highway safety performance criteria in support of vehicle weight and dimension regulations: candidate criteria & recommended thresholds," Working Draft for Consultation, 1999.

[2] S. Takano and M. Nagai, "Dynamics control of large vehicles for rollover prevention," in *Proceedings of the Vehicle Electronics Conference*, pp. 85–89, Tottori, Japan, 2001.

[3] M. H. Kim, J. H. Oh, J. H. Lee, and M. C. Jeon, "Development of rollover criteria based on simple physical model of rollover event," *International Journal of Automotive Technology*, vol. 7, no. 1, pp. 51–59, 2006.

[4] X. Wu, X. Ge, and H. Huang, "Study on vehicle rollover avoidance," in *Proceedings of the International Conference on Measuring Technology and Mechatronics Automation (ICMTMA '10)*, pp. 681–684, Changsha, China, March 2010.

[5] G. Phanomchoeng and R. Rajamani, "New rollover index for the detection of tripped and untripped rollovers," *IEEE Transactions on Industrial Electronics*, vol. 60, no. 10, 2013.

[6] R. Ghandour, A. Victorino, A. Charara, and D. Lechner, "Risk indicators anticipation based on the vehicle dynamics anticipation to avoid accidents," in *Proceedings of the IEEE Intelligent Vehicles Symposium*, pp. 93–98, Alcala de Henares, España, 2012.

[7] C. Zong, Z. Yu, Q. Miao, and B. Zhang, "Research on rollover warning algorithm of heavy commercial vehicle based on prediction model," in *Proceedings of the International Conference on Computer, Mechatronics, Control and Electronic Engineering (CMCE '10)*, pp. 432–435, Changchun, China, August 2010.

[8] T. Zhu and C. Zong, "An advanced methodology for rollover warning of heavy duty truck based on Kalman filter state estimation," in *Proceedings of the International Conference on Intelligent Computation Technology and Automation (ICICTA '10)*, pp. 466–469, Changsha, China, May 2010.

[9] H. Imine, L. M. Fridman, and T. Madani, "Steering control for rollover avoidance of heavy vehicles," *IEEE Transactions on Vehicular Technology*, vol. 61, no. 8, pp. 3499–3509, 2012.

[10] J. Kang, J. Yoo, and K. Yi, "Driving control algorithm for maneuverability, lateral stability, and rollover prevention of 4WD electric vehicles with independently driven front and rear wheels," *IEEE Transactions on Vehicular Technology*, vol. 60, no. 7, pp. 2987–3001, 2011.

[11] M. Islam and C. Ha, "Road vehicle rollover avoidance using active steering controller," in *Proceedings of the 14th International Conference on Computer and Information Technology (ICCIT '11)*, pp. 298–302, Dhaka, Bangladesh, December 2011.

[12] S. Yim, "Design of a preview controller for vehicle rollover prevention," *IEEE Transactions on Vehicular Technology*, vol. 60, no. 9, pp. 4217–4226, 2011.

[13] S. Solmaz, "Switched stable control design methodology applied to vehicle rollover prevention based on switched suspension settings," *IET Control Theory & Applications*, vol. 5, no. 9, pp. 1104–1112, 2011.

[14] H. Horisberger and P. R. Belanger, "Regulators for linear, time invariant plants with uncertain parameters," *IEEE Transactions on Automatic Control*, vol. 21, no. 5, pp. 705–708, 1976.

[15] G. Chesi, A. Garulli, A. Tesi, and A. Vicino, "Robust stability of time-varying polytopic systems via parameter-dependent homogeneous Lyapunov functions," *Automatica*, vol. 43, no. 2, pp. 309–316, 2007.

[16] A. BenAbdallah, M. A. Hammami, and J. Kallel, "Robust stability of uncertain piecewise-linear systems: LMI approach," *Nonlinear Dynamics*, vol. 63, no. 1-2, pp. 183–192, 2011.

[17] S. Solmaz, M. Corless, and R. Shorten, "A methodology for the design of robust rollover prevention controllers for automotive vehicles: part 1-differential braking," in *Proceedings of the 45th IEEE Conference on Decision and Control (CDC '06)*, pp. 1739–1744, San Diego, Calif, USA, December 2006.

[18] T. Zhu, B. Wu, and C. Zong, "The design of rollover prevention system in heavy vehicles based on an robust method," in *Proceedings of the 2nd Conference on Power Electronics and Intelligent Transportation System (PEITS '09)*, pp. 98–101, Shenzhen, China, December 2009.

[19] M. Akar and A. D. Dere, "A switching rollover controller coupled with closed-loop adaptive vehicle parameter identification," *IEEE Transactions on Intelligent Transportation Systems*, no. 99, pp. 1–7, 2014.

[20] C. Larish, D. Piyabongkarn, V. Tsourapas, and R. Rajamani, "A new predictive lateral load transfer ratio for rollover prevention systems," *IEEE Transactions on Vehicular Technology*, vol. 62, no. 7, pp. 2928–2936, 2013.

[21] J. Pathompong and O. Toshiyuki, "Stabilization of vehicle rollover by nonlinear model predictive control," in *Proceedings of the SICE Annual Conference*, pp. 1568–1573, Nagoya, Japan, 2013.

[22] N. Bouton, R. Lenain, B. Thuilot, and P. Martinet, "An active anti-rollover device based on predictive functional control: application to an all-terrain vehicle," in *Proceedings of the IEEE International Conference on Robotics and Automation (ICRA '09)*, pp. 1309–1314, Kobe, Japan, May 2009.

[23] H. B. Pacejka, "Vehicle modeling," in *Tyre and Vehicle Dynamics*, chapter 4, p. 150, Butterworth-Heinemann, Oxford, UK, 3rd edition, 2012.

[24] U. Kiencke and L. Nielsen, "Vehicle modeling," in *Automotive Control Systems for Engine Driveline and Vehicle*, chapter 8, p. 360, Springer, San Diego, Calif, USA, 2nd edition, 2005.

[25] S. Solmaz, *Topics in automotive rollover prevention: robust and adaptive switching strategies for estimation and control [Ph.D. dissertation]*, Hamilton Institute, Kildare, Ireland, 2007.

[26] T. Pancake, M. Corless, and M. Brockman, "Analysis and control of polytopic uncertain/nonlinear systems in the presence of bounded disturbance inputs," in *Proceedings of the American Control Conference*, pp. 159–163, Chicago, Ill, USA, June 2000.

[27] M. R. Licea and I. Cervantes, "On the predictive rollover detection in wheeled vehicles," in *Proceedings of the 38th Annual Conference on IEEE Industrial Electronics Society*, pp. 2410–2415, Montreal, Canada, 2012.

[28] H. Gardner, "Evaluation of uncertainties in environmental radiological assessment models," Tech. Rep. NUREG/CR-3332, U.S. Nuclear Regulatory Commission, Washington, DC, USA, 1983.

Hindawi Publishing Corporation
Mathematical Problems in Engineering
Volume 2014, Article ID 513283, 20 pages
http://dx.doi.org/10.1155/2014/513283

Research Article

Invariant Hough Random Ferns for Object Detection and Tracking

Yimin Lin,[1,2] Naiguang Lu,[1,2] Xiaoping Lou,[2] Fang Zou,[3] Yanbin Yao,[3] and Zhaocai Du[3]

[1] *Institute of Optical Communication & Optoelectronics, Beijing University of Posts & Telecommunications (BUPT), No. 10 Xitucheng Road, Haidian District, Beijing 100876, China*
[2] *School of Instrumentation Science & Optoelectronics Engineering, Beijing Information Science & Technology University (BISTU), No. 12 Qinghe Xiaoying East Road, Haidian District, Beijing 100192, China*
[3] *Beijing Aeronautical Manufacturing Technology Research Institute, Beijing 100024, China*

Correspondence should be addressed to Yimin Lin; linyimin2012@hotmail.com and Naiguang Lu; nglv2002@sina.com

Received 8 December 2013; Revised 6 February 2014; Accepted 13 February 2014; Published 8 April 2014

Academic Editor: Ilse C. Cervantes

Copyright © 2014 Yimin Lin 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 introduces an invariant Hough random ferns (IHRF) incorporating rotation and scale invariance into the local feature description, random ferns classifier training, and Hough voting stages. It is especially suited for object detection under changes in object appearance and scale, partial occlusions, and pose variations. The efficacy of this approach is validated through experiments on a large set of challenging benchmark datasets, and the results demonstrate that the proposed method outperforms state-of-the-art conventional methods such as bounding-box-based and part-based methods. Additionally, we also propose an efficient clustering scheme based on the local patches' appearance and their geometric relations that can provide pixel-accurate, top-down segmentations from IHRF back-projections. This refined segmentation can be used to improve the quality of online object tracking because it avoids the drifting problem. Thus, an online tracking framework based on IHRF, which is trained and updated in each frame to distinguish and segment the object from the background, is established. Finally, the experimental results on both object segmentation and long-term object tracking show that this method yields accurate and robust tracking performance in a variety of complex scenarios, especially in cases of severe occlusions and nonrigid deformations.

1. Introduction

Object detection and tracking have become central topics in computer vision research, and recent approaches have shown considerable progress under challenging situations such as changes in object appearance, scale, occlusions, and pose variations, [1, 2]. In this paper, we just focus on three themes, which are object representation, detection, and tracking, and propose a novel framework that can be used for part-based object detection in images and online object tracking through videos.

For feature description, the common local binary feature (LBF) [3] is computed in the image intensity or gradient domain yielding successful detection results for specific objects. However, the original LBF cannot be robust to rotation variations. Traditionally, the rotation problem has been addressed from a multiclass perspective by using classifiers repeatedly trained at different orientations [4]. Unfortunately, these methods suffer from two weaknesses. Firstly, the computational cost for both the training and detecting stages increases with the number of classifiers, and secondly, the use of multiple classifiers increases the number of false positives [5]. In this paper, our object representation is related to several paradigms, such as SIFT and SURF descriptor [6, 7]. They assigned one or more orientations to each patch based on image gradient directions. Thus the matching operations are simply performed on image data that has been transformed relative to the assigned orientation and achieve invariance to these transformations. We extend the same idea to LBF and propose to compute a LBF in the polar coordinates instead, since the coordinates of the

descriptor are easy to rotate with a certain polar angle relative to the patch orientation. Therefore, our rotation invariant features have demonstrated remarkable results for object categorization under the challenging conditions, such as rotation variations and cluttered background.

Recently for part-based object detection, a popular way to address occlusions and nonrigid deformations is to combine the two ideas of appearance codebooks [8, 9] and generalized Hough transform [10, 11]. Such codebooks are used to classify the local appearance of interest points into scattered fragments of visual words that represent an object class [12]. The Hough transform was initially developed to detect analytically defined shapes, such as lines, circles, and ellipses. But up to now, the generalized Hough transform can be used to detect arbitrary shapes (i.e., shapes having no simple analytical form) [13]. It has been successfully adapted to the problem of part-based object detection since it is robust to partial occlusions and slightly deformed shapes [14]. Moreover it is tolerant to noise and can find multiple occurrences of a shape during the same processing pass. The main disadvantage is that it requires a lot of storage and extensive computation. Although it is inherently parallelizable, it has been reported that the Hough voting efficiency during the object categorization can be improved by a highly efficient classifier [15]. Therefore, in this paper, we apply the random ferns classifier (RFC) [3] to Hough transform for improving the search speed and reducing the need of a large space for data store. In addition, our Hough voting is performed in a rotation and scale invariant Hough space since each support point shares a stable polar angle and a scalable displacement related to the object's center.

Nowadays, visual object tracking has been formulated as online tracking by detection problem [16]. For the purpose of separating the target from background in individual frames, this method involves the continuous application of an object detection algorithm, where a target object is discriminated by a classifier. In order to handle the lack of prior knowledge and appearance changes, there is an essential need for an online learning algorithm incrementally updating the object template and retraining the classifier over time [17]. The most straightforward method, which replaces the template every frame with the image region believed to be the target, is found to suffer from gradual drift of the target out of the template, eventually resulting in the loss of the target. This phenomenon is referred to as template drifting problem [18], where template drift is due to the accumulation of small errors introduced in the location of the template each time when the template is updated. In this paper we propose a template segmentation algorithm based on a clustering scheme which groups the binary masks according to the appearances and geometric relations of object parts. Then we use back-projection to locate the hypothesis's support, which gives a rough localization of object parts. This support that has valid geometric relations guides intensity matching with the clustering patches. Therefore, the candidate template can be separated from the background pixelwise and still stays firmly attached to the original object.

In a word, the main contributions of this paper include the following.

(1) The rotation invariant LBF based on polar coordinates is shown to be invariant to image rotation. In particular, our object representation integrating with the intensity and gradient information is robust across a substantial range of rotation variations, addition of noise, and changes in illumination.

(2) The IHRF framework where Hough transform is combined with RFC provides an efficient way for object detection regardless of partial occlusions and nonrigid deformations. Specially, the Hough voting is robust to changing orientation and scale due to the stable polar angle and scalable displacement between the support point and object's center.

(3) A refined top-down segmentation algorithm based on a clustering scheme is proposed to guide a precisely segmentation process after the back-projection. The most important property of the clustering scheme is that all clusters are compact and only contain image patches that are visually similar since the clustering criterion relies on the geometric relations of object parts. As a result, this algorithm is able to separate the object precisely from a cluttered background.

(4) On the basis of IHRF, we present an online tracking approach that is able to prevent drifting problem because of the stable object detection and refined segmentation results, which enables more robust training of the classifier.

The rest of the paper is organized as follows. Firstly, Section 2 reviews the work related to Hough transform for object detection and online object tracking. Next, we introduce the IHRF framework in Section 3 and show how this could be applied for online tracking in Section 4. The experimental results including a comparison to state-of-the-art are given in Section 5. Finally, contributions and suggestions for future research are discussed in Section 6.

2. Related Work

The problem of object detection in images is known to be very challenging and needed to address several difficult issues such as large intraclass object variations, changes in object pose and illumination, cluttered backgrounds, and partial occlusions, [19]. Currently, there are two existing approaches which are sliding window [20–22] and part-based methods [23–25]. The latter one is more competent for solving these problems since many object categories are poorly represented by bounding boxes. Furthermore, it achieves excellent performance for occluded and deformable objects since object is represented as an assembly of local parts and flexible spatial relations between them [26, 27]. One specific subtype of this part-based detecting model is the implicit shape model (ISM) [28], which is a well-known approach based on the generalized Hough transform technique. During training, the ISM learns a model of the spatial occurrence distributions of local patches with respect to the object's center. During testing, this learned model is used to cast probabilistic votes to the location of object's center by the generalized Hough

transform. Many modified approaches that are related to ISM have been proposed [29–31].

A drawback of such methods is that matching the patches with the codebooks during testing is computationally expensive due to the large number of codebooks. To overcome this, Gall et al. [1] proposed a Hough Forest for object detection that employs random forests to learn the patches in a supervised manner. Hough Forests have been shown to outperform the sliding window classifiers and are inherently capable of multiclass detection. These advantages inspired a series of applications and extensions [32, 33]. However, so far Hough Forests also have some limitations. For instance, the Hough voting step implies considerable computational effort since the computational complexity of matching a patch against a tree is logarithmic to the number of leaves. In addition, they do not include the top-down segmentation capabilities that were available in their ISM predecessor [26].

Visual object tracking, whose goal is to estimate the states (positions or regions) of a target corresponding from one frame to the next, is one of the most important issues in computer vision. It is used in a wide range of applications including automated security and surveillance, human computer interaction, augmented reality, traffic control, and vision navigation, [34, 35]. Although visual object tracking has been studied for several decades and much progress has been made in recent years, it remains a very challenging problem due to a variety of factors that affect the performance of a tracking algorithm, such as the loss of information caused by the projection of the 3D world on a 2D image, noise in images, background clutters, illumination and scale variations, partial or full occlusions, complex object motion, camera motion, and real-time processing requirements. Nowadays there exists no single tracking approach that can successfully handle all of the above scenarios [36].

Recently, visual object tracking has been formulated as an online tracking by detection problem [16]. Although state-of-the-art online approaches as [37–39] perform well in certain scenarios, the update of the error appearance degrades the model and can lead to a significant drift. Therefore several novel algorithms, such as multiple instance learning [40] or the combination of tracking and detection [41], became very robust against the drifting problem. As the author suggested, a part-based model could potentially reduce the amount of drift by better aligning the tracker location with the object [42]. In addition, inspired by works such as [9, 26, 43], Godec et al. [15] proposed an online tracker based on the GrabCut segmentation algorithm [44] to find the object boundaries. It delivers a more precise description of the object and avoids the drifting problem to some extent, while it often leads to poor segmentations of objects since GrabCut algorithm sometimes splits them into multiple regions or merges them with parts of the background.

3. Invariant Hough Random Ferns

Random fern descriptors, also called LBF, consist of some logical pairwise comparisons of the intensity or gradient levels of randomly selected pixels in the input images [45]. However, such comparisons are not robust to rotation and scale variations because every pairwise pixel is randomly generated offline while remaining fixed in runtime. Under these variations, we propose a rotation and scale invariant descriptor that demonstrates a high degree of stability, which we show in Sections 3.1 and 3.2. In addition, our proposed IHRF consists of random ferns that are trained to learn a mapping from a densely sampled rotation and scale invariant LBF to their corresponding votes in a Hough space (see Section 3.3). In Section 3.4, we will also show how the back-projected hypothesis's support can be used to infer a pixelwise figure-ground segmentation of the object by using a clustering scheme.

Figure 1 illustrates the procedure for object detection and segmentation based on our IHRF. We formulate the rotation invariant and multiscale object detection problem as a probabilistic Hough voting procedure. For this example, the IHRF is trained on 328 horse images and masks are obtained from the Weizmann Horse database [46]. When presented with the test image, the system extracts 100 interest points (Figure 1(a)) within the horse masks to generate the dense scanning windows (Figure 1(b)). Those local patches then cast probabilistic votes containing object centroid locations, which are collected in the voting space (Figure 1(c)). As the visualization of this space in Figure 1(d) shows, the system searches for local maxima in the voting space and returns the correct detection as the strongest hypothesis. By back-projecting the contributing votes, we retrieve the hypothesis's support in the image (Figure 1(g)) and roughly separate the object from the background. To yield precise segmentation, we cluster the local masks according to their spatial and visual similarity, as shown in Figures 1(e) and 1(f). After matching the clusters, we deliver a more precise description (Figure 1(i)) of the object based on the masks (Figure 1(h)). All the key steps are described in detail in the following sections.

3.1. Rotation Invariant Local Binary Feature. Rotation invariant descriptors are useful when objects of the same class can appear in different poses. The original LBF formulation can consist of randomly selected pairwise comparisons of the image values, such as the intensity, color, and gradient of the input images. To make the illumination and intraclass variations more robust, refer to [47], we use the following 16 feature channels: 3 color channels of the Lab color space, 4 absolute values of the first- and second-order derivatives in the x and y directions, and 9-bin histogram of gradients as feature channels.

More specifically, suppose $I_n(x, y)$ is the nth $(n = 1, \ldots, 16)$ feature channel obtained from an image patch centered at pixel locations x and y. Each fern applies a series of binary tests to the pairwise pixels as follows:

$$f(i, j, n) = \begin{cases} 1, & I_n(x_i, y_i) > I_n(x_j, y_j), \\ 0, & I_n(x_i, y_i) \leq I_n(x_j, y_j), \end{cases} \quad (1)$$

where (x_i, y_i) and (x_j, y_j) are random pairwise pixels locations and each comparison returns 0 or 1. n is also randomly chosen. Generally, the number of selected pairwise pixels S maps an image patch to a 2^S-dimensional space of binary

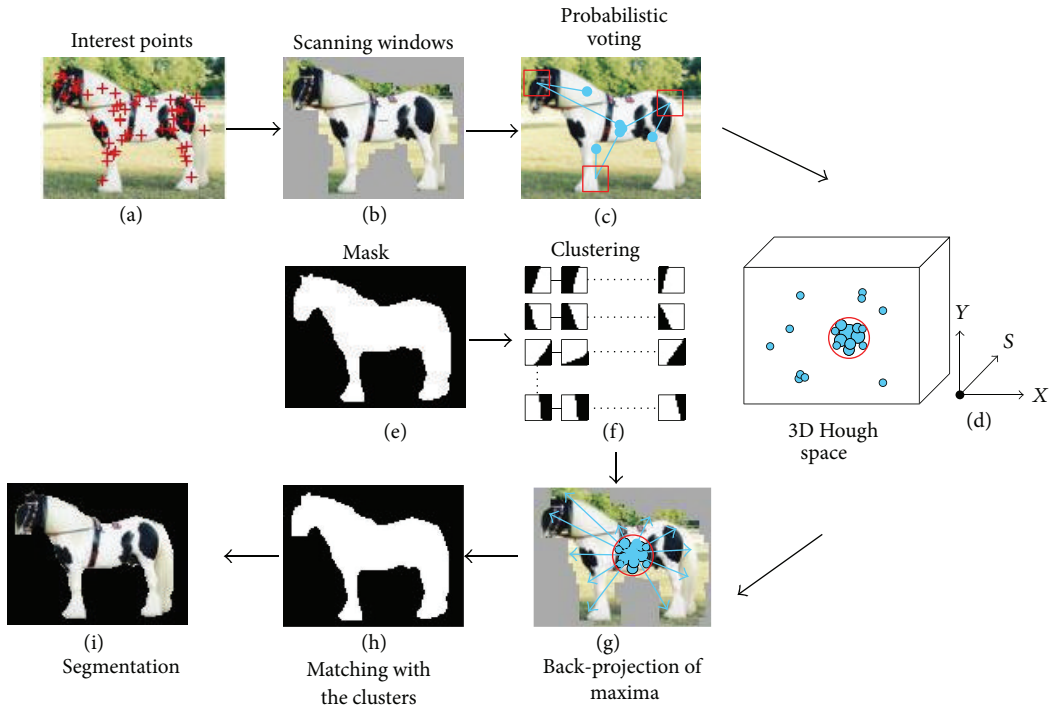


Figure 1: The part-based object detection and segmentation procedure.

descriptors in each fern. According to (1), we use the abbreviated form f for $f(i, j, n)$ and LBF can be computed as

$$F_m = \{f_1, f_2, \ldots, f_S\}, \tag{2}$$

where F_m is the mth fern and f_i means the ith binary feature (as (1)). Note that $m = 1, \ldots, K$, where K is the number of ferns. Therefore, the entire set of random ferns can be denoted by $F = \{F_1, F_2, \ldots, F_K\}$. A trade-off between performance and memory can be made by changing the number of ferns K and their sizes S. For example, $K = 1$ and $S = 4$; we suppose the feature $F = \{1101\}$. If $K = 2$ and $S = 4$, we get the feature $F = \{0101, 1010\}$.

Because the original LBF is not robust to rotation variations, here we present a novel rotation invariant descriptor for detecting objects in a specific category that may appear in images under different rotations. The differences are shown in Figure 2.

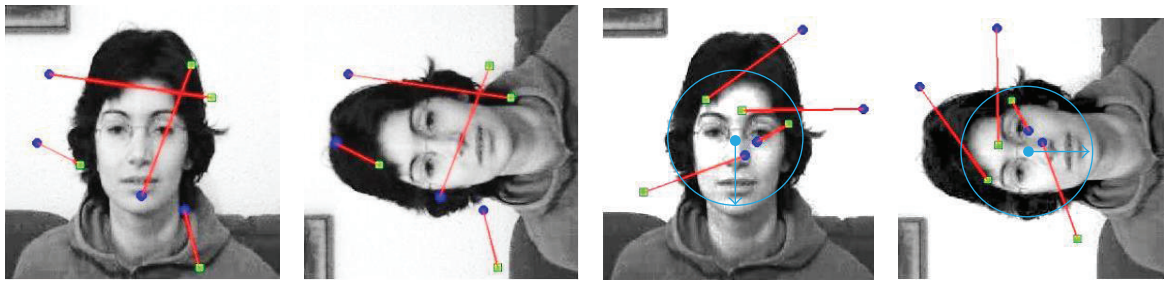
Figures 2(a) and 2(c) are the original image and Figures 2(b) and 2(d) are the result of a 90-degree in-plane rotation. Four random pairwise pixels are connected by the red lines shown in Figure 2. Thus, the LBF in Figure 2(a) is 1010, which is obviously different from Figure 2(b)'s LBF (1100). The reason is that the intensity distributions changed due to the rotation variations and fixed pairwise pixels, as shown in Figures 2(a) and 2(b).

Inspired by SIFT and SURF descriptors [6, 7], we propose a rotation invariant local binary feature (RILBF) based on the maximum gradient orientation (MGO) of the local image region. Therefore, an orientation histogram is formed from the gradient orientations of sample points within the region. For instance, the histograms of gradient (HoG) [21] for the original and rotational images are calculated as shown in Figure 3. Note that the orientation histogram has 72 bins covering the 360-degree range of orientations.

As can be clearly seen from Figure 3, peaks in the orientation histogram correspond to the dominant directions of the local gradients. Therefore, the MGO of the original and rotational images are 270° and 0°, respectively. The pairwise pixels can be changed according to the MGO, which is shown as the cyan arrows originating from the center of the circle in Figures 2(c) and 2(d). Considering the MGO, it is possible to precisely predict where each pairwise pixel in an original image should appear in the transformed image. To correctly measure repeatability and positional accuracy, we define the random pairwise pixels locations in a polar coordinate system, as shown in Figure 4.

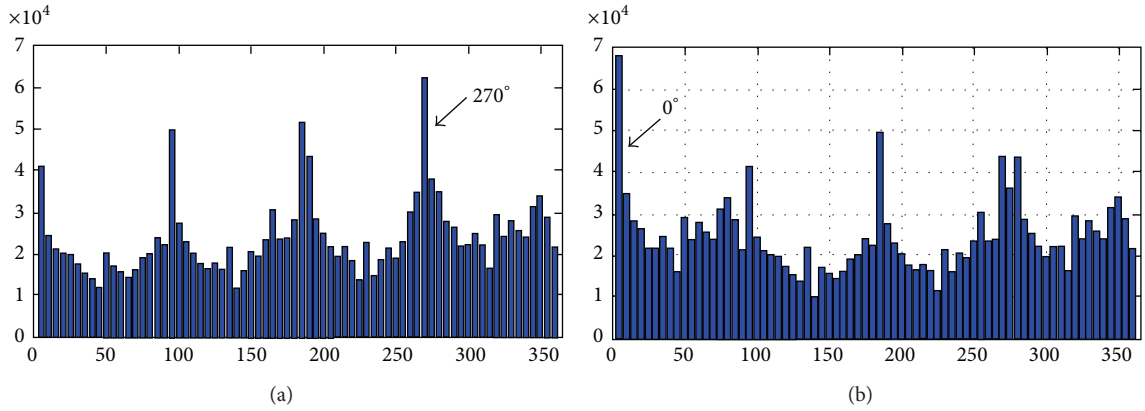
Generally, the polar coordinate system is a 2D coordinate system where each point on a plane is determined by a distance R from a fixed pole and a polar angle θ from a fixed polar axis. The polar coordinates R and θ can be converted to the Cartesian coordinates x and y by using the trigonometric functions sine and cosine:

$$\begin{aligned} x &= R\cos\theta, \\ y &= R\sin\theta. \end{aligned} \tag{3}$$



(a) LBF = {1010} (b) LBF = {1100} (c) RILBF = {0111} (d) RILBF = {0111}

FIGURE 2: The results of LBF and RILBF on rotated images. (a) and (b) show the unequal LBF and (c) and (d) are the equal RILBF, where the orientations are indicated by the arrow from the center of the circle.



(a) (b)

FIGURE 3: The HoG of the original and rotated images. (a) MGO = 270° and (b) MGO = 0°.

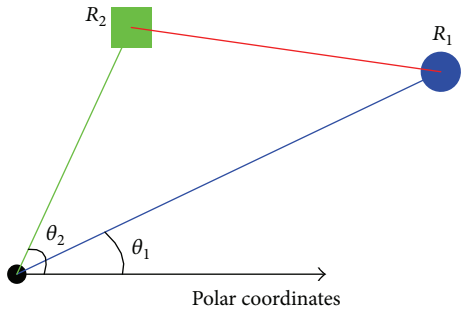


FIGURE 4: The polar coordinate system for pairwise pixels.

To achieve orientation invariance, the pairwise pixels used in (1) can be calculated by the polar coordinates, where gradient orientations are rotated relative to the MGO = θ_m:

$$\begin{aligned} x_1 &= R_1 \cos(\theta_1 + \theta_m), \\ y_1 &= R_1 \sin(\theta_1 + \theta_m), \\ x_2 &= R_2 \cos(\theta_2 + \theta_m), \\ y_2 &= R_2 \sin(\theta_2 + \theta_m). \end{aligned} \tag{4}$$

Note that the fixed pole is located at the center of image and the fixed polar axis has the same direction as MGO, as the cyan arrows shown in Figures 2(c) and 2(d). This allows pairwise pixels to be matched correctly under arbitrary orientation change between the two images. For example, both Figures 2(c) and 2(d) result in the same representation where the RILBF $F(\theta_{m1})$ and $F(\theta_{m2})$ (here θ_{m1} and θ_{m2} are the MGO) are always equal to 0111.

Therefore, by assigning a consistent orientation to each LBF based on local image properties, the RILBF can be represented simply relative to this orientation and therefore achieve invariance to image rotation. Furthermore, by reserving typical features and reducing redundancy features, the generalization performance and training efficiency of the classifier are guaranteed.

3.2. Scale Invariant Scanning Grid Pyramid. To detect the position of the object, the detector scans the input image by a scanning window and, for each patch, determines the presence or absence of the object. The scanning window needs to be resized at different scales because the search often involves comparing objects that have been resized. To handle scale variations in RILBF, we maintain the same polar angle of the pairwise pixels but apply different scale ratios to radius R. Therefore, refer to [7]; the scale space is analyzed by changing the scanning window size rather than iteratively reducing the image size, as in the pyramid structure shown in Figure 5.

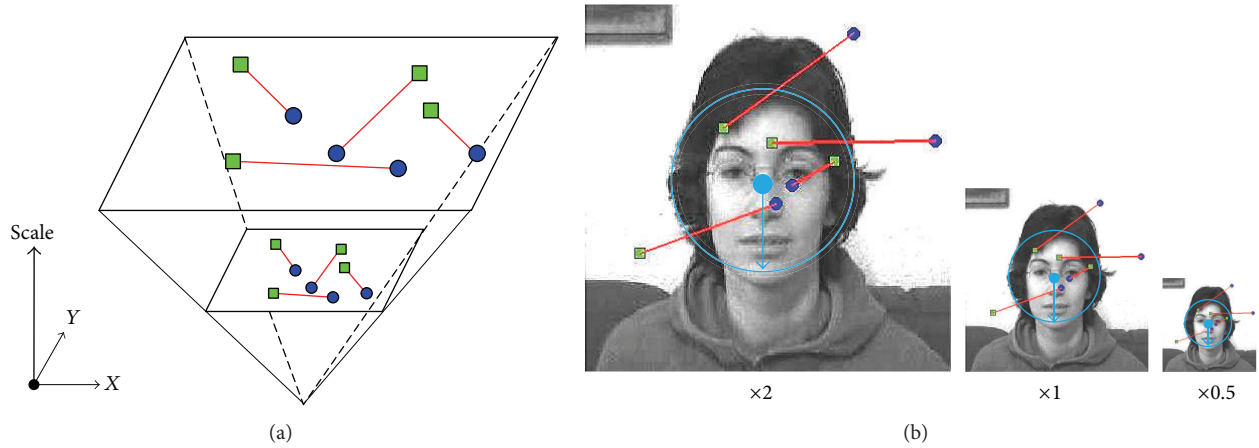


FIGURE 5: The pyramid structure of the scanning window for scale variations. (a) The pyramid local binary features. (b) The pyramid images for RILBF.

As shown in Figure 5, the distributions of the pairwise pixels are changed according to the different scale ratios but are preserved as the same profile. In this paper, the square region of the scanning window is 16 × 16 pixels. This is considered the initial scale layer, which we will refer to as scale $s = 1$. The following layers are obtained by gradually changing the radius R according to the scales $s = 1.2^n$, where n is the index range $[-2, 4]$. We generate all possible shifts of an initial bounding box with the following parameters: horizontal step = 10% of width, vertical step = 10% of height. Therefore, all these scanning windows in different scale ratios make up a scanning grid pyramid, as presented in [41]. Note that, as we do not have to downsample the image, no aliasing occurs as in Gall's work [1].

3.3. Part-Based Object Detection on the Hough Space. Our IHRF consists of a set of random ferns that are trained to learn a mapping from densely sampled RILBF to their corresponding votes in a Hough space. The Hough space encodes the hypothesis h for a part-based object centroid position in different scale space.

Let Γ denote the mapping from the input appearance $I(x, y)$ of the local image patch P_Y, which is represented by a RILBF F centered at Y and where the MGO is θ_{md} (it is obtained during object detection), to the probabilistic Hough vote for the hypothesis h:

$$\bigcup_{x_i, y_i \in P_Y, \theta_{md}} I(x_i, y_i) \xrightarrow{\Gamma} p\left(\frac{h}{F}, \theta_{md}, Y\right). \tag{5}$$

Learning the mapping Γ and using it for part-based object detection are described in Sections 3.3.1 and 3.3.2, respectively.

3.3.1. Training the Random Ferns Classifier. Random ferns are of great interest in the computer vision domain because of their speed, parallelization characteristics, and robustness to noisy training data. They are used for various tasks, such as key-point recognition [3] and image classification [48]. Özuysal et al. [3] argued that the vital element of RFC is the independence of the base ferns, which can be enforced by generating different pixel comparisons from the same image patch, which has been presented in Section 3.1. When they are applied to a large number of input vectors of the same class C, the output of each fern is a frequency distribution histogram, which is shown in Figure 6. In the histogram, the horizontal axis represents a 2^S-dimensional space of binary descriptors, and the vertical axis shows the number of times the binary code appeared in a class C, also called class conditional probability (CCP) $p(F_i/C)$, where $i \in [1, K]$.

Random ferns replace the trees in random forests [49] by nonhierarchical ferns and pool their answers in a naive Bayesian manner to yield better results and improve classification rates in terms of the number of classes. As discussed in Section 3.1, the set of RILBF (F, θ_{mt}) located in a local patch with MGO θ_{mt} (it is obtained during classifier training) is regarded as a class. Thus, a randomly selected patch detected in another image will be assigned to the most likely class $\widehat{C}$ by evaluating the posterior probability:

$$\widehat{C} = \underset{k}{\operatorname{argmax}}\, p\left(\frac{c_k}{F}, \theta_{mt}\right), \tag{6}$$

where $k = 1; 2; \ldots; H$, c_k is the set of classes. and $c_k \in C$. According to seminaive Bayes [50] and (2), (6) is equivalent to a joint CCP for binary representations in each fern as

$$\widehat{C} = \underset{k}{\operatorname{argmax}} \prod_{L=1}^{K} p\left(\frac{(F_L, \theta_{mt})}{c_k}\right). \tag{7}$$

For a given test input, simply apply the binary representations accounting for the ferns and look up the corresponding probability distribution over class label, as shown in Figure 6. Finally, the RFC selects the class with the highest posterior probability as the categorized results. RFC is a remarkable classification algorithm that randomly selects and trains a collection of ferns. Then, classifying new inputs involves only simple look-up operations.

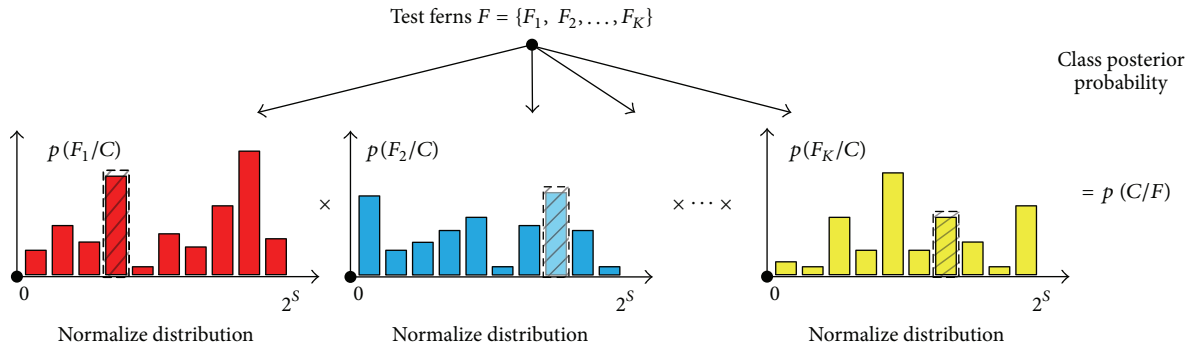


Figure 6: The classification using a RFC, where × is the symbol of multiplication.

For training, a set of training examples is provided for each class C. For each positive class, we assume that an object center within a bounding box including the whole object is provided, and we randomly divide it into several local patches. Then, we need to record the mutual geometric constraints between them. As the green rectangle in Figure 7 shows, the displacement vector d from the object center to the center of a local patch in its polar coordinate system is recorded as

$$d = \{R_C, \theta_C\}, \tag{8}$$

where R_C is the displacement between the two centers and θ_C is the orientation, which is rotated relative to the MGO θ_{mt} of the local patch. We refer to the ISM representation [9] as

$$\mathrm{ISM}(c_o) = \left(p\left(\frac{F}{c_o}\right), \theta_{mt}, d\right). \tag{9}$$

As a result, each fern in the IHRF consists of the ISM of each local patch belonging to the object class c_o. It is noted that the size of an object used for training can be represented by a scale factor $s = 1$. For the negative instances, we simply record their own class labels and the pseudodisplacements.

This allows the classifier to exploit the available training data more efficiently because image patches representing the same object but in a different configuration (i.e., rotated or scaled) can be considered the same types of information. During classification, we do not need to classify multiple scaled and rotated versions of the image, and it results in a lower complexity.

3.3.2. Probabilistic Voting on Hough Space. In the following, we cast the voting procedure into a probabilistic framework. For detection in a scanning patch, the extracted RILBF is passed through every fern in the random ferns, and then the potential object centroid positions stored in the ferns are used to cast votes to the Hough space. According to this,

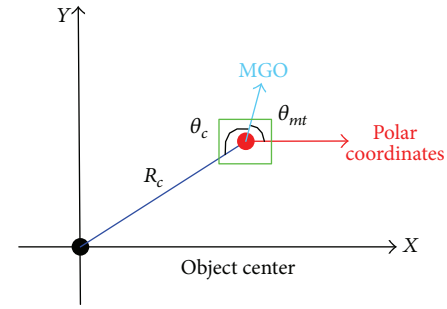


Figure 7: Displacement vector, where the green rectangle is the positive instance. The polar axis of the local patch has the same direction as the image x.

the posterior probability (5) can be simply decomposed as follows:

$$\begin{aligned} &p\left(\frac{h}{F}, \theta_{md}, Y\right) \\ &= p\left(\frac{h(c, X, S_V)}{F}, \theta_{md}, Y\right) \\ &= \sum_{k=1}^{H} p\left(\frac{h(c, X, S_V)}{c_k}, \theta_{md}, F, Y\right) p\left(\frac{c_k}{F}, \theta_{md}, Y\right), \end{aligned} \tag{10}$$

where $h(c, X, S_V)$ is the hypothesis for the object belonging to the class $c \in C$ with position X and scale factor S_V. As shown in (6), we have evaluated the patch's probability independent of their location. In addition, the first term in (10) can be treated as independent of RILBF F because we have mapped the unknown patch to a hypothesis h. Thus, (10) can be reduced to

$$\begin{aligned} &p\left(\frac{h}{F}, \theta_{md}, Y\right) \\ &= \sum_{k=1}^{H} p\left(\frac{h(c, X, S_V)}{c_k}, \theta_{md}, Y\right) p\left(\frac{c_k}{F}, \theta_{md}\right) \\ &= p\left(\frac{h(c, X, S_V)}{c_k} = c, \theta_{md}, Y\right) p\left(\frac{c_k}{F} = c, \theta_{md}\right), \end{aligned} \tag{11}$$

where the first term is the probabilistic Hough vote for an object position X in different scale space S_V, which is based on the class label and its ISM. More specifically, the first term votes for the following object position as follows:

$$\begin{aligned} X_x &= Y_x - S_V X_d, \\ X_y &= Y_y - S_V Y_d, \end{aligned} \tag{12}$$

where the subscripts x and y indicate the image position in the x and y directions, respectively. According to the displacement vector d, which has been stored in the ISM, the connection vector (X_d, Y_d), which is relative to the current position (Y_x, Y_y), can be presented as

$$\begin{aligned} X_d &= R_C \cos\left(\theta_C + \theta_{md}\right), \\ Y_d &= R_C \sin\left(\theta_C + \theta_{md}\right). \end{aligned} \tag{13}$$

When casting votes for the object center (X_x, Y_y), the object scale ratios S_V are selected according to the description in Section 3.2 and are treated as a third dimension in the Hough voting space. Therefore, the distribution of the first term in (11) can be approximated by a sum of Dirac functions δ_d for all the displacement vector set $D = \{d_i\}_{i=1,2,\ldots,N}$ (N is the number of displacement vectors which are obtained after training) as follows:

$$\begin{aligned} &p\left(\frac{h\left(c, X, S_V\right)}{c_k} = c, \theta_{md}, Y\right) \\ &= \frac{1}{N}\sum_{d\in D}\delta_d\left(\frac{Y_x - X_x}{S_V} - X_d, \frac{Y_y - X_y}{S_V} - Y_d\right). \end{aligned} \tag{14}$$

Thus, the vote distribution in (14) is obtained by casting a vote for each stored observation from the learned ISM. For all the ferns, the second term in (11), which is calculated by (7), is averaged as follows:

$$p\left(c_k = \frac{c}{F}, \theta_{md}\right) = \frac{1}{K}\sum_{L=1}^{K} p\left(\frac{\left(F_L, \theta_{md}\right)}{c_k} = c\right). \tag{15}$$

Note that the accumulation of the probabilities in (15) is nonprobabilistic, but the results of summation are preferred over multiplication in (7) because this approach is more stable in practice [1]. To integrate the votes coming from the scanning grid pyramid of the input image Ω, we accumulate them into the Hough image H:

$$\begin{aligned} HI &= p\left(\frac{h}{\Omega}\right) = \sum_{Y\in\Omega} p\left(\frac{h}{F}, \theta_{md}, Y\right) \\ &= \frac{1}{KN}\sum_{Y\in\Omega}\sum_{L=1}^{K}\sum_{d\in D} p\left(\frac{\left(F_L, \theta_{md}\right)}{c_k} = c\right) \\ &\quad \times \delta_d\left(\frac{Y_x - X_x}{S_V} - X_d, \frac{Y_y - X_y}{S_V} - Y_d\right). \end{aligned} \tag{16}$$

As a result, the value $p(h/\Omega)$ serves as a confidence measure for the hypothesis h. After all the votes are cast, a global search for the local maxima obtains the position of the object center as a nonparametric probability density estimate.

3.4. Object Segmentation. By back-projecting the contributing votes, we retrieve the hypothesis's support in the image, which shows the rough profile on the depicted object. However, this is not a precise segmentation yet. Therefore, we propose a segmentation approach to improve recognition again by allowing the system to focus its efforts on object pixels and discard misleading influences from the background.

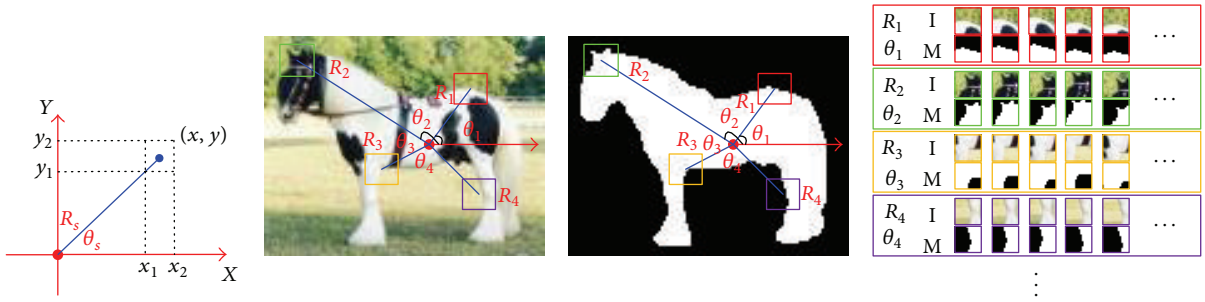
3.4.1. Back-Projection for Object Detection. In addition to the hypothesis voting capabilities, the IHRF can also be applied in reverse to detect the positions of their support. The location of a local maximum in Hough image HI encodes scale, class c_o, and ISM of the object. More specifically, given a local maximum at position S_m, we define the support of the strongest hypothesis as the sample set

$$\begin{aligned} B\left(D_l, S_m\right) &= \bigcup_{l\in D_l}\left\{S_{V,l}, \theta_{md,l}, \mathrm{ISM}_l\left(c_o\right) \mid S_m\right\} \\ &= \bigcup_{l\in D_l}\left\{S_{V,l}, \theta_{md,l}, p_l\left(\frac{F}{c_o}\right), \theta_{mt,l}, d_l \mid S_m\right\}, \end{aligned} \tag{17}$$

which contains the patch entries of all local samples D_l (they are obtained after Hough voting) that have voted for the center position S_m. By using their corresponding voting vectors d_l and MGO $\theta_{md,l}$, we can back-project the original position of samples l onto the image space. In this way, we obtain a sparse point set of positions supposedly belonging to the object that voted for the center position S_m.

3.4.2. Refined Object Segmentation Based on a Clustering Scheme. The back-projected hypothesis's support already provides a rough indication of where the object is in the image. However, the sampled local patches still contain the background structure, as shown in Figure 1(g). We can actually express the a priori known object content without background in terms of a binary mask according to the extracted image patch—for example, see Figure 1(e). Thus, we know more about the pure interpretation of the matched patches for the target object and use it to segment the object from the background by a clustering scheme.

To produce this top-down segmentation, our approach clusters the local masks according to the spatial distribution as the first constraint shown in Figure 8(a), where the coordinates' origins demonstrate the object center and we generate uniformly spaced coordinates with a special interval Δ in both the x and y directions. Then we group the local training patches, with positions $\{x, y\}$ that belong to the $([x_1 x_2], [y_1 y_2]$, where $x_2 - x_1 = \Delta$, $y_2 - y_1 = \Delta)$ range, as a cluster. Specifically, we use a polar coordinate $\{R, \theta\}$ to satisfy the rotation requirement in IHRF, which has a polar axis that overlaps the x-axis. As a result, the image and mask information can be clustered and refer to the training patches' polar parameters $\{R_i, \theta_i\}$, as shown in Figures 8(b)–8(d).



(a) (b) Image (c) Mask (d)

FIGURE 8: The clustering scheme based on the spatial constraint and appearance similarity. (a) is the spatial constraint; (b) and (c) are the image and its mask, respectively; and (d) is the clusters for object with the two constraints.

The second constraint is the appearance similarity for matching of the support's patch, which is estimated in a similarity metric as

$$D(C_1, C_2) = \frac{\sum_{p \in C_1, q \in C_2} \mathrm{NCC}(p, q)}{|C_1| \times |C_2|} > T, \tag{18}$$

where C_1 and C_2 are the two inquiring patches, p and q are the color or grayscale values, T is the similarity threshold, and $|*|$ represents the number of pixels in the matching patch. In addition, NCC is the similarity between two patches and is measured by a normalized correlation coefficient, defined as

$$\mathrm{NCC}(p, q) = \frac{\sum_i (p_i - \bar{p}_i)(q_i - \bar{q}_i)}{\sqrt{\sum_i (p_i - \bar{p}_i)^2 \sum_i (q_i - \bar{q}_i)^2}}. \tag{19}$$

In this paper, the object's appearance is represented by a 16×16 normalized image patch, which is resampled from an image within the original object bounding box regardless of the aspect ratio.

Our clustering scheme can be split into two stages: storing the patches and their masks using the spatial constraint and matching the similar patch using similarity metric. This approach guarantees that only the grouped patches are spatially and visually similar and provide the binary mask to segment the foreground from the background, which is evident from the refined results shown in Figures 1(h) and 1(i).

Furthermore, the spatial and appearance similarity can perform rotation invariance according to the IHRF and (17). Thus, the invariant cluster indexes of the inquiring patch can be looked up from

$$\theta_S = \begin{cases} \theta_{mt} + \theta_C - 180^\circ, & \text{in the 1th or 2th quadrant,} \\ \theta_{mt} + \theta_C + 180^\circ, & \text{in the 3th or 4th quadrant,} \end{cases} \tag{20}$$

while R_s is still equal to R_C. Then, the local patch should be rotated by the MGO deviation angle $\Delta\theta = \theta_{md} - \theta_{mt}$, after matching by (18) and finding the local mask by the similarity, which is above T. Finally, we rotate the binary mask by the inverse angle $\Delta\theta$ to segment the object more precisely without the effects of rotation.

To this end, the training data is segmented by the two constraints, and the local foreground-background masks are stored at the same time. When a maximum is detected in the voting space, the local segmentation masks are used to infer a global segmentation for the detection in the image.

4. IHRF Based Online Object Tracking

Up to now, we have defined all the parts that are necessary to perform object detection and segmentation in an IHRF framework. We can extend this method to handle the online tracking task as well. Recently, online learning frameworks have been designed for long-term tracking of an unknown moving object. The key defining characteristic of online learning is that it can use current true label feedback to update its hypothesis for future predictions, which are close to the true labels. This framework is able to adapt and learn in difficult changing situations because of its continual feedback and update.

During long-term tracking, the main challenge is to avoid drifting problems while still being adaptive to significant occlusions, scale variations, and changes in the object's appearance and deformation. Fine segmentation delivers a more precise description of the object and is used to decrease the effect from the background in the online learning stage [15]. As long as an appropriate classifier exists, the online learning framework will learn to predict correct labels. Therefore, the key point is to use the precise segmentation of a moving object, which has been produced in Section 3.4. We then use this segmentation (essentially a binary mask) to accurately update our classifier, which allows learning of extensive object variations during tracking.

In this paper, the block diagram for online object tracking is illustrated in Figure 9. The main components of the model can be characterized as follows.

(1) The RILBF and scanning pyramid element performs full scanning of the local patches to represent object and background appearance. It provides rotation and scale invariance and discriminable features to significantly increase the classification accuracy.

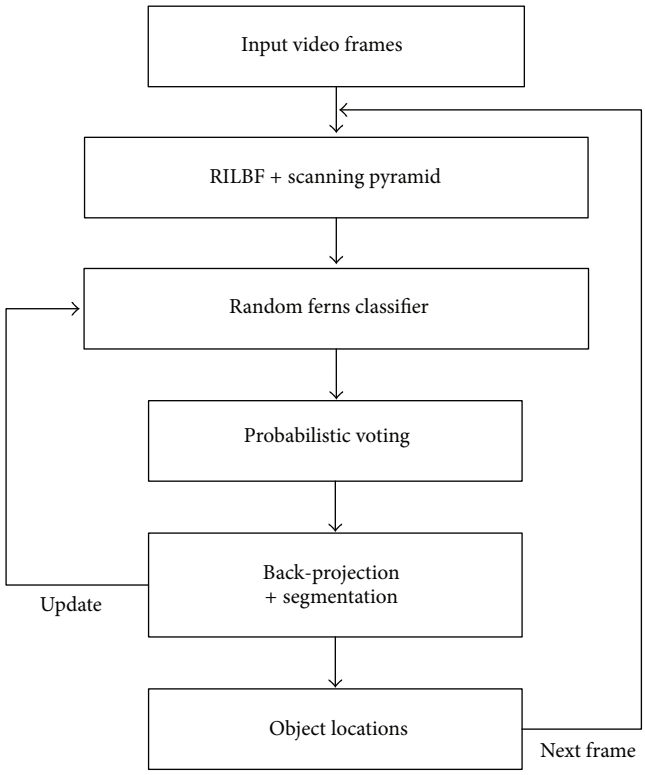


FIGURE 9: Block diagram of the online learning framework based on IHRF for object tracking.

(2) The RFC decides about the presence or absence of the object. Once an unknown template occurs, it will bring this new variation for retraining the classifier.

(3) The probabilistic voting model estimates the object's locations under the local maxima assumption that encodes the hypothesis's scale and ISM information.

(4) The back-projection and segmentation model integrates both of the hypothesis's support and binary mask. It provides a precise object mask without background noise to retrain the classifier and extract the object features in the next frame.

The implementation of these components was described as an IHRF in Section 3.

Based on the block diagram (see Figure 9), a processing flowchart for online object tracking can be summarized as follows.

Step 1 (initialization in the first frame). It initializes the object RILBF as a positive sample defined by the user in the first frame and then trains the original RFC associated with some random selected background representation as a negative sample.

Step 2 (a cyclic thread for tracking in the remaining video stream). It uses the IHRF to detect the object by a bounding box and estimate its motion between consecutive frames. In this model, a motion constraint, such as a temporal and spatial structure [41], restricts the potential states of the object.

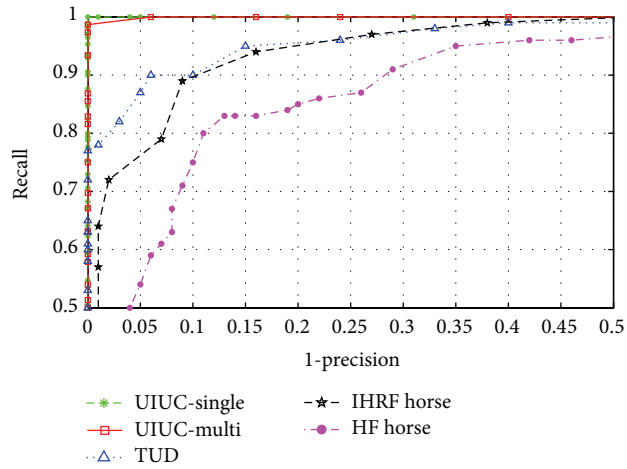


FIGURE 10: RPC on four datasets. All curves are generated by IHRF, except that of HF Horse, which is produced by a Hough Forest approach on Weizmann Horse.

Step 3 (providing an update and mask procedure in each frame). According to the motion constraints and the precise binary mask, it distinguishes the object's appearance from the background and updates the classifier for all of the variances.

5. Experiments and Results

In this experimental section, we present two types of comparative experiments: object detection and object tracking, which are presented in Sections 5.1 and 5.2, respectively. In order to evaluate our method's performance and compare it with state-of-the-art approaches, we apply these methods to several challenging datasets. We also adhere to the experimental protocols and detection accuracy criteria established for each of the datasets in previous works. All experiments have been done on a standard 3.2 GHz PC with 2 Gigabytes of RAM.

5.1. Object Detection. For object detection, the settings were as follows: the RFC consists of $K = 10$ ferns and we pick $S = 13$ pairwise pixels for RILBF. In a multiscale setting, seven scale ratios $S_V = 1.2^n$, where n is the index range $[-2, 4]$, were used to handle the variety of scales in the test data. The interval $\Delta = 20$ pixels and the similarity threshold $T = 0.5$ are used for clustering.

According to previous works, we evaluated the IHRF on several challenging datasets including UIUC cars, TUD pedestrians, and Weizmann Horse. The recall-precision curve (RPC) [51] (see Figure 10) is generated by changing the probability threshold on the hypotheses vote strength. In Table 1, we also provide a performance comparison with the best previously published results [1]. Obviously, IHRF outperforms the previous methods and achieves the best results.

Table 1: Performance of different methods on the four datasets at recall-precision equal error rate (EER).

Methods	UIUC-Single	UIUC-Multi	TUD	Horse
ISM [9]	97.5%	95%	80%	—
Efficient Subwindow Search [22]	98.5%	98.6%	—	—
Hough Forest (HF) [1]	98.5%	98.6%	86.5%	83%
Mutch and Lowe [52]	99.9%	90.6%	—	—
IHRF	100%	98.7%	90%	89%

Table 2: Comparison of the two methods on the three datasets using SR and ALA.

	Methods	UIUC-Single	UIUC-Multi	TUD
SR	HF	95%	91%	95%
	IHRF	95%	95%	98%
ALA	HF	0.80	0.70	0.76
	IHRF	0.85	0.75	0.78

To define the performance more precisely, we compare the object position with ground truth using two evaluation protocols based on bounding-box overlap [36].

(1) Successful rate (SR) is equal to the number of correct positions divided by the number of test images. The correct position means that the overlap score OS = $A \cap B/A \cup B$ between the bounding box of detection and its ground truth is larger than 50 percent.

(2) Average localization accuracy (ALA) is an average overlap score calculated from all the test images.

The overlap score results on the three datasets are compared for the IHRF and HF approaches, as illustrated in Figure 11, where the vertical and horizontal axis are the overlap score and corresponding image number, respectively. Note that each image in these datasets contains only one object since we only focus on single object detection in this work.

As a result, the quantitative evaluations compared with Hough Forest confirm that IHRF performs well on both ALA and SR, as demonstrated in Table 2. More details are discussed as follows.

5.1.1. UIUC Cars. The UIUC car dataset [53] contains two types of car images. The first is the UIUC single-scale (UIUC-Single) test set, which consists of 170 images containing side views of cars of approximately the same size. Another is the UIUC multiscale (UIUC-Multi) test set, which consists of 108 images containing car side views at multiple scales. We trained the IHRF using the available 400 positive and 400 negative training images.

Our IHRF approach achieved an impressive 100% EER for UIUC-Single and 98.7% EER for UIUC-Multi, thus exactly outperforming the state-of-the-art performance. Table 1 also shows that the IHRF considerably outperformed the Hough-based ISM approach [9] and Efficient Subwindow Search approach [22] as well as the Mutch and Lowe's method [52]. In addition, our method is both simpler and more powerful than Hough Forest [1] since naive Bayesian scheme in ferns outperforms the averaging of posteriors used to combine the output of the decision trees [3].

As the overlap score results show in Figures 11(a) and 11(b), the Hough Forest approach is sometimes equal to zero since it is more vulnerable to missing object because of the cluttered background or the multiple scales influences. However, our method can still detect the object with a tolerant error. Table 2 further confirms that our method (SR = 95%, ALA = 0.85 for UIUC-Single and SR = 95%, ALA = 0.75 for UIUC-Multi) slightly outperforms Hough Forest on both SR and ALA evaluation scheme.

Some examples of those detection cases are displayed in Figure 12. In this experiment, the results show that the IHRF not only detects object despite partial occlusion but also is often even able to deal with the scale variations. For an image from the UIUC car dataset, our method only requires 0.4 seconds (no less than 200×150 pixel resolution).

Both the datasets include samples of partially occluded cars, cars with low contrast within the cluttered background, and challenging illumination. However, the shape of the objects remains rigid, which makes the detection task easier. Therefore, we will assess the performance of our method on more challenging datasets that include highly nonrigid object transformations as follows.

5.1.2. TUD Pedestrian. In this section, we apply our approach to pedestrian detection in crowded street scenes using the TUD pedestrian dataset. This is a highly challenging test set that consists of 400 training images containing crowded street scenes [54]. For this experiment, we followed the experimental protocol of [47] to train our detector and tested it on 100 pedestrian images.

The performance of the different methods is shown in Table 1. For TUD pedestrians, our method performed EER = 90%, which is significantly better than both the Hough Forest [1] and ISM-based methods [9]. Figure 11(c) and Table 2 also demonstrate that our method (SR = 98%, ALA = 0.78) is still competitive and performs better than Hough Forest approach.

In order to give a better impression of our method's performance, Figure 13 shows obtained detection results on example images from the test set. As can be seen from those examples, the proposed method can reliably detect and localize such deformable classes as pedestrians in crowded scenes and with severe overlaps. Our method requires no more than 1 second for a 400×320 pixel image from the TUD dataset.

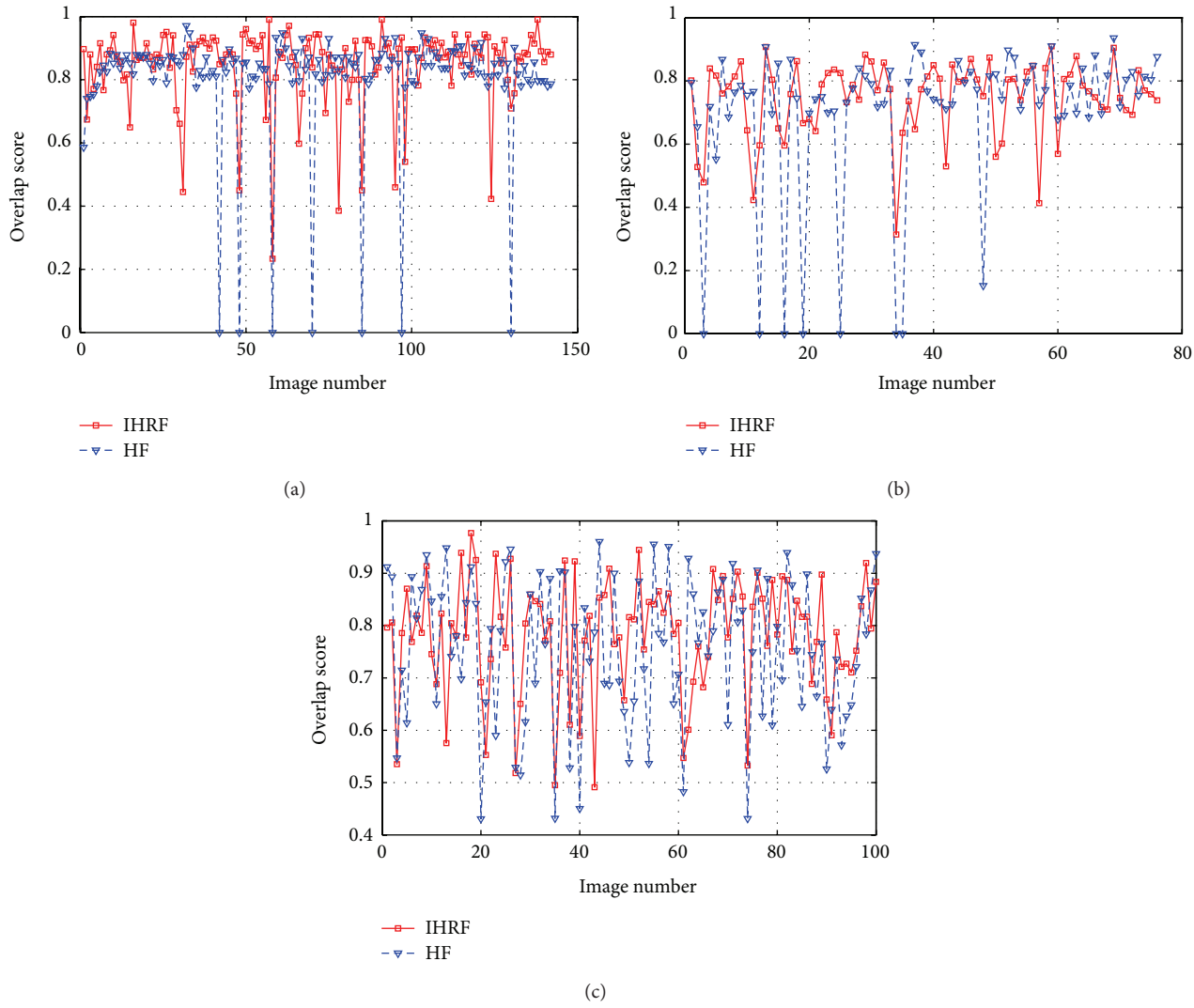


Figure 11: The overlap score results of detecting a single object with IHRF and HF on the three datasets. (a) UIUC-Single, (b) UIUC-Multi, and (c) TUD.

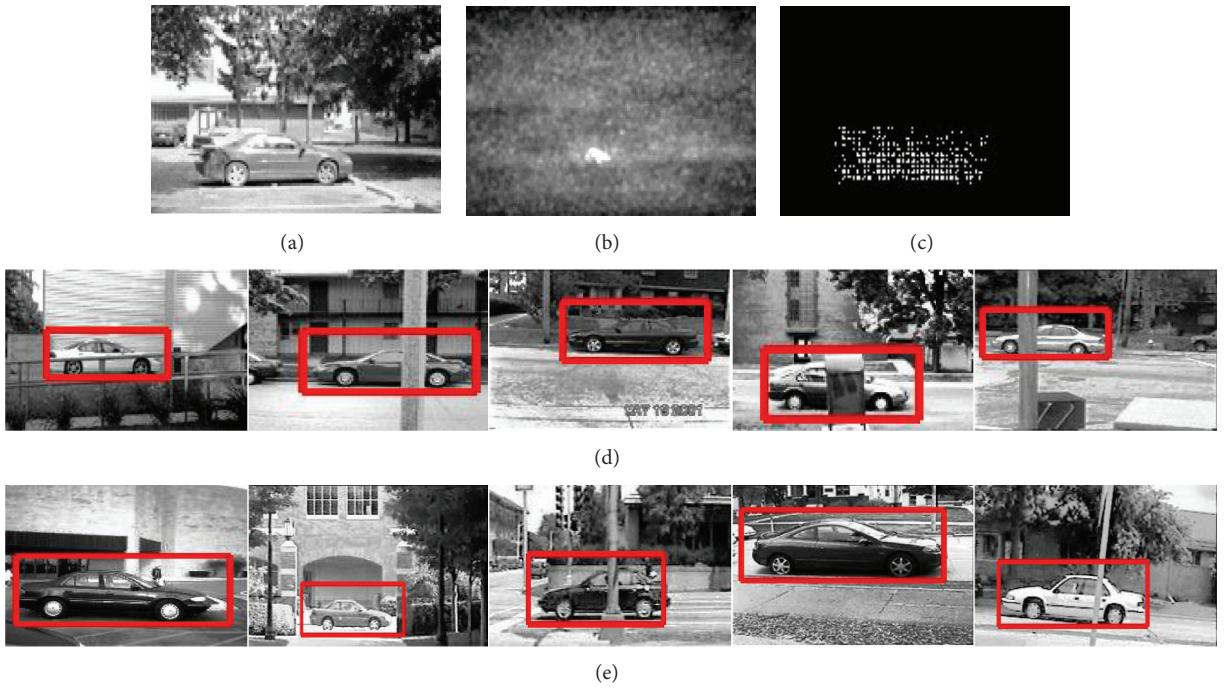
5.1.3. Weizmann Horse. Finally, we also assessed the IHRF and Hough Forests performance on the Weizmann Horses dataset [46], which comprises 328 multiscale side views of horses in cluttered environments under varying scale and strongly varying poses. We split the training testing as suggested in [1] by using 100 horse images and 100 background images for training and the rest of 228 horse images and 228 background images for testing.

Figure 10 shows the RPC, and Table 1 shows a comparison of our method's EER obtained by IHRF and Hough Forest on the same Horse dataset. As can be seen from those results, our method achieves good detection results with an EER performance of 89%, which presents a significant improvement over previous results. Our detection results are shown in Figure 14. Note that the training and testing stages adhere to the original horse images without two improvements, as described in [47].

5.2. Online Object Tracking. For the online object tracking task, we first separate the moving object from the background on a more fine-grained level to obtain more accurate training data. In fact, this is a stable and effective way to avoid the drifting problem. Therefore in this section, the experiments are divided into two parts. First, we perform segmentation experiments demonstrating three specific properties of our approach, and second we present results on public available sequences for comparison with other tracking approaches.

5.2.1. Refined Segmentation. The goal of this section is to illustrate three properties of the proposed segmentation approach: rotation and scale invariance, occlusion, and deformation handling capability. Therefore, multiple challenging datasets have been provided to verify our method's performance.

(1) Rotation and Scale Invariant Performance. We have presented a new object detection and segmentation approach based on IHRF to detect objects that may appear in the image under different orientations and scales. In contrast to other works that address this problem using multiple

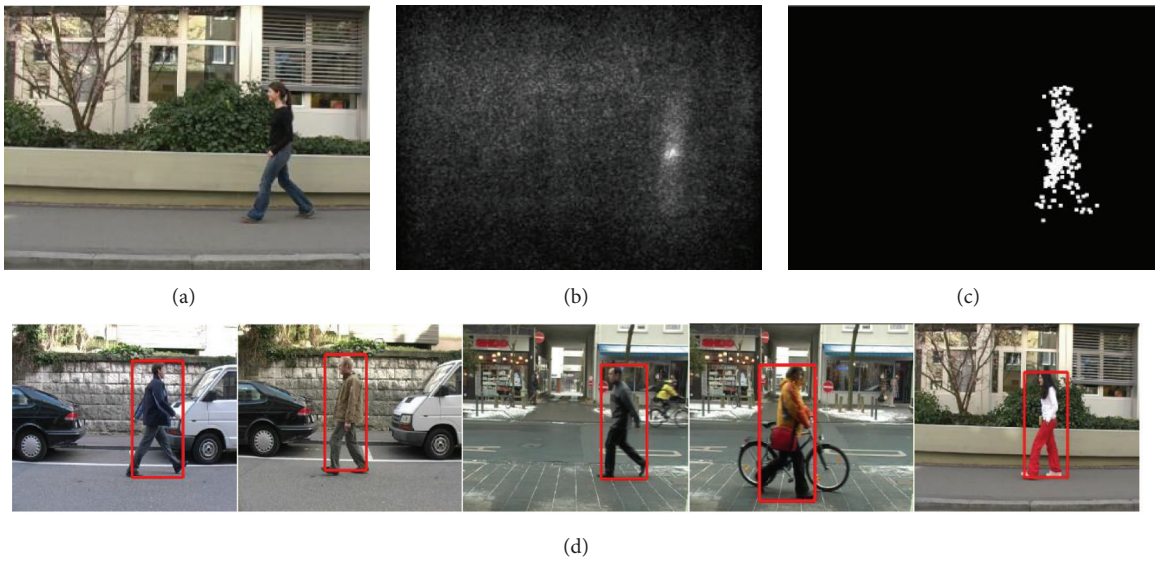


(a) (b) (c)

(d)

(e)

Figure 12: Object detection obtained by an offline trained IHRF for cars. (a) Original image; (b) Hough voting image; (c) support of the strongest hypothesis. The detection results on UIUC car datasets, such as (d) UIUC-Single and (e) UIUC-Multi.



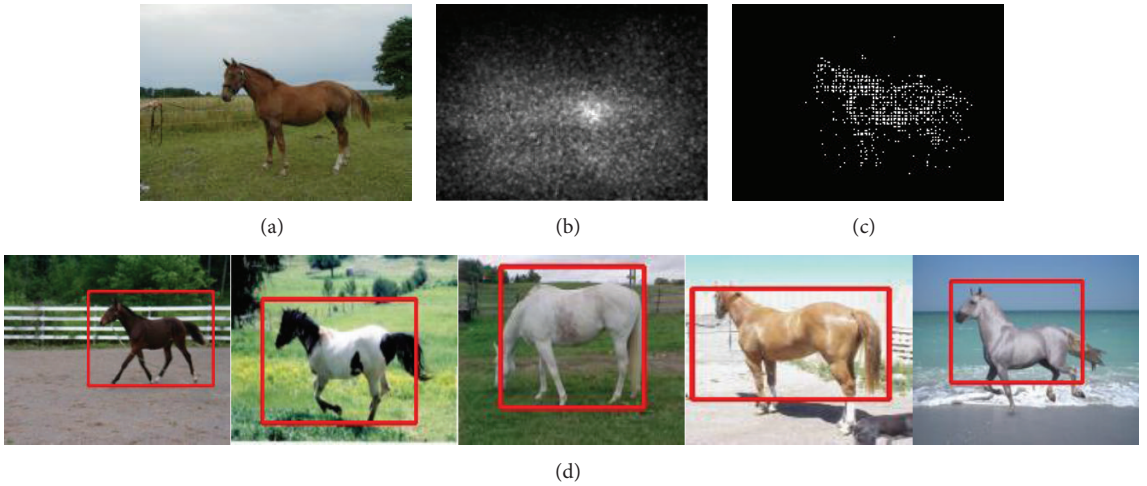
(a) (b) (c)

(d)

Figure 13: Object detection obtained by an offline trained IHRF for pedestrians. (a) Original image; (b) Hough voting image; (c) support of the strongest hypothesis; (d) some detection results on TUD pedestrian dataset.

classifiers or multiple orientation samples, we assign a consistent orientation (MGO) to each patch based on local image properties, and then the RILBF descriptor can be represented relative to this orientation and achieve invariance to image rotation. We also use a rotation invariant back-projection to locate the support of our detection and guide a refined segmentation process that precisely separates the object from the background. In addition, the scanning windows are resized at different scales since we maintain the same polar angle of the pairwise pixels but apply different scale ratios to their radius R.

The following experiment is used to demonstrate the IHRF's capability of rotation invariance. We train our IHRF on our own training image and a hand-segmented face mask; see Figure 15(a). Then we sample the image under 2D rotations in 30-degree steps that leads to 12 samples. Figure 15(b) depicts their segmentation results using the IHRF. As can be seen from Figure 15(b), the IRHF approach



(a) (b) (c)

(d)

Figure 14: Object detection obtained by an offline trained IHRF for horses. (a) Original image; (b) Hough voting image; (c) support of the strongest hypothesis; (d) some detection results on Weizmann Horse dataset.

can adapt to the face rotation variation and yields a stable profile regardless of the rotation. Similarly using the IHRF, the localization of face in different scales can be observed in Figure 15(c).

(2) Occlusion and Deformation Handling Performance. In the domain of part-based object detection, detectors can always improve detection results for partial or self-occlusion and nonrigid deformations. We have demonstrated that the IHRF voting and segmenting models allow us to detect objects reliably even under partial occlusions and heavy nonrigid deformation. Therefore, we process some example detections in occlusion and deformation configurations and the corresponding top-down segmentations can be seen in Figure 16.

Figures 16(a) and 16(b) show human faces that are occluded by a book or a hat [42, 55]. IHRF is highly robust to facial occlusion because a large number of small patches without occlusion still have a high probability of voting for the supporting points. Those results confirm that our method still works in the presence of occlusion and cluttered backgrounds. Figure 16(c) depicts a walker with limited pose variations. In this case, several parts of the object maintain a stable geometric configuration vote for the center of the object. This clearly shows that our approach delivers reliable results even if the object undergoes heavy deformations in a complex background.

5.2.2. Online Tracking. This section reports on a set of quantitative experiments comparing our system (denoted as IHRFT) with other relevant algorithms, which include online boosting (OLB) [56], tracking-learning-detection (TLD) [41], online learning tracker (OLT) [57], and Hough-based GrabCut tracker (HGT) [15].

Unfortunately, no public framework is available for comparing tracking techniques. Hence, we decided to process publically available sequences, as shown in Table 3. The first two experiments (David and Girl) evaluate our system on face-tracking sequences that are commonly used in the literature [58, 59]. In both of these experiments, a human face is moving with challenging conditions such as lighting, scale, and pose changes. In addition, the Coke Can and Tiger sequences [42] contain frequent occlusions and fast motion as well as challenging out-of-plane rotations, cluttered backgrounds, changing illumination conditions, and partial occlusions. For the purpose of tracking deformed object during runtime, we have also collected three more challenging videos (Bike, Motor [15], and Diving [60]) that show different ranges of complexity and highly nonrigid deformations.

Using these datasets, each tracking task has been initialized by manually marking the target object in the first frame, and each tracker, respectively, tracks the object until the end of a sequence. In order to illustrate the performances of the trackers, the produced trajectory is then compared with ground truth using two evaluation protocols: SR and ALA, which have been adopted in Section 5.1. The parameters of our system in this experiment are the same as those defined in Section 5.1. Note that this parameter has been set empirically and its value is not critical. The overlap score results in each sequence are compared with our system and other approaches on SR (Table 3) and ALA (Table 4).

In Table 3, the results in the first two rows show that the four approaches perform well on saturated sequences, except for the OLB. However, the results for Coke Can and Tiger show that HGT may be more vulnerable to missing object in cases with cluttered backgrounds. Furthermore, the TlD and OLT show poor performance in the last three datasets because they utilize a bounding-box-based tracker that is not designed to cope with the amount of deformations in these videos. In particular, when tracking a diving woman, HGT cannot adapt to the gradual nonrigid deformations since the GrabCut segmentation algorithm fails when there are similar colors in the background [15]. In contrast, the IHRFT is not only capable of robustly tracking moving objects of interest through all these challenging sequences but also superior

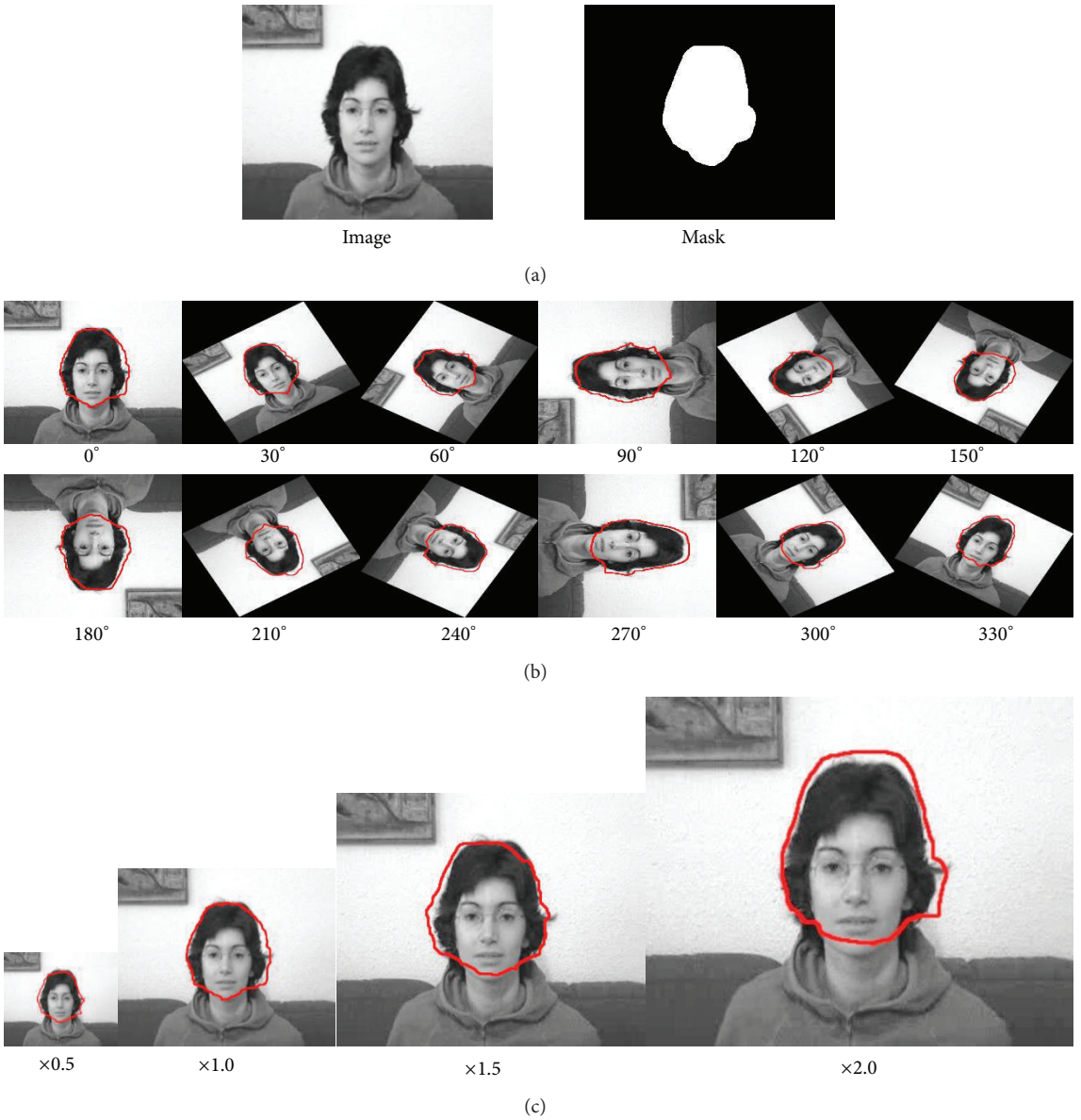


FIGURE 15: Rotation and scale invariance. (a) The training image and mask. Segmentation results under different (b) orientations and (c) scale factors.

TABLE 3: Comparison of the tracker on SR.

Sequence	OLB	TLD	OLT	HGT	IHRFT
David	24%	**100%**	98.7%	98.3%	**100%**
Girl	24.5%	91.7%	86.1%	86.6%	**97.1%**
Coke	90.7%	81.3%	92%	27.3%	**99%**
Tiger	44.8%	88.7%	89.6%	49.2%	**98.6%**
Bike	94%	74.5%	80.2%	99.1%	**100%**
Motor	41.1%	16%	50.3%	**97.6%**	93.2%
Diving	16.3%	22%	24%	34.9%	**71.5%**
Mean	47.9%	67.7%	74.4%	70.4%	**94.2%**

Bold font indicates the best overlap score obtained by one of the trackers in each video.

Table 4: Comparison of the tracker on ALA.

Sequence	OLB	TLD	OLT	HGT	IHRFT
David	0.40	**0.83**	0.79	0.81	**0.83**
Girl	0.35	0.71	0.72	0.67	**0.80**
Coke	0.66	0.59	0.61	0.46	**0.81**
Tiger	0.35	0.68	0.74	0.36	**0.77**
Bike	0.64	0.53	0.55	0.82	**0.86**
Motor	0.26	0.60	0.59	0.61	**0.75**
Diving	0.22	0.27	0.28	0.21	**0.63**
Mean	0.41	0.60	0.61	0.56	**0.78**

Bold font indicates the best accuracy obtained by one of the trackers in each video.

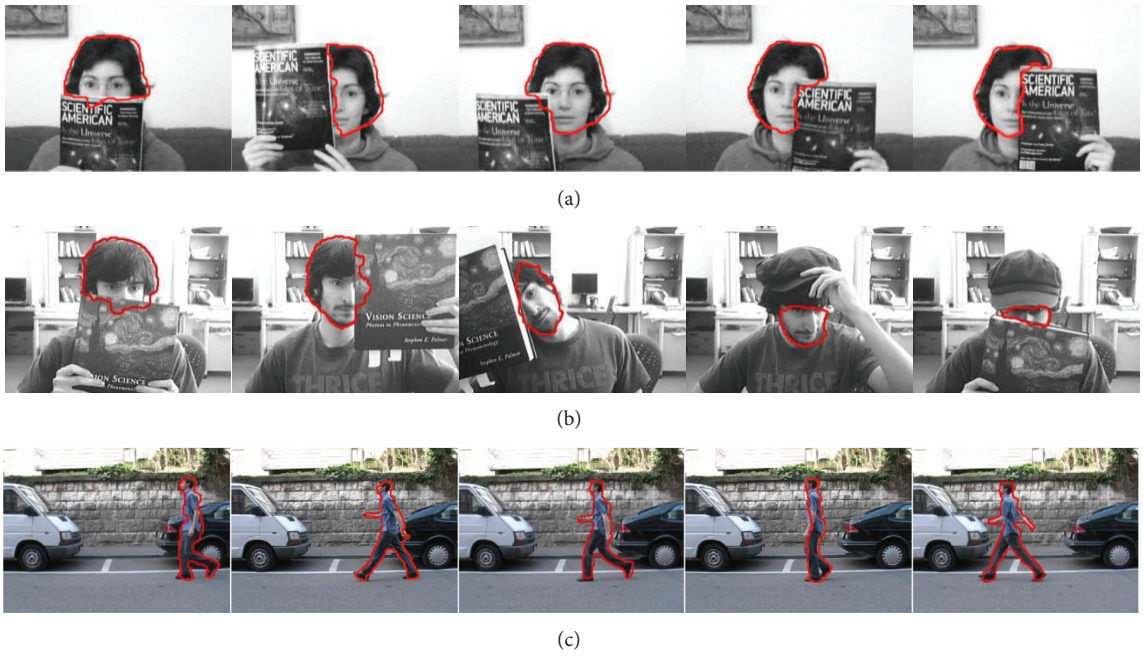


Figure 16: Object segmentation under occlusion and deformation configurations. (a) Face occlusion 1 [55]. (b) Face occlusion 2 [42]. (c) TUD pedestrian.

to the other approaches. Specifically, for the performances measured by SR, our average score outperforms the second best tracker by more than 19.8 percent.

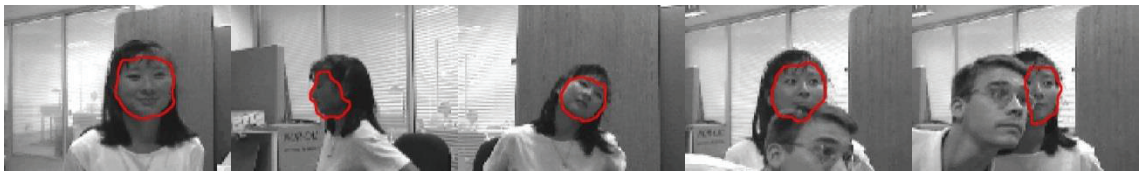
In addition, the quantitative evaluations (in Table 4) compared with other approaches on ALA confirm that IHRFT achieves the best scores in all the sequences, where the average score is 17 percent higher than the second best results. These successful achievements rely on a highly accurate segmentation algorithm. Figure 17 shows some selected frames of the sequences and our tracking results. Even though the challenging conditions present in the sequences, such as cluttered backgrounds, changing illumination conditions, partial occlusions, and non-rigid deformations, the segmentations are still reliable and can serve as a basis for later update stages to further improve the classifier's performance.

6. Conclusions

We have proposed an IHRF approach for part-based object detection and online tracking. It relies on rotation and scale invariant descriptors based on RFC that are able to cast probabilistic votes within the Hough transform framework. The matching operations are simply performed on local patches that have been transformed relative to their assigned orientation, scale, and location, thereby providing invariance to these transformations. Such IHRF can be efficiently used

Frame 11 Frame 108 Frame 165 Frame 266 Frame 388

(a)

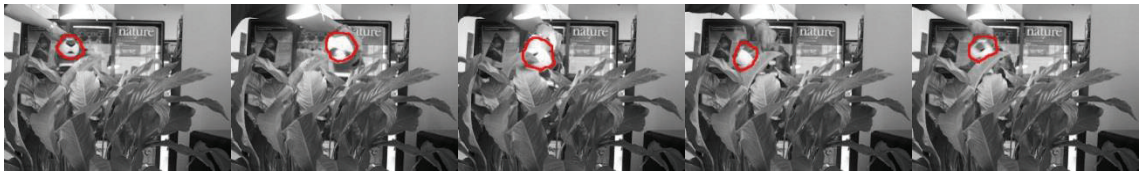


Frame 15 Frame 85 Frame 228 Frame 355 Frame 370

(b)

Frame 13 Frame 32 Frame 95 Frame 182 Frame 270

(c)

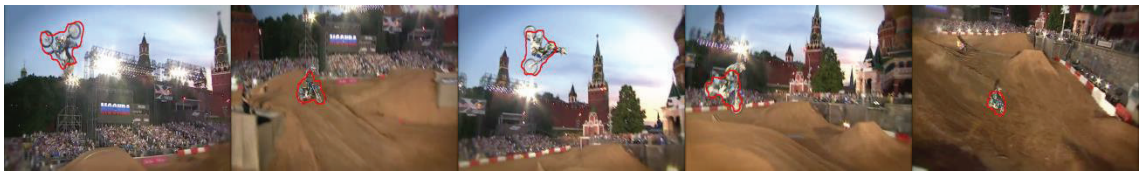


Frame 15 Frame 122 Frame 264 Frame 350 Frame 363

(d)

Frame 5 Frame 95 Frame 125 Frame 170 Frame 216

(e)



Frame 3 Frame 37 Frame 106 Frame 119 Frame 152

(f)

Frame 9 Frame 88 Frame 167 Frame 195 Frame 212

(g)

Figure 17: Snapshots of tracking results.

to detect instances of classes in large challenging images with an accuracy that is superior to previous methods. This approach also allows for time-efficient and space-saving implementation compared with related techniques.

In addition, based on the hypothesis's support determined by the back-projection, we have provided an efficient clustering scheme to guide a segmentation process which precisely separates the object from the background. This top-down segmentation delivers a more precise description of the object and is used to decrease the noise in the online learning stage for object tracking. Therefore, our online tracking method has been validated using several datasets under challenging conditions, such as cluttered background, partial occlusions, and nonrigid deformations. The tracking results show our tracker achieves good object location and segmentation performance in difficult real-world scenes and outperforms state-of-the-art methods.

Our approach applies the RFC to build IHRF which are then used for the object detection. The success of this achievement is highly conditioned on the IHRF performance in 2D images. Therefore, in future work, an interesting extension would be to apply the IHRF to the problem of 3D shape recognition and registration.

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 Open Project of Beijing Key Laboratory on Measurement and Control of Mechanical and Electrical System (no. KF20121123206) and Key Laboratory of Modern Measurement and Control Technology (BISTU), Ministry of Education; Program for Changjiang Scholars and Innovative Research Team in University (PCSIRT, the IRT1212); Funding Project for Academic Human Resources Development Institutions of Higher Learning under the Jurisdiction of Beijing Municipality (no. PHR201106130); and Funding Project of Beijing Municipal Science & Technology Commission (no. Z121100001612011).

References

[1] J. Gall, A. Yao, N. Razavi, L. van Gool, and V. Lempitsky, "Hough forests for object detection, tracking, and action recognition," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 33, no. 11, pp. 2188–2202, 2011.

[2] B. Leibe, K. Schindler, N. Cornelis, and L. van Gool, "Coupled object detection and tracking from static cameras and moving vehicles," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 30, no. 10, pp. 1683–1698, 2008.

[3] M. Özuysal, M. Calonder, V. Lepetit, and P. Fua, "Fast keypoint recognition using random ferns," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 32, no. 3, pp. 448–461, 2010.

[4] C. Huang, H. Ai, Y. Li, and S. Lao, "Vector boosting for rotation invariant multi-view face detection," in *Proceedings of the 10th IEEE International Conference on Computer Vision (ICCV '05)*, vol. 1, pp. 446–453, October 2005.

[5] M. Villamizar, F. Moreno-Noguer, J. Andrade-Cetto, and A. Sanfeliu, "Efficient rotation invariant object detection using boosted random ferns," in *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR '10)*, pp. 1038–1045, June 2010.

[6] D. G. Lowe, "Distinctive image features from scale-invariant keypoints," *International Journal of Computer Vision*, vol. 60, no. 2, pp. 91–110, 2004.

[7] H. Bay, A. Ess, T. Tuytelaars, and L. van Gool, "Speeded-up robust features (SURF)," *Computer Vision and Image Understanding*, vol. 110, no. 3, pp. 346–359, 2008.

[8] R. Marée, P. Geurts, J. Piater, and L. Wehenkel, "Random subwindows for robust image classification," in *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR '05)*, vol. 1, pp. 34–40, June 2005.

[9] B. Leibe, A. Leonardis, and B. Schiele, "Robust object detection with interleaved categorization and segmentation," *International Journal of Computer Vision*, vol. 77, no. 1–3, pp. 259–289, 2008.

[10] S. Maji and J. Malik, "Object detection using a max-margin hough transform," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR '09)*, pp. 1038–1045, June 2009.

[11] O. Barinova, V. Lempitsky, and P. Kholi, "On detection of multiple object instances using Hough Transforms," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 34, no. 9, pp. 1773–1784, 2012.

[12] J. Shotton, A. Blake, and R. Cipolla, "Multiscale categorical object recognition using contour fragments," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 30, no. 7, pp. 1270–1281, 2008.

[13] D. H. Ballard, "Generalizing the Hough transform to detect arbitrary shapes," *Pattern Recognition*, vol. 13, no. 2, pp. 111–122, 1981.

[14] R. Okada, "Discriminative generalized hough transform for object detection," in *Proceedings of the 12th International Conference on Computer Vision (ICCV '09)*, pp. 2000–2005, October 2009.

[15] M. Godec, P. M. Roth, and H. Bischof, "Hough-based tracking of non-rigid objects," *Computer Vision and Image Understanding*, vol. 117, pp. 1245–1256, 2012.

[16] M. D. Breitenstein, F. Reichlin, B. Leibe, E. Koller-Meier, and L. van Gool, "Online multiperson tracking-by-detection from a single, uncalibrated camera," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 33, no. 9, pp. 1820–1833, 2011.

[17] Z. Kalal, J. Matas, and K. Mikolajczyk, "Pn learning: bootstrapping binary classifiers by structural constraints," in *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR '10)*, pp. 49–56, usa, June 2010.

[18] I. Matthews, T. Ishikawa, and S. Baker, "The template update problem," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 26, no. 6, pp. 810–815, 2004.

[19] C. Galleguillos and S. Belongie, "Context based object categorization: a critical survey," *Computer Vision and Image Understanding*, vol. 114, no. 6, pp. 712–722, 2010.

[20] P. Viola and M. J. Jones, "Robust real-time face detection," *International Journal of Computer Vision*, vol. 57, no. 2, pp. 137–154, 2004.

[21] N. Dalal and B. Triggs, "Histograms of oriented gradients for human detection," in *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR '05)*, vol. 1, pp. 886–893, June 2005.

[22] C. H. Lampert, M. B. Blaschko, and T. Hofmann, "Beyond sliding windows: object localization by efficient subwindow search," in *Proceedings of the 26th IEEE Conference on Computer Vision and Pattern Recognition (CVPR '08)*, pp. 1–8, June 2008.

[23] P. F. Felzenszwalb, R. B. Girshick, D. McAllester, and D. Ramanan, "Object detection with discriminatively trained part-based models," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 32, no. 9, pp. 1627–1645, 2010.

[24] C. Gu, J. J. Lim, P. Arbeláez, and J. Malik, "Recognition using regions," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR '09)*, pp. 1030–1037, June 2009.

[25] P. Wohlhart, M. Donoser, P. M. Roth et al., "Detecting partially occluded objects with an implicit shape model random field," in *Asian Conference on Computer Vision*, vol. 7724, pp. 302–315, Springer, 2013.

[26] K. Rematas and B. Leibe, "Efficient object detection and segmentation with a cascaded Hough Forest ISM," in *Proceedings of the IEEE International Conference on Computer Vision Workshops (ICCV '11)*, pp. 966–973, November 2011.

[27] J. Kwon and K. M. Lee, "Highly non-rigid object tracking via patch-based dynamic appearance modeling," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 35, no. 10, pp. 2427–2441, 2013.

[28] B. Leibe and B. Schiele, "Interleaving object categorization and segmentation," in *Proceedings of the British Machine Vision Conference*, pp. 759–768, 2003.

[29] P. Yarlagadda, A. Monroy, and B. Ommer, "Voting by grouping dependent parts," in *Proceedings of the 11th European Conference on Computer Vision*, pp. 197–210, 2010.

[30] N. Razavi, J. Gall, P. Kohli et al., "Latent Hough transform for object detection," in *Proceedings of the 13th European Conference on Computer Vision*, pp. 312–325, 2012.

[31] A. Lehmann, B. Leibe, and L. van Gool, "Fast PRISM: branch and bound hough transform for object class detection," *International Journal of Computer Vision*, vol. 94, no. 2, pp. 175–197, 2011.

[32] V. Kumar and I. Patras, "A discriminative voting scheme for object detection using hough forests," in *Proceedings of the British Machine Vision Conference*, pp. 1–10, 2010.

[33] P. Wohlhart, S. Schulter, M. Köstinger et al., "Discriminative hough forests for object detection," in *Proceedings of the British Machine Vision Conference*, pp. 1–11, 2012.

[34] A. Yilmaz, O. Javed, and M. Shah, "Object tracking: a survey," *ACM Computing Surveys*, vol. 38, no. 4, pp. 1–45, 2006.

[35] H. Yang, L. Shao, F. Zheng, L. Wang, and Z. Song, "Recent advances and trends in visual tracking: a review," *Neurocomputing*, vol. 74, no. 18, pp. 3823–3831, 2011.

[36] Y. Wu, J. Lim, and M. Yang, "Online object tracking: a benchmark," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR '13)*, 2013.

[37] S. Schulter, C. Leistner, P. M. Roth et al., "On-line hough forests," in *Proceedings of the British Machine Vision Conference*, pp. 1–11, 2011.

[38] J. Gall, N. Razavi, and L. V. Gool, "On-line adaption of class-specific codebooks for instance tracking," in *Proceedings of the British Machine Vision Conference*, pp. 55. 1–55. 12, 2010.

[39] H. Grabner and H. Bischof, "On-line boosting and vision," in *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR '06)*, vol. 1, pp. 260–267, June 2006.

[40] B. Babenko, M. Yang, and S. Belongie, "Robust object tracking with online multiple instance learning," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 33, no. 8, pp. 1619–1632, 2011.

[41] Z. Kalal, K. Mikolajczyk, and J. Matas, "Tracking-learning-detection," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 34, no. 7, pp. 1409–1422, 2012.

[42] B. Babenko, M. Yang, and S. Belongie, "Visual tracking with online multiple instance learning," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR '09)*, pp. 983–990, June 2009.

[43] H. Riemenschneider, S. Sternig, M. Donoser et al., "Hough regions for joining instance localization and segmentation," in *Proceedings of the 13th European Conference on Computer Vision*, pp. 258–271, 2012.

[44] C. Rother, V. Kolmogorov, and A. Blake, "GrabCut: interactive foreground extraction using iterated graph cuts," *ACM Transactions on Graphics*, vol. 23, no. 3, pp. 309–314, 2004.

[45] M. Özuysal, P. Fua, and V. Lepetit, "Fast keypoint recognition in ten lines of code," in *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR '07)*, pp. 1–8, June 2007.

[46] E. Borenstein and S. Ullman, "Combined top-down/bottom-up segmentation," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 30, no. 12, pp. 2109–2125, 2008.

[47] 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 (CVPR '09)*, pp. 1022–1029, June 2009.

[48] A. Bosch, A. Zisserman, and X. Muñoz, "Image classification using random forests and ferns," in *Proceedings of the IEEE 11th International Conference on Computer Vision (ICCV '07)*, pp. 1–8, October 2007.

[49] V. Lepetit and P. Fua, "Keypoint recognition using randomized trees," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 28, no. 9, pp. 1465–1479, 2006.

[50] F. Zheng and G. I. Webb, "A comparative study of semi-naive bayes methods in classification learning," in *Proceedings of the 4th Australasian Data Mining Conference*, pp. 141–156, 2005.

[51] J. Davis and M. Goadrich, "The relationship between Precision-Recall and ROC curves," in *Proceedings of the 23rd International Conference on Machine Learning*, pp. 233–240, June 2006.

[52] J. Mutch and D. G. Lowe, "Multiclass object recognition with sparse, localized features," in *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR '06)*, vol. 1, pp. 11–18, June 2006.

[53] S. Agarwal, A. Awan, and D. Roth, "Learning to detect objects in images via a sparse, part-based representation," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 26, no. 11, pp. 1475–1490, 2004.

[54] M. Andriluka, S. Roth, and B. Schiele, "People-tracking-by-detection and people-detection-by-tracking," in *Proceedings of the 26th IEEE Conference on Computer Vision and Pattern Recognition (CVPR '08)*, pp. 1–8, June 2008.

[55] A. Adam, E. Rivlin, and I. Shimshoni, "Robust fragments-based tracking using the integral histogram," in *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR '06)*, vol. 1, pp. 798–805, June 2006.

[56] H. Grabner and H. Bischof, "On-line boosting and vision," in *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR '06)*, vol. 1, 5, pp. 260–267, June 2006.

[57] Y. Lin, N. Lu, X. Lou et al., "Online learning of a cascaded classifier designed for multi-object tracking," in *Proceedings of the 11th IEEE International Conference on Electronic Measurement & Instruments*, pp. 1069–1075, 2013.

[58] D. A. Ross, J. Lim, R. Lin, and M. Yang, "Incremental learning for robust visual tracking," *International Journal of Computer Vision*, vol. 77, no. 1–3, pp. 125–141, 2008.

[59] S. Birchfield, "Elliptical head tracking using intensity gradients and color histograms," in *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR '98)*, pp. 232–237, June 1998.

[60] J. Kwon and K. M. Lee, "Tracking of a non-rigid object via patch-based dynamic appearance modeling and adaptive basin hopping monte carlo sampling," in *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops (CVPR '09)*, pp. 1208–1215, June 2009.

Hindawi Publishing Corporation
Mathematical Problems in Engineering
Volume 2014, Article ID 781654, 13 pages
http://dx.doi.org/10.1155/2014/781654

Research Article

Analysis and Optimization of Resource Control in High-Speed Railway Wireless Networks

Shengfeng Xu,[1,2] Gang Zhu,[1] Chao Shen,[1,2] Yan Lei,[1] and Zhangdui Zhong[1]

[1] *State Key Laboratory of Rail Traffic Control and Safety, Beijing Jiaotong University, Beijing 100044, China*
[2] *The State Key Laboratory of Integrated Services Networks, Xidian University, Xi'an 710126, China*

Correspondence should be addressed to Chao Shen; shenchao@bjtu.edu.cn

Received 28 November 2013; Revised 21 February 2014; Accepted 26 February 2014; Published 6 April 2014

Academic Editor: Ilse C. Cervantes

Copyright © 2014 Shengfeng Xu 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 considers a joint optimal design of admission control and resource allocation for multimedia services delivery in high-speed railway (HSR) wireless networks. A stochastic network optimization problem is formulated which aims at maximizing the system utility while stabilizing all transmission queues under the average power constraint. By introducing virtual queues, the original problem is equivalently transformed into a queue stability problem, which can be naturally decomposed into three separate subproblems: utility maximization, admission control, and resource allocation. A threshold-based admission control strategy is proposed for the admission control subproblem. And a distributed resource allocation scheme is developed for the mixed-integer resource allocation subproblem with guaranteed global optimality. Then a dynamic admission control and resource allocation algorithm is proposed, which is suitable for distributed implementation. Finally, the performance of the proposed algorithm is evaluated by theoretical analysis and numerical simulations under realistic conditions of HSR wireless networks.

1. Introduction

With the rapid development of high-speed railway (HSR) around the world, the wireless communication in HSR networks plays an important role in the recent years [1]. On the one hand, more and more data related with the railway controlling information needs to be transmitted between the train and the ground such that the safety can be guaranteed and the transportation efficiency can be significantly improved. On the other hand, the passengers in the train have an increasingly high demand on multimedia services. However, these requirements on high throughput impose a great challenge over the HSR communication designs due to the fast-varying channel, train penetration loss, and so forth.

There have been some recent works to improve the throughput in HSR wireless networks. A two-hop HSR network architecture was proposed in [2] to provide high data-rate services. A HSR communication system based on radio over fiber technology was proposed in [3], which can increase the system throughput and help to reduce the number of handoffs. Multi-input multi-output (MIMO) antennas were employed to improve the throughput performance of the HSR wireless networks [4, 5]. However, these works were carried out only to improve the throughput performance in HSR wireless networks. Since the buffering is involved at network devices, for example, content servers, it is necessary to consider not only the throughput performance but also the queue stability in HSR wireless networks.

Admission control and resource allocation, as critical parts of radio resource management, play important roles in improving the throughput and ensuring queue stability. In the literature, the energy constrained control algorithm was proposed in [6] to stabilize the queue and maximize the throughput by Lyapunov optimization theory. Paper [7] studied the joint scheduling and admission control problem in a single user scenario and an online learning algorithm was proposed based on the Markov decision process approach and stochastic control theory. However, these existing schemes designed for general communication systems are not easily extended to the scenario considered in this paper, due to the following reasons: (1) in HSR wireless networks, the channel condition cannot remain at the same

level because of the fast-varying distance between the base station and the train, which causes that the power control along the time has a large influence on system transmission performance [8]; (2) many types of services with different quality-of-service (QoS) requirements and priorities should be supported [9], which makes the admission control and resource allocation for multiple services more challenging.

In HSR wireless networks, few studies have been conducted on admission control and resource allocation. A scheduling and resource allocation mechanism was presented in [10] to maximize the service rate in HSR networks with a cell array architecture. In [11], a multidimensional resource allocation strategy was proposed in downlink orthogonal frequency-division multiplexing (OFDM) system for HSR communications. The optimal resource allocation problem in a cellular/infostation integrated HSR network was investigated in [12], which considered the intermittent network connectivity and multiservice demands. In a relay-assisted HSR network, [13] studied delay-aware fair downlink service scheduling problem with heterogeneous packet arrivals and delay requirements for the services. Paper [14] proposed an effective admission control scheme considering different service priorities for HSR communications with MIMO antennas. However, to the best of our knowledge, the joint admission control and resource allocation problem under the average power constraint in HSR wireless networks is still an open problem.

The main contribution of this paper is a stochastic optimization framework for transmitting multimedia services in HSR wireless networks, which focuses on the joint admission control and resource allocation problem under the average power constraint. Firstly, the joint admission control and resource allocation problem is formulated as a stochastic optimization problem, and then the problem is transformed into a queue stability problem with the help of virtual queues. By the drift-plus-penalty approach [15], the transformed problem can be decomposed into three separate subproblems: utility maximization, admission control, and resource allocation. The former two subproblems are easy to handle and the distributed solutions can be obtained directly, while the mixed-integer resource allocation subproblem is transformed into a single variable problem and a distributed packet loading scheme is developed with guaranteed global optimality. We further propose a dynamic admission control and resource allocation algorithm, which is suitable for distributed implementation in HSR wireless networks. Finally, we present the analysis of algorithm performance by theoretical derivations and simulations under realistic conditions for HSR wireless networks.

1.1. Relation to Prior Work. The Lyapunov drift theory has a long history in the field of discrete stochastic processes and Markov chains [16]. It can be used to directly analyze the characteristics of the control policies in the stochastic stability sense and plays important roles in the dynamic control strategies in queuing networks [17]. Stabilizing queuing networks by minimizing Lyapunov drift was pioneered by Tassiulas and Ephremides in [18]. The Lyapunov drift theory was then extended to the Lyapunov optimization theory [6], which enables optimization of time averages of general network utilities subject to queue stability. A general framework for solving the stochastic network optimization problem based on Lyapunov optimization theory was developed in [15]. This framework has been extended to minimizing a drift-plus-penalty expression in [6, 7, 17, 19, 20] for joint queue stability and time average utility optimization. For the engineering applications of Lyapunov optimization theory, interested readers are referred to the aforementioned references for the details.

Our approach in the present paper treats the joint admission control and resource allocation problem associated with average power constraint using Lyapunov drift and Lyapunov optimization theory from [15]. This is the first time, to the best of our knowledge, that the Lyapunov optimization theory is extended into the HSR wireless networks. Considering the features of HSR wireless networks, the Lyapunov optimization theory is successfully applied for solving the joint admission control and resource allocation problem in HSR wireless network.

1.2. Outline of Paper. The rest of the paper is organized as follows. Section 2 describes the system model. The problem formulation and transformation are provided in Section 3. A distributed dynamic admission control and resource allocation algorithm is proposed in Section 4. Some numerical results and discussions are shown in Section 5. Finally, conclusions are drawn in Section 6.

Notations. In this paper, $\mathbb{E}[\cdot]$ denotes expectation. $\lfloor x \rfloor = \max\{n \in \mathbb{Z} \mid n \leq x\}$. $\max[x, y]$ and $\min[x, y]$ mean the maximum and minimum between x and y, respectively.

2. System Model

In this paper, a two-hop HSR wireless network architecture is considered, as shown in Figure 1, which consists of a backbone network, K content servers (CSs), several base stations (BSs), a relay station (RS), and some users in the train. The BSs deployed along the rail line can provide continuous data packets delivery. The distributed CSs connected to the BSs via wireline links are deployed in the backbone network to offload the data traffic [21]. The RS with powerful antennas installed on the top of the train is used for communicating with the BSs so that the large train penetration loss can be well resolved. The RS is further connected to the access points (APs) which can be accessed by the users inside the train. Thus, there is a two-hop wireless link, consisting of the BS-RS link and the AP-Users link. If the users on the train request multimedia services during a trip, the data packets of the requested services are then delivered from the corresponding CS to the RS via a BS. Suppose that the data transmission rate in the AP-Users link is sufficiently large; hence the data packet can be successfully received if it has been delivered to the RS.

2.1. Time-Distance Mapping. Consider a train traveling from an origin station to a destination station within the time

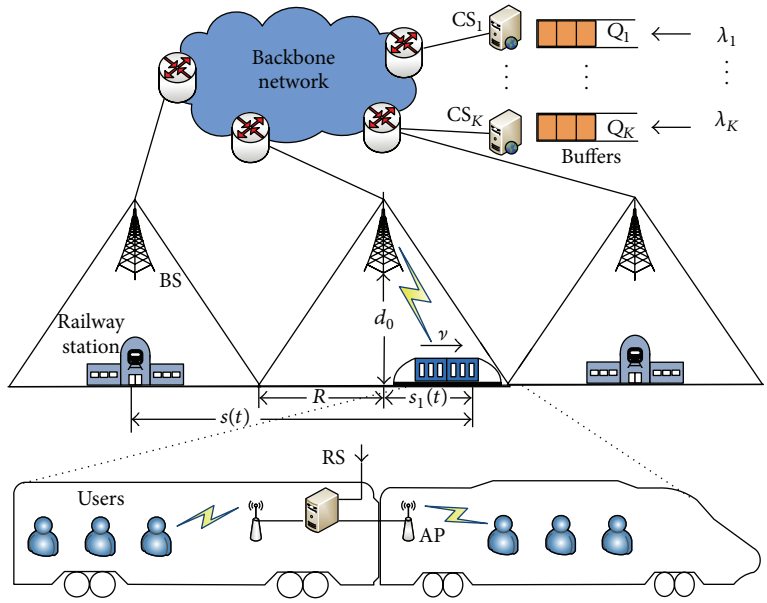


FIGURE 1: System model.

duration $[0, T]$. The whole time is divided into slots of equal duration T_s. Without loss of generality, we assume that the train starts at the centre of the first cell and the train moving speed during the slot t keeps constant, denoted by $v(t)$; thus the traveled distance until slot t is given by $s(t) = \sum_{\tau=0}^{t} v(\tau)T_s$. The train location between two adjacent BSs at slot t is $s_1(t) = s(t) \bmod 2R$, where R is the cell radius. Define a time-distance mapping function $d(t)$ as the distance between BS and RS at slot t; that is, $d(t) : [0, T] \rightarrow [d_0, d_{\max}]$, where $d_{\max} = \sqrt{R^2 + d_0^2}$ and d_0 is the distance between each BS and the rail line as shown in Figure 1. The mapping function $d(t)$ can be expressed by

$$d(t) = \begin{cases} \sqrt{s_1(t)^2 + d_0^2}, & \text{if } 0 \le s_1(t) < R, \\ \sqrt{(2R - s_1(t))^2 + d_0^2}, & \text{if } R \le s_1(t) < 2R. \end{cases} \quad (1)$$

Here we assume that the distance $d(t)$ does not change within slot t since T_s is very small.

2.2. Physical Layer Model. For HSR wireless networks, the channel condition cannot remain at the same level due to the fast-varying distance between BS and RS. Only the line-of-sight (LOS) path in the BS-RS link is available at most of the time, which was confirmed by engineering measurements [22, 23]. The service provided by the independent identical distributed (i.i.d.) fading channels is a deterministic time-linear function, just like the AWGN channel [24]. Therefore, the wireless channel in the BS-RS link can be assumed to be an additive white Gaussian noise channel (AWGN) with LOS path loss [8]. At the same time, the power control along the travel time in HSR wireless networks is important. Denote by $P(t)$ the transmit power of the BS at slot t, which is limited by the maximum value $P_{\max}$ and average value P_{av}. With the help of mapping function $d(t)$ and according to Shannon's theorem [25], the transmission rate of the wireless channel between BS and RS at slot t can be expressed by

$$R(t) = W \log_2 \left(1 + \frac{P(t)}{N(t)}\right) \text{ bits/s}, \quad (2)$$

where $N(t) = WN_0 d^{\alpha}(t)$, W is the system bandwidth, N_0 is the noise power spectral density, and α is the path loss exponent. Suppose that the packets have equal size of L bits; hence the link capacity $C(t)$ at slot t can be denoted as the maximum number of packets; that is, $C(t) = \lfloor R(t)T_s/L \rfloor$. Note that the maximum capacity $C_{\max}$ can be obtained when $d(t) = d_0$ and $P(t) = P_{\max}$.

2.3. Service Model. Assume that there are K types of services in the HSR wireless networks and the service type set is denoted by $\mathcal{K} \triangleq \{1, \ldots, K\}$. We further assume that CS_k is equipped with a buffer and can provide service k, for $k \in \mathcal{K}$. Let $\mathbf{A}(t) = [A_1(t), \ldots, A_K(t)]^T$ represent the packet arrival vector, where $A_k(t)$ denotes the number of new arrival packets of service k at slot t. The packet arrival process for each service is assumed to be i.i.d. across slots. Suppose that, in general, $A_k(t)$ follows a truncated Poisson distribution $f_k(b)$ with average arrival rate $\lambda_k = \mathbb{E}[A_k(t)]$ for service k, and the distribution $f_k(b)$ can be written as

$$f_k(b) = \exp(-\lambda_k) \frac{\lambda_k^b}{b!}, \quad b = 0, \ldots, B_k, \quad (3)$$

where B_k denotes the maximum number of arrival packets per slot for service k and can be found assuming $f_k(B_k) \rightarrow 0$.

Let $\mathbf{Q}(t) = [Q_1(t), \ldots, Q_K(t)]^T$ represent the vector of current queue backlogs, where $Q_k(t)$ denotes the number of packets at the beginning of slot t in the buffer of CS_k. The dynamics of each buffer are controlled by admission

control (AC) and resource allocation (RA) actions. Specifically, at each slot, the AC action determines the number of packets from the newly arriving packets to be stored into the buffer. And the RA action determines the number of packets removed from the buffer for transmission. Let $r_k(t) \in [0, A_k(t)]$ and $\mu_k(t) \in [0, Q_k(t)]$ denote the AC action and RA action for service k at slot t, respectively. Thus, the queue dynamics can be characterized by

$$Q_k(t+1) = Q_k(t) - \mu_k(t) + r_k(t), \quad \forall k \in \mathcal{K}. \tag{4}$$

Notice that for any slot t, without AC actions, $r_k(t) = A_k(t)$. Here we assume that the arrival packets at slot t can only be transmitted at slot $t+1$.

3. Problem Formulation and Transformation

3.1. Problem Formulation. In this paper, the objective of the joint AC and RA problem is to maximize a sum of utility functions under time average constraints by designing a dynamic algorithm over a trip of the train. We define that $\phi_k(r_k)$ is a utility function to present throughput benefit for service k, which is nondecreasing concave continuous with r_k. Throughout this work, the following notation for the long-term time average expectation of any quantity z is defined:

$$\overline{z} := \lim_{t \to \infty} \frac{1}{t} \sum_{\tau=0}^{t-1} \mathbb{E}[z(\tau)]. \tag{5}$$

In particular, $\overline{Q}_k$ represents the average queue backlog in the buffer of CS_k and $\overline{P}$ represents the average power consumption along the travel time. Here we introduce the definitions of queue stability as follows [15].

Definition 1. A single queue $Q(t)$ is mean rate stable if $\lim_{t \to \infty}(\mathbb{E}[Q(t)]/t) = 0$.

Definition 2. A single queue $Q(t)$ is strongly stable if $\overline{Q} < \infty$.

From Definition 2, a queue is strongly stable if it has a bounded time average backlog. Strong stability implies mean rate stability according to [15]. Throughout this paper, we use the term "stability" to refer to strong stability. Define $\Omega(t) \triangleq (\mathbf{A}(t), C(t))$ as the observed network event at slot t. For each slot t, observing the event $\Omega(t)$ and the queue state $\mathbf{Q}(t)$, the AC actions $r_k(t)$ and RA actions $\mu_k(t)$ should be made for $k \in \mathcal{K}$. The joint AC and RA problem is formulated as

$$\text{(P1)} \quad \text{maximize} \quad \sum_{k \in \mathcal{K}} \phi_k(\overline{r}_k) \tag{6a}$$

$$\text{subject to} \quad \overline{P} \le P_{\mathrm{av}}, \quad P(t) \le P_{\max}, \quad \forall t \in [0, T], \tag{6b}$$

$$\overline{Q}_k < \infty, \quad \forall k \in \mathcal{K}, \tag{6c}$$

$$0 \le \sum_{k \in \mathcal{K}} \mu_k(t) \le C(t), \quad \forall t \in [0, T], \tag{6d}$$

$$\mu_k(t) \in [0, Q_k(t)], \quad \forall t \in [0, T], \tag{6e}$$

$$r_k(t) \in [0, A_k(t)], \quad \forall t \in [0, T], \tag{6f}$$

$$\text{variables } r_k(t), \mu_k(t), P(t), \quad \forall k \in \mathcal{K}, \quad t \in [0, T], \tag{6g}$$

where (6b) corresponds to the power constraint and (6c) corresponds to the queue stability constraints for all queues. Problem (P1) is a stochastic optimization problem [15], but it cannot be solved efficiently owing to the difficulty from the objective function (6a) and the average power constraint in (6b). In order to better characterize the problem (P1) and develop an efficient algorithm, we consider the problem transformation, which consists of two steps, that is, objective function transformation and average power constraint transformation as presented in the following subsections.

3.2. Objective Function Transformation. Since problem (P1) involves maximizing a function of time averages, it is hard to handle. Based on the dynamic stochastic optimization theory [15], it can be transformed into an equivalent problem that involves maximizing a single time average of a function. This transformation is achieved through the use of auxiliary variables $\gamma_k(t)$ and corresponding virtual queues $Z_k(t)$ with queue evolutions:

$$Z_k(t+1) = \max[Z_k(t) - r_k(t), 0] + \gamma_k(t), \quad \forall k \in \mathcal{K}, \tag{7}$$

where the initial condition is assumed that $Z_k(0) = 0$, $\forall k \in \mathcal{K}$. Intuitively, the auxiliary variables $\gamma_k(t)$ can be viewed as the "arrivals" of virtual queues $Z_k(t)$, while $r_k(t)$ can be viewed as the service rate of such virtual queues.

Then, we consider the following transformed problem:

$$\text{(P2)} \quad \text{maximize} \quad \sum_{k \in \mathcal{K}} \overline{\phi_k(\gamma_k)} \tag{8a}$$

$$\text{subject to} \quad \overline{\gamma}_k \le \overline{r}_k, \quad \forall k \in \mathcal{K}, \tag{8b}$$

$$0 \le \gamma_k(t) \le B_k, \quad \forall k \in \mathcal{K}, \quad t \in [0, T], \tag{8c}$$

$$\text{(6b)–(6f)}, \tag{8d}$$

$$\text{variables } \gamma_k(t), r_k(t), \mu_k(t), P(t), \quad \forall k \in \mathcal{K}, \quad t \in [0, T]. \tag{8e}$$

Constraint (8b) corresponds to the stability of the virtual queue $Z_k(t)$, since $\overline{\gamma}_k$ and $\overline{r}_k$ are regarded as the time-averaged arrival rate and the time-averaged service rate for the virtual queue $Z_k(t)$, respectively. Specifically, from (7) we can obtain that $Z_k(t+1) \ge Z_k(t) - r_k(t) + \gamma_k(t)$. By summing this inequality over time slots $\tau \in \{0, 1, \ldots, t-1\}$ and then dividing the result by t, we have that $(Z_k(t) - Z_k(0))/t + (1/t)\sum_{\tau=0}^{t-1} r_k(\tau) \ge (1/t)\sum_{\tau=0}^{t-1} \gamma_k(t)$. With $Z_k(0) = 0$, taking

expectations of both sides yields that $\lim_{t\to\infty}(\mathbb{E}[Z_k(t)]/t) + \overline{r}_k \geq \overline{\gamma}_k$. If the virtual queues $Z_k(t)$ are mean rate stable, then $\lim_{t\to\infty}(\mathbb{E}[Z_k(t)]/t) = 0$, so that constraint (8b) can be satisfied. Notice that we will prove the strong stability of the virtual queues $Z_k(t)$ in Lemma 7 later.

Lemma 3. *Problem (P1) and problem (P2) are equivalent.*

Proof. The proof of Lemma 3 follows [26] and a sketch of the proof is provided in Appendix A. □

3.3. Average Power Constraint Transformation. To handle the average power constraint in (6b), we define a virtual queue $X_k(t)$ for each $k \in \mathcal{K}$, which has the following dynamic update equation:

$$X_k(t+1) = \max\left[X_k(t) - P_{\text{av}}, 0\right] + P(t), \tag{9}$$

where $P(t)$ and P_{av} can be viewed as the "arrivals" and "offered service" at slot t, respectively.

Based on [15, Chapter 4], if the virtual queue $X_k(t)$ is mean rate stable for $k \in \mathcal{K}$, that is, $\lim_{t\to\infty}(\mathbb{E}[X_k(t)]/t) = 0$, then the average power constraint $\overline{P} \leq P_{\text{av}}$ can be satisfied. This holds because if the backlog in the virtual queue is stabilized, it must be the case that the time average arrival rate (corresponding to $\overline{P}$) is not larger than the service rate (corresponding to P_{av}). Therefore, the average power constraint in (6b) can be transformed into a single queue stability problem.

4. The Distributed Dynamic AC and RA Algorithm

In this section, the dynamic stochastic optimization approach is applied to solve problem (P2), which seeks to maximize the sum of time-averaged utility functions subject to queue stability constraints. Firstly, the problem (P2) is decomposed into three separate subproblems by the drift-plus-penalty approach. Then a distributed dynamic AC and RA algorithm is proposed. Finally, the performance of the proposed algorithm is analyzed by theoretical derivations.

4.1. Lyapunov Drift. Define $\mathbf{X}(t)$ and $\mathbf{Z}(t)$ as a vector of all virtual queues $X_k(t)$ and $Z_k(t)$ for $k \in \mathcal{K}$, respectively. We denote by $\mathbf{\Theta}(t)$ the combined vector of all virtual queues and all actual queues; namely,

$$\mathbf{\Theta}(t) \triangleq \left[\mathbf{Q}^T(t), \mathbf{X}^T(t), \mathbf{Z}^T(t)\right]^T. \tag{10}$$

The quadratic Lyapunov function is defined as [15]

$$L(\mathbf{\Theta}(t)) \triangleq \frac{1}{2}\sum_{k\in\mathcal{K}}\left(Q_k(t)^2 + X_k(t)^2 + Z_k(t)^2\right). \tag{11}$$

Then the one-slot conditional Lyapunov drift $\Delta(\mathbf{\Theta}(t))$ at slot t is given by

$$\Delta(\mathbf{\Theta}(t)) = \mathbb{E}\left[L(\mathbf{\Theta}(t+1)) - L(\mathbf{\Theta}(t)) \mid \mathbf{\Theta}(t)\right], \tag{12}$$

which admits the following lemma.

Lemma 4. *Under any AC actions and RA actions at slot t, and for any value of $\mathbf{\Theta}(t)$, we have*

$$\Delta(\mathbf{\Theta}(t)) \leq D + \mathbb{E}\left[G(t) \mid \mathbf{\Theta}(t)\right], \tag{13}$$

where D is a finite constant defined by

$$D \triangleq \frac{1}{2}\sum_{k\in\mathcal{K}}\left[P_{\max}^2 + P_{av}^2 + 3B_k^2 + C_{\max}^2\right], \tag{14}$$

and $G(t)$ is defined by

$$\begin{aligned} G(t) \triangleq & \sum_{k\in\mathcal{K}} Q_k(t)\left[r_k(t) - \mu_k(t)\right] \\ & + \sum_{k\in\mathcal{K}} Z_k(t)\left(\gamma_k(t) - r_k(t)\right) \\ & + \sum_{k\in\mathcal{K}} X_k(t)\left[P(t) - P_{av}\right]. \end{aligned} \tag{15}$$

Proof. The proof of Lemma 4 is provided in Appendix B. □

4.2. The Drift-Plus-Penalty Expression. Instead of directly minimizing the upper bound $\mathbb{E}[G(t)]$ by taking AC actions and RA actions, we desire to jointly stabilize all queues and maximize the sum of utility $\sum_{k\in\mathcal{K}}\phi_k(\gamma_k(t))$. The drift-plus-penalty theory in [6] approaches this by greedily minimizing the following "drift-plus-penalty" expression:

$$\mathbb{E}\left[G(t) - V\sum_{k\in\mathcal{K}}\phi_k\left(\gamma_k(t)\right)\right], \tag{16}$$

where $V \geq 0$ is a parameter that represents the weight on how much we emphasize the sum utility maximization.

We observe that the objective function in (16) is of separable structure, which motivates us to determine the auxiliary variables $\gamma_k(t)$ and AC actions $r_k(t)$ as well as RA actions $\mu_k(t)$ in an alternative optimization fashion. The overall minimization problem (16) is decomposed into three separate subproblems. Specifically, isolating the $\gamma_k(t)$ variables from (16) gives the following *utility maximization* subproblem:

$$\max_{\{\gamma_k(t)\}} \quad \sum_{k\in\mathcal{K}}\left[V\phi_k\left(\gamma_k(t)\right) - Z_k(t)\gamma_k(t)\right] \tag{17a}$$

$$\text{s.t.} \quad 0 \leq \gamma_k(t) \leq B_k, \quad \forall k \in \mathcal{K},\ t \in [0,T]. \tag{17b}$$

Similarly, isolating the AC actions $r_k(t)$ from (16) leads to the following *admission control* subproblem:

$$\max_{\{r_k(t)\}} \quad \sum_{k\in\mathcal{K}}\left[\left(Z_k(t) - Q_k(t)\right)r_k(t)\right] \tag{18a}$$

$$\text{s.t.} \quad 0 \leq r_k(t) \leq A_k(t), \quad \forall k \in \mathcal{K},\ t \in [0,T]. \tag{18b}$$

Also, isolating the RA actions $\mu_k(t)$ from (16) gives the following *resource allocation* subproblem:

$$\max_{\{\mu_k(t)\},P(t)} \quad \sum_{k\in\mathcal{K}} \left[Q_k(t)\,\mu_k(t) - X_k(t)\,P(t)\right] \tag{19a}$$

$$\text{s.t.} \quad 0 \le \mu_k(t) \le Q_k(t), \quad \mu_k(t) \in \mathbb{N}, \quad \forall k \in \mathcal{K}, \quad t \in [0,T], \tag{19b}$$

$$\sum_{k\in\mathcal{K}} \mu_k(t) \le C(t), \quad t \in [0,T], \tag{19c}$$

$$P(t) \le P_{\max}, \quad t \in [0,T], \tag{19d}$$

where $P(t)$ is related to $\mu_k(t)$ since a larger $\mu_k(t)$ requires more power consumption. These separate subproblems can be computed in a decentralized fashion, as stated below.

4.3. Utility Maximization. The utility maximization subproblem ((17a) and (17b)) can be decoupled into K separate maximization problems. Specifically, CS_k keeps track of $Z_k(t)$ and determines the optimum $\gamma_k(t)$ by solving the following problem:

$$\max_{\gamma_k(t)} \quad V\phi_k(\gamma_k(t)) - Z_k(t)\,\gamma_k(t) \tag{20a}$$

$$\text{s.t.} \quad 0 \le \gamma_k(t) \le B_k, \quad t \in [0,T]. \tag{20b}$$

Notice that the key point to solve (20a) and (20b) is the choice of the utility function, which is contingent on the purpose of the networking application or the prerogative of HSR network designer. For example, in order to represent the maximum desired delivery ratio for each service, the piecewise linear utility function can be considered for service k as follows:

$$\phi_k(\gamma_k) = \nu_k \min\left[\gamma_k, x_k\lambda_k\right], \tag{21}$$

where $\nu_k > 0$ and $x_k > 0$ represent the priority and the maximum desired delivery ratio of service k, respectively. In general, $0 \le x_k \le 1$ and $x_k\lambda_k \le B_k$ for $k \in \mathcal{K}$. Thus the optimal solution to problem (20a) and (20b) is given by

$$\gamma_k(t) = \begin{cases} x_k\lambda_k, & \text{if } Z_k(t) \le V\nu_k, \\ 0, & \text{otherwise.} \end{cases} \tag{22}$$

Alternatively, the following strictly concave function can serve as the utility function for service k:

$$\phi_k(\gamma_k) = \ln(1 + \nu_k\gamma_k), \tag{23}$$

which can be regarded as an accurate approximation of the proportionally fair utility function if the same ν_k is selected with a large value for all $k \in \mathcal{K}$. In this case, the optimal solution to problem (20a) and (20b) can be obtained by

$$\gamma_k(t) = \left[\frac{V}{Z_k(t)} - \frac{1}{\nu_k}\right]_0^{B_k}, \tag{24}$$

where the operation $[y]_0^a$ is equal to y if $0 < y < a$, 0 if $y \le 0$, and a if $y \ge a$.

4.4. Admission Control. The admission control subproblem ((18a) and (18b)) can be also decoupled into K separate maximization problems. Specifically, CS_k chooses the AC action $r_k(t)$ by solving the following optimization problem:

$$\max_{r_k(t)} \quad \left(Z_k(t) - Q_k(t)\right) r_k(t) \quad \text{s.t.} \quad 0 \le r_k(t) \le A_k(t), \quad t \in [0,T]. \tag{25}$$

It is immediate to see that the optimal solution depends on the queue backlog of $Z_k(t)$ and $Q_k(t)$, which is given by

$$r_k(t) = \begin{cases} A_k(t), & \text{if } Z_k(t) \ge Q_k(t), \\ 0, & \text{otherwise.} \end{cases} \tag{26}$$

We note that (26) is a simple threshold-based admission control strategy. On the one hand, when the queue backlog $Q_k(t)$ is not larger than the threshold $Z_k(t)$, then all the newly arriving packets are admitted into the buffer in CS_k. Essentially, this not only reduces the value of $Z_k(t+1)$ so as to push $\overline{\gamma}_k$ closer to $\overline{r}_k$, but also increases the average throughput $\overline{r}_k$ so as to improve the utility. On the other hand, when the queue backlog $Q_k(t)$ is larger than the threshold $Z_k(t)$, then all the newly arriving packets will be dropped to ensure the network stability. Finally, we emphasize that the AC actions for all services are made in a distributed manner with only local queue backlog information and packet arrival information.

4.5. Resource Allocation. The resource allocation subproblem ((19a)–(19d)) at slot t can be explicitly expressed as

$$\max_{\{\mu_k(t)\},P(t)} \quad M(t) \triangleq \sum_{k\in\mathcal{K}} \left[Q_k(t)\,\mu_k(t) - X_k(t)\,P(t)\right] \tag{27a}$$

$$\text{s.t.} \quad 0 \le \mu_k(t) \le Q_k(t), \quad \mu_k(t) \in \mathbb{N}, \quad \forall k \in \mathcal{K}, \tag{27b}$$

$$\sum_{k\in\mathcal{K}} \mu_k(t) \le C(t) = \left\lfloor \frac{T_s W \log_2(1 + P(t)/N(t))}{L} \right\rfloor, \tag{27c}$$

$$P(t) \le P_{\max}. \tag{27d}$$

The problem (27a)–(27d) is a mixed-integer programming (MIP) problem, including a continuous variable $P(t)$ and K integer variables $\mu_k(t)$, which cannot be solved efficiently [27]. The main difficulty of problem (27a)–(27d) comes from the integer nature of $\mu_k(t)$. However, we will show that problem (27a)–(27d) can be transformed into a single variable problem, which is easy to handle. In the sequel of this subsection, we will omit the time index for brevity.

Firstly, as for constraint (27c), when the optimal RA actions are achieved, it can be shown that

$$\sum_k \mu_k = C = \frac{1}{\eta}\log_2\left(1 + \frac{P}{N}\right), \tag{28}$$

where $\eta = L/T_sW > 0$. Otherwise we can reduce the value of C and P such that the objective function can be further

maximized without any violation of the constraints in (27b)–(27d). From (28), we have the following power consumption of C:

$$P = N\left(2^{\eta C} - 1\right), \tag{29}$$

and constraints (27b) and (27d) further imply that

$$0 \le C \le \min\left(\sum_{k\in\mathcal{K}} Q_k, \widetilde{C}_{\max}\right), \tag{30}$$

where $\widetilde{C}_{\max} \triangleq (1/\eta)\log_2(1 + P_{\max}/N)$.

Secondly, the resource allocation subproblem (27a)–(27d) can be equivalently transformed into a single variable problem as follows:

$$\max_{C\in\mathbb{N}} \quad \widehat{M}(C) \triangleq g_1(C) - g_2(C) \tag{31a}$$

$$\text{s.t.} \quad (30), \tag{31b}$$

where $g_1(C)$ is given by

$$g_1(C) \triangleq \max_{\{\mu_k\}} \quad \sum_{k\in\mathcal{K}} Q_k\mu_k \tag{32a}$$

$$\text{s.t.} \quad 0 \le \mu_k \le Q_k, \quad \mu_k \in \mathbb{N}, \quad \forall k \in \mathcal{K}, \tag{32b}$$

$$\sum_{k\in\mathcal{K}} \mu_k = C, \tag{32c}$$

and $g_2(C)$ is given by

$$g_2(C) \triangleq \sum_{k=1}^{K} X_k P = \sum_{k=1}^{K} X_k\left(2^{\eta C} - 1\right) N = \rho\left(2^{\eta C} - 1\right), \tag{33}$$

with $\rho \triangleq N\sum_{k=1}^{K} X_k$.

Now let us focus on the problem (32a)–(32c) with any given C. Clearly, the maximum objective value of (32a)–(32c) can always be achieved by allocating link capacity C to the services in the descending order of their backlogs, which is similar to the max-weight algorithm in [15]. Hence, we sort all the services in descending order of Q_k and denote the ordered set by $\{k_1, k_2, \ldots, k_K\}$. For convenience, we define $f(m) = \sum_{n=0}^{m} Q_{k_n}(t)$ for $m = 0, 1, \ldots, K$, where $Q_{k_0}(t) = 0$. One can see that $f(m)$ is an increasing function of m, and $0 \le C \le f(K)$ from (30). Therefore the optimal solutions to the problem (32a)–(32c) are given by

$$\mu_{k_n} = \begin{cases} Q_{k_n}, & \text{if } 1 \le n < m, \\ C - f(n-1), & \text{if } n = m, \\ 0, & \text{if } m < n \le K, \end{cases} \tag{34}$$

where $m \in \{1, \ldots, K\}$ such that $C \in (f(m-1), f(m)]$ if $C > 0$; otherwise $\mu_{k_n} = 0$ for all n.

Next, let us focus on problem (31a)-(31b). Indeed, we have the following lemma.

Lemma 5. *$\widehat{M}(C)$ is a unimodal function of C over $[0, f(K)]$.*

Proof. On the one hand, since $\Delta g_1(C) = g_1(C+1) - g_1(C) = Q_{k_m}$ for $f(m-1) \le C(t) < f(m)$, $\forall m \in [1, K]$, $\Delta g_1(C)$ is a monotonically nonincreasing function of C. On the other hand, since $\Delta g_2(C) = \rho(2^{\eta} - 1)2^{\eta C}$, $\Delta g_2(C)$ is a monotonically increasing function of C. Therefore $\Delta\widehat{M}(C) = \Delta g_1(C) - \Delta g_2(C)$, which is a monotonically decreasing function of C. For any $C \in [1, f(K) - 1]$, $\Delta\widehat{M}(C-1) > \Delta\widehat{M}(C)$, which implies $\widehat{M}(C) > (1/2)(\widehat{M}(C-1) + \widehat{M}(C+1))$, so $\widehat{M}(C)$ is concave on $[0, f(K)]$. Based on [28], if $\widehat{M}(C)$ is concave, then $\widehat{M}(C)$ is unimodal. □

Based on Lemma 5, since $\widehat{M}(C)$ is a unimodal function of C over $[0, \min\{\widetilde{C}_{\max}, f(K)\}]$, the golden section search method [29] can be used to obtain the global optimal solution to the problem (31a)-(31b). However, this method requires the knowledge of all queue backlog information. When the center controller is not available, a distributed resource allocation scheme is highly desirable. Relying on the insights from Lemma 5, we propose a distributed resource allocation scheme, where each CS can communicate with all other CSs, and the network resources are allocated packet by packet.

The proposed packet loading resource allocation scheme is detailed in Algorithm 1. For each slot, each CS exchanges the backlog information with all other CSs and the order of the backlogs is obtained by all CSs (step 2). Then the packets of the services are fetched from the corresponding CS in descending order of their backlogs. When one CS fetches a new packet (step 5), $\widehat{M}(C)$ is calculated (step 6). This will be repeated until the optimal condition in step 7 is satisfied, which implies that $\widehat{M}(C-1)$ is the maximum or constraint (30) is violated, and thus step 8 should be performed since the last packet cannot be transmitted. When one CS empties its buffer, it should send the value information of C to the next CS (step 12), and the next CS should continue the resource allocation. Here we remark that the optimal solutions can be achieved by the proposed scheme, which can be readily proved by Lemma 5.

4.6. Distributed Dynamic AC and RA Algorithm. Based on the above three separate subproblems, we propose a distributed dynamic AC and RA algorithm as shown in Algorithm 2. All system parameters should be initialized before the trip begins. At each slot, each CS solves three subproblems in steps 4, 5, and 6. At the end of each slot, the queue vector $\mathbf{\Theta}(t+1)$ is updated according to (4), (7), and (9). This algorithm will be repeated when the train travels from the origin station to the destination station.

Remark 6 (utility-backlog tradeoff). Based on [15], the achieved utility differs from optimality by $\mathcal{O}(1/V)$, in the sense that

$$\phi_1^* - \lim_{t\to\infty}\inf \sum_{k\in\mathcal{K}} \phi_k(\bar{r}_k) \le \mathcal{O}\left(\frac{1}{V}\right), \tag{35}$$

where ϕ_1^* is the maximal utility for the problem (P1). It implies that the proposed algorithm can achieve a utility which is arbitrarily close to ϕ_1^* by increasing V. In addition, the actual

(1) Initialize $\mu_k = 0$ for $k \in \mathcal{K}$; $\widehat{M}(0) = 0$; $C = 0$;
(2) Sort services in descending of Q_k and obtain the ordered set $\{k_1, k_2, \ldots, k_K\}$;
(3) **for** $n = 1$ to K **do**
(4) **for** $\mu_{k_n} = 1$ to Q_{k_n} **do**
(5) $C := C + 1$;
(6) Calculate $\widehat{M}(C)$;
(7) **if** $\widehat{M}(C) < \widehat{M}(C-1)$ or $C > \widetilde{C}_{\max}(t)$ **then**
(8) $\mu_{k_n} := \mu_{k_n} - 1$, $C := C - 1$;
(9) go to Step 14;
(10) **end if**
(11) **end for**
(12) Send C to the next CS;
(13) **end for**
(14) **return** P by (29), and $\mu_k \, \forall k \in \mathcal{K}$.

Algorithm 1: Packet loading resource allocation scheme for problem (27a) – (27d).

(1) Initialize V, $C_{\max}$, $\boldsymbol{\Theta}(0) = \mathbf{0}$;
(2) **while** $t \in [0, T]$ **do**
(3) **for** $k = 1$ to K **do**
(4) Obtain $\gamma_k(t)$ by solving (20a) and (20b);
(5) Obtain AC actions $r_k(t)$ by solving (25);
(6) Obtain RA actions $\mu_k(t)$ by solving problem (27a) – (27d);
(7) Update $\boldsymbol{\Theta}(t+1)$ according to (4), (7) and (9);
(8) **end for**
(9) **end while**.

Algorithm 2: Distributed dynamic AC and RA algorithm for (P2).

queue backlog of each service grows linearly with V, which is given by

$$\lim_{t \to \infty} \sup \overline{Q}_k \leq \frac{D}{\epsilon} + \mathcal{O}(V), \tag{36}$$

where $\epsilon > 0$ is a parameter and D is defined in (14). Therefore, the above expressions (35) and (36) present a utility-backlog tradeoff of $[\mathcal{O}(1/V), \mathcal{O}(V)]$.

Recalling the utility functions (21) and (23), we observe that they have the maximum right derivatives $\nu_k > 0$ over the interval $0 \leq \gamma_k(t) \leq B_k$. Based on this observation, we obtain the boundedness property on the virtual queue $Z_k(t)$ in the following lemma.

Lemma 7. *If the utility function $\phi_k(\gamma_k)$ has maximum right derivatives $\nu_k > 0$, then the backlog of virtual queue $Z_k(t)$ satisfies*

$$0 \leq Z_k(t) \leq V\nu_k + B_k, \quad \forall t \in [0, T], \tag{37}$$

provided that this inequality holds for $Z_k(0)$.

Proof. The proof of Lemma 7 is provided in Appendix C. □

Table 1: Parameters in simulation.

Parameter	Description	Value
P_{av}	Average power constraint	35 W
B	System bandwidth	5 MHz
L	Packet size	240 bits
T_s	Slot duration	1 ms
α	Pathloss exponent	4
$P_{\max}$	Maximum transmit power	45 W
v	Constant moving speed	100 m/s
R	Cell radius	1.5 km
d_0	Distance between BS and rail	50 m
K	Number of services	6

5. Numerical Results and Discussions

In this section, we implement the proposed distributed dynamic AC and RA algorithm using MATLAB and present simulation results to illustrate the performance of it. We use the piecewise linear utility functions (21) for all services and summarize the simulation parameters in Table 1. The packet size L is set to 240 bits according to [12, 30], and the slot duration T_s is set to 1 ms according to [31]. A single simulation runs the proposed algorithm when the train moves through a cell (30,000 slots).

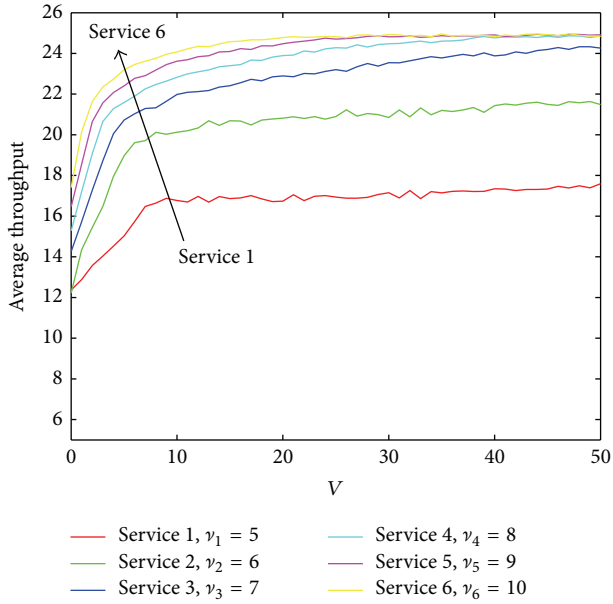


Figure 2: Average throughput with different V.

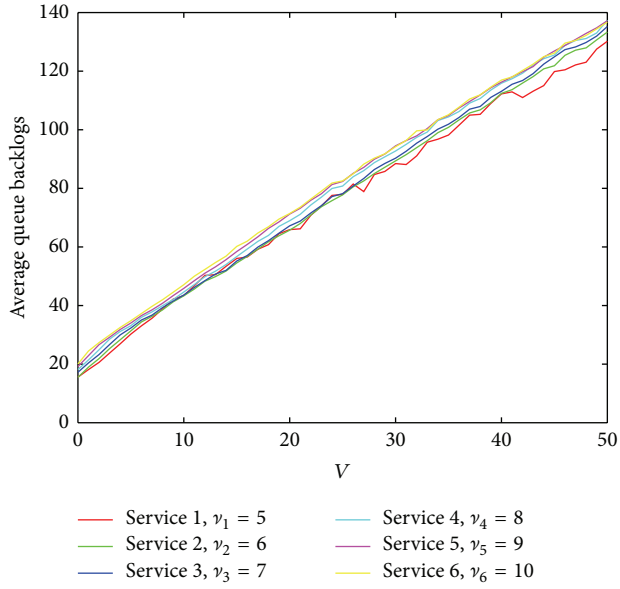


Figure 3: Average queue backlog with different V.

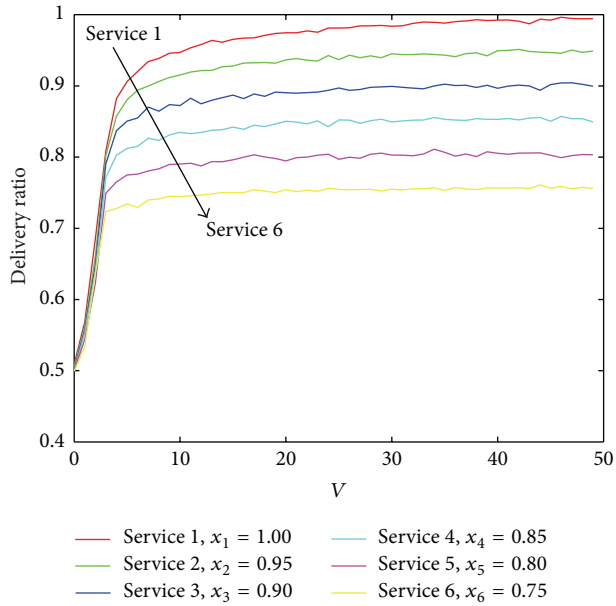


Figure 4: Delivery ratio with different V.

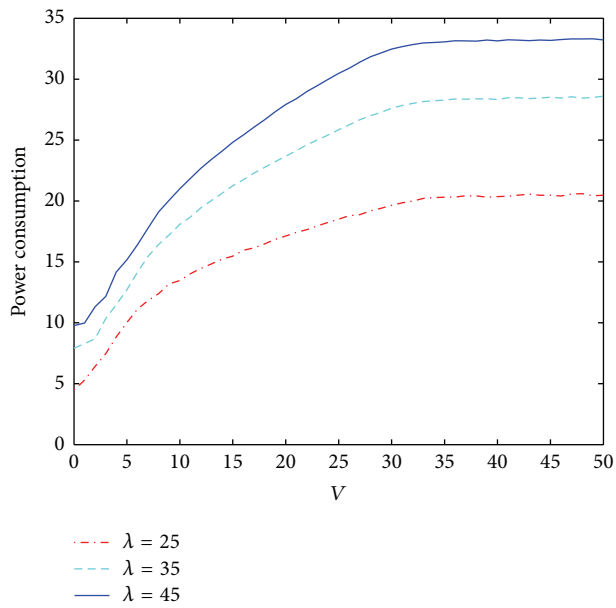


Figure 5: Power consumption with different V.

Figures 2 and 3 explore the throughput-backlog tradeoff with different V. In the simulations, we use the same parameters $\lambda_k = 25$ and $x_k = 1$ and different priorities for all services. As shown in Figure 2, the average throughput for each service increases as V is increased and the service with high priority gets the large average throughput. Figure 3 presents that the average queue backlogs of all services are linearly increasing with V, which demonstrates the $\mathcal{O}(V)$ behavior in (36). Furthermore, the proposed algorithm can ensure that the average queue backlogs of the services with different priorities are almost the same.

Figure 4 illustrates the achieved delivery ratios for the services with different maximum desired delivery ratios and same arrival rate $\lambda_k = 25$ as well as priority $\nu_k = 10$. It can be observed that the large V will result in the improvement of the delivery ratio performance. This can be explained as follows: since a larger V gives a higher priority on throughput, more packets will be admitted into the buffers, which causes the higher delivery ratio performance. In addition, the delivery ratio for each service is close to its maximum desired delivery ratio when V is larger than 40, which implies that the

proposed algorithm can archive different maximum desired delivery ratios when a large V is chosen.

Figure 5 compares the average power consumption under different arrival rate conditions. In this simulation, we set the same parameters $\nu_k = 10$, $x_k = 1$, and $\lambda_k = \lambda$ for all services.From this figure, we can see that the average power consumption increases as V is increased. This is exactly what happens. A larger V results in more packets admitted into buffers, while transmitting these packets will cost more power. As for the same V, a larger λ will cause more power consumption, since more packets will be admitted into buffers based on (26). In addition, the average power consumption can satisfy the average power constraint when the arrival rate is $\lambda_k = 25$, which is reasonably set in the previ-ous simulations.

Figure 6 describes the backlog update processes of virtual queues $Z_k(t)$ for three types of services. In the simulation, we set the same parameters $\nu_k = 10$, $x_k = 1$, and $\lambda_k = 35$ for all services and $V = 100$. From the figure, we can see $0 \le Z_k(t) \le V\nu_k + B_k$ for all services at all slots, which illustrates the boundedness property on queue backlogs in Lemma 7.

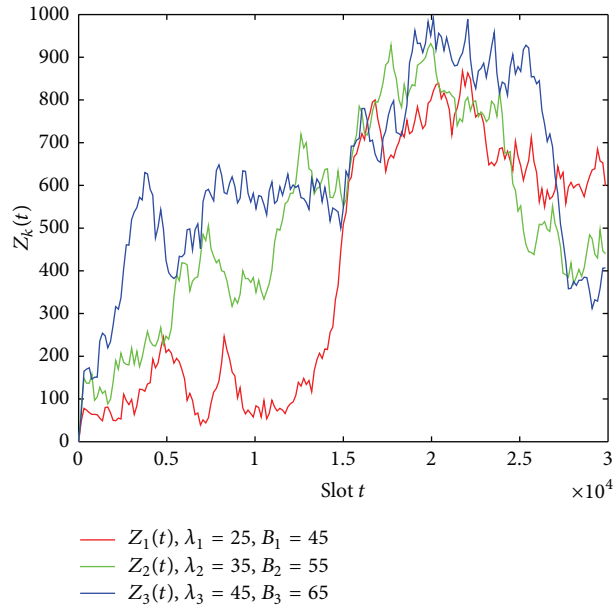


FIGURE 6: The backlog update processes of virtual queues $Z_k(t)$.

6. Conclusion

In this paper, we formulate the joint admission control and resource allocation problem under average power constraint for multimedia services delivery in HSR wireless networks. With the help of virtual queues, the original stochastic optimization problem is transformed into a queue stability problem, which is decomposed into three separate subproblems by the drift-plus-penalty approach. It is worth noting that the optimal solution to the resource allocation subproblem can be obtained by the packet loading resource allocation scheme. Based on the stochastic optimization technique, the dynamic admission control and resource allocation algorithm is proposed, which is suitable for distributed implementation in HSR wireless networks. Furthermore, the performance of the proposed algorithm is analyzed theoretically and validated by numerical simulations under realistic conditions for HSR wireless networks. For future work, we will further investigate the effects of the nonmentioned parts in the communication system, such as frame error check blocks and adaptive channel equalizers.

Appendices

A. Proof of Lemma 3

Let ϕ_1^* and ϕ_2^* be the optimal utility of problems (P1) and (P2), respectively.

First, to prove $\phi_1^* \ge \phi_2^*$, let $\alpha_2^*(t)$ be an optimal solution achieving ϕ_2^* in problem (P2), which includes $\gamma_k^*(t)$, $r_k^*(t)$, and $\mu_k^*(t)$ at slot t. Since $\phi(\cdot)$ is concave, based on Jensen's inequality, we have

$$\sum_{k\in\mathcal{K}} \phi\left(\overline{\gamma_k^*}\right) \ge \sum_{k\in\mathcal{K}} \overline{\phi(\gamma_k^*)} = \phi_2^*. \tag{A.1}$$

In addition, since the solution $\alpha_2^*(t)$ satisfies constraint (8b) and $\phi(\cdot)$ is nondecreasing, we further have

$$\sum_{k\in\mathcal{K}} \phi\left(\overline{r_k^*}\right) \ge \sum_{k\in\mathcal{K}} \phi\left(\overline{\gamma_k^*}\right). \tag{A.2}$$

Since the constraints in problem (P2) include all of the desired constraints of the original problem (P1), $\alpha_2^*(t)$ is a feasible solution for problem (P1) which gives a utility that is not larger than ϕ_1^*. Thus we conclude that

$$\phi_1^* \ge \sum_{k\in\mathcal{K}} \phi\left(\overline{r_k^*}\right) \ge \phi_2^*. \tag{A.3}$$

Next, to prove $\phi_1^* \le \phi_2^*$, let $\alpha_1^*(t)$ be an optimal solution achieving ϕ_1^* for problem (P1), which includes $r_k^*(t)$ and $\mu_k^*(t)$ at slot t. Since $\alpha_1^*(t)$ satisfies constraints (6b)–(6f), it also satisfies constraints (8d) of the problem (P2). Further, for all $k \in \mathcal{K}$, we set $\gamma_k(t) = \overline{r_k^*}$ at all time t which can satisfy constraints (8b) and (8c). Thus, such choice of $\gamma_k(t)$ together with the solution $\alpha_1^*(t)$ forms a feasible solution for the problem (P2). By definition, $\overline{\phi(\gamma_k)} = \lim_{t\to\infty}(1/t)\sum_{\tau=0}^{t-1}\phi(\gamma_k(\tau)) = \lim_{t\to\infty}(1/t)\sum_{\tau=0}^{t-1}\phi(\overline{r_k^*}) = \phi(\overline{r_k^*})$. Therefore, we get

$$\phi_2^* \ge \sum_{k\in\mathcal{K}} \overline{\phi\left(\gamma_k\right)} = \sum_{k\in\mathcal{K}} \phi\left(\overline{r_k^*}\right) = \phi_1^*. \tag{A.4}$$

From the above analysis, we can conclude that $\phi_1^* = \phi_2^*$ based on (A.3) and (A.4) and that an optimal solution for the problem (P2) can be directly turned into an optimal solution for the problem (P1).

B. Proof of Lemma 4

Recall the evolution equation (9) for the queue $X_k(t)$ and by squaring this equation, we obtain

$$X_k(t+1)^2 - X_k(t)^2 \tag{B.1a}$$

$$= \left(\max\left[X_k(t) - P_{\text{av}}, 0\right] + P(t)\right)^2 - X_k(t)^2 \tag{B.1b}$$

$$\leq \left[P(t)^2 + P_{\text{av}}^2\right] + 2X_k(t)\left[P(t) - P_{\text{av}}\right], \tag{B.1c}$$

where in the final inequality we have used the following facts: $(\max[X_k(t) - P_{\text{av}}, 0])^2 \leq (X_k(t) - P_{\text{av}})^2$ and $\max[X_k(t) - P_{\text{av}}, 0] \leq X_k(t)$.

Similarly, it can be shown for $k \in \mathcal{K}$ that

$$\begin{aligned} &Z_k(t+1)^2 - Z_k(t)^2 \\ &\quad \leq \left[\gamma_k(t)^2 + r_k(t)^2\right] + 2Z_k(t)\left[\gamma_k(t) - r_k(t)\right] \\ &Q_k(t+1)^2 - Q_k(t)^2 \\ &\quad \leq \left[r_k(t)^2 + \mu_k(t)^2\right] + 2Q_k(t)\left[r_k(t) - \mu_k(t)\right]. \end{aligned} \tag{B.2}$$

Based on (12) and (B.1a)–(B.2), we have

$$\Delta(\boldsymbol{\Theta}(t)) = \mathbb{E}\left[\frac{1}{2}\sum_{k\in\mathcal{K}}\left[X_k(t+1)^2 - X_k(t)^2 + Q_k(t+1)^2 - Q_k(t)^2 + Z_k(t+1)^2 - Z_k(t)^2\right] \mid \boldsymbol{\Theta}(t)\right] \tag{B.3a}$$

$$\leq \mathbb{E}\left[\frac{1}{2}\sum_{k\in\mathcal{K}}\left[P(t)^2 + P_{\text{av}}^2 + \gamma_k(t)^2 + r_k(t)^2 + r_k(t)^2 + \mu_k(t)^2\right] \mid \boldsymbol{\Theta}(t)\right] + \mathbb{E}\left[\sum_{k\in\mathcal{K}}\left[X_k(t)\left[P(t) - P_{\text{av}}\right] + Z_k(t)\left[\gamma_k(t) - r_k(t)\right] + Q_k(t)\left[r_k(t) - \mu_k(t)\right]\right] \mid \boldsymbol{\Theta}(t)\right] \tag{B.3b}$$

$$\leq D + \mathbb{E}\left[G(t) \mid \boldsymbol{\Theta}(t)\right], \tag{B.3c}$$

where $G(t)$ is defined by (15) and the last inequality can be obtained as follows. For any slot t, any possible packet arrival vector $\mathbf{A}(t)$, and any possible $P(t)$ as well as RA actions that can be taken, we have

$$\begin{aligned} &\mathbb{E}\left[\frac{1}{2}\sum_{k\in\mathcal{K}}\left[P(t)^2 + P_{\text{av}}^2 + \gamma_k(t)^2 + 2r_k(t)^2 + \mu_k(t)^2\right] \mid \boldsymbol{\Theta}(t)\right] \\ &\quad \leq \mathbb{E}\left[\frac{1}{2}\sum_{k\in\mathcal{K}}\left[P_{\max}^2 + P_{\text{av}}^2 + 3B_k^2 + C_{\max}^2\right] \mid \boldsymbol{\Theta}(t)\right] \\ &\quad = \frac{1}{2}\sum_{k\in\mathcal{K}}\left[P_{\max}^2 + P_{\text{av}}^2 + 3B_k^2 + C_{\max}^2\right] = D, \end{aligned} \tag{B.4}$$

where the inequality holds based on $P(t) \leq P_{\max}$, $r_k(t) \leq B_k$, and $\mu_k(t) \leq C_{\max}$ as well as (8c), and the equality holds since the constant in the square bracket is independent of queue vector $\boldsymbol{\Theta}(t)$ at slot t.

C. Proof of Lemma 7

We prove this lemma by induction. Assume that $Z_k(t) \leq V\nu_k + B_k$ for slot t (it holds by assumption at slot $t = 0$); then we prove it also holds for slot $t + 1$. Firstly, we consider the case $Z_k(t) \leq V\nu_k$. From the queue update equation (7), we can see that this queue can increase by at most B_k at each slot, and thus we have $Z_k(t+1) \leq V\nu_k + B_k$, proving the result for this case.

Secondly, we consider the case $V\nu_k < Z_k(t) \leq V\nu_k + B_k$. For each slot t, CS_k decides $\gamma_k(t)$ to maximize the following expression:

$$V\phi\left(\gamma_k(t)\right) - Z_k(t)\gamma_k(t). \tag{C.1}$$

Based on the property of the maximum derivative, for any $\gamma_k(t) \geq 0$, we obtain

$$V\phi\left(\gamma_k(t)\right) - Z_k(t)\gamma_k(t) \leq V\phi(0) + V\nu_k\gamma_k(t) - Z_k(t)\gamma_k(t) \tag{C.2a}$$

$$= V\phi(0) + \gamma_k(t)\left[V\nu_k - Z_k(t)\right] \tag{C.2b}$$

$$\leq V\phi(0), \tag{C.2c}$$

where the equality holds if and only if $\gamma_k(t) = 0$. Then the algorithm will choose $\gamma_k(t) = 0$ to maximize expression (C.1), and we can obtain

$$Z_k(t+1) = \max\left[Z_k(t) - r_k(t), 0\right] \leq Z_k(t) \leq V\nu_k + B_k. \tag{C.3}$$

Thus, $Z_k(t+1) \leq V\nu_k + B_k$ is satisfied for these two cases, which completes the proof.

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 the Fundamental Research Funds for the Central Universities (Grant no. 2014YJS026), the Key Projects of State Key Lab of Rail Traffic Control and Safety (no. RCS2012ZZ004 and no. RCS2010ZT011), the Opening Project of The State Key Laboratory of Integrated Services Networks, Xidian University (Grant no. ISN14-09), and the Key Grant Project of Chinese Ministry of Education (no. 313006).

References

[1] G. Barbu, "E-Train—broadband communication with moving trains," Tech. Rep., Technology State of the Art, June 2010.

[2] L. Tian, J. Li, Y. Huang, J. Shi, and J. Zhou, "Seamless dual-link handover scheme in broadband wireless communication systems for high-speed rail," *IEEE Journal on Selected Areas in Communications*, vol. 30, no. 4, pp. 708–718, 2012.

[3] J. Wang, H. Zhu, and N. J. Gomes, "Distributed antenna systems for mobile communications in high speed trains," *IEEE Journal on Selected Areas in Communications*, vol. 30, no. 4, pp. 675–683, 2012.

[4] Y. S. Zhao, X. Li, Y. Li, and H. Ji, "Resource allocation for high-speed railway downlink MIMO-OFDM system using quantum-behaved particle swarm optimization," in *Proceedings of the IEEE International Conference on Communications*, pp. 936–940, 2013.

[5] W. Luo, X. Fang, M. Cheng, and Y. Zhao, "Efficient multiple-group multiple-antenna (MGMA) scheme for high-speed railway viaducts," *IEEE Transactions on Vehicular Technology*, vol. 62, no. 6, pp. 2558–2569, 2013.

[6] L. Georgiadis, M. J. Neely, and L. Tassiulas, "Resource allocation and cross-layer control in wireless networks," *Foundations and Trends in Networking*, vol. 1, no. 1, pp. 1–144, 2006.

[7] K. T. Phan, T. Le-Ngoc, M. van ser Schaar, and F. Fu, "Optimal scheduling over time-varying channels with traffic admission control: structural results and online learning algorithms," *IEEE Transactions on Wireless Communications*, vol. 12, no. 9, pp. 4434–4444, 2013.

[8] Y. Q. Dong, P. Y. Fan, and K. B. Letaief, "High speed railway wireless communications: efficiency v.s. fairness," *IEEE Transactions on Vehicular Technology*, vol. 63, no. 2, pp. 925–930, 2013.

[9] D. Pareit, E. V. de Velde, D. Naudts, and J. Bergs, "A novel network architecture for train-to-wayside communication with quality of service over heterogeneous wireless networks," *EURASIP Journal on Wireless Communications and Networking*, vol. 2012, article 114, 2012.

[10] O. B. Karimi, J. Liu, and C. Wang, "Seamless wireless connectivity for multimedia services in high speed trains," *IEEE Journal on Selected Areas in Communications*, vol. 30, no. 4, pp. 729–739, 2012.

[11] Y. S. Zhao, X. Li, Z. X. Liang, Y. Li, and H. Ji, "Multidimensional resource allocation strategy for high-speed railway MIMO-OFDM system," in *Proceedings of the IEEE Global Communications Conference (Globecom '12)*, pp. 1653–1657, 2012.

[12] H. Liang and W. Zhuang, "Efficient on-demand data service delivery to high-speed trains in cellular/infostation integrated networks," *IEEE Journal on Selected Areas in Communications*, vol. 30, no. 4, pp. 780–791, 2012.

[13] S. F. Xu, G. Zhu, C. Shen, and Y. Lei, "Delay-aware fair scheduling in relay-assisted high-speed railway networks," in *Proceedings of the International Conference on Communications and Networking in China*, pp. 17–21, 2013.

[14] Y. S. Zhao, X. Li, and H. Ji, "Radio admission control scheme for high-speed railway communication with MIMO antennas," in *Proceedings of the IEEE International Conference on Communications*, pp. 5005–5009, 2012.

[15] M. J. Neely, *Stochastic Network Optimization with Application to Communication and Queueing Systems*, Morgan & Claypool, 2010.

[16] S. Meyn and R. L. Tweedie, *Markov Chains and Stochastic Stability*, Cambridge University Press, Cambridge, UK, 2nd edition, 2009.

[17] Y. Cui, V. K. N. Lau, R. Wang, H. Huang, and S. Zhang, "A survey on delay-aware resource sontrol for wireless systems-large deviation theory, stochastic Lyapunov drift, and distributed stochastic learning," *IEEE Transactions on Information Theory*, vol. 58, no. 3, pp. 1677–1701, 2012.

[18] L. Tassiulas and A. Ephremides, "Stability properties of constrained queueing systems and scheduling policies for maximum throughput in multihop radio networks," *IEEE Transacations on Automatic Control*, vol. 37, no. 12, pp. 1936–1948, 1992.

[19] B. L. Niu, V. W. S. Wong, and R. Schober, "Downlink scheduling with transmission strategy selection for multi-cell MIMO systems," *IEEE Transactions on Wireless Communications*, vol. 12, no. 2, pp. 736–747, 2013.

[20] Z. Zhou, F. M. Liu, H. Jin, B. Li, B. C. Li, and H. B. Jiang, "On arbitrating the power-performance tradeoff in SaaS clouds," in *Proceedings of the IEEE INFOCOM*, pp. 872–880, 2013.

[21] S. Spagna, M. Liebsch, and R. Baldessari, "Design principles of an operator-owned highly distributed content delivery network," *IEEE Communications Magazine*, vol. 51, no. 4, pp. 132–140, 2013.

[22] L. Liu, C. Tao, J. Qiu et al., "Position-based modeling for wireless channel on high-speed railway under a viaduct at 2.35 GHz," *IEEE Journal on Selected Areas in Communications*, vol. 30, no. 4, pp. 834–845, 2012.

[23] R. He, Z. Zhong, B. Ai, and J. Ding, "An empirical path loss model and fading analysis for high-speed railway viaduct scenarios," *IEEE Antennas and Wireless Propagation Letters*, vol. 10, pp. 808–812, 2011.

[24] Y. Dong, Q. Wang, P. Fan, and K. B. Letaief, "The deterministic time-linearity of service provided by fading channels," *IEEE Transactions on Wireless Communications*, vol. 11, no. 5, pp. 1666–1675, 2012.

[25] C. E. Shannon, "A mathematical theory of communication," *ACM SIGMOBILE Mobile Computing and Communications Review*, vol. 5, no. 1, pp. 3–55, 2001.

[26] D. Bethanabhotla, G. Caire, and M. J. Neely, "Utility optimal scheduling and admission control for adaptive video streaming in small cell networks," in *Proceedings of the IEEE International Symposium on Information Theory (ISIT '13)*, 2013.

[27] G. L. Nemhauser and L. A. Wolsey, *Integer and Combinatorial Optimization*, John Wiley & Sons, New York, NY, USA, 1988.

[28] E. M. J. Bertin and R. Theodorescu, "Some characterizations of discrete unimodality," *Statistics and Probability Letters*, vol. 2, no. 1, pp. 23–30, 1984.

[29] S. L. S. Jacoby, J. S. Kowalik, and J. T. Pizzo, *Iterative Methods for Nonlinear Optimization Problems*, Prentice-Hall, 1972.

[30] D. H. Ho and S. Valaee, "Information raining and optimal link-layer design for mobile hotspots," *IEEE Transactions on Mobile Computing*, vol. 4, no. 3, pp. 271–284, 2005.

[31] R. Atat, E. Yaacoub, M. Alouini, and A. Dayya, "Heterogeneous LTE/802. 11a mobile relays for data rate enhancement and energy-efficiency in high speed trains," in *Proceedings of the IEEE Global Communications Conference (Globecom '12)*, pp. 421–425, 2012.

Hindawi Publishing Corporation
Mathematical Problems in Engineering
Volume 2014, Article ID 981845, 14 pages
http://dx.doi.org/10.1155/2014/981845

Research Article

The Development and Verification of a Novel ECMS of Hybrid Electric Bus

Jun Wang, Qing-nian Wang, Peng-yu Wang, and Xiao-hua Zeng

State Key Laboratory of Automotive Simulation and Control, Jilin University, Changchun 130025, China

Correspondence should be addressed to Peng-yu Wang; wpy123456789000@sohu.com

Received 20 December 2013; Revised 16 January 2014; Accepted 21 January 2014; Published 18 March 2014

Academic Editor: Ali Davoudi

Copyright © 2014 Jun 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 presents the system modeling, control strategy design, and hardware-in-the-loop test for a series-parallel hybrid electric bus. First, the powertrain mathematical models and the system architecture were proposed. Then an adaptive ECMS is developed for the real-time control of a hybrid electric bus, which is investigated and verified in a hardware-in-the-loop simulation system. The ECMS through driving cycle recognition results in updating the equivalent charge and discharge coefficients and extracting optimized rules for real-time control. This method not only solves the problems of mode transition frequently and improves the fuel economy, but also simplifies the complexity of control strategy design and provides new design ideas for the energy management strategy and gear-shifting rules designed. Finally, the simulation results show that the proposed real-time A-ECMS can coordinate the overall hybrid electric powertrain to optimize fuel economy and sustain the battery SOC level.

1. Introduction

In recent years, vehicle fuel consumption and air pollution emissions have attracted growing attention. In order to solve these problems, a tremendous amount of effort is directed toward hybrid power vehicle's driving systems that have a significant potential in fuel saving and emissions reduction. Meanwhile, a large number of hybrid electric vehicles have become available in the markets, which were considered to be the most promising vehicles to replace conventional engine-driven vehicles.

The improvements in fuel economy and the reductions in emissions of hybrid electric vehicles (HEV) mainly depend upon the energy management strategy (EMS); therefore, substantial research efforts have been carried out. The research methods can be classified into three categories: first, rule-based controls such as logic threshold control and finite state machine [1, 2]; second, intelligent control algorithms such as model predictive control [3], fuzzy logic [4, 5], and neural networks [6]; third, optimal theory methods such as minimum theory, deterministic dynamic programming (DP) [7, 8], and stochastic DP (SDP) [9].

Rule-based control strategy is also named as the baseline control, which is a steady state optimization method through engineering experience and a simple analysis of the efficiencies of components such as engine, motor, and battery.

Intelligent control algorithm depends on experts' knowledge to be coded into control rules, and this method has good robustness and does not need to build complex control model. However, both the rule-based control and the fuzzy logic control strategy need to be predetermined and can only be optimized for a specific drive cycle [10–12].

Optimal energy management control strategy includes DP, SDP, and Pontryagin minimum principle (PMP) [13]. DP control algorithm is often used to obtain global optimal solutions for various types of HEV under the certain drive cycles, but it should know the future drive cycle information in advance. Therefore, DP control algorithm is impossible to be applicable in real-time control system, which is often used as a reference for energy management strategy design [14]. In order to optimize the torque distribution in real-time, SDP control algorithm has been proposed, which applies the current road and traffic information to predict future

Table 1: Main specifications of the vehicle.

Item	Value
Curb/gross weight (kg)	11500/18000
Frontal area (m^2)	7.21
Drag coefficient	0.585
Wheel base (m)	6.1
Dynamic rolling radius (m)	0.512

route information for EMS so as to acquire the ideal results. However, trip-based control approach is a major drawback of curse of dimensionality, which limits their application for real-time implementation.

The equivalent consumption minimization strategy (ECMS) is the most commonly used optimization method for real-time HEV energy management [15–17]. It is considered as suboptimal control method for HEV, since the fuel economy deviation between ECMS and DP control algorithm was verified to be less than 1.2% [18]. Therefore, in this study, ECMS is chosen as the online energy management strategy to optimize torque distribution and gear-shift rules.

Then the novel ECMS was verified in the hardware-in-the-loop (HIL) simulator for its high real-time performance and high precision. Not only can the HIL simulation verify the effectiveness of the control strategy, but it can also optimize its control parameters [19, 20]. Hardware-in-the-loop simulation is the trusted method to put ECU functions, ECU bus communication, and integrated ECUs to the test. The tests are performed in a simulated environment, meaning that the HIL simulator makes the ECU believe that it is located in a real vehicle driving somewhere. This way the simulator can test the ECU's reaction to specific situations, and you can move tedious, expensive, and sometimes even dangerous driving tests from the actual vehicle into the laboratory [21]. In this way, the control functions and performance of the HCU of the series-parallel hybrid bus can be tested and improved.

2. System Architecture

The HEV control algorithm is to regulate the operation of powertrain system to achieve the optimal fuel economy and emissions performance under the different drive cycles, while meeting drivability requirements and sustaining the battery SOC level [22]. Moreover, the development of HEV control algorithm is based on the specific powertrain configuration. Basic parameters of the hybrid electric bus (HEB) are listed in Table 1.

The proposed HEV powertrain consists of a 4-cylinder Deutz diesel engine, an M1 motor, an M2 motor, an automated mechanical transmission (AMT), a torque coupler, and an electronically controlled clutch. The main specifications of the powertrain components are listed in Table 2. In this powertrain, the engine is directly connected to the input shaft of the clutch, the M1 motor is attached to the input shaft of the AMT, and the M2 motor couples the shaft via the torque coupler at a constant gear ratio. The novel hybrid driving system is to coordinate engine output power by M2 motor in the low-speed conditions and to coordinate engine output power by M1 motor in the high-speed conditions, so as to improve the fuel economy and keep dynamic performance. According to the structural characteristics of hybrid driving system, the drive mode can be divided into pure electric mode, series mode, engine-driven alone, parallel mode 1 (engine and M2 motor combined driving), and parallel mode 2 (engine and M1 motor combined driving). This series-parallel HEB powertrain was shown in Figure 1.

This paper provided adaptive equivalent minimum fuel consumption control algorithm through driving cycle recognition to update the equivalent fuel consumption coefficient and then to get the optimal torque distribution and gear-shift rules. Figure 2 describes the TCU and HCU integration control flowchart.

T_{req} is the torque requirement, V_{veh} is the vehicle actual velocity, V_{tar} is the target velocity, T_e is the engine torque, T_{m1} is the M1 motor torque, T_{m2} is the M2 motor torque, $i_{g(n)}$ is the transmission ratio, and R_{cluth} is the clutch signal.

3. Powertrain Models

3.1. Engine Model. The engine operation status is very complicated which makes it hard to establish a precise simulation model by maths model and theoretical formula. In this paper, the engine model is simplified as static maps that are used to calculate the engine torque output and the fuel consumption. But the engine BSFC map may only give steady-state fuel consumption when engine is operating at normal temperature, since the engine may often operate in transient conditions that would consume more fuel than steady-state engine fuel consumption. So it should be corrected to obtain the transient-state fuel consumption according to the experimental data [23]:

$$\begin{aligned} T_e &= T_e\left(a_e, \omega_e\right) - T_{\text{fl}}\left(\omega_e\right), \\ g_e &= f_{\text{fuel}}\left(T_e, \omega_e\right), \\ g_{e,f} &= g_e\left[1 + f_1\left(t_{em}\right) + f_2'\left(\alpha_e'\right)\right], \end{aligned} \tag{1}$$

where a_e is the actual load of engine, ω_e is the engine speed, T_{fl} is the engine resistant torque, g_e is the engine steady fuel consumption, f_1 is the fuel consumption increase rate due to engine warm-up, f_2' is the fuel consumption increase rate due to the engine load increasing rate, t_{em} is the engine operating temperature, α_e' is the engine load increasing rate, and $g_{e,f}$ is the engine transient fuel consumption.

3.2. Electric Motor Model. Electric motor is used to provide electric propelling power and energy recycling for fuel economy improvement. This paper mainly concerns motor external characteristics: maximum motor torque, minimal motor torque, and efficiency map. A simple lookup table is

Table 2: Vehicle basic parameters.

Powertrain	Parameters	Product model
Engine power rating	121 kW/2500 r/min	Deutz, BF4M1013FC
M1 motor rated power/peak power	30/50 kW	Shanghai eDrive TYC-168-260-8-C
M2 motor rated power/peak power	58/116 kW	ENOVA, M10000DA
Battery voltage/capacity	336 V/80 Ah	Chunlan, DY336-40
Six speed gearbox transmission ratios	7.285, 4.193, 2.485, 1.563, 1, and 0.847	FAW, CA5-85 AMT
Main reducer gear ratio	6.333	—
Coupler gear ratio	3.8	—

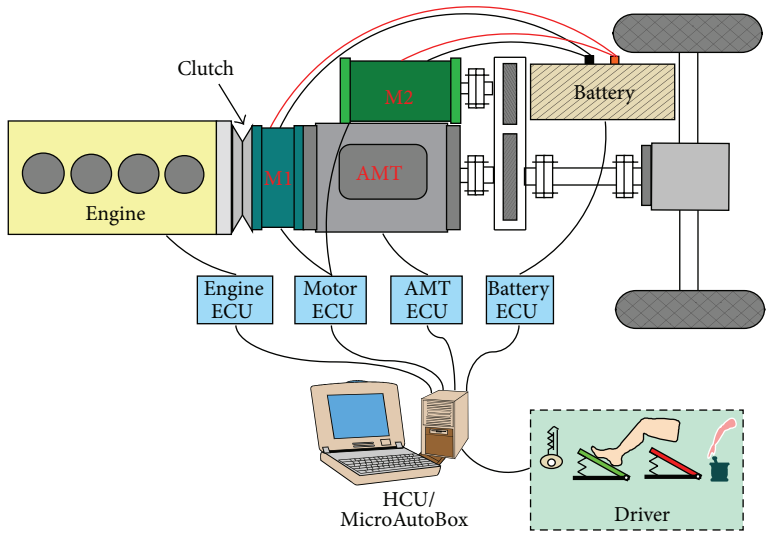


Figure 1: Schematic of a series-parallel HEB powertrain.

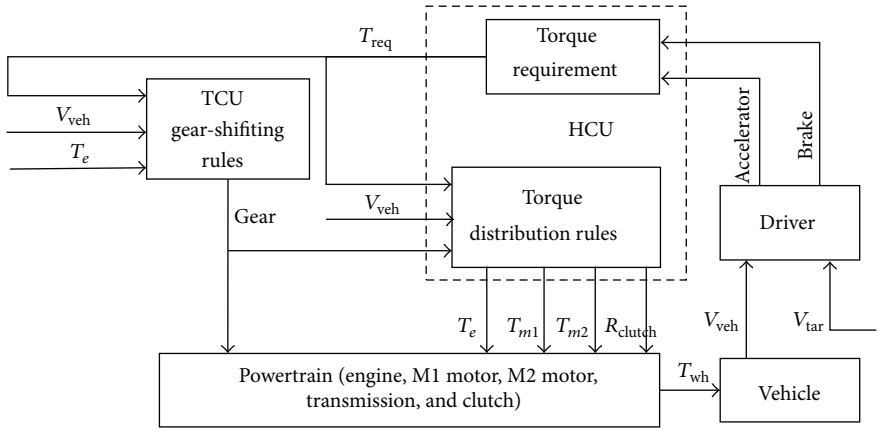


Figure 2: The integrated control flowchart of gear-shifting rules and energy management.

adopted to characterize the motor efficiency depending on the motor speed and torque and terminal voltage [24]:

$$T_m = T_{m\_\max}(n_m, U_m) \cdot a_m,$$

$$\eta_m = \eta(n_m, T_m),$$

$$P_m = \begin{cases} \dfrac{T_m \cdot n_m}{\eta_m}; & (T_m \geq 0) \\ T_m \cdot n_m \cdot \eta_m; & (T_m < 0), \end{cases}$$

$$I_m = \frac{P_m}{U_m}, \tag{2}$$

where $T_{m\_\max}$ is the motor maximum torque, a_m represents percentage of full load torque, n_m, T_m, and U_m are the motor speed, motor torque, and motor's terminal voltage, η_m is the motor efficiency, I_m is the motor current, and P_m is the motor power.

3.3. Battery Model. There were a lot of battery modeling approaches that have been introduced; among them, PNGV battery model [25] can accurately simulate the dynamic characteristics of the battery, but this modeling focus is the powertrain dynamics in a system level rather than the internal battery dynamics; hence the Rint model [26] was used to calculate the battery open-circuit voltage, the internal resistance, and the battery SOC.

$$\begin{aligned} U_b &= U_0\,(\mathrm{SOC}) \cdot N_b, \\ R_b &= R_0\,(\mathrm{SOC}) \cdot N_b, \\ \mathrm{SOC} &= \mathrm{SOC}_0 + \frac{1}{3600 \cdot C}\int_{t_0}^{t} I_b dt, \\ I_b &= \frac{U_b - \sqrt{U_b^2 - 4R_bP_b}}{2R_b}, \end{aligned} \tag{3}$$

where N_b is the number of battery cells , U_b is the battery open-circuit voltage, R_b is the battery internal resistance, I_b is the battery current, P_b is the battery power, SOC is the battery state of charge, SOC_0 is the battery initial SOC, and C is the battery capacity.

3.4. Clutch Model. The clutch model is composed of three modes: sliding mode, engagement mode, and disengagement mode. The operating mode of the clutch is determined by the displacement of clutch release S_c and the relative speed of the clutch $n_{c,\mathrm{rel}}$. The transmitted torque of the clutch T_{cl} can be calculated by the following equation [27]:

$$T_{\mathrm{cl}} = \begin{cases} -F_{C,\mathrm{act}} \cdot N_C \cdot \mu_{C,\mathrm{st}} \cdot r_{c,m}; \\ \qquad ((S_c < 1)\,\&\&\,(n_{c,\mathrm{rel}} = 0)) \\ F_{C,\mathrm{act}} \cdot N_C \cdot r_{c,m} \cdot \mathrm{sign}\,(n_{c,\mathrm{rel}}) \\ \quad \cdot \left[\mu_{C,\mathrm{sl}} + (\mu_{C,\mathrm{st}} - \mu_{C,\mathrm{sl}}) \cdot e^{(|n_{c,\mathrm{rel}}| \cdot C_C/(\mu_{C,\mathrm{st}} - \mu_{C,\mathrm{sl}}))}\right]; \\ \qquad ((S_c < 1)\,\&\&\,(n_{c,\mathrm{rel}} \neq 0)) \\ 0; \quad (S_c = 1), \end{cases} \tag{4}$$

where $S_c = 1$ is the disengagement mode, $(S_c < 1)\&\&(n_{c,\mathrm{rel}} \neq 0)$ is the sliding mode, $(S_c < 1)\&\&(n_{c,\mathrm{rel}} = 0)$ is the engagement mode, $\mu_{C,\mathrm{sl}}$ is the sliding friction coefficient of the clutch, $\mu_{C,\mathrm{st}}$ is the static friction coefficient of the clutch, C_C is the friction gradient, $r_{c,m}$ is the effective friction radius, and N_C is the number of friction surfaces .

3.5. AMT Model. Manual transmission has the highest overall efficiency and the simplest structure among all types of transmissions. AMT has a similar efficiency to manual transmission; it is essentially a manual transmission with an add-on electronic control unit that automates the gear-shifting operations. There are three operation states defined in the AMT model, neutral state, speed testing state, and engaged state. Speed testing state is adopted to check the minimal time duration when the speed discrepancy is under a certain threshold that is used to avoid gear shifting frequently.

If the gear is engaged,

$$\begin{aligned} \omega_I &= \omega_{II} i_{g(G_{\mathrm{cur}})}, \\ T_{io} &= T_e + T_{m1}, \\ T_{o,\mathrm{tran}} &= T_{io} i_{g(G_{\mathrm{cur}})} \eta_{g(G_{\mathrm{cur}})} + T_{m2} i_{m2}, \end{aligned} \tag{5}$$

where ω_{II}, ω_I are the output shaft speed and the input shaft speed, $i_{g(G_{\mathrm{cur}})}$ is the gear ratio of current gear number, i_{m2} is the main reducer gear ratio, $T_{o,\mathrm{tran}}$ is the torque of output shaft, and $\eta_{g(G_{\mathrm{cur}})}$ is the gear efficiency.

The neutral state and speed testing state are actually the same physical state; their dynamics can be described by the following equations:

$$\begin{aligned} \omega_I\,(t) &= \omega_{I0} + \int_{t_0}^{t} \frac{T_{io}}{J_e + J_{m1} + J_c} dt, \\ T_{o,\mathrm{tran}} &= T_{m2} i_{m2}, \end{aligned} \tag{6}$$

where ω_{I0} is the speed of input shaft when gear changed from being engaged to neutral at the last time. J_e, J_{m1}, and J_c are the moment of inertia of engine, M1 motor, and clutch.

3.6. Driveline and Vehicle Model. Consider

$$F_{\mathrm{drv}} = \frac{T_t i_0 \eta - T_{\mathrm{br}}}{r_{\mathrm{wh}}} - F_{\mathrm{loss}}, \tag{7}$$

where T_t is the driving torque, F_{drv} is the joint force applied on vehicle, i_0 and η are the gear ratio and efficiency of efficiency of the drivetrain. T_{br} is the mechanical brake torque applied on wheels, r_{wh} is the wheel rolling radius, and F_{loss} is the vehicle resistance described by a function of the vehicle speed, the road grade, and the other vehicle parameters, respectively [28]:

$$F_{\mathrm{loss}} = Mg \sin\theta + \mu Mg \cos\theta + 0.5\rho C_d A_f V^2, \tag{8}$$

where M is the vehicle mass, g is the acceleration of gravity, ρ is the air density, C_d is the coefficient of air drag, A_f is the frontal area, V is the vehicle speed, θ is the road grade, and μ is the rolling friction coefficient.

4. Adaptive Online-Optimal Controller Design

The equivalent consumption minimization strategy (ECMS) is the most commonly used optimization method for real-time HEV energy management. It is considered as suboptimal control method for HEV control problems. Therefore, in this study, ECMS is chosen as the online energy management strategy.

This control method can not only coordinate control of gear shifting and motor assist, but also simplify

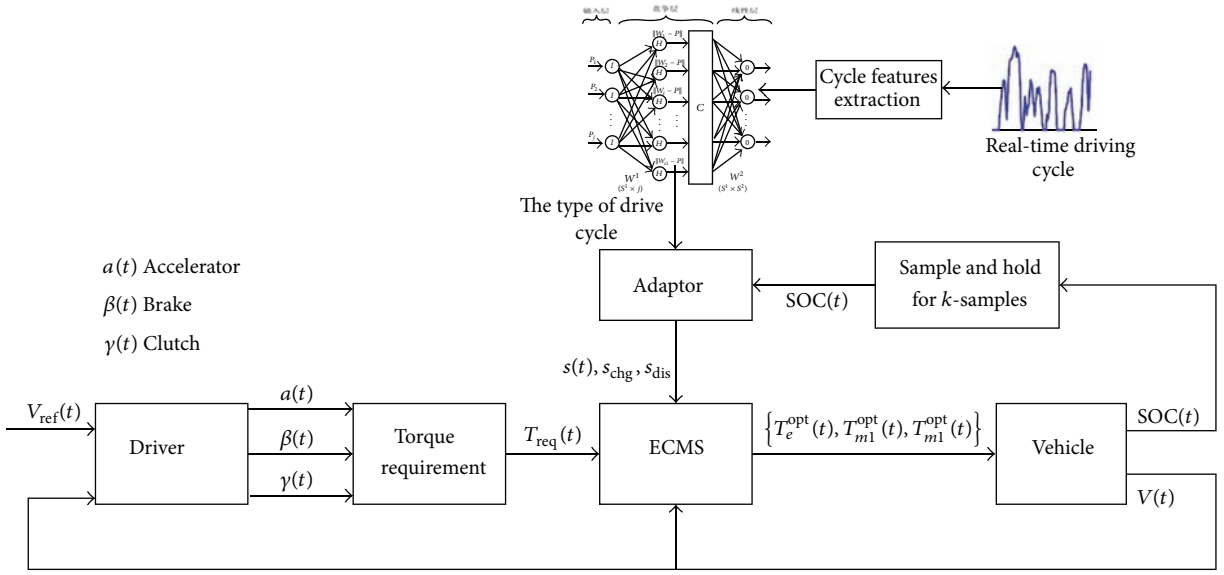


Figure 3: The control flowchart of adaptive ECMS (A-ECMS).

the complexity of the torque distribution in parallel mode. The optimizing searching method was adopted to obtain optimal torque distribution rules and gear-shift rules for real-time control. Nevertheless, the equivalent fuel consumption factor is one of the most important influencing factors, so a large number of scholars have conducted in-depth discussions relevant to equivalent fuel conversion factor. The control flowchart of this paper was described in Figure 3.

4.1. Equivalent Fuel Consumption Factor. Equivalent fuel consumption factor is the key for the implementation of ECMS control strategy. Because motor driving power consumed should be supplied by engine fuel consumed, the energy stored in the battery cannot be directly converted from the energy of fuel. So the fixed conversion factor or adaptive equivalent fuel consumption coefficient was proposed; the fixed conversion factor can be acquired by SAEJ1711 standards equivalent factor (33440 Wh/gal), and this method is too simple that it ignores internal resistance of the battery energy consumption, motor operating efficiency, and so on [29]. The adaptive equivalent fuel consumption coefficient calculation method has fairly ideal performance; Zhao and Stobart [30] presented minimum equivalent fuel consumption based on fuzzy control, considering the battery SOC and motor instantaneous power to acquire the real-time optimal conversion coefficient, but the computation quantity of this algorithm is still too large that it is difficult to apply in the real vehicle. In this study, a simple and accurate adaptive equivalent fuel consumption coefficient calculation method was proposed:

$$g_m = \lambda \cdot s(t) \cdot s_{\text{dis}} \frac{P_m}{\overline{\eta}_m(P_m) \cdot \overline{\eta}_{\text{batt}}(P_m) \cdot H_{\text{LHV}}} + (1-\lambda) \cdot s(t) \cdot s_{\text{chg}} \cdot \overline{\eta}_m(P_m) \cdot \overline{\eta}_{\text{batt}}(P_m) \frac{P_m}{H_{\text{LHV}}},$$

$$\lambda = \frac{1 + \text{sign}(P_m)}{2}, \tag{9}$$

where $H_{\text{LHV}} = 42.6 \times 10^3$ J/g is the diesel combustion characteristic parameters.

$\overline{\eta}_{\text{batt}}$, $\overline{\eta}_m$, respectively, are the average operating efficiency of engine and the motor under the specific drive cycle:

$$\overline{\eta}_{\text{batt}} = \frac{\sum_{j=1}^{k} (\eta_{\text{batt}}(j) \cdot P_m(j))}{\sum_{j=1}^{k} P_m(j)}, \qquad \overline{\eta}_m = \frac{\sum_{j=1}^{k} (\eta_m(j) \cdot P_m(j))}{\sum_{j=1}^{k} P_m(j)}, \tag{10}$$

where K is the number of sampling points and η_m, η_{batt}, P_m, and P_{batt} are the motor efficiency, battery efficiency, motor power, and battery power, respectively.

Battery SOC equivalent conversion factors are

$$s(t) = s_0 + K_p \cdot \left(\text{SOC}_{\text{obj}} - \text{SOC}(t)\right). \tag{11}$$

Target battery $\text{SOC}_{\text{obj}} = 0.6$, s_0, K_p adjusts the parameter, $s_0 = 1$, and $K_p = 1.82$.

s_{chg}, s_{dis} are the equivalent charge and discharge coefficients.

4.2. Objective Function. The energy management controller aims to obtain minimum fuel consumption by coordinated

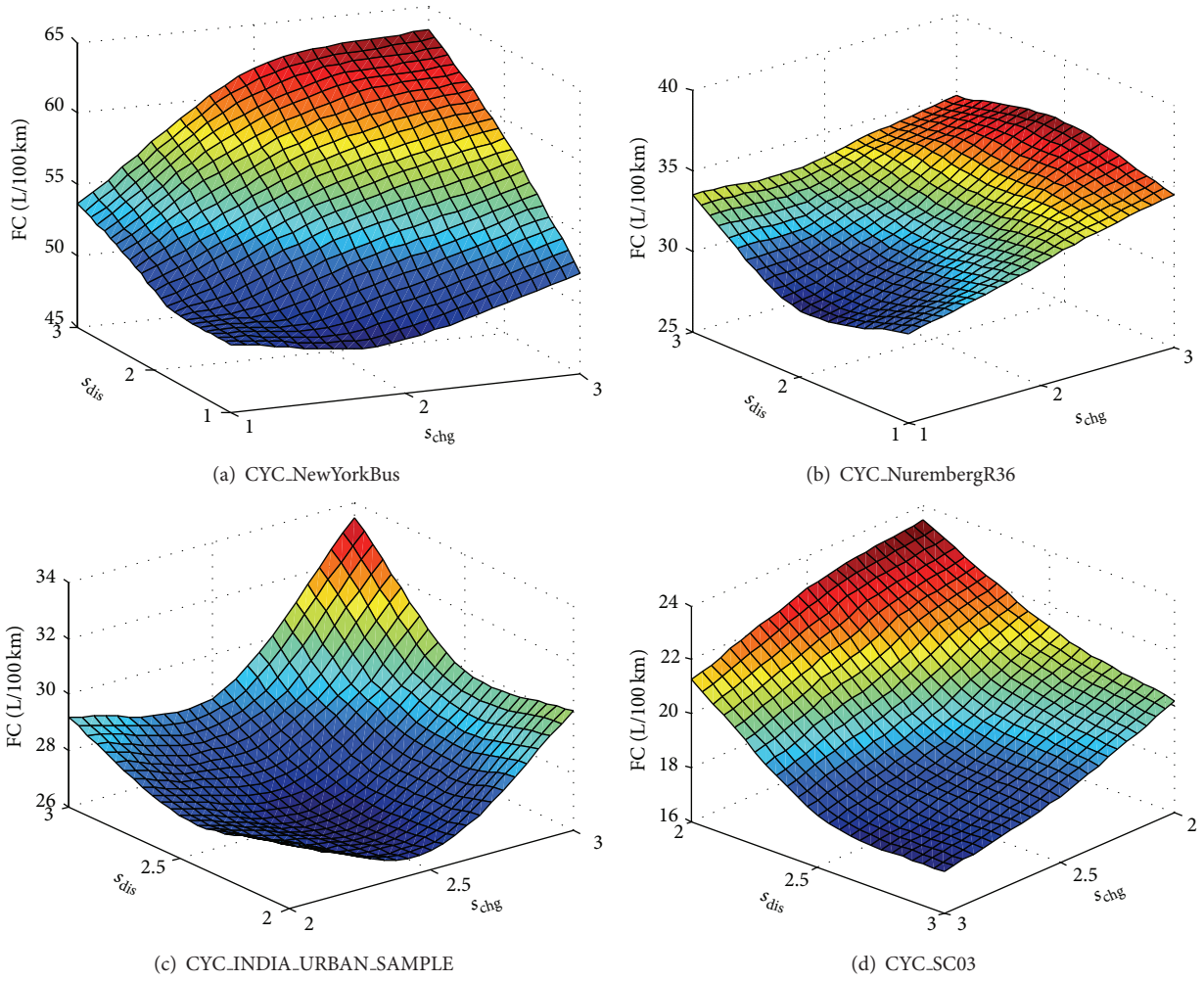


Figure 4: Equivalent charge and discharge coefficients with fuel economy.

control of engine power, M1 motor power, M2 motor power, and transmission speed position:

$$J = g_e + g_{m1} + g_{m2},$$

$$\left\{P_e^{\text{opt}}(t), P_{m1}^{\text{opt}}(t), P_{m2}^{\text{opt}}(t), G_{\text{ear}}^{\text{opt}}\right\} = \arg \min_{\{P_e(t), P_{m1}(t), P_{m2}(t)\}} J\left(P_{\text{req}}(t) \geq 0\right), \tag{12}$$

where g_e is the engine instantaneous fuel consumption, g_{m1} is the M1 motor equivalent instantaneous fuel consumption, g_{m2} is the M2 motor equivalent instantaneous fuel consumption, P_e^{opt} is the optimal engine power output, P_{m1}^{opt} is the optimal M1 motor power output, P_{m2}^{opt} is the optimal M2 motor power output, $G_{\text{ear}}^{\text{opt}}$ is the optimal gear, and P_{req} is the power requirement.

4.3. Constraints. In order to extend the battery life and maximum use of charge and discharge power in the reasonable range, they need to limit battery SOC within a certain range:

$$\begin{aligned} P_{\text{req}}(t) &= P_e(t) + P_{m1}(t) + P_{m2}(t), \\ \text{SOC}_{\min} < \text{SOC}(t) &< \text{SOC}_{\max}, \\ 0 \leq P_e(t) &\leq P_{e\_\max}(t), \\ P_{m1\_\min}(t) \leq P_{m1}(t) &\leq P_{m1\_\max}(t), \\ P_{m2\_\min}(t) \leq P_{m2}(t) &\leq P_{m2\_\max}(t). \end{aligned} \tag{13}$$

4.4. Adaptive Equivalent Charge and Discharge Coefficients. According to the simulation results demonstrate that different drive cycles correspond to different optimal equivalent charge and discharge coefficients. By choosing different charge and discharge coefficients to simulate calculation,

Table 3: Typical driving cycle optimal coefficients.

Typical driving cycle	S_{chg}	S_{dis}
CYC_NewYorkBus	2.25	2.13
CYC_NurembergR36	2.39	2.19
CYC_INDIA_URBAN_SAMPLE	2.74	2.45
CYC_SC03	2.92	2.67

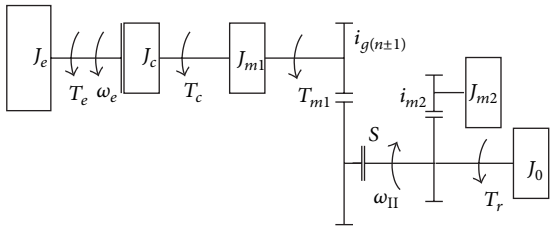


Figure 5: Vehicle dynamics model.

the curved surface fitting shows that choosing a smaller equivalent coefficient means that electric drive is more energy efficient in the low speed. In the high way, selecting a larger equivalent coefficient is more efficient.

From Figure 4 fitting surfaces, you can get the lowest fuel consumption with different charge and discharge coefficient under each drive cycle, the optimal charge and discharge coefficients were shown in Table 3. Then, selecting optimal charge and discharge coefficient based on the drive cycle recognition result to select optimal charge and discharge coefficient. Driving cycle recognition algorithm is based on the learning vector quantization neural network, through selecting the best identification parameters, the optimal identification cycle, and forecast period to identify the closest types of drive cycle [31, 32].

5. Analytical Solution to the Minimization Problem

5.1. Gear-Shifting Rules Based on the Minimum Equivalent Fuel Consumption. The simplified physical model of Figure 5 was presented for the analysis of powertrain system. According to the mathematical model of vehicle drive system, the establishment of kinetic equations unified a formula as follows:

$$\left(J_e + J_c + J_{m1} + \frac{J_{m2} i_{m2}^2 + J_o}{i_{g(n)}^2}\right)\dot{\omega}_e = T_e + T_{m1} + \frac{T_{m2} i_{m2} - T_r}{i_{g(n)}}, \tag{14}$$

where J_{m1} is the equivalent moment of inertia of M1 motor, J_{m2} is the equivalent moment of inertia of M2 motor, J_e is the equivalent moment of inertia of engine, J_c is the equivalent moment of inertia of clutch, J_o is the equivalent moment of inertia of the wheels and vehicle, and i_{m2} is the torque coupler transmission ratio.

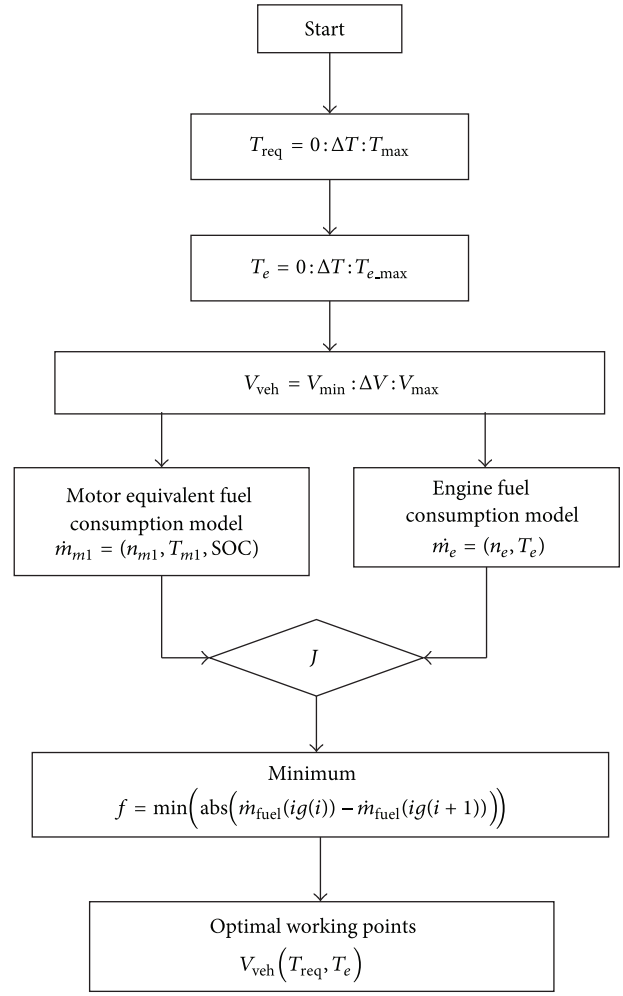


Figure 6: The flowchart of gear-shift rules optimization computation.

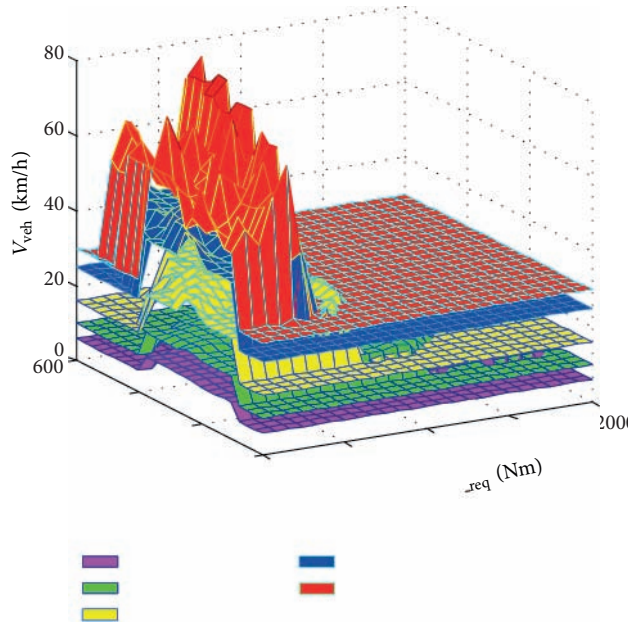


Figure 7: Up-shift surfaces at the 80% of battery SOC.

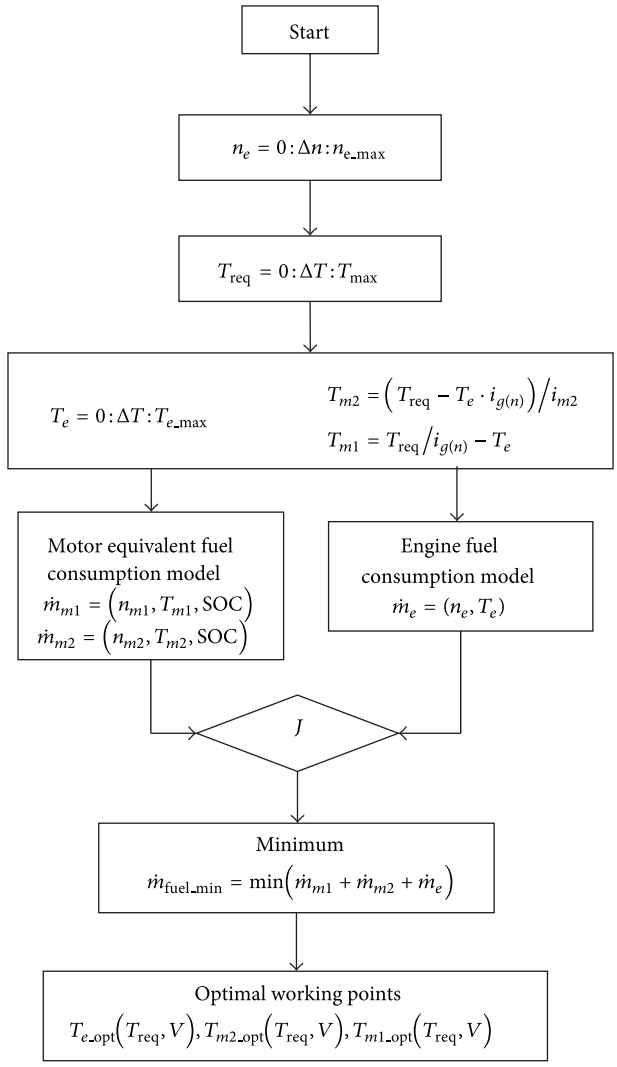


Figure 8: The flowchart of instantaneous optimal torque distribution.

To get the lowest fuel consumption by searching for the optimal torque distribution and gear-shifting rules, we need to simplify the dynamic model:

$$\begin{aligned} \left(T_e + T_{m1}\right) \cdot i_{g(n)} &= T_r - T_{m2} i_{m2}, \\ \frac{\omega_e}{i_{g(n)}} &= \frac{\omega_{m1}}{i_{g(n)}} = \frac{\omega_{m2}}{i_{m2}}, \end{aligned} \tag{15}$$

where ω_{m1} is the M1 motor speed, ω_{m2} is the M2 motor speed, and $i_{g(n)}$ is the transmission gear ratio.

The flowchart of gear-shift rules optimization computation was shown in Figure 6, through the mesh generation of engine torque and torque requirement to search for the optimal shift points and fit these points into shift surfaces and then to code into the controller and optimize the efficiency of the engine and the motor.

Figure 7 represents up-shift surface from 1st gear to 2nd gear, 2nd gear to 3rd gear, and so on at the 80% of battery SOC:

$$T_{equ\_req} = T_{req} - T_{m2} \cdot i_{m2}. \tag{16}$$

Down-shift rules were to ensure the less frequent gear shifting, but they would cause high fuel consumption (FC). Most of them choose a reasonable negative offset velocity to make a compromise.

5.2. Energy Management Control Strategy Based on the Minimum Equivalent Fuel Consumption. This series-parallel drive system included parallel mode 1 (engine and M2 motor) in the low speed and parallel mode 2 (engine and M1 motor) in the high speed, by searching for the best torque allocation rules into the real-time controller. Control strategy will no longer distinguish electric, pure engine-driven, and parallel mode, but through the use of three-dimensional

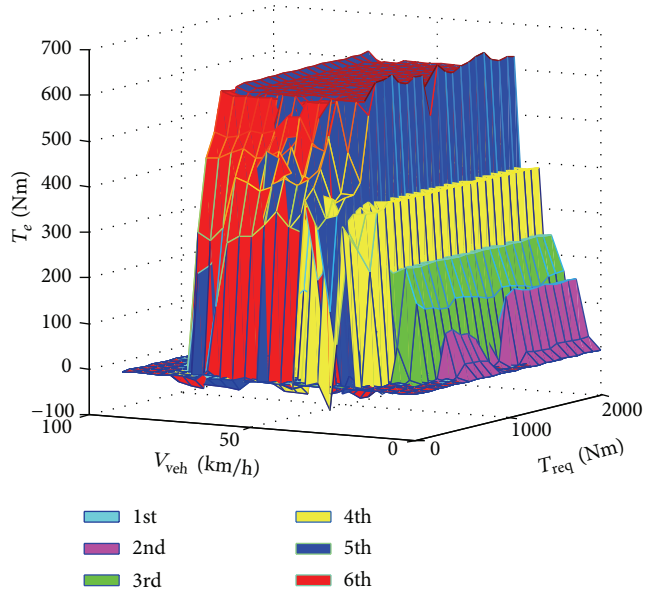


Figure 9: Parallel mode 1 at the 80% of battery SOC.

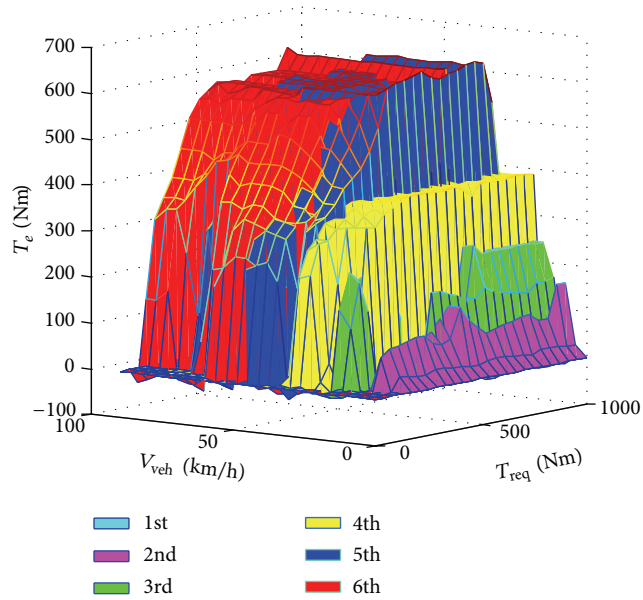


Figure 10: Parallel mode 2 at the 80% of battery SOC.

interpolation method to distribute engine torque and electric motor torque. This strategy not only can avoid the complexity of the drive mode and optimize torque distribution, but also can solve the problem of tedious debugging to achieve good simulation results.

The flowchart of instantaneous optimal torque distribution was shown in Figure 8, through the mesh generation of engine torque, engine speed, and torque requirement to search for the optimal torque distribution. Figure 8 proposed the offline optimal torque distribution in the specific drive cycle and then extracted the offline optimal control rules into the real-time controller.

The battery SOC is divided into many different threshold values, which corresponds with the different equivalent fuel consumption coefficients to repetitive computation. Figures 9 and 10 show the optimal engine torques in the different gear at the 80% of battery SOC.

6. Simulation and Experiments Validation

HIL simulation is helpful to verify the real time of control strategy and reduce the vehicle debug cycle by verifying that CAN bus communication works well in the hardware-in-loop

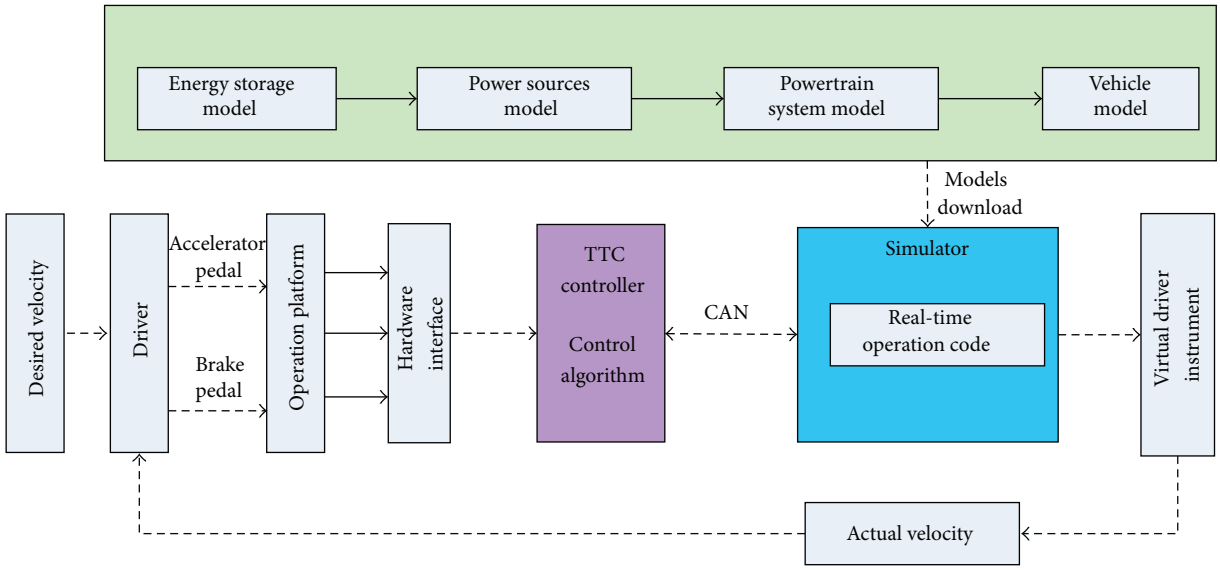


Figure 11: Hardware in loop real-time simulation architecture.

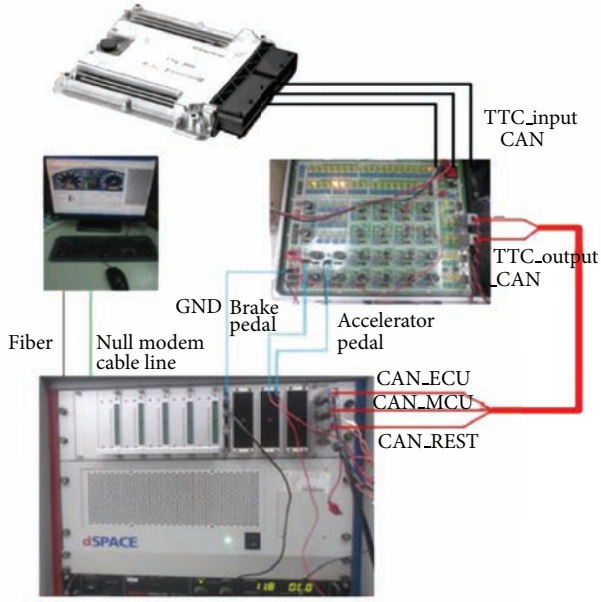
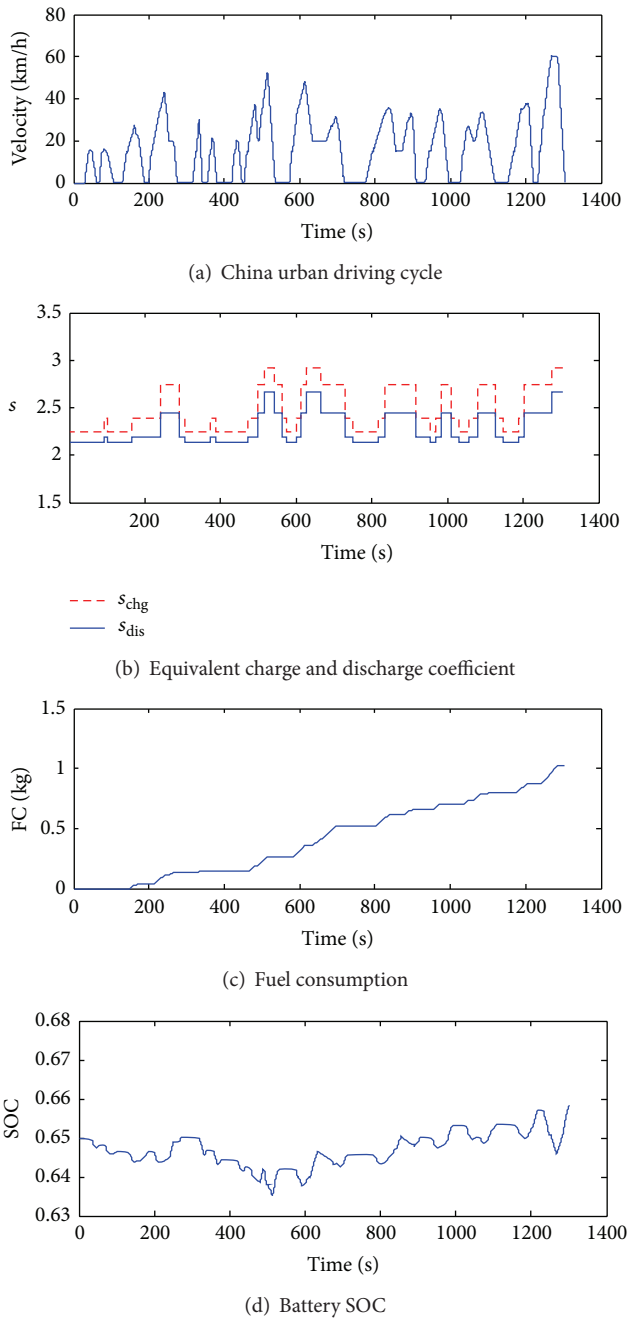


Figure 12: Simulator test bench for the HCU of the hybrid bus.

system. In this research, dSPACE/simulator was selected as HIL simulator for its high real-time performance and high simulation precision. The HEV powertrain model was developed and then vehicle model was coded and downloaded into simulator. TTC controller was chosen as the hybrid vehicle controller unit (HCU) and then control algorithm was compiled and downloaded into TTC controller by automatic code generation techniques [33]. Realizing the real-time communication between the virtual vehicle simulator and the real HCU is via CAN messages and analog signals. Simulator test bench was displayed in Figure 12.

In order to maintain the high-fidelity virtual model in the hardware-in-the-loop simulation system and to simulate the real vehicle test. The HIL simulation model should make a compromise on real-time requirements and complexity of the model. In this study, The HEV real-time simulation model includes hybrid vehicle powertrain components, driving environment, and driver. The HEV control strategy model includes torque requirements, torque distribution, torque coordinating, and gear-shifting controlling. Real-time simulation architecture was presented in Figure 11.

Figure 13(a) is the China urban driving cycle; Figure 13(b) shows that the charge and discharge coefficients are updated in real time by drive cycle recognition algorithm to ensure the suboptimal torque distribution and suboptimal gear-shifting rules. In Figure 13(d), the control strategy can ensure that the



(a) China urban driving cycle

(b) Equivalent charge and discharge coefficient

(c) Fuel consumption

(d) Battery SOC

Figure 13: Simulation results of HEB based on the A-ECMS.

battery operates within the upper and lower bounds and can sustain the battery SOC over the China urban driving cycle while minimizing fuel consumption.

The distribution maps for the operating points of the engine, M1 motor, and M2 motor are in Figure 14, which demonstrate that the hybrid bus can be driven in pure electric mode during low-speed driving, in parallel mode 1 during low-speed and accelerating driving, and in parallel mode 2 during high speed under the real-time A-ECMS controller above. Furthermore, the engine can avoid operating in the idle region. The M2 motor in the hybrid bus should propel the vehicle during low-speed driving and absorb regenerative energy during braking. It can also be used to regulate the peak and valley load of the engine through power assist and electricity generation. The M1 motor in the hybrid bus should propel the vehicle during high-speed driving and regulate the engine operating points as well, so that most of the operating points of the engine can be moved to the high-efficiency region. Consequently, the fuel consumption of the hybrid bus can be significantly reduced.

Table 4: Fuel consumption in the China urban driving cycle.

Control strategy	Fuel consumption (L/100 km)	Improvement (%)
Conventional	38.92	—
Logic threshold	26.51	31.89
A-ECMS	23.13	40.57

Table 4 gives the fuel consumption simulation results for the hybrid bus with different control strategies based on the same prototype vehicle. As is shown in Table 4, hybrid bus using the proposed real-time A-ECMS control strategy can achieve 12.75% and 40.57% lower fuel consumption compared to the hybrid bus employing the logic-threshold control strategy and the conventional ICE bus, respectively.

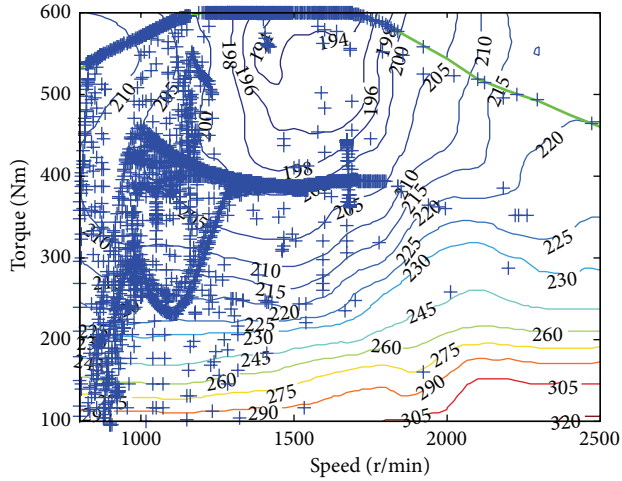
In Figure 15, we can get that the different initial SOC value is in accordance with the fuel economy. The battery acts as a buffer for load balancing and different battery SOC values will lead to different energy losses for the different battery internal resistances. The fuel economy is 23.13 L and 22.82 L per 100 km at the initial SOC value of 0.6 and 0.7, respectively, that are better performance than 24.62 L per 100 km at the initial SOC value of 0.5.

7. Conclusion

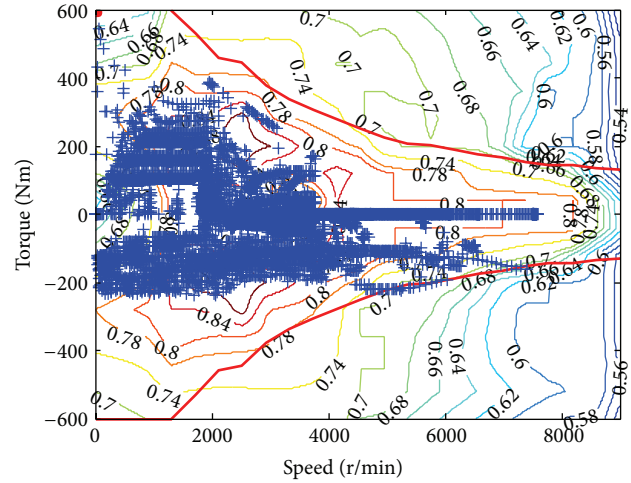
A real-time ECMS was developed for a series-parallel hybrid electric bus, and a HIL simulation system was constructed for energy management strategy investigation and verification in this study. The EMS design goal is minimizing fuel consumption while meeting drivability requirements and sustaining the battery SOC level; the A-ECMS was proposed to coordinate the relationship between the gear shifting and motor assist. At the same time, in order to realize adaptive control under different drive cycles, drive cycle recognition was presented to update the charge and discharge coefficients so as to achieve better fuel economy in each drive cycle.

The HIL simulation results demonstrated that the fuel consumption of the hybrid bus with real-time A-ECMS control strategy can be reduced by 12.75% and 40.57% compared to the hybrid electric bus with the logic-threshold control strategy and the conventional ICE bus, respectively.

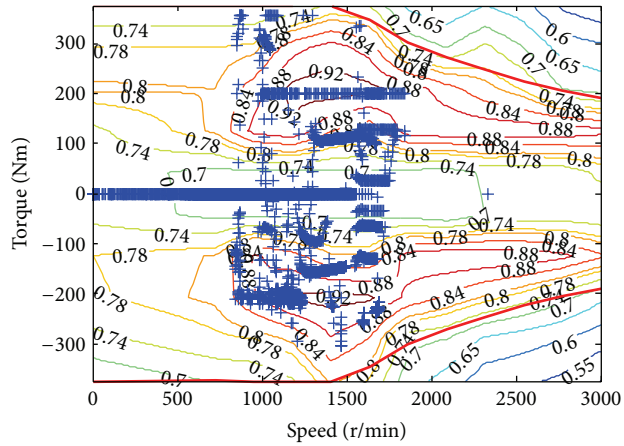
The real-time optimization control strategy plays an important role in the vehicle fuel economy improvement; the successful application of A-ECMS control strategy in the series-parallel hybrid electric bus provides good support for the vehicle experiment. In the next step in our work, we will have a trial that implements the proposed A-ECMS control strategy in a real controller with a real-world application.



(a) Engine operating points distribution



(b) M2 motor operating points distribution



(c) M1 motor operating points distribution

Figure 14: Power source operating points distribution.

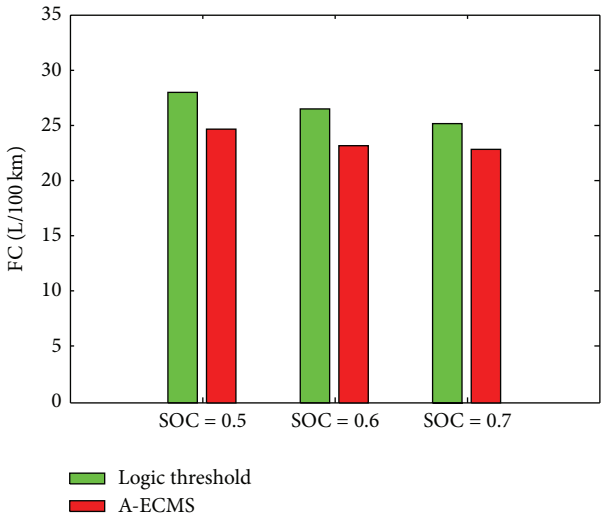


Figure 15: Fuel consumption under different initial SOC.

Nomenclature

ECMS:	Equivalent consumption minimization strategy
EMS:	Energy management strategy
HEV:	Hybrid electric vehicle
DP:	Dynamic programming
SDP:	Stochastic dynamic programming
PMP:	Pontryagin minimum principle
HIL:	Hardware-in-the-loop
ECU:	Electronic control unit
HCU:	Hybrid vehicle controller unit
HEB:	Hybrid electric bus
AMT:	Automated mechanical transmission
SOC:	State of charge
TCU:	Transmission control unit
BSFC:	Brake specific fuel consumption
PNGV:	Partnership for a New Generation of Vehicles

A-ECMS: Adaptive equivalent consumption minimization strategy
CAN: Controller area network.

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 School of Automotive Engineering, Changchun, Jilin, China, as well as the National Natural Science Foundation of China for their supports on this project (Grant no. 51075179) and the National High Technology Research and Development Program of China (863 Program) for their supports on this project (Grant no. 2011AA11A210).

References

[1] S. B. Han, Y. H. Chang, E. Y. Lee, Y. J. Chung, and B. Suh, "Emissions simulation in a 7000 kg-grade diesel hybrid electric vehicle," *International Journal of Automotive Technology*, vol. 11, no. 1, pp. 105–110, 2010.

[2] A. Boukehili, Y. T. Zhang, Q. Zhao, C. Q. Ni, H. F. Su, and G. J. Huang, "Hybrid vehicle power management modeling and refinement," *International Journal of Automotive Technology*, vol. 13, no. 6, pp. 987–998, 2012.

[3] S. Bogosyan, M. Gokasan, and D. J. Goering, "A novel model validation and estimation approach for hybrid serial electric vehicles," *IEEE Transactions on Vehicular Technology*, vol. 56, no. 4, pp. 1485–1497, 2007.

[4] W. Haikun, F. Lixin, Z. Yu, and L. He, "Modelling of the fuel consumption for passenger cars regarding driving characteristics," *Transportation Research D*, vol. 13, pp. 479–482, 2008.

[5] A. Rousseau, S. Saglini, M. Jakov, D. Gray, and K. Hardy, "Trade-offs between fuel economy and NOx emissions using fuzzy logic control with a hybrid CVT configuration," *International Journal of Automotive Technology*, vol. 4, no. 1, pp. 47–55, 2003.

[6] B.-C. Chen, Y.-Y. Wu, Y.-L. Wu, and C.-C. Lin, "Adaptive power split control for a hybrid electric scooter," *IEEE Transactions on Vehicular Technology*, vol. 60, no. 4, pp. 1430–1437, 2011.

[7] O. Dingel, N. Pini, I. Trivic et al., "Benchmarking Hybrid Concepts: On-Line versus Off-Line Fuel Economy Optimization for Different Hybrid Architectures," SAE Technical Paper 2013-24-0084, 2013.

[8] K. L. Butler, M. Ehsani, and P. Kamath, "A matlab-based modeling and simulation package for electric and hybrid electric vehicle design," *IEEE Transactions on Vehicular Technology*, vol. 48, no. 6, pp. 1770–1778, 1999.

[9] 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.

[10] K. Chen, A. Bouscayrol, A. Berthon, P. Delarue, D. Hissel, and R. Trigui, "Global modeling of dfferent vehicles," *IEEE Vehicular Technology Magazine*, vol. 4, no. 2, pp. 80–89, 2009.

[11] V. T. Minh and A. A. Rashid, "Automatic control of clutches and simulations for parallel hybrid vehicles," *International Journal of Automotive Technology*, vol. 13, no. 4, pp. 645–651, 2012.

[12] 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.

[13] G. Paganelli, S. Delprat, T. Guerra et al., "Equivalent consumption minimization strategy for parallel hybrid powertrains," in *Proceedings of the IEEE Vehicular Technlogy Conference*, pp. 2076–2081, 2002.

[14] B. Geng, J. K. Mills, and D. Sun, "Energy management control of microturbine-powered plug-in hybrid electric vehicles using the telemetry equivalent consumption minimization strategy," *IEEE Transactions on Vehicular Technology*, vol. 60, no. 9, pp. 4238–4248, 2011.

[15] O. Dingel, N. Pini, I. Trivic et al., "Benchmarking Hybrid Concepts: On-Line versus Off-Line Fuel Economy Optimization for Different Hybrid Architectures," SAE Technical Paper 2013-24-0084.

[16] J.-P. Gao, G.-M. G. Zhu, E. G. Strangas, and F.-C. Sun, "Equivalent fuel consumption optimal control of a series hybrid electric vehicle," *Proceedings of the Institution of Mechanical Engineers D*, vol. 223, no. 8, pp. 1003–1018, 2009.

[17] C. Musardo, G. Rizzoni, Y. Guezennec, and B. Staccia, "A-ECMS: an adaptive algorithm for hybrid electric vehicle energy management," *European Journal of Control*, vol. 11, no. 4-5, pp. 509–524, 2005.

[18] X. Ye, Z. Jin, X. Hu, Y. Li, and Q. Lu, "modeling and control strategy development of a parallel hybrid electric bus," *International Journal of Automotive Technology*, vol. 14, no. 6, pp. 971–985, 2013.

[19] V. Sezer, M. Gokasan, and S. Bogosyan, "A novel ECMS and combined cost map approach for high-efficiency series hybrid electric vehicles," *IEEE Transactions on Vehicular Technology*, vol. 60, no. 8, pp. 3557–3570, 2011.

[20] M. Sivertsson, "Adaptive control using map-based ECMS for a PHEV," in *Proceedings of the IFAC Workshop on Engine and Powertrain Control, Simulation and Modeling (ECOSM' 12)*, pp. 357–362, 2012.

[21] R. Chen, L. Mi, and W. Tan, "A new hardware-in-the-loop test system for electronic control unit of dual-clutch transmission vehicle," *Advanced Materials Research*, vol. 490–495, pp. 13–18, 2012.

[22] S. Onori, L. Serrao, and G. Rizzoni, "Adaptive equivalent consumption minimization strategy for hybrid electric vehicles," in *Proceedings of the Dynamic Systems and Control Conference (DSCC '10)*, pp. 499–505, September 2010.

[23] H.-W. He, R. Xiong, and Y.-H. Chang, "Dynamic modeling and simulation on a hybrid power system for electric vehicle applications," *Energies*, vol. 3, no. 11, pp. 1821–1830, 2010.

[24] G. Rizzoni, L. Guzzella, and B. M. Baumann, "Unified modeling of hybrid electric vehicle drivetrains," *IEEE/ASME Transactions on Mechatronics*, vol. 4, no. 3, pp. 246–257, 1999.

[25] V. H. Johnson, "Battery performance models in ADVISOR," *Journal of Power Sources*, vol. 110, no. 2, pp. 321–329, 2002.

[26] P. Nelson, I. Bloom, K. Amine, and G. Henriksen, "Design modeling of lithium-ion battery performance," *Journal of Power Sources*, vol. 110, no. 1, pp. 437–444, 2002.

[27] L. Wang, Y. Zhang, C. Yin, H. Zhang, and C. Wang, "Hardware-in-the-loop simulation for the design and verification of the control system of a series-parallel hybrid electric city-bus," *Simulation Modelling Practice and Theory*, vol. 25, pp. 148–162, 2012.

[28] M. Ehsani, K. M. Rahman, and H. A. Toliyat, "Propulsion system design of electric and hybrid vehicles," *IEEE Transactions on Industrial Electronics*, vol. 44, no. 1, pp. 19–27, 1997.

[29] O. Simona, S. Lorenzo, and R. Giorgio, "Adaptive equivalent consumption minimization strategy for hybrid electric vehicles," in *Proceedings of the ASME Dynamic Systems and Control Conference*, pp. 499–505, 2010.

[30] D. Zhao and R. Stobart, "Real-Time Optimal Energy Management of Heavy Duty Hybrid electric Vehicles," SAE Technical Paper 2013-01-1748.

[31] C.-C. Lin, S. Jeon, H. Peng, and J. M. Lee, "Driving pattern recognition for control of hybrid electric trucks," *Vehicle System Dynamics*, vol. 42, no. 1-2, pp. 41–58, 2004.

[32] R. Langari and J.-S. Won, "Intelligent energy management agent for a parallel hybrid vehicle. Part I: system architecture and design of the driving situation identification process," *IEEE Transactions on Vehicular Technology*, vol. 54, no. 3, pp. 925–934, 2005.

[33] J. Wu, C. Dufour, and L. Sun, "Hardware-in-the-loop testing of hybrid vehicle motor drives at Ford Motor Company," in *Proceedings IEEE Vehicle Power and Propulsion Conference (VPPC '10)*, September 2010.

Hindawi Publishing Corporation
Mathematical Problems in Engineering
Volume 2014, Article ID 142304, 8 pages
http://dx.doi.org/10.1155/2014/142304

Research Article

Vehicle Type Recognition in Sensor Networks Using Improved Time Encoded Signal Processing Algorithm

Yan Wang,[1] Xi Wu,[2] Xiaohua Li,[3] and Jiliu Zhou[3]

[1] *College of Electronics and Information Engineering, Sichuan University, Chengdu 610065, China*
[2] *College of Computer Science, Chengdu University of Information Technology, Chengdu 610065, China*
[3] *College of Computer Science, Sichuan University, Chengdu 610065, China*

Correspondence should be addressed to Jiliu Zhou; zhoujl@scu.edu.cn

Received 23 September 2013; Revised 15 January 2014; Accepted 30 January 2014; Published 6 March 2014

Academic Editor: Ilse C. Cervantes

Copyright © 2014 Yan 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.

Vehicle type recognition is a demanding application of wireless sensor networks (WSN). In many cases, sensor nodes detect and recognize vehicles from their acoustic or seismic signals using wavelet based or spectral feature extraction methods. Such methods, while providing convincing results, are quite demanding in computational power and energy and are difficult to implement on low-cost sensor nodes with limitation resources. In this paper, we investigate the use of time encoded signal processing (TESP) algorithm for vehicle type recognition. The conventional TESP algorithm, which is effective for the speech signal feature extraction, however, is not suitable for the vehicle sound signal which is more complex. To solve this problem, an improved time encoded signal processing (ITESP) is proposed as the feature extraction method according to the characteristics of the vehicle sound signal. Recognition procedure is accomplished using the support vector machine (SVM) and the k-nearest neighbor (KNN) classifier. The experimental results indicate that the vehicle type recognition system with ITESP features give much better performance compared with the conventional TESP based features.

1. Introduction

Along with the development of communication technology, wireless sensor network (WSN) is playing an increasingly important role in our daily life. Recent advances in wireless communications, electronics, and ubiquitous computing, in combination with intensive research on the field of WSN, have changed the way we interact with the physical environment [1–3]. Vehicle type recognition, which can be used in intrusion detection, transportation, and border monitoring, is a significant and demanding application of WSN. In most cases, the usage of fast Fourier transform (FFT) [4], wavelet transform (WT) [5, 6], and Hilbert-Huang transform (HHT) [7] to extract the frequency or time-frequency features of the signals acquired from the acoustic and seismic sensors is a common approach for the vehicle recognition. These methods, while providing convincing results, are quite demanding in computational power and energy and are difficult to implement on low-cost sensor nodes with limited resources. In WSN, the sensor nodes usually process signals locally to come to a decision rather than transmitting the measurements. Due to network bandwidth limitation and energy consumption of sensor nodes, we usually wish to use low-complexity and low-energy consumption algorithms to recognize the vehicle type [8]. In this paper, we investigate the use of time-domain encoding and feature extraction methods and propose an improved time encoded signal processing (ITESP) algorithm. The conventional time encoded signal processing (TESP) algorithm, which uses a symbol table with 29 characters to encode the time-domain information of the signal, has performed well in speech recognition [9]. The advantage of the method is its computational simplicity and low memory requirements. However, it turns out that TESP is not suitable for the vehicle sound signal which is more complex. To solve this problem, the improved time encoded signal processing feature extraction method is proposed for the vehicle sound signal recognition. Figure 1 shows the flow chart of vehicle type recognition system based on ITESP algorithm.

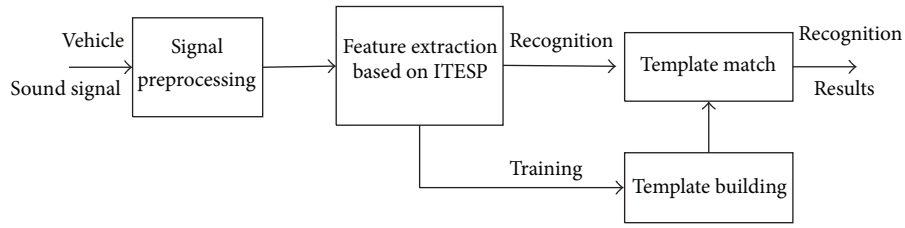


Figure 1: Flow chart of vehicle type recognition system based on ITESP algorithm.

The rest of the paper is organized as follows. In Section 2, the methodologies of TESP and ITESP algorithms are discussed, respectively. Section 3 describes the simulation experiments of feature extraction based on ITESP algorithm. The experimental results on the recognition performance of the feature extraction methods are discussed in Section 4. Finally, Section 5 presents discussion and conclusion.

2. Methodology

2.1. TESP Algorithm. TESP is a digital language that originated as a means of coding signals for speech recognition, and it describes signal waveforms according to its real and complex zeros based on a mathematical waveforms representation. TESP quantisation procedure has been developed to encode signals according to the period between two consecutive zero-crossings and the shape of the curve thus contained [10]. This period is named an epoch. The TESP procedure can be described using four simple steps.

Step 1. Divide the signal into successive epochs.

Step 2. Characterize each epoch with two descriptors, duration and shape, as follows.

(i) Duration (D) which is the number of samples between two successive real zeros and provides information on the fundamental frequency of the waveform.

(ii) Shape (S) which is the number of local minima (for a positive epoch) or the number of local maxima (for a negative epoch). The shape of an epoch contains harmonic information of the signal.

Step 3. Map each epoch, from its corresponding D/S descriptors, to a predefined symbol table.

The encoding procedure results in the mapping of every epoch of the waveform in a two-dimensional space with dimensions of $\max(D) \times \max(S)$. This bidimensional space can get very large and depends on the bandwidth and the complexity of the signal. To reduce the number of descriptors needed, a quantization method is used to create a one-dimensional symbol stream from the two-dimensional space. D/S pairs are mapped to a character using a symbol table created beforehand, to approximate the D/S space using fewer characters [11].

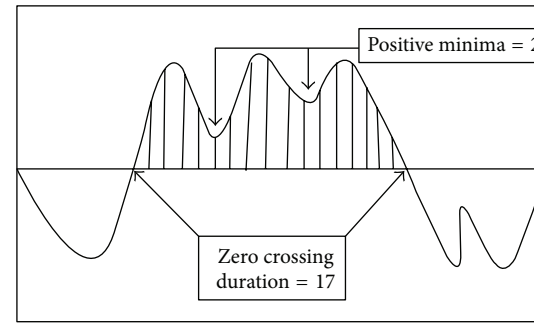


Figure 2: TESP single epoch with $D = 17$, $S = 2$.

Step 4. Create a fixed-dimensions matrix containing the appearance probability of each symbol in the entire waveform. This matrix will be used for the recognition task.

To make the methodology more clear, here we give an example [10]. Figure 2 shows an epoch encoded into its TESP parameters where $D = 17$ and $S = 2$.

The encoding of a waveform using the aforementioned coding scheme results in a one-dimensional symbol stream. This symbol stream can be further manipulated to create a one-dimensional M-size matrix (where M is the total number of the characters in the symbol table) which contains the number of appearances of each symbol in the symbol stream, called S-matrix. It can be created using the following expression:

$$S(i) = \sum_{j=1}^{N} f(x(j)), \quad 1 \le i \le M, \tag{1}$$

where $S(i)$, j, and N represent the ith element of S-matrix, the jth epoch of the signal, and the number of total epochs in the waveform, respectively. $x(j)$ is the symbol describing the jth epoch:

$$f(x(j)) = \begin{cases} 1, & x(j) = i, \\ 0, & \text{otherwise.} \end{cases} \tag{2}$$

2.2. ITESP Algorithm. In conventional TESP algorithm, the standard symbol table used in Step 3, which contains 29 characters, has been found to be sufficient for speech signals description, but may not be suitable for the vehicle sound signal which is more complex. In this paper, according to

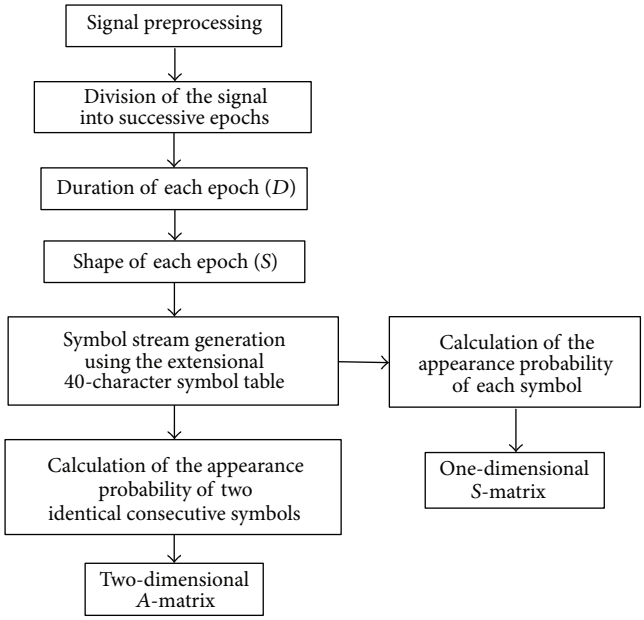


Figure 3: Flow chart of feature extraction based on ITESP algorithm.

the characteristics of the vehicle sound signals, an extensional symbol table with 40 characters is designed, and then based on the symbol stream which is encoded by the symbol table, the one-dimensional S-matrix is constructed using the appearance probability of each symbol. Meanwhile, using the appearance probability of the two identical consecutive symbols, the two-dimensional A-matrix is constructed as well in order to obtain more accurate features of the signal. The flow chart of feature extraction procedure based on ITESP algorithm is shown in Figure 3.

3. Simulation Experiment of Feature Extraction

3.1. Data Acquisition. In this paper, the sound signals of two typical vehicles types (wheeled vehicles and tracked vehicles) are selected as samples to evaluate the performance of the feature extraction method. The wheeled vehicles sound signals were recorded during a real world WSN experiment at Chengdu, China, and the data set was gathered from 15 microphone sensors which were deployed at three different roads. All the signals studied were sampled at 22050 Hz and quantized with 8 bits per sample. Since most of the tracked vehicles are military vehicles which are difficult for acquisition in real environment, the sound signals of tracked vehicle were downloaded from the sensor website. The sensor website indicates that the tracked vehicle data is gathered from two different sensor types: geophone and microphone, each sampled at 4096 Hz with 16-bit accuracy. Figure 4 shows the comparison between the two types of vehicle signal in time domain.

Table 1: Standard symbol table using 29 characters.

D	S					
	0	1	2	3	4	5
1	1					
2	2	2				
3	3	3	3			
4	4	4	4	4		
5	5	5	5	5	5	
6	6	6	6	6	6	6
⋮	⋮	⋮	⋮	⋮	⋮	⋮
34	24	25	26	27	28	29
35	24	25	26	27	28	29

3.2. Signal Preprocessing. To reduce the complexity of data processing, the downsampling frequency of 4096 Hz is firstly employed. Moreover, the vehicle sound signals must be filtered before the encoding procedure for three main reasons as follows.

(1) To minimize the number of symbols needed for the symbol table by keeping only the important frequency range of the signal. In this way the dimensions of the S-matrix are minimized.

(2) To eliminate high frequency "flicker" on the waveform which can be translated to local minima or maxima inside an epoch, thus increasing its S descriptor.

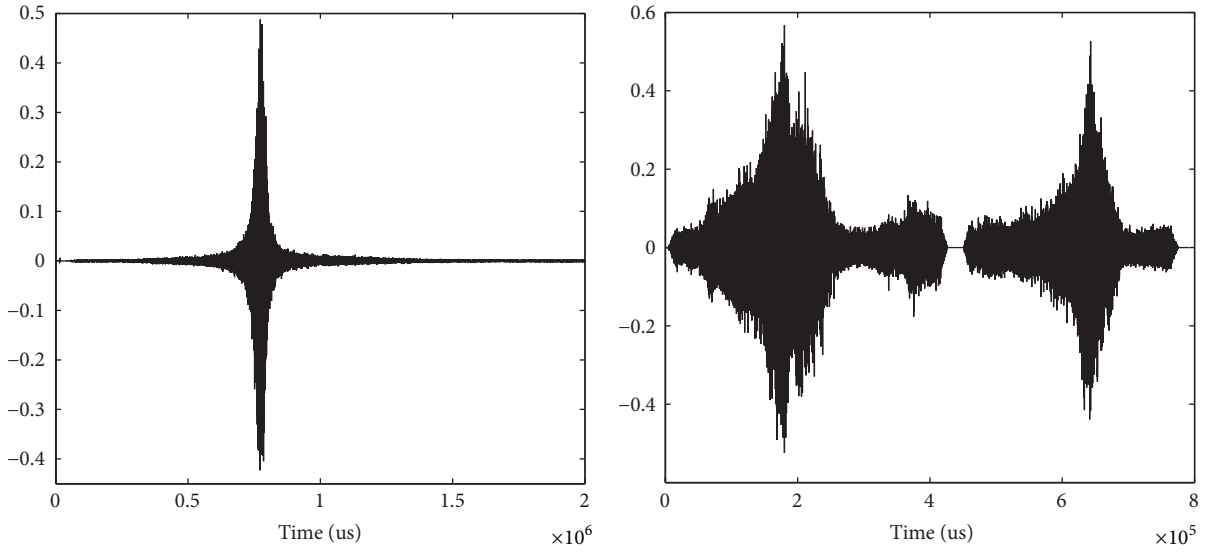
(3) To prevent the introduction of quantization noise.

By analyzing the main noise source of the sound signals, we find that the frequencies of the vehicle sounds are mainly below 800 Hz. Therefore, an 800 Hz low-pass Butterworth filter is employed accordingly. Furthermore, frequency components below 50 Hz for the vehicle sound signals are not very important in recognition task [12]. Such low frequency signals can be ignored without decreasing performance, which can significantly reduce the maximum value of D descriptor.

3.3. Feature Extraction Based on ITESP

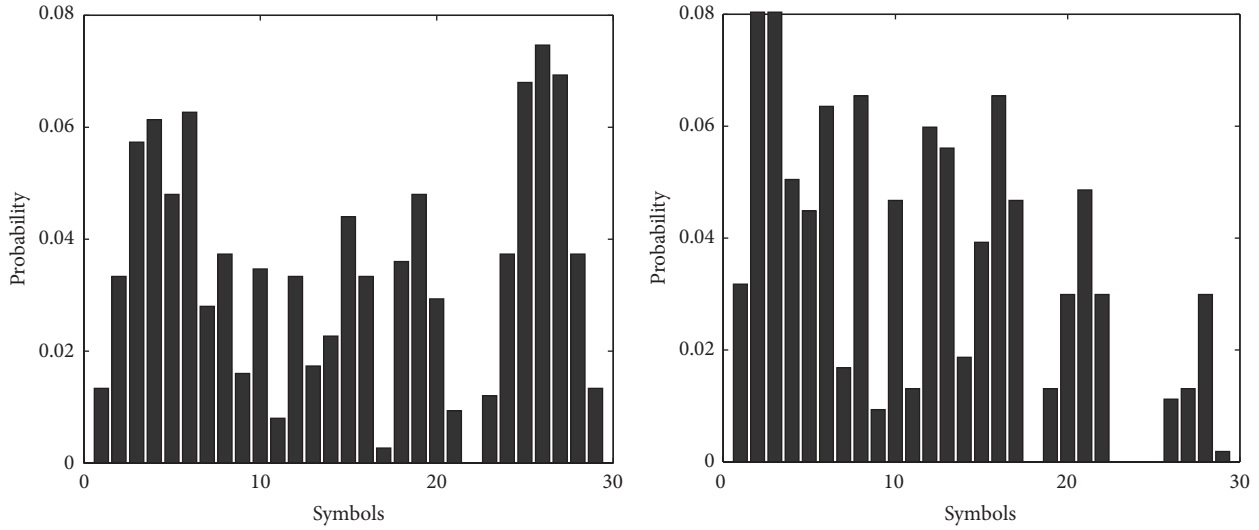
3.3.1. One-Dimensional S-Matrix. The recognition performance and the S-matrix length depend on the symbol table used. In most cases of speech recognition, a standard 29-character symbol table shown in Table 1 is employed, allowing for a maximum D of 35 and a maximum S of 5. The standard symbol table is optimized for speech signals. However, its performance in vehicle sound signals should be examined. Figure 5 shows the S-matrices of the two types of vehicle sound signal based on 29-character symbol table.

As can be seen from Figure 5, the differences between the appearance probabilities for the two-type sound signals are small and scattered, which could not be used to obtain high recognition rate in theory. The main reason of the undesirable S-matrices is that the standard 29-character symbol table is not suitable for the vehicle sound signal which is quite different from the speech signal in frequency distribution.



(a) Wheeled vehicle (b) Tracked vehicle

Figure 4: Time-domain graphs of the two types of vehicle sound signal.



(a) Wheeled vehicle (b) Tracked vehicle

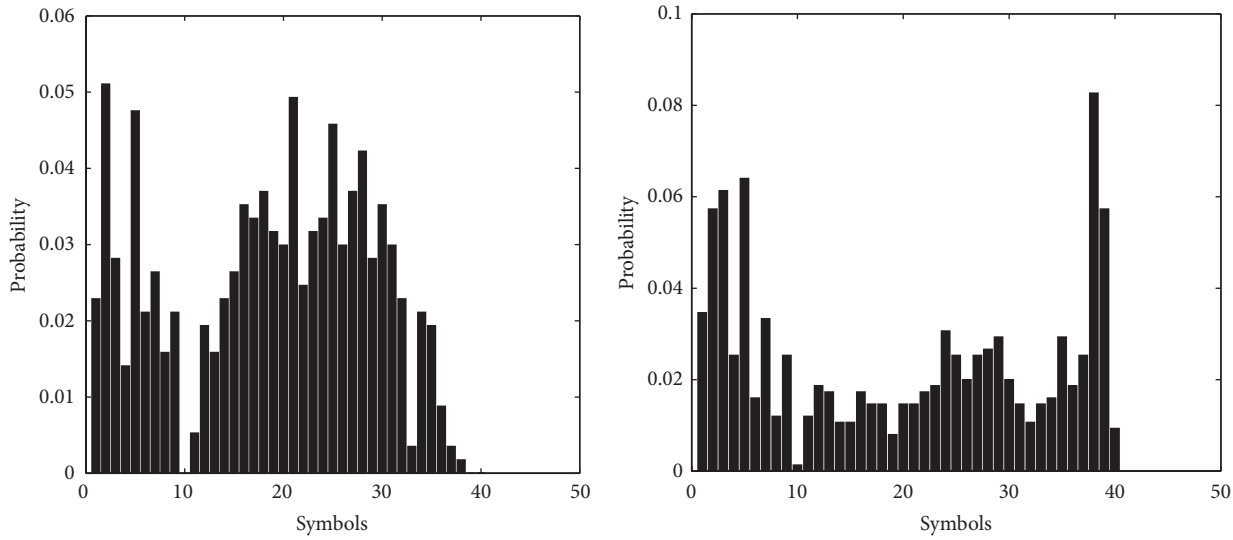
Figure 5: *S*-Matrices of the two types of vehicle sound signal based on 29-character symbol table.

Compared with the low-frequency part of speech signal, its counterpart of vehicle sound signal is lower, which means more D is needed to describe the waveforms. Besides, the vehicle sound signal contains more harmonic component; consequently, the numeric values of S should be a little larger. According to the characteristics of the vehicle sound signal, an extensional symbol table using 40 characters is designed to obtain more effective S-matrix. Table 2 shows the extensional 40-character symbol table.

Compared with Table 1, the characters in Table 2 change more frequently with the increase of D and S, resulting in higher separability of different D/S descriptors. For the vehicle sound signal that contains more harmonic information, using the extensional symbol table can obtain more time-domain features of the signal. The S-matrices of the two types of vehicle sound signal based on 40-character symbol table are shown in Figure 6.

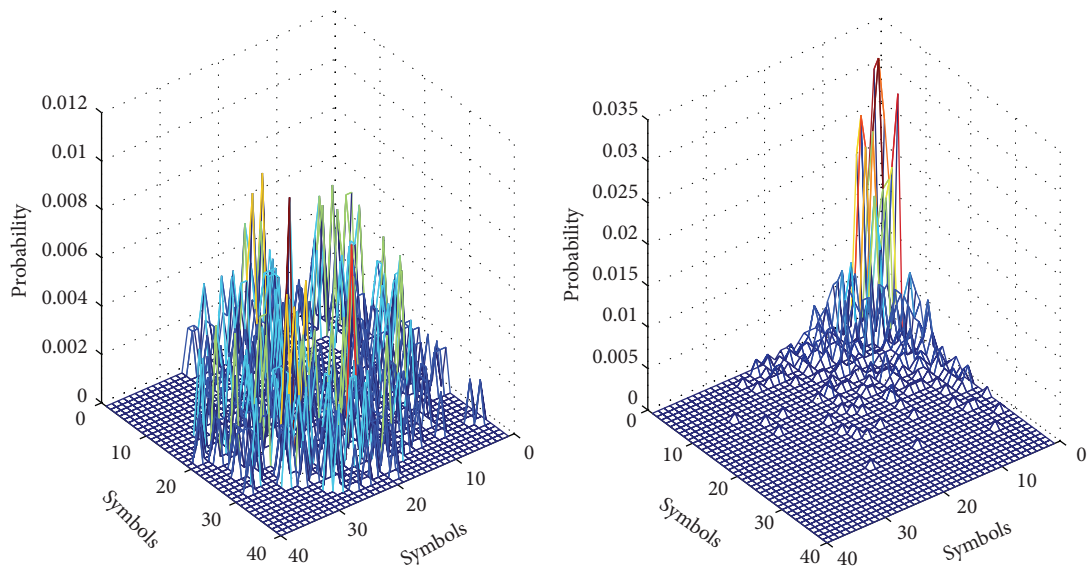
Figure 6 indicates that, after encoding the signal using the 40-character symbol table, the S-matrix possesses enough difference from the different signals. Thus, we can assume that S-matrices abstracted from wheeled vehicles and tracked vehicles are different enough to enable recognition.

3.3.2. Two-Dimensional A-Matrix. Using the extensional 40-character symbol table, the sound signal is encoded and a one-dimensional symbol stream is generated accordingly.



(a) Wheeled vehicle (b) Tracked vehicle

Figure 6: *S*-Matrices of the two types of vehicle sound signal based on 40-character symbol table.



(a) Wheeled vehicle (b) Tracked vehicle

Figure 7: *A*-Matrices features distribution of the two types of vehicle sound signal.

In Section 3.3.1, the *S*-matrix is obtained by calculating the appearance probability of each symbol. In order to obtain more accurate time-domain features of the signal, we use the appearance probability of the two identical consecutive symbols to construct the two-dimensional *A*-matrix. Figure 7 shows the *A*-matrices features distribution of the two types of vehicle sound signal, where the x-axis and y-axis represent the symbols while the z-axis represents the appearance probability.

As can be seen from Figure 7, the two-dimensional *A*-matrices of the wheeled and tracked vehicle sound signals present obvious feature distribution differences from each other. Compared with the one-dimensional *S*-matrix, *A*-matrix not only shows the probability features of each symbol but also presents the spatial probability features of the symbols. Thus, more accurate time-domain features of the signal are obtained, which can further improve the recognition performance.

4. Recognition Experimental Results

After the extraction of the time-domain features from *S*-matrix and *A*-matrix based on ITESP, respectively, the classifier is employed to evaluate the performance of the proposed

Table 2: Extensional symbol table using 40 characters.

D	S					
	0	1	2	3	4	5
1	1					
2	1	1				
3	2	2	2			
4	3	3	3	4		
5	4	4	5	5	5	
6	5	5	6	6	6	6
⋮	⋮	⋮	⋮	⋮	⋮	⋮
36	34	35	36	37	38	39
37	35	36	37	38	39	40

algorithms. In this paper, we design two classifiers for comparison, that is, k-nearest neighbor (KNN) classifier and support machine vector (SVM).

4.1. k-Nearest Neighbor Algorithm. In pattern recognition, the KNN algorithm is a nonparametric method for classifying objects based on closest training examples in the feature space [13]. KNN is a type of instance-based learning, or lazy learning, where the function is only approximated locally and all computation is deferred until classification. The KNN algorithm is amongst the simplest of all machine learning algorithms: an object is classified by a majority vote of its neighbors, with the object being assigned to the class most common amongst its k-nearest neighbors (k is a positive integer, typically small). If $k = 1$, then the object is simply assigned to the class of that single nearest neighbor. KNN algorithm has been extensively studied in the field of computational geometry and used in many applications [14].

4.2. Support Vector Machine. SVM, originally introduced by Vapnik, has been shown to be effective in learning linear and nonlinear decision boundaries and is successfully used in many applications [15–18]. SVM performs classification by constructing an N-dimensional hyperplane that optimally separates the data into two categories. SVM models are closely related to neural networks. In fact, a SVM model using a sigmoid kernel function is equivalent to a two-layer, perceptron neural network. SVM has been used very successfully in recent years as a substitute to neural networks. A basic SVM can handle only two-class classification. To use SVM in multiclass classification, the problem must be broken down into several two-class classification tasks. The effectiveness of SVM depends on the selection of kernel, the kernel's parameters, and soft margin parameter C. A common choice is a Gaussian kernel, which has a single parameter σ. In this paper, the SVM algorithm is implemented by using the Gaussian kernel defined by

$$k(x, y) = \exp\left(\frac{\|x - y\|^2}{2\sigma^2}\right), \tag{3}$$

where σ is a user-defined variance parameter. The best combination of C and σ is often selected by a grid search with exponentially growing sequences of C and σ. Typically, each combination of parameter choices is checked using cross validation, and the parameters with best cross validation accuracy are picked. The final model, which is used for testing and for classifying new data, is then trained on the whole training set using the selected parameters.

4.3. Recognition Results. In the recognition experiments, for the wheeled vehicle type, a heavy wheeled truck and a sedan car are recorded while moving on the roads for multiple times. The sound signals of tracked vehicle type we downloaded from the sensor website also involve different types. After that, we selected 80 signals for each vehicle type as the recognition experimental dataset which is randomly divided into two subsets: 30 for training and the other 50 for validation to study the recognition performance. The recognition rate is defined as the percentage ratio of the number of vehicle sounds correctly recognized to the total number of sounds considered for recognition. To evaluate the performance of the proposed ITESP algorithm, we take the conventional TESP algorithm based on 29-character symbol table for comparison. The computational time is measured on a laptop with an Intel i3-2310M 2.1 GHz processor using MATLAB commands *tic* and *toc*. Tables 3 and 4 show the comparison of the recognition results using different feature extraction methods based on KNN classifier and SVM classifier, respectively.

From Tables 3 and 4, we can see that, for the two types of classifiers, the SVM exhibits better performance than KNN in both recognition rate and computational time, indicating that SVM classifier is more suitable for vehicle sound recognition system. On the other hand, for the three feature extraction methods, the ITESP algorithm (including 40-character based S-matrix and 40-character based A-matrix) obtains much higher recognition rate than TESP algorithm. Compared with the recognition rate based on the conventional TESP algorithm (51%), the recognition rates using the proposed S-matrix and A-matrix are up to 84% and 87%, respectively. However, the computational time based on A-matrix (2.37 s) is longer than that of S-matrix (0.79 s). Therefore, the selection between the S-matrix and the A-matrix as the feature extraction method should depend on the actual demand of the recognition system for the recognition rate or the computational time.

However, as to the entire vehicle type recognition WSN system, the sensor network lifetime mainly depends on the energy consumption due to the difficulty in charging batteries [19]. Next, in the same sensor network, we compare the energy performances of the network that using the time-domain signal processing method based on ITESP algorithm and using the time-frequency domain signal processing method based on wavelet transform (WT). The initial energy of the network is set to 200 J. The comparison of the residual energy with the network lifetime is shown in Figure 8.

Figure 8 shows that the energy consumption ratio of the WSN system using WT is much quicker than that using the ITESP algorithm, indicating that the vehicle type recognition

Table 3: Comparison of the recognition results using three feature extraction methods based on KNN classifier.

KNN classifier	Feature extraction methods		
	TESP algorithm	ITESP algorithm	
	29-character based *S*-matrix	40-character based *S*-matrix	40-character based *A*-matrix
Wheeled vehicles	48%	76%	88%
Tracked vehicles	54%	84%	82%
Average recognition rate	52%	80%	85%
Average computational time	0.63 s	0.81 s	2.55 s

Table 4: Comparison of the recognition results using three feature extraction methods based on SVM classifier.

SVM classifier	Feature extraction methods		
	TESP algorithm	ITESP algorithm	
	29-character based *S*-matrix	40-character based *S*-matrix	40-character based *A*-matrix
Wheeled vehicles	52%	82%	90%
Tracked vehicles	50%	86%	84%
Average recognition rate	51%	84%	87%
Average computational time	0.82 s	0.79 s	2.37 s

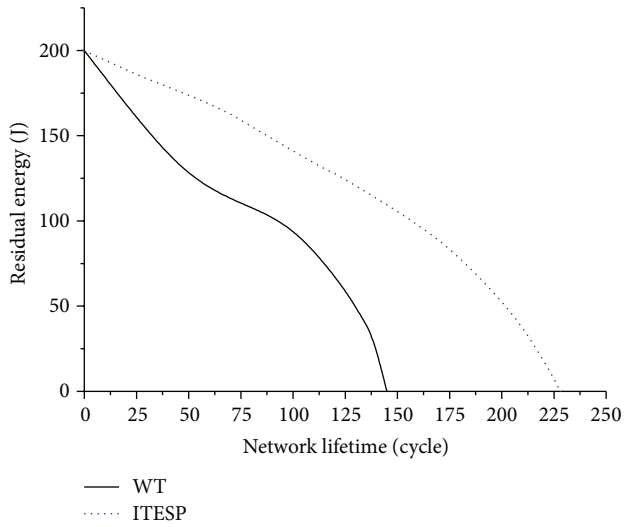


Figure 8: Comparison of the residual energy with the network lifetime.

WSN system based on ITESP can save network energy and extend the life cycle of network more effectively.

5. Conclusion

In this paper, we presented a promising method for vehicle type recognition, using the improved time encoded signal processing (ITESP) for signal encoding and support vector machine for classification. Aiming at the problem that the conventional TESP algorithm was effective for the speech signal feature extraction but not suitable for the vehicle sound signal, we designed an extensional 40-character symbol table, and then used the symbol stream which was encoded by the symbol table to construct one-dimensional *S*-matrix and two-dimensional *A*-matrix, respectively, as the time-domain features of the sound signal. Experimental results indicated that the ITESP methods provided higher recognition rates between two types of vehicles (wheeled type and tracked type) using their sound signature. Compared with the feature extraction method based on time-frequency domain analysis (WT), the ITESP algorithm needs less energy consumption. Our future work will focus on the recognition performance using more vehicle types, hardware implementation of the classifiers on the prototype sensor nodes, and field testing of the system.

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 the Grants from the National Science Foundation of China (nos. 61272448, 61302028), STMSP project (no. 2012RZ0005), and the Foundation of Sichuan University Early Career Researcher Award (nos. 2012SCU11036, 2012SCU11070).

References

[1] P. Mehta, D. Chander, M. Shahim, K. Tejaswi, S. N. Merchant, and U. B. Desai, "Distributed detection for landslide prediction using wireless sensor network," in *Proceedings of the 1st International Global Information Infrastructure Symposium (GIIS '07)*, pp. 195–198, IEEE, 2007.

[2] P. K. Sahoo, K.-Y. Hsieh, and J.-P. Sheu, "Boundary node selection and target detection in wireless sensor network," in *Proceedings of the IFIP International Conference on Wireless and Optical Communications Networks (WOCN '07)*, pp. 1–5, IEEE, 2007.

[3] H. Shi, W. Wang, and K. Ngaiming, "Energy dependent divisible load theory for wireless sensor network workload allocation," *Mathematical Problems in Engineering*, vol. 2012, Article ID 235289, 16 pages, 2012.

[4] P. Gil-Jimenez, H. Gomez-Moreno, P. Siegmann, S. Lafuente-Arroyo, and S. Maldonado-Bascón, "Traffic sign shape classification based on Support Vector Machines and the FFT of the signature of blobs," in *Proceedings of the IEEE Intelligent Vehicles Symposium (IV' 07)*, pp. 375–380, 2007.

[5] Y. Wang and Z. Li, "Study on feature extraction method in border monitoring system using optimum wavelet packet decomposition," *AEU—International Journal of Electronics and Communications*, vol. 66, no. 7, pp. 575–580, 2012.

[6] P. Chakraborty, S. Kumar, R. Ghosh, K. Sardana, and A. Akula, "Computational techniques for classification of military vehicles using seismic signatures," in *Proceedings of the 3rd International Conference on Computing Communication & Networking Technologies (ICCCNT '12)*, pp. 1–4, 2012.

[7] P.-W. Shan and M. Li, "Nonlinear time-varying spectral analysis: HHT and MODWPT," *Mathematical Problems in Engineering*, vol. 2010, Article ID 618231, 14 pages, 2010.

[8] M. Duarte and H. Hu, "Vehicle classification in distributed sensor networks," *Journal of Parallel and Distributed Computing*, vol. 64, no. 7, pp. 826–838, 2004.

[9] E. Lupu, Z. Feher, and P. Pop, "Speaker verification rate study using the TESPAR coding method," in *Proceedings of the COST 276 Workshop on Information and Knowledge Management for Integrated Media Communication*, 2002.

[10] S. Abdusslam, M. Ahmed, P. Raharjo, F. Gu, and A. Ball, "Time encoded signal processing and recognition of incipient bearing faults," in *Proceedings of the 17th International Conference on Automation & Computing (ICAC '11)*, pp. 289–293, 2011.

[11] G. Mazarakis and J. Avaritsiotis, "Vehicle classification in sensor networks using time-domain signal processing and Neural Networks," *Microprocessors and Microsystems*, vol. 31, no. 6, pp. 381–392, 2007.

[12] S. M. Chen, D. F. Wang, and J. M. Zan, "Interior noise prediction of the automobile based on hybrid FE-SEA method," *Mathematical Problems in Engineering*, vol. 2011, Article ID 327170, 20 pages, 2011.

[13] J. Ryu, J.-J. Choi, B.-D. Hahn, D.-S. Park, W.-H. Yoon, and K.-Y. Kim, "Sintering and piezoelectric properties of KNN ceramics doped with KZT," *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, vol. 54, no. 12, pp. 2510–2515, 2007.

[14] S. Tan, "An effective refinement strategy for KNN text classifier," *Expert Systems with Applications*, vol. 30, no. 2, pp. 290–298, 2006.

[15] C. Liao and T. Chang, "Applying ranking SVM in query relaxation," in *Proceedings of the 30th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR '07)*, pp. 763–764, Association for Computing Machinery, 2007.

[16] S. H. Yen and C. J. Wang, "SVM based watermarking technique," *Tamkang Journal of Science and Engineering*, vol. 9, no. 2, pp. 141–150, 2006.

[17] Z. Guo, Y. Dong, J. Wang, and H. Lu, "The forecasting procedure for long-term wind speed in the Zhangye area," *Mathematical Problems in Engineering*, vol. 2010, Article ID 684742, 17 pages, 2010.

[18] C. Lardeux, P.-L. Frison, J.-P. Rudant, J.-C. Souyris, C. Tison, and B. Stoll, "Use of the SVM classification with polarimetric SAR data for land use cartography," in *Proceedings of the IEEE International Conference on Geoscience and Remote Sensing Symposium (IGARSS '06)*, pp. 493–496, 2006.

[19] W. Yang and H. Shi, "Sensor selection schemes for consensus based distributed estimation over energy constrained wireless sensor networks," *Neurocomputing*, vol. 87, pp. 132–137, 2012.

Hindawi Publishing Corporation
Mathematical Problems in Engineering
Volume 2014, Article ID 757282, 12 pages
http://dx.doi.org/10.1155/2014/757282

Research Article

H_∞ Controller Design for Asynchronous Multirate Sampled-Data Systems

Xiaoqiang Sun and Weijie Mao

Department of Control Science and Engineering, Zhejiang University, Hangzhou 310027, China

Correspondence should be addressed to Weijie Mao; wjmao@zju.edu.cn

Received 29 October 2013; Revised 22 December 2013; Accepted 6 January 2014; Published 23 February 2014

Academic Editor: Ilse C. Cervantes

Copyright © 2014 X. Sun and W. Mao. 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 considers the analysis and synthesis of a linear discrete asynchronous multirate sampled-data system. An H_∞ controller based on an observer is proposed, which guarantees the stability of the closed system and makes the H_∞ norm of the closed system less than a given attenuation level. To improve the performance further, a tradeoff strategy is applied. That is, the exogenous signals sampled at different rates are lifted to an appropriate signal rate, while the endogenous signals are not lifted for avoiding the causal constraint and the dimension multiplied again. The controller is obtained by solving the corresponding matrix inequality, which can be calculated by Matlab. Finally, an example is presented to demonstrate the validity of these methods.

1. Introduction

Sampled-data control systems, which provide digital control techniques, and emerged with the development of computers, have been long studied by many researchers [1, 2]. Multirate sampled-data control systems are required when the signals of interest are sampled at different rates compared with others. There are two primary reasons for this requirement. First, it has been shown that multirate sampling can improve the closed-loop control performance significantly. Second, technological or economic constraints in some real applications necessitate the use of control schemes where sensor measurements and control inputs have to be performed at different sampling rates. Figure 1 shows an ideal model of a standard multirate sampled-data control system. For example, in hybrid electric vehicles, a hierarchical controller is designed to decide the torque demands of motor, generator, internal combustion engine, and mechanical brake according to the driver's torque demand, speed of vehicle, and battery's state of charge (SOC), where the SOC is estimated by battery management system (BMS) and the speed of vehicle is fed by sensor [3]. As a large-scale and complex system, the sensors and control inputs of subsystems are difficult to be performed at the same sampling rates; thus a multirate sampled-data system will be a better choice for modeling.

Multirate systems have been studied since the late 1950s, resulting in the development and application of various sampling methods. Classical approaches include the lifting method, H_2/H_∞ control [4, 5], LMI method [6], prediction control, LQG control, iterative control [7], and parameter estimation [8]. Among these, the lifting method [9] is a vital tool for studying multirate systems, whereby such systems can be changed into single-rate systems. This method can improve the performance of a plant to some extent [10]. However, during the lifting process the dimensions of the system will be multiplied and it is increasingly difficult to analyze the new system effectively, especially systems with very different sampling rates, which can lead to the "model explosion" problem. H_2/H_∞ control is a useful method for analyzing systems and for synthesizing control problems based on the lifting technique, and LMI tools can be used to solve the inequalities [11]. By the way in real applications, fast-input and slow-output is a good way of improving the performance of the system, and this mechanism and its results have been discussed in previous studies [12, 13].

Most of lifting methods require that all signals should be sampled and held during the same time slot [1, 2, 9, 10, 14, 15].

That is, at a specific k,

$$S\left(k\overline{T}\right) = \operatorname{diag}\left\{\left[I_1 \;\; 0 \;\; \cdots \;\; 0\right]_{\overline{T}_1}, \left[I_2 \;\; 0 \;\; \cdots \;\; 0\right]_{\overline{T}_2}, \ldots, \left[I_p \;\; 0 \;\; \cdots \;\; 0\right]_{\overline{T}_p}\right\},$$

$$H\left(k\widetilde{T}\right) = \operatorname{diag}\left\{\left[I_1 \;\; \cdots \;\; I_1\right]'_{\widetilde{T}_1}, \left[I_2 \;\; \cdots \;\; I_2\right]'_{\widetilde{T}_2}, \ldots, \left[I_q \;\; \cdots \;\; I_q\right]'_{\widetilde{T}_q}\right\}, \tag{1}$$

where I_i, $i = 1, 2, \ldots, p(q)$, are unit matrices and $\overline{T}$, $\overline{T}_i$, $\widetilde{T}$ and $\widetilde{T}_i$, will be defined in Section 2. However, in most actual situations, this condition cannot be satisfied because the overall system lacks a slot when the entire signal can be sampled or held, and there is a class of multirate systems that each output of them has its own frequency of measurement and each input has its own frequency of updating [16–19]. A system with this type of sampling mechanism is referred to as asynchronous system in this paper. The traditional lifting technique is not applicable for the sampling frequencies of the sampling and holding elements are incommensurate [4, 20].

LQG control based on an observer is an effective method for reducing the dimensions of a multirate system without using the lifting technique [14, 21]. An LQG optimal controller derived is periodic and can guarantee the overall closed-loop stability. To compensate for the skewed input signal caused by slow-output feedback, only the input signal is supplemented by the observer. However, in the previous results, the controller and observer are designed separately, which may result in conservativeness.

Based on the methods mentioned above, the design of a controller for a multirate sampled-data system that does not multiply the dimensions necessitates the synthesis of various appropriate measures. In this paper, we propose an observer-based H_∞ method for analysis and synthesis of linear discrete asynchronous multirate sampled-data systems. In contrast to the traditional lifting method, the control inputs and measured outputs are not lifted to avoid the dimension multiplication and causal constraints. However, the exogenous system signals are lifted rationally to change the system into a fast-input-slow-output system. The controller can be obtained by solving an inequality with LMI tools of Matlab. Although the inequality is nonlinear, it can be calculated linearly under a constraint [22]. The approach proposed in this paper is different from the observer-based LQG method in [21]. Here, the controller and observer are obtained simultaneously by solving a matrix inequality under the same optimal performance constraint, thus reducing the conservativeness and improving the system performance. In addition, the concept of asynchronism in this paper is different from those in [23, 24], and so forth, where the sampling at sensors is assumed to be asynchronized with time-varying delay and it falls into the framework of time-delay systems, not the multirate sampled-data systems.

The remainder of this paper is organized as follows. Section 2 provides a description of the multirate system and its background, including lemmas and definitions. In Section 3, the main results obtained using the synthesis methods are presented. In Section 4, a simulated numerical example is provided to demonstrate the effectiveness of the method. Finally, Section 5 concludes the paper.

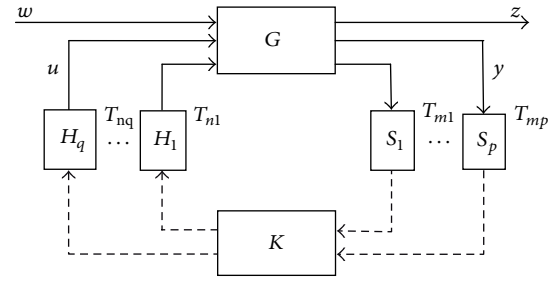


Figure 1: Multirate sampled-data system.

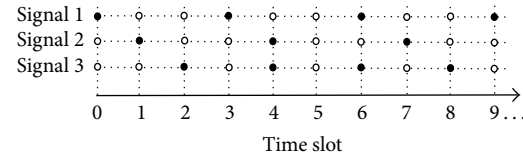


Figure 2: Asynchronous signals from the sampler or the holder.

2. System Descriptions

Definition 1. A multirate sampled-data system is called an asynchronous system if the signals from its sampler or holder cannot be obtained simultaneously during any time slot.

For example, as shown in Figure 2, Signal 1 and Signal 2 are the signals of the sampler and the holder with period 3, respectively, while Signal 3 is the signal with period 2. Thus, they cannot be obtained simultaneously during any time slot because the odd period signals appear alternately. In Figure 2, the filled circles indicate that the signal is sampled or held in the slot, whereas the hollow cylinders indicate that the signal is not sampled or held. Therefore, this system cannot satisfy the assumption mentioned in most previous studies where all signals can be sampled or held during the same time slot.

We consider a linear discrete asynchronous multirate sampled-data system:

$$\begin{aligned} x(k+1) &= Ax(k) + Bu(k) + M_1 w(k), \\ z(k) &= Cx(k) + M_2 w(k), \\ y(k) &= S(k) z(k), \end{aligned} \tag{2}$$

where $k \in \mathbb{N}$, $x(k) \in \mathbb{R}^n$, $u(k) \in \mathbb{R}^q$, $z(k) \in \mathbb{R}^p$, $w(k) \in \mathbb{R}^l$, $x(0) = x_0$, $|M_2| \neq 0$, $S(k)$ is a sampling device, $y(k)$ is the output measured vector, and x_0 is the initial state.

Assume that the ith component $u_i(k)$, $i = 1, 2, \ldots, q$, of the input vector $u(k)$ can be modified every $\widetilde{T}_i$ time instant

$\widetilde{T}_i \in \mathbb{Z}^+$. Then, $u(k)$ is the output of the following discrete-time period system, which is defined as the input-holding mechanism:

$$\begin{aligned} v(k+1) &= (I - H(k))\, v(k) + H(k)\, r(k), \\ u(k) &= (I - H(k))\, v(k) + H(k)\, r(k), \end{aligned} \tag{3}$$

where

$$\begin{aligned} H(k) &= \operatorname{diag}\left\{h_1(k), h_2(k), \ldots, h_q(k)\right\}, \\ h_i(k) &= \begin{cases} 1 & k = j\widetilde{T}_i + \widetilde{\tau}_i, \\ 0 & k \neq j\widetilde{T}_i + \widetilde{\tau}_i, \end{cases} \quad j \text{ integer}, \end{aligned} \tag{4}$$

where $v(k) := \left[v_1(k)\ \ v_2(k)\ \cdots\ v_q(k)\right]'$ is a new state variable and $r(k) := \left[r_1(k)\ \ r_2(k)\ \cdots\ r_q(k)\right]'$ is a new input variable. The integer $0 \le \widetilde{\tau}_i < \widetilde{T}_i$ $(i = 1, 2, \ldots, q)$ describes the skew of the inputs-holding mechanism.

Then, matrix $H(\cdot)$ and system (2) have a period of $\widetilde{T}$, and

$$\widetilde{T} := \underset{i=1,2,\ldots,q}{\text{l.c.m.}} \left\{\widetilde{T}_i\right\}. \tag{5}$$

Similarly, the output sampling mechanism is given as

$$S(k) = \operatorname{diag}\left\{v_1(k), v_2(k), \ldots, v_p(k)\right\}, \tag{6}$$

where

$$v_i(k) = \begin{cases} 1 & k = j\overline{T}_i + \overline{\tau}_i, \\ 0 & k \neq j\overline{T}_i + \overline{\tau}_i, \end{cases} \quad j \text{ integer}, \tag{7}$$

where $0 \le \overline{\tau}_i < \overline{T}_i$ $(i = 1, 2, \ldots, p)$ indicates the skew of the outputs mechanism.

Matrix $S(\cdot)$ has a $\overline{T}$-period, and

$$\overline{T} := \underset{i=1,2,\ldots,p}{\text{l.c.m.}} \left\{\overline{T}_i\right\}. \tag{8}$$

After combining (2) and (3), we obtain a new sampling-holding system:

$$\begin{aligned} \xi(k+1) &= \Phi(k)\,\xi(k) + \Gamma(k)\, r(k) + \Psi w(k), \\ z(k) &= \Upsilon\xi(k) + M_2 w(k), \\ y(k) &= \Delta(k)\,\xi(k) + \Omega(k)\, w(k), \end{aligned} \tag{9}$$

where

$$\begin{aligned} \xi(k) &:= \begin{bmatrix} x(k) \\ v(k) \end{bmatrix}, \qquad \Phi(k) := \begin{bmatrix} A & B(I - H(k)) \\ 0 & I - H(k) \end{bmatrix}, \\ \Gamma(k) &:= \begin{bmatrix} B \\ I \end{bmatrix} H(k), \qquad \Psi := \begin{bmatrix} M_1 \\ 0 \end{bmatrix}, \\ \Upsilon &:= \begin{bmatrix} C & 0 \end{bmatrix}, \qquad \Delta(k) := \begin{bmatrix} S(k)C & 0 \end{bmatrix}, \\ \Omega(k) &:= S(k) M_2. \end{aligned} \tag{10}$$

Based on the $\widetilde{T}$ of matrix $H(\cdot)$ and the $\overline{T}$ of matrix $S(\cdot)$, system (9) has the period of

$$T := \text{l.c.m.}\left\{\widetilde{T}, \overline{T}\right\}. \tag{11}$$

The state observer of system (2) is

$$\widehat{x}(k+1) = A\widehat{x}(k) + Bu(k) + L(k) \cdot \left[y(k) - S(k) C\widehat{x}(k)\right], \tag{12}$$

where $L(k)$ asymptotically stabilizes system (12).

The observer-based controller is chosen as

$$r(k) = K(k)\,\widehat{\xi}(k), \tag{13}$$

where $\widehat{\xi}(k) := \left[\begin{smallmatrix} \widehat{x}(k) \\ v(k) \end{smallmatrix}\right]$.

The state error is defined as

$$e(k) := x(k) - \widehat{x}(k). \tag{14}$$

According to (2), (12) and (14), the following equation can be obtained:

$$\begin{aligned} e(k+1) = {} & [A - L(k) S(k) C]\, e(k) \\ & + \left(M_1 - L(k) S(k) M_2\right) w(k). \end{aligned} \tag{15}$$

Combining (9), (13), and (15) yields

$$\begin{aligned} \begin{bmatrix} \xi(k+1) \\ e(k+1) \end{bmatrix} &= \widetilde{A} \begin{bmatrix} \xi(k) \\ e(k) \end{bmatrix} + \widetilde{B} w(k), \\ z(k) &= \widetilde{C} \begin{bmatrix} \xi(k) \\ e(k) \end{bmatrix} + \widetilde{D} w(k), \end{aligned} \tag{16}$$

where

$$\begin{aligned} \widetilde{A} &:= \begin{bmatrix} \Phi(k) + \Gamma(k) K(k) & -\Gamma(k) K(k) \begin{bmatrix} I_2 \\ 0 \end{bmatrix} \\ 0 & A - L(k) S(k) C \end{bmatrix}, \\ \widetilde{B} &:= \begin{bmatrix} \Psi \\ M_1 - L(k) S(k) M_2 \end{bmatrix}, \\ \widetilde{C} &:= \begin{bmatrix} \Upsilon & 0 \end{bmatrix}, \qquad \widetilde{D} := M_2. \end{aligned} \tag{17}$$

Lemma 2 (see [21, 25]). *If*

(i) *the pair* (A, B) *is stabilizable,*

(ii) *the pair* (A, M_1) *is stabilizable,*

(iii) *the pair* (A, C) *is detectable,*

(iv) *for any pair of distinct eigenvalues of* A, λ_i *and* λ_j, $|\lambda_i| \ge 1$, $|\lambda_j| \ge 1$, *it follows that* $\lambda_i^{\widetilde{T}} \neq \lambda_j^{\widetilde{T}}$,

(v) *there is no eigenvalue* λ *of* A, $\lambda \neq 1$, *such that* $\lambda^{\widetilde{T}} = 1$,

then the pairs $(\Phi(k), \Gamma(k))$ *and* $(\Phi(k), \Psi)$ *are controllable, and the pair* $(\Phi(k), \Upsilon)$ *is detectable.*

Lemma 3 (see [10] (the Bounded Real Lemma)). *For a given scalar $\gamma > 0$, if there is a matrix $X = X' > 0$ that satisfies the matrix inequality*

$$\begin{bmatrix} X & \widetilde{A}X & \widetilde{B} & 0 \\ X\widetilde{A}' & X & 0 & X\widetilde{C}' \\ \widetilde{B}' & 0 & \gamma I & \widetilde{D}' \\ 0 & \widetilde{C}X & \widetilde{D} & \gamma I \end{bmatrix} > 0, \tag{18}$$

then the system (16) *is asynchronously stable and it achieves a specified attenuation level $\gamma > 0$, such that $\|T_{zw}\|_\infty < \gamma$.*

3. Main Results

For the linear discrete asynchronous multirate sampled-data system (9), the exogenous parts of the system are lifted to improve its performance. To avoid multiplying the dimensions of the control inputs and measured outputs, we do not lift them, thereby avoiding causal constraints of the traditional lifting method.

If we assume that there is an underlying clock with a base period of τ, then system (9) represents a fast discrete system with a subperiod of τ/m for the discrete process described in [9]. The sampler S_a samples the continuous signal w for a discrete a-period and the holder H_b holds the output signal z during period b. The rates of the sampler and the holder can be set at different flexible rates; that is, $a \neq b$. If we select (13) as the new input, $K(k)$ is the controller of system (9).

These concepts are described in Figure 3, where

$$m := \text{l.c.m.}\{a, b\}. \tag{19}$$

Theorem 4. *Consider the sampled-data system* (9) *with the lifting process shown in Figure 3. For a given attenuation level $\gamma > 0$, if there exist matrices $X_1(k) = X_1'(k) > 0$, $X_2(k) = X_2'(k) > 0$, $K(k)$, and $L(k)$ ($k = 0, \ldots, N$ denotes the iteration time), which are suboptimal solutions of the inequalities below, then the closed system of* (9) *with the controller* (13) *is asynchronously stable and satisfies the H_∞ norm constraint $\|T_{zw}\|_\infty < \gamma$:*

$$\begin{bmatrix} X_1(k) & 0 & A_d + B_{2d}K(k) & -B_{2d}K(k)\begin{bmatrix} I_2 \\ 0 \end{bmatrix} & \underline{B}_1 & 0 \\ * & X_2(k) & 0 & A - L(k)S(k)C & \left(M_1 - L(k)S(k)M_2\right)\left[I_l \;\; 0 \;\; \cdots \;\; 0\right]_m & 0 \\ * & * & X_1^{-1}(k) & 0 & 0 & \underline{C}_1' + K'(k)\underline{D}_{12}' \\ * & * & * & X_2^{-1}(k) & 0 & -\left[I_2 \;\; 0\right]K'(k)\underline{D}_{12}' \\ * & * & * & * & \gamma I & \underline{D}_{11}' \\ * & * & * & * & * & \gamma I \end{bmatrix} > 0. \tag{20}$$

Proof. After lifting system (9) according to Figure 3, the transform function of a completely discrete system with a period of τ/m, which is partly contained within the dot-dash line, is

$$\begin{aligned} \underline{G} &= \begin{bmatrix} L_m & 0 \\ 0 & I \end{bmatrix}\begin{bmatrix} S_f & 0 \\ 0 & S(k) \end{bmatrix}\begin{bmatrix} G_{11} & G_{12} \\ G_{21} & G_{22} \end{bmatrix}\begin{bmatrix} H_f & 0 \\ 0 & H(k) \end{bmatrix}\begin{bmatrix} L_m^{-1} & 0 \\ 0 & I \end{bmatrix} \\ &= \begin{bmatrix} L_m S_f G_{11} H_f L_m^{-1} & L_m S_f G_{12} H(k) \\ S(k) G_{21} H_f L_m^{-1} & S(k) G_{22} H(k) \end{bmatrix}. \end{aligned} \tag{21}$$

Based on the above, we can derive minimal state-space representations $\widehat{g}(\lambda)$ for $\underline{G}$ as follows based on the transfer function theory given in [9, 15]. However, in contrast to the theory, H and S are not blocks of unit matrices because sampling and holding are not synchronous; that is, the entire signal cannot be sampled or held at the same time:

$$\widehat{g}(\lambda) = \begin{bmatrix} P_{11} & P_{12} \\ P_{21} & P_{22} \end{bmatrix}. \tag{22}$$

Next, we calculate the minimal state-space realization of (22) as follows.

(a) Transfer Function for P_{11}. Note that $P_{11} = L_m S_f G_{11} H_f L_m^{-1}$, which comes directly from the theory in [8], and its corresponding state model is

$$\widehat{P}_{11}(\lambda) = \left[\begin{array}{c|c} A_d & \underline{B}_1 \\ \hline \underline{C}_1 & \underline{D}_{11} \end{array}\right], \tag{23}$$

where

$$A_d := \begin{bmatrix} A^m & \sum_{j=1}^{m} A^{m-j} B (I - H(k)) \\ 0 & I - H(k) \end{bmatrix},$$

$$\underline{B}_1 := \begin{bmatrix} A^{m-1}M_1 & A^{m-2}M_1 & \cdots & M_1 \\ 0 & 0 & \cdots & 0 \end{bmatrix},$$

$$\underline{C}_1 := \begin{bmatrix} C & 0 \\ CA & 0 \\ \vdots & \vdots \\ CA^{m-1} & 0 \end{bmatrix},$$

$$\underline{D}_{11} := \begin{bmatrix} M_2 & 0 & \cdots & 0 \\ CM_1 & M_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ CA^{m-2}M_1 & CA^{m-3}M_1 & \cdots & M_2 \end{bmatrix}. \tag{24}$$

(b) Transfer Function for P_{12}. As $H(k)$ in $P_{12} = L_m S_f G_{12} H(k)$ is not an m-blocks unit matrix, the theory cannot be applied directly. However, note that

$$S_f H(k) = L_m^{-1} \begin{bmatrix} I \\ I \\ \vdots \\ I \end{bmatrix}_m H(k). \tag{25}$$

Then,

$$P_{12} = L_m S_f G_{12} H_f L_m^{-1} \begin{bmatrix} I \\ I \\ \vdots \\ I \end{bmatrix}_m H(k). \tag{26}$$

The corresponding state model can be obtained as

$$\widehat{P}_{12}(\lambda) = \left[\begin{array}{c|c} A_d & B_{2d} \\ \hline \underline{C}_1 & \underline{D}_{12} \end{array}\right], \tag{27}$$

where A_d and $\underline{C}_1$ have been defined previously and

$$B_{2d} := \begin{bmatrix} \sum_{i=1}^{m} A^{m-i} B \\ I \end{bmatrix} H(k),$$

$$\underline{D}_{12} := \begin{bmatrix} 0 \\ 0 + \sum_{j=m}^{m} CA^{m-j} B \\ \vdots \\ 0 + \sum_{j=2}^{m} CA^{m-j} B \end{bmatrix} H(k). \tag{28}$$

(c) Transfer Function for P_{21}. Similarly, because $S(k)$ is also not an m-blocks unit matrix, P_{21} can be changed to

$$P_{21} = S(k) \begin{bmatrix} I & 0 & \cdots & 0 \end{bmatrix}_m L_m S_f G_{21} H_f L_m^{-1}. \tag{29}$$

Thus, the state model can be derived as

$$\widehat{P}_{21}(\lambda) = \left[\begin{array}{c|c} A_d & \underline{B}_1 \\ \hline C_2 & \underline{D}_{21} \end{array}\right], \tag{30}$$

where A_d and $\underline{B}_1$ were defined previously and

$$C_2 := \Delta(k),$$
$$\underline{D}_{21} := \begin{bmatrix} \Omega(k) & 0 & \cdots & 0 \end{bmatrix}. \tag{31}$$

(d) Transfer Function for P_{22}. Like to parts *b* and *c*, the state description of P_{22} is

$$\widehat{P}_{22}(\lambda) = \left[\begin{array}{c|c} A_d & B_{2d} \\ \hline C_2 & D_{22} \end{array}\right], \tag{32}$$

where A_d, B_{2d}, and C_2 were defined previously, while

$$D_{22} := 0. \tag{33}$$

Finally, formula (23) is represented using the state-space form as follows:

$$\begin{aligned} \xi(k+1) &= A_d \xi(k) + \underline{B}_1 \underline{w}(k) + B_{2d} r(k), \\ \underline{z}(k) &= \underline{C}_1 \xi(k) + \underline{D}_{11} \underline{w}(k) + \underline{D}_{12} r(k), \\ y(k) &= C_2 \xi(k) + \underline{D}_{21} \underline{w}(k), \end{aligned} \tag{34}$$

where the underlining below z and w indicates that the signals are lifted.

According to Lemma 2, the controller $K(k)$ exists and can stabilize the system.

Let $\wp(k) := \left[\xi'(k) \;\; e'(k)\right]'$, and after combining (13), (15), and (34), we obtain

$$\begin{aligned} \wp(k+1) &= A_c \wp(k) + B_c \underline{w}(k), \\ \underline{z}(k) &= C_c \wp(k) + D_c \underline{w}(k), \end{aligned} \tag{35}$$

where

$$A_c := \begin{bmatrix} A_d + B_{2d} K(k) & -B_{2d} K(k) \begin{bmatrix} I_2 \\ 0 \end{bmatrix} \\ 0 & A - L(k) S(k) C \end{bmatrix},$$

$$B_c := \begin{bmatrix} \underline{B}_1 \\ (M_1 - L(k) S(k) M_2) \begin{bmatrix} I_l & 0 & \cdots & 0 \end{bmatrix}_m \end{bmatrix},$$

$$C_c := \left[\underline{C}_1 + \underline{D}_{12} K(k) \;\; -\underline{D}_{12} K(k) \begin{bmatrix} I_2 \\ 0 \end{bmatrix}\right], \tag{36}$$

$$D_c := \underline{D}_{11}.$$

In the light of [5], we set

$$X(k) := \begin{bmatrix} X_1(k) & 0 \\ 0 & X_2(k) \end{bmatrix}. \tag{37}$$

By substituting X in Lemma 3 with $X(k)$, we can obtain

$$\begin{bmatrix} X_1(k) & 0 & (A_d + B_{2d}K(k))X_1(k) & -B_{2d}K(k)\begin{bmatrix} I_2 \\ 0 \end{bmatrix} X_2(k) & \underline{B}_1 & 0 \\ * & X_2(k) & 0 & (A - L(k)S(k)C)X_2(k) & (M_1 - L(k)S(k)M_2)[I_l \ 0 \ \cdots \ 0]_m & 0 \\ * & * & X_1(k) & 0 & 0 & X_1(k)\left(\underline{C}_1' + K'(k)\underline{D}_{12}'\right) \\ * & * & * & X_2(k) & 0 & -X_2(k)[I_2 \ 0]K'(k)\underline{D}_{12}' \\ * & * & * & * & \gamma I & \underline{D}_{11}' \\ * & * & * & * & * & \gamma I \end{bmatrix} > 0. \tag{38}$$

Then, according to [26], after multiplying the left- and right-hand sides of inequality (38) by $\mathrm{diag}\{I, I, X_1^{-1}(k), X_2^{-1}(k), I, I\}$, we can obtain inequality (20).

According to quadratic stability theory, the system should be stable at each step k. Thus, for a discrete linear time-invariant system, the eigenvalues of the closed system should lie within a unit circle.

From Theorem 4, if matrix inequalities (20) can be solved, there exists a state feedback H_∞ controller that can asymptotically stabilize closed system (35) of enlarged system (34), which satisfies the performance index. Since system (34) is lifted from the original system (9) and we know that the lifting operator is an isometric isomorphism, that is, the feedback stability and system norm can be preserved, then system (16), which is the closed system of (9), is asynchronously stable and it satisfies the H_∞ norm constraint $\|T_{zw}\|_\infty < \gamma$. □

The synchronous case can be regarded as a special example of the asynchronous case. In this case, the elements of sampling and holder are unit matrices with the assumption that the whole signals of sampling and holding can be obtained at the beginning of every period; then the system (9) turns into a linear discrete synchronous multirate sampled-data system with period τ as follows:

$$\begin{aligned} \xi(k+1) &= \Phi\xi(k) + \Gamma r(k) + \Psi w(k), \\ z(k) &= \Upsilon\xi(k) + M_2 w(k), \\ y(k) &= \Delta\xi(k) + \Omega w(k). \end{aligned} \tag{39}$$

As H and S are unit matrices, that is, constant now, Φ, Γ, Δ, and Ω related to H and S are also constant. Theorem 4 can be applied to system (39) straightforwardly.

Corollary 5. *Consider the sampled-data system* (39) *with the lifting process shown in Figure 3. For a given attenuation level* $\gamma > 0$, *if there exist matrices* $X_1 = X_1' > 0$, $X_2 = X_2' > 0$, K, *and* L, *which are suboptimal solutions of the inequalities below, then the closed system of* (39) *with the controller* (13) *is synchronously uniformly stable and it satisfies the* H_∞ *norm constraint* $\|T_{zw}\|_\infty < \gamma$:

$$\begin{bmatrix} X_1 & 0 & A_d + B_{2d}K & -B_{2d}K\begin{bmatrix} I_2 \\ 0 \end{bmatrix} & \underline{B}_1 & 0 \\ * & X_2 & 0 & A - LC & (M_1 - LM_2)[I_l \ 0 \ \cdots \ 0]_m & 0 \\ * & * & X_1^{-1} & 0 & 0 & \underline{C}_1' + K'\underline{D}_{12}' \\ * & * & * & X_2^{-1} & 0 & -[I_2 \ 0]K'\underline{D}_{12}' \\ * & * & * & * & \gamma I & \underline{D}_{11}' \\ * & * & * & * & * & \gamma I \end{bmatrix} > 0. \tag{40}$$

Remark 6. Let us compare the designed observer (12) with that in [21], where the observer (predictor) is

$$\begin{aligned} \hat{z}(k+1 \mid k) = {} & \overline{A}\hat{z}(k \mid k-1) + \overline{B}u(k) \\ & + L(k)\cdot\left[\overline{\zeta}(k) - \overline{C}(k)\hat{z}(k \mid k-1)\right], \end{aligned}$$

$$\overline{\zeta}(k) = \sum_{i=1}^{\overline{T}^*} y(k-i), \quad \overline{T}^* = \max_{i=1,\ldots,p}\left\{\overline{T}_i\right\}. \tag{41}$$

The sum of outputs is used to avoid the singularity but may cause more noise. The observer applied in this paper is

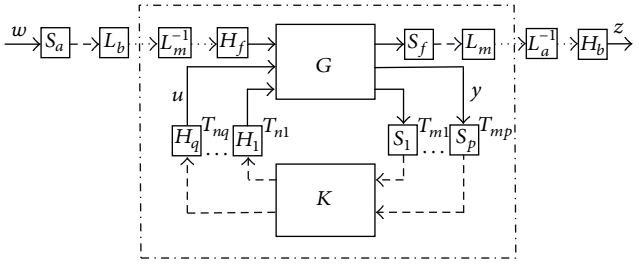


Figure 3: System lifting and controller designing.

simpler; that is, just the current output is introduced in the observer. This will result in better tracking performance and smaller state ripple than [21], where the past outputs other than current output are used for feedback.

Remark 7. In Theorem 4 and Corollary 5, due to the existence of inverse matrices X_1^{-1} and X_2^{-1}, inequalities (20) and (40) are not LMIs and cannot be directly solved by the LMI tool. However, notice that the matrices and their inverse matrices appear in pairs; we can solve the inequalities by iteration on LMIs according to Cone Complementarity Linearization Algorithm [22].

4. Numerical Example with Simulation

Consider the original plant described by

$$\begin{aligned} \dot{x}^c(t) &= A^c x^c(t) + B^c u^c(t) + M_1^c w^c(t), \\ z^c(t) &= C^c x^c(t) + M_2^c w^c(t), \end{aligned} \tag{42}$$

where

$$A^c = \begin{bmatrix} 0 & 1 \\ -6 & -5 \end{bmatrix}, \quad B^c = \begin{bmatrix} 0 & 1 \\ 1 & 1.2 \end{bmatrix}, \quad M_1^c = \begin{bmatrix} 1 \\ 0.6 \end{bmatrix},$$
$$C^c = \begin{bmatrix} 0.2 & 0.1 \end{bmatrix}, \quad M_2^c = 0.01. \tag{43}$$

If the fast discretization formulae follow $A = e^{A^c(\tau/m)}$, $B = \int_0^{\tau/m} e^{A^c\sigma} d\sigma B^c$, $M_1 = \int_0^{\tau/m} e^{A^c\sigma} d\sigma M_1^c$, $C = C^c$, and $M_2 = M_2^c$, then system (2) can be obtained. The initial state is $x_0 = [30 \;\; 20]$.

If $\widetilde{T}_1 = \widetilde{T}_2 = 2$, the input-holding mechanism can be described as

$$H(0) = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \quad H(1) = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}. \tag{44}$$

If $\overline{T}_1 = 3$ and $\overline{T}_2 = 1$, the output-sampling matrices are assumed to be

$$S(0) = 1, \quad S(1) = 0, \quad S(2) = 0. \tag{45}$$

Therefore, this is a linear discrete asynchronous dual-rate sampled-data system [18], which applies a fast-input-slow-output mechanism for $\widetilde{T} < \overline{T}$.

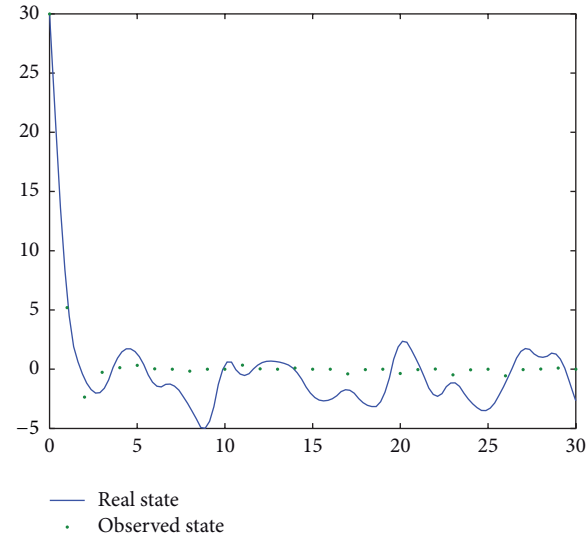


Figure 4: Plant state $x_1(k)$.

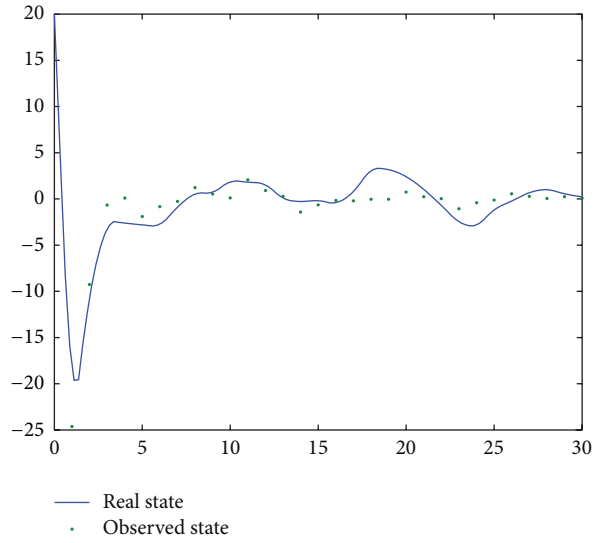


Figure 5: Plant state $x_2(k)$.

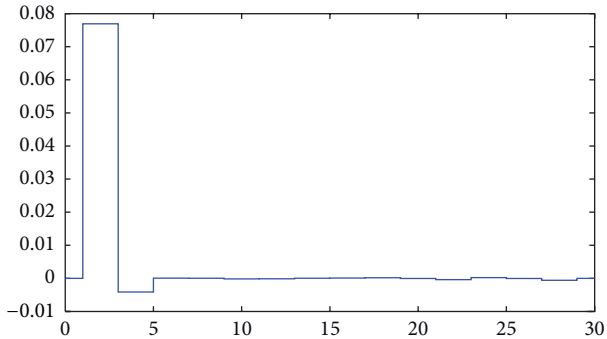


Figure 6: Control signal $u_1(k)$.

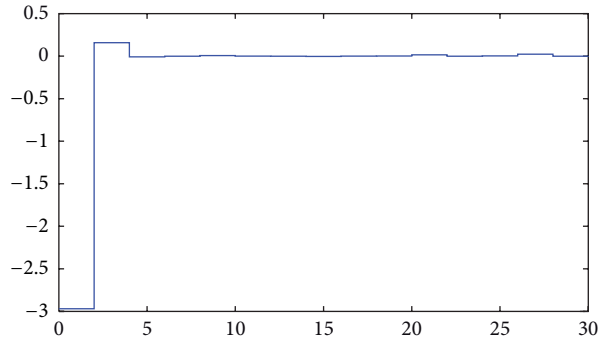


Figure 7: Control signal $u_2(k)$.

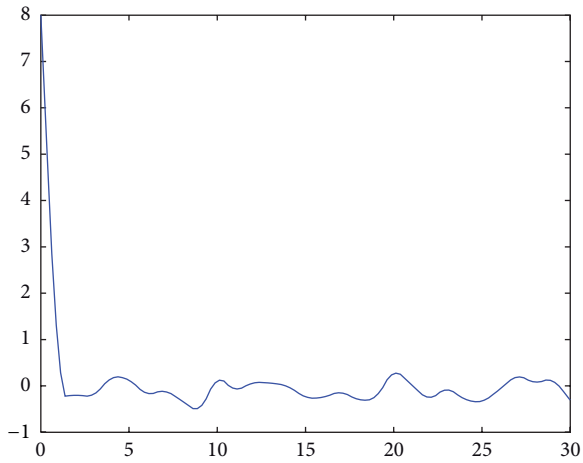


Figure 8: Output.

For comparison, we introduce a previous method [21], where the system is not lifted according to the pure LQG theory, and also let $w_1 = w_2 = w$ be Gaussian white noise; amplitude is increased to 10, which is much bigger to the system. The results obtained using this method are shown in Figures 4, 5, 6, 7, and 8, where $\tau = 1.2$ s.

Next, the synthesis method described in Theorem 4 is applied to the same plant. The simulations obtain the transients of the system shown in Figures 9, 10, 11, 12, and 13, where $\gamma = 0.2$, $a = 1$, and $b = 1$, while $m = 1$ (monorate).

A comparison of Figures 4–8 with Figures 9–13 shows clearly that the H_∞ method performs better than LQG; the controlled states are consistent with the observed states with smaller ripples, especially state $x_2(k)$, the biggest ripple of which is almost half of the LQG method. And the output of H_∞ method has better performance of disturbance attenuation than LQG method.

When $\tau = 0.6$ s and $m = 1$ (monorate), applying Theorem 4 again, we can get Figures 14, 15, 16, 17, and 18. Then the lifting number is modified to $m = 2$ (multirate); the results shown in Figures 19, 20, 21, 22, and 23 are obtained.

Comparing Figures 14–18 with Figures 19–23, when the lifting amplitude increases from 1 to 2, the controlled states are almost consistent with the observed states with smaller ripples; here the biggest ripple of state $x_2(k)$ is almost half of

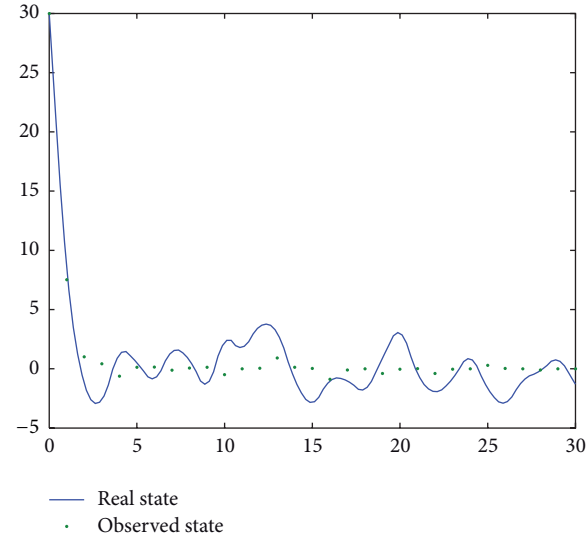


Figure 9: Plant state $x_1(k)$.

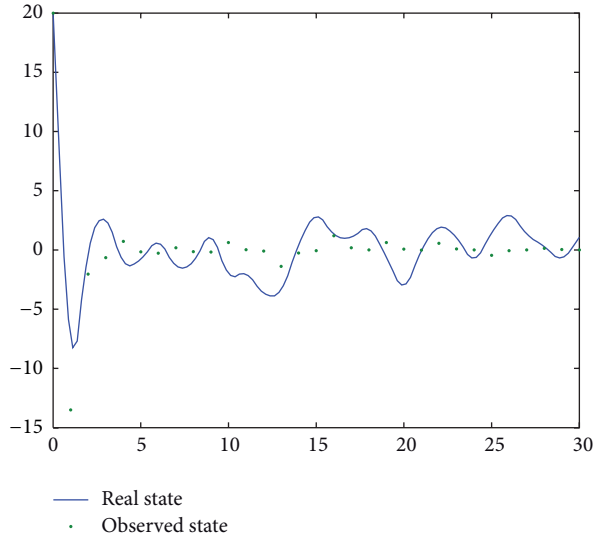


Figure 10: Plant state $x_2(k)$.

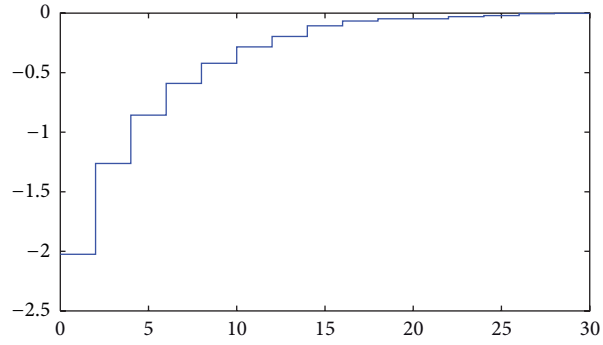


Figure 11: Plant control signal $u_1(k)$.

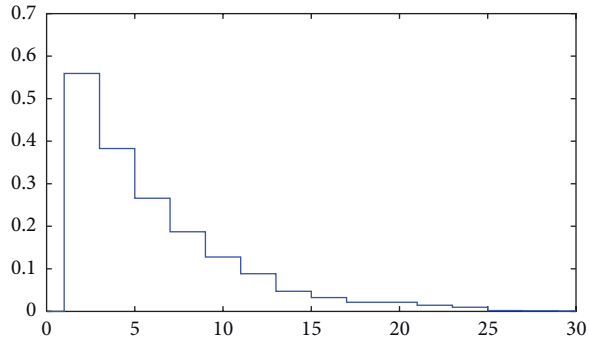


Figure 12: Plant control signal $u_2(k)$.

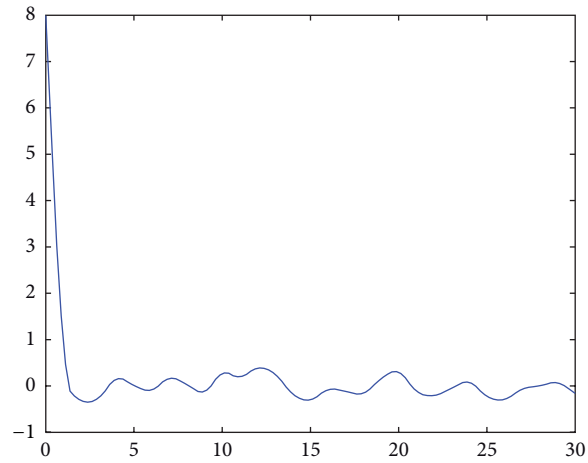


Figure 13: Output.

Table 1: Minimum H_∞ attenuation level.

τ	$m = 1$	$m = 2$	$m = 4$
1.2 sec	0.0974	0.0959	0.0942
0.6 sec	0.0904	0.0863	0.0856

the monorate. And the output of the multirate case has better performance of disturbance attenuation than the monorate case.

Furthermore, we change the lifting amplitude to $m = 4$, adjust the underlying period of system (2) to $\tau = 0.6$ and can draw the similar pictures; here we omit them but only list their minimum H_∞ attenuation levels to compare them with different cases; see Table 1.

The inequality acquired using the theorem is almost linear, except in a few terms. Although these terms are just the inverses of their corresponding terms, they could easily be calculated by LMI tool of Matlab under a constraint. Thus, it is far easier to solve the controller and observer under the same H_∞ constraint comparing with most of the existing methods, which use highly complex formulae.

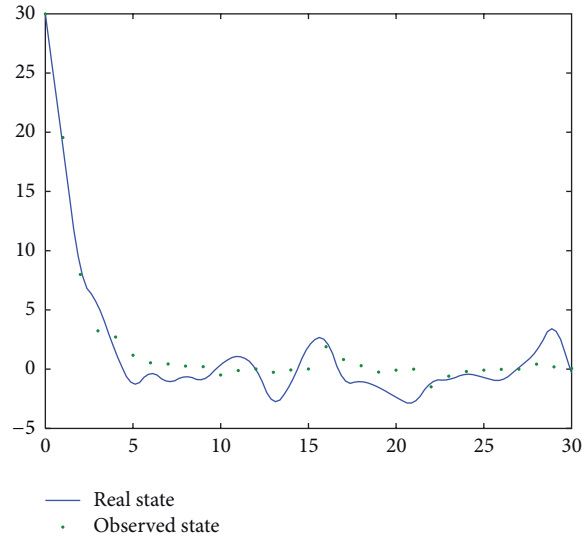


Figure 14: Plant state $x_1(k)$.

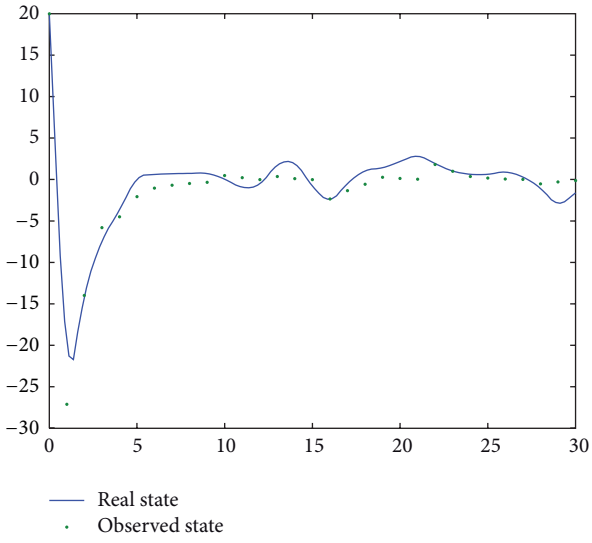


Figure 15: Plant state $x_2(k)$.

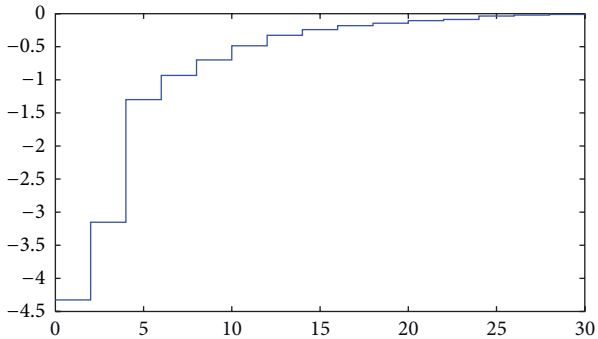


Figure 16: Plant control signal $u_1(k)$.

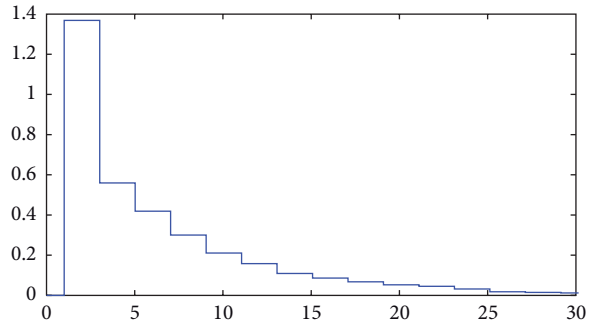


Figure 17: Plant control signal $u_2(k)$.

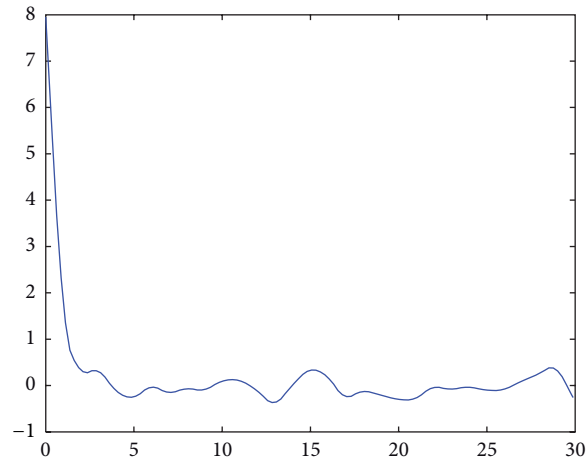


Figure 18: Output.

5. Conclusion

Motivated by a desire to provide a better solution for the design of a controller for a discrete-time asynchronous multirate sampled-data system, we developed synthesis methods in this paper to improve performance and guarantee closed-loop transient behavior via observer-based H_∞ control. Furthermore, an extended lifting technique was adopted to change the original system into a fast-input-slow-output discrete system, thereby improving its performance. Finally, we presented an example that demonstrates the effectiveness of these methods in controlling a discrete-time asynchronous multirate sampled-data system, whose performance is thus improved considerably.

In this paper, we focus on the standard H_∞ control of asynchronous multirate sampled-data system, and the input noise discussed here is Gaussian white for simplicity. As is well known, practical input signals often have finite frequency (FF), so it is reasonable to consider the H_∞ performance in finite frequency range for sampled-data systems. As in [27–31], where Kalman-Yakubovic-Popov (KYP) lemma is used to establish the equivalence between a frequency domain inequality (FDI) and a linear matrix inequality, it is a valuable work to extend the results in this paper to the FF framework based on KYP lemma. This will be the direction for future development of this paper.

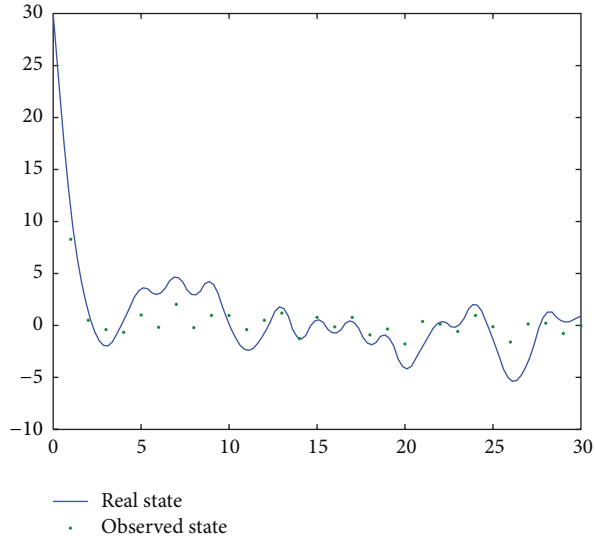


Figure 19: Plant state $x_1(k)$.

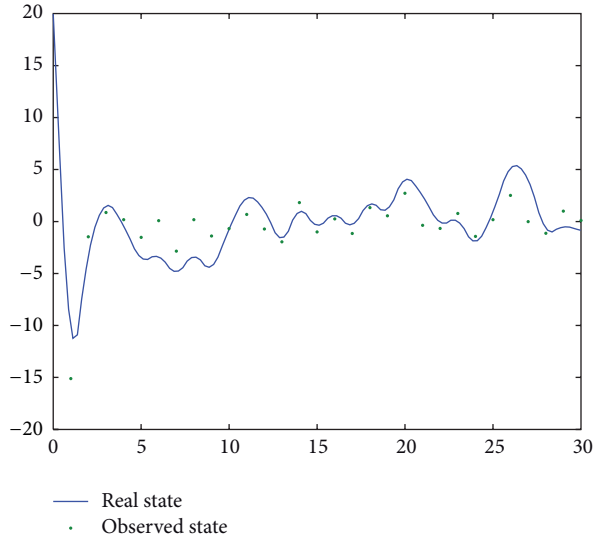


Figure 20: Plant state $x_2(k)$.

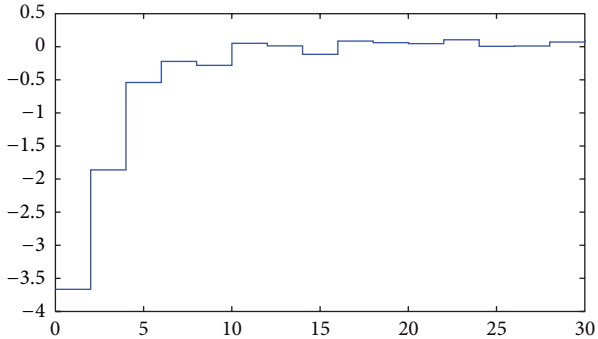


Figure 21: Plant control signal $u_1(k)$.

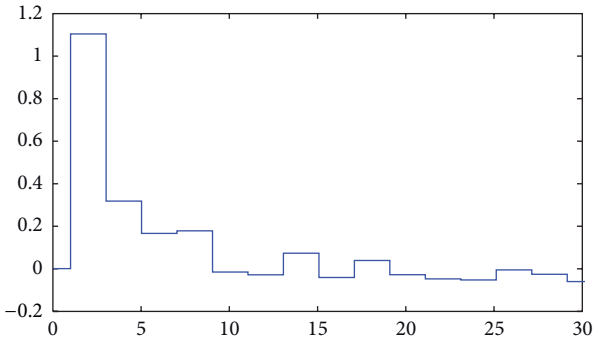


Figure 22: Plant control signal $u_2(k)$.

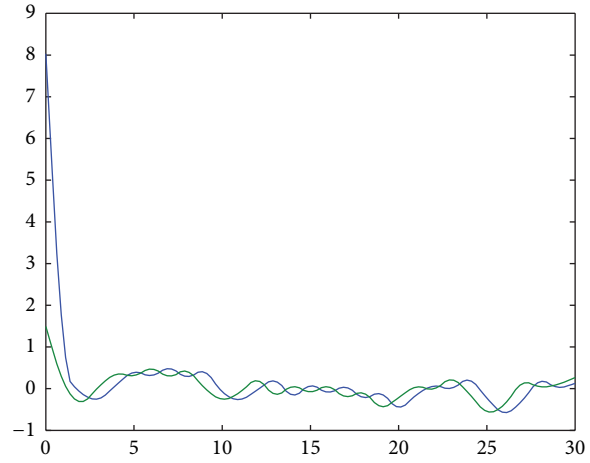


Figure 23: Output.

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 Zhejiang Provincial Natural Science Foundation of China under Grant no. LR12F03002 and National Natural Science Foundation of China (NSFC) under Grant no. 61074045.

References

[1] M. F. Sågfors, H. T. Toivonen, and B. Lennartson, "State-space solution to the period multirate H_∞ control problem: a lifting approach," *IEEE Transaction on Automatic Control*, vol. 37, no. 5, pp. 2345–2350, 2000.

[2] X. Zhao, Y. Yao, and J. Ma, "Overview of multirate sampled-data control system," *Journal of Harbin Institute of Technology*, vol. 9, no. 4, pp. 405–410, 2002.

[3] Y. Fukada, "Slip-angle estimation for vehicle stability control," *Vehicle System Dynamics*, vol. 32, no. 4, pp. 375–388, 1999.

[4] T. Chen and L. Qiu, "H_∞ design of general multirate sampled-data control systems," *Automatica*, vol. 30, no. 7, pp. 1139–1152, 1994.

[5] Z. Wang, B. Huang, and H. Unbehauen, "Robust H_∞ observer design of linear state delayed systems with parametric uncertainty: the discrete-time case," *Automatica*, vol. 35, no. 6, pp. 1161–1167, 1999.

[6] S. Boyd, L. El Ghaoui, E. Feron, and V. Balakrishnan, *Linear Matrix Inequalities in Systems and Control Theory*, vol. 15 of *SIAM Studies in Applied Mathematics*, SIAM, Philadelphia, Pa, USA, 1994.

[7] S. Patra and A. Lanzon, "A closed-loop data based test for robust performance improvement in iterative identification and control redesigns," *Automatica*, vol. 48, no. 10, pp. 2710–2716, 2012.

[8] J. Ding, Y. Shi, H. Wang, and F. Ding, "A modified stochastic gradient based parameter estimation algorithm for dual-rate sampled-data systems," *Digital Signal Processing*, vol. 20, no. 4, pp. 1238–1247, 2010.

[9] T. W. Chen and B. Francis, *Optimal Sampled-Data Control Systems*, Springer, London, UK, 1995.

[10] L. Shen and M. J. Er, "Multiobjectives design of a multirate output controller," in *Proceedings of the IEEE International Conference on Control Applications (CCA '99)*, pp. 193–198, August 1999.

[11] M. Chilali and P. Gahinet, "H_∞ design with pole placement constraints: an LMI approach," *IEEE Transactions on Automatic Control*, vol. 41, no. 3, pp. 358–367, 1996.

[12] S. Mo, X. Chen, J. Zhao, J. Qian, and Z. Shao, "A two-stage method for identification of dual-rate systems with fast input and very slow output," *Industrial and Engineering Chemistry Research*, vol. 48, no. 4, pp. 1980–1988, 2009.

[13] M. De la Sen and S. Alonso-Quesada, "Model matching via multirate sampling with fast sampled input guaranteeing the stability of the plant zeros: extensions to adaptive control," *IET Control Theory & Applications*, vol. 1, no. 1, pp. 210–225, 2007.

[14] M. F. Sågfors and H. T. Toivonen, "H_∞ and LQG control of asynchronous sampled-data systems," *Automatica*, vol. 33, no. 9, pp. 1663–1668, 1997.

[15] S. López-lópez, *Optimal H_∞ Design of Causal Multirate Controllers and Filters*, Mechanical and Aerospace Engineering, University of California, Irvine, Calif, USA, 2010.

[16] U. Borison, "Self-tuning regulators for a class of multivariable systems," *Automatica*, vol. 15, no. 2, pp. 209–215, 1979.

[17] V. S. Ritchey and G. F. Franklin, "A stability criterion for asynchronous multirate linear systems," *IEEE Transactions on Automatic Control*, vol. 34, no. 5, pp. 529–535, 1989.

[18] B. Shen, Z. Wang, and X. Liu, "Sampled-data synchronization control of dynamical networks with stochastic sampling," *IEEE Transactions on Automatic Control*, vol. 57, no. 10, pp. 2644–2650, 2012.

[19] S. Longo, G. Herrmann, and P. Barber, "Robust scheduling of sampled-data networked control systems," *IEEE Transactions on Control Systems Technology*, vol. 20, no. 6, pp. 1613–1621, 2011.

[20] P. Voulgaris, "Control of asynchronous sampled data systems," *IEEE Transactions on Automatic Control*, vol. 39, no. 7, pp. 1451–1455, 1994.

[21] P. Colaneri, R. Scattolini, and N. Schiavoni, "LQG optimal control of multirate sampled-data systems," *IEEE Transactions on Automatic Control*, vol. 37, no. 5, pp. 675–682, 1992.

[22] L. El Ghaoui, F. Oustry, and M. AitRami, "A cone complementarity linearization algorithm for static output-feedback and related problems," *IEEE Transactions on Automatic Control*, vol. 42, no. 8, pp. 1171–1176, 1997.

[23] A. Seuret, "Stability analysis of networked control systems with asynchronous sampling and input delay," in *Proceedings of the American Control Conference (ACC '11)*, pp. 533–538, July 2011.

[24] B. Tavassoli, "Stability of nonlinear networked control systems over multiple communication links with asynchronous sampling," *IEEE Transaction on Automatic Control*, 2013.

[25] L. Yang and J. M. Li, "Sufficient and necessary conditions of controllability and observability of a class of linear switching system," *Systems Engineering and Electronics*, vol. 25, no. 5, pp. 588–590, 2003.

[26] W.-J. Mao, "Observer-based energy decoupling of linear time-delay systems," *Journal of Zhejiang University*, vol. 37, no. 5, pp. 499–503, 2003.

[27] T. Iwasaki and S. Hara, "Generalized KYP lemma: unified frequency domain inequalities with design applications," *IEEE Transactions on Automatic Control*, vol. 50, no. 1, pp. 41–59, 2005.

[28] H. Gao and X. Li, "H_∞ filtering for discrete-time state-delayed systems with finite frequency specifications," *IEEE Transactions on Automatic Control*, vol. 56, no. 12, pp. 2935–2941, 2011.

[29] X. Li, H. Gao, and C. Wang, "Generalized Kalman-Yakubovich-Popov lemma for 2-D FM LSS model," *IEEE Transactions on Automatic Control*, vol. 57, no. 12, pp. 3090–3103, 2012.

[30] X. Li and H. Gao, "Robust finite frequency H_∞ filtering for uncertain 2-D systems: the FM model case," *Automatica*, vol. 49, no. 8, pp. 2446–2452, 2013.

[31] X. W. Li and H. J. Gao, "A heuristic approach to static output-feedback controller synthesis with restricted frequency-domain specifications," *IEEE Transaction on Automatic Control*, 2013.

Hindawi Publishing Corporation
Mathematical Problems in Engineering
Volume 2013, Article ID 785383, 9 pages
http://dx.doi.org/10.1155/2013/785383

Research Article

An Improved Ant Colony Algorithm and Its Application in Vehicle Routing Problem

Min Huang and Ping Ding

School of Software Engineering, South China University of Technology, Guangzhou, Guangdong 510006, China

Correspondence should be addressed to Min Huang; minh@scut.edu.cn

Received 6 August 2013; Revised 16 October 2013; Accepted 18 October 2013

Academic Editor: Ali Davoudi

Copyright © 2013 M. Huang and P. Ding. 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.

Optimal path planning is an important issue in vehicle routing problem. This paper proposes a new vehicle routing path planning method which adds path weight matrix and save matrix. The method uses a new transition probability function adding the angle factor function and visibility function, while setting penalty function in a new pheromone updating model to improve the accuracy of the route searching. Finally, after each cycle, we use 3-opt method to update the optimal solution to optimize the path length. The results of comparison also confirm that this method is better than the traditional ant colony algorithm for vehicle routing path planning method. The result of computer simulation confirms that the method can plan a more rational rescue path focused on the real traffic situation.

1. Introduction

Vehicle routing problem (VRP) is one of combinatorial optimization problems and VRP can be turned into other application problems; for example, Chen and Wang have turned VRP into tour planning problem [1]. A typical VRP can be described as follows: a warehouse that offers services to different positions of customers at the lowest cost path planning. It has real economic significance in many fields, such as transportation scheduling, routing, railway transportation, and other practical problems. Currently, heuristic algorithm is the main method for solving vehicle routing problem. Heuristic algorithm can be divided into simple heuristic algorithm, two-phase heuristic algorithm, and artificial intelligence methods.

Vehicle routing problem has successfully applied in many areas, such as Li et al. [2] and Malekly et al. [3]. They solve the uncertain and ambiguous problem in vehicle routing problem using fuzzy set theories. Toth and Vigo [4] considered VRP as a significant part in logistic handing and grouped the methods that have been found to solve the problem of VRP. Kim et al. [5] divided the waste collection business into different areas and proposed methods to solve these problems in VRP. Many researchers have proposed heuristics or metaheuristics algorithm for efficiently solving the vehicle routing problem. Semet and Taillard [6] designed the taboo algorithm considering time window and different types of vehicles, which mainly used methods to generate the initial solution, then optimize the initial solution with taboo search algorithm. It is a local neighborhood search for an extension. Cordeau and Laporte [7] studied simulated annealing algorithm for the VRP, they proposed a simulated annealing method which is suitable for solving the vehicle routing problem and shows the advantage of accuracy and the speed of search convergence. Genetic algorithm is good for solving combinatorial optimization problems. Niazy and Badr [8] and Zulvia et al. [9] use genetic algorithm (GA) encoding to solve VRP problem. Niazy and Badr solved the complexity of CVRP with the goal of minimizing the total distances using the cellular genetic algorithm (CGA). Multiple ant colony algorithms are proposed by Gambardella et al. [10] and Bell and McMullen [11] to solve the time window of the multiobjective VRP problem. Ant colony optimization [12–14] (ACO) was successfully applied to emergency rescue [15, 16], which is an ideal heuristic algorithm and capable of intelligent search and global optimization.

In this paper, we apply ant colony optimization algorithm to the path planning of emergency rescue vehicle. This paper

Table 1: Distance map and delivery of CVRP example.

Number	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
x	0	3	−2	−2	0	−3	3	−4	3	−3	2	1	−1	−2	2	4
y	0	2	1	−2	3	−3	−1	−1	4	0	1	3	−1	3	−2	3
Delivery	0	1.8	0.8	1.5	2.5	0.2	2.4	0.5	0.6	0.9	1.3	2	0.6	0.2	1.4	0.9

will focus on the shortcomings of ant colony algorithm. Our proposed algorithm adds the angle factor, path weight matrix, and save matrix in the transfer rules and optimizes visibility function and penalty function; thus, it avoids falling into local optimum and speeds up the convergence of the algorithm. At the same time, pheromone update rules with local update and global update, update the global optimal solution only use the best ant to improve the algorithm's global search. Finally, after each cycle update the optimal solution in accordance with 3-opt method [17, 18] to shorten the length of the rescue path. Use benchmark test [19] to compare the results with the literature [20] in the algorithm to verify the effectiveness of the algorithm.

2. Problem Description and Mathematical Model

Disasters such as earthquakes and floods are sudden and unpredictable, which cause serious sufferings for people. How to find an effective and short rescue path for the affected nodes in the limited time after the disasters become the focus thing? Before making vehicle routing problem path planning, we first establish reasonable free space model. Capacitated vehicle routing problem [21, 22] (CVRP) was defined as a transportation and a distribution problem with the constrained of vehicle load, travel time or distance. Research on this model is the longest, and the results are achieved the most. Meanwhile, there are a large number of heuristic algorithms for solving this problem. Thus, we established CVRP model for disaster emergency response.

Vehicle routing problem can be modeled by describing as the following. There is a rescue center a_0 to deliver goods to k disaster areas. We suppose that the ith area sent by the rescue center's m vehicle, where each vehicle can carry W ton relief supplies, demands g_i ($i = 1, 2, \ldots, k$). The goods are sent to each disaster area and finally back to the rescue center. If we know $g_i \leq W_m$, then we look for the shortest length that meet the demand of the affected node. It can be described as an undirected graph $G = (V, E)$, where G represents the layout of affected area after the occurrence of disasters, V is the vertices set corresponding with affected spot and rescue center, and E stands for any edge of two nodes.

In this paper we define several constraints as follows:

(a) the start and the end of each route are at the rescue center a_0;

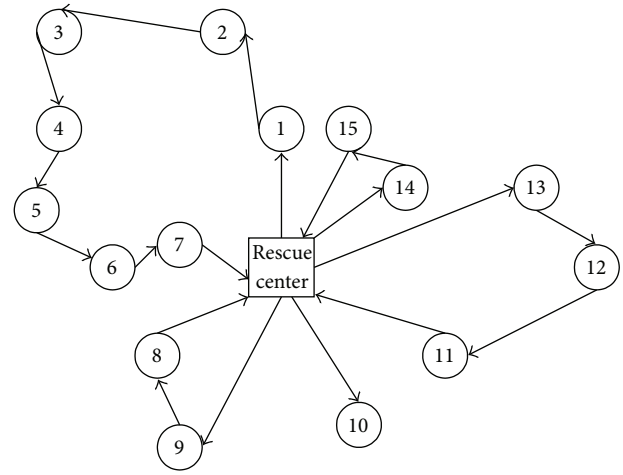


Figure 1: A real example of CVRP.

(b) the total amount of relief supplies loaded by vehicle must not exceed the maximum affected node of vehicle load;

(c) each affected node is visited once by only one vehicle in the limited time.

Figure 1 shows a work environment instance of the vehicle path planning. In this example, a rescue center a_0, the number of vehicles is 5 ($b = 5$, b means the number of vehicles); and the affected node is 15 ($v = 15$). Table 1 shows the affected nodes' coordinates and each area's relief supplies demands of Figure 1, and the rescue center coordinates are $(0, 0)$. The units of axes are meters, x represents the x-coordinate, and y represents the y-coordinate. The numbers $0, 1, 2, \ldots, 15$ sign each area.

When choosing the right path, the more spacious, smooth road is more likely to be selected; thus, road grade, traffic congestion levels, traffic, and other factors make up how to try to meet various requirements to plan a trip with a more reasonable path problem. When disaster occurs, how timely and reasonable given the rescue path is essential. Then the vehicle path planning is necessary to consider the following aspects: path length, road conditions, traffic smoothness, single time constraint that means each area can withstand the maximum rescue time, and weather conditions. Pheromones substance such as more than one rescue center and total time and length constraints needs to be pointed out in the paper. According to the [16] proposed method to calculate weights

matrix based on AHP, as is shown in Table 2, the number 0 node represents rescue center.

3. Brief Introduction of Ant Colony Optimization Algorithm

3.1. Ant Colony Optimization Algorithm. Ant colony optimization algorithm and the traditional ant colony algorithm are different in these three aspects as follows. That traditional ant colony algorithm is of slow convergence and is easy to fall into local optimum and stagnation in the vehicle path planning applications, while improved ant colony algorithm solves these problems in the following aspects.

(1) State transition probability function: the transfer rules that ants select the next affected node are modified, which is calculated as follows:

$$p_{ij}^{k}(t) = \begin{cases} \dfrac{\left[\tau_{ij}(t)\right]^{a} * D^{\beta} * \left[w_{ij}(t)\right]^{\gamma} * \left[U_{ij}(t)\right]^{\gamma} * \zeta_{ij}(j)}{\sum_{k \in \text{allowed}_k} \left[\tau_{ik}(t)\right]^{a} * D^{\beta} * \left[w_{ik}(t)\right]^{\gamma} * \left[U_{ik}(t)\right]^{\gamma} * \zeta_{ik}(k)}, & \text{if } j \in \text{allowed}_k \\ 0, & \text{otherwise,} \end{cases} \tag{1}$$

where τ_{ij} represents pheromones substance on the path edge (i, j), D stands for visibility, p_{ij}^{k} is transition probabilities of the kth ant from the affected node i to affected node j, the user parameters α and β are adjustable parameters that determine the relative influence of the pheromone and the visibility in the transition probabilities, γ means the relative importance of the path weights in the transition probability, and allowed_k represents the neighbor affected node of affected node i that kth ant has not yet visited. Here introducing the path weight value w_{ij} that stands for the weight value between affected nodes and the rescue center, $D = 1/(d_{ij} + d_{j,a_0})$ replaces the traditional visibility function; that is, $\eta_{ij} = 1/d_{ij}$, where the value d_{ij} means the distance between i and j and the value d_{j,a_0} is the distance between the affected node j and the rescue centre node. To modify visibility function is to enhance the ant's perception of rescue centre node in the current affected node, which will guide the ants to move and therefore avoid falling into local optimum search results and reduce the search time. Where U_{ij} stands for the save distances between nodes, both the affected nodes and the rescue centre. The save matrix calculated by the formula:

After adding this factor, it may avoid out the connection of a direct connection with the affected nodes, without through disaster rescue center node, meanwhile shortening the length of the path. Shown in Figure 2, $\zeta_{ij}(j)$ is an angle factor, which is the angle between the vector $[\overrightarrow{\text{Node}_{i^1}\text{Node}_i}, \overrightarrow{\text{Node}_i\text{Node}_j}]$; Node_{i^1} stands for the previous node before visited affected node i, the value Node_i that the black point represents the current visiting affected node i, Node_j represents the next visiting affected node j, red arrow direction equals with the opposite $\overrightarrow{\text{Node}_{i^1}\text{Node}_i}$, yellow arrow direction equals with $\overrightarrow{\text{Node}_i\,\text{Node}_j}$, and $\zeta_{ij}(j)$ is calculated as (2). When θ is between $[0, \pi/2]$, $\zeta_{ij}(j)$ is a constant 0.00001, which is for escape when an ant is faced with a dead-end street. Consider

$$\zeta_{ij}(j) = \begin{cases} \dfrac{\theta}{\pi}, & \text{if } \dfrac{\pi}{2} \le \theta \le \pi, \\ 0.00001, & \text{if } 0 \le \theta \le \dfrac{\pi}{2}. \end{cases} \tag{2}$$

$$j = \begin{cases} \arg\max\limits_{u \in JJ_i^k} \left\{ \tau_{iu}(t)^{\alpha} * D^{\beta} * \left[w_{iu}(t)\right]^{\gamma} * \left[U_{iu}(t)\right]^{\gamma} \right\}, & \text{if } q \le q_0, \\ JJ, & \text{if } q > q_0. \end{cases} \tag{3}$$

Ants choose the next affected node j in accordance with (3).

Where JJ is determined by (1), q is a random variable number between 0 and 1, q_0 is an adjustable parameter which is between 0 and 1, and here we define q_0 as 0.05. When (3) satisfied the condition $q > q_0$, it chooses the next access node in accordance with (1).

There are two rules in global update. One is the local optimal iteration and the other is the global optimal iteration.

(2) Update the local pheromone substances in accordance with the following:

$$\tau_{ij}(t+1) = \rho \cdot \tau_{ij}(t) + (1-\rho)\,\tau_{ij}(0). \tag{4}$$

In the progress of searching paths, according to state transition rules, ants apply (4) to choose route whose pheromone should be updated. After all ants move a step, local pheromones update once. Where $\tau_{ij}(0)$ represents the initial value of pheromone of edge (i, j), usually a small

Table 2: The path weight matrix.

W_{ij}	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
0	0	14	9	15	20	10	5	13	16	12	13	7	8	2	10	5
1	14	0	8	12	9	7	15	4	11	15	8	10	12	3	5	12
2	9	8	0	7	19	14	5	17	12	16	1	2	13	10	11	6
3	15	12	7	0	9	15	12	10	7	14	9	8	14	11	16	2
4	20	9	19	9	0	8	11	14	17	8	17	13	21	4	5	8
5	10	7	14	15	8	0	9	13	10	6	10	11	12	21	7	14
6	5	15	5	12	11	9	0	7	18	21	5	7	9	3	16	23
7	13	4	17	10	14	13	7	0	15	10	9	12	15	21	9	9
8	16	11	12	7	17	10	18	15	0	8	3	4	12	19	21	10
9	12	15	16	14	8	6	21	10	8	0	7	23	16	10	9	13
10	13	8	1	9	17	10	5	9	3	7	0	11	12	21	7	3
11	7	10	2	8	13	11	7	12	4	23	11	0	9	3	16	2
12	8	12	13	14	21	12	9	15	12	16	12	9	0	21	9	17
13	2	3	10	11	4	21	3	21	19	10	21	3	21	0	21	8
14	10	5	11	16	5	7	16	9	21	9	7	16	9	21	0	16
15	5	12	6	2	8	14	23	9	10	13	3	2	17	8	16	0

positive number to prevent ant colony from falling into local optimal solution prematurely, ρ is a user-defined coefficient, which stands for the degree of global pheromone evaporation.

(3) Update the Global pheromone substances with the following:

$$\tau_{ij}(t+n) = \rho \cdot \tau_{ij}(t) + (1-\rho)\Delta\tau_{ij}. \tag{5}$$

$$\Delta\tau_{ij} = \begin{cases} \dfrac{1}{L}, & \text{if edge } (i,j) \text{ is part of the best} \\ & \text{ant's route at this iteration,} \\ 0, & \text{otherwise,} \end{cases} \tag{6}$$

where $\tau_{ij}(t+n)$ represents pheromone of the best path (i, j) from time t to time $t+n$ after all of the ants complete a trip, which is unlike traditional ant colony algorithm. In (6), L represents length of the best moving path corresponding to ants crawling at this iteration. After iteration, update the global pheromone by (5).

Meanwhile, to take advantage of convergence of the algorithm, this paper adds incentive factor γ in pheromone global updating. It is calculated by the following:

$$\gamma = \frac{L^{+} - L}{T}. \tag{7}$$

T is a positive constant whose value can be up to the same order of magnitude as current working space. In this paper set $T = 1000$. Thus, it reduced the complexity of the search, and improved the accuracy and convergence speed of the route search. L^{+} is the shortest path length for moving till previous cycle. If the path length of optimal solution in this cycle is shorter than optimal solution having been found, then $\gamma > 0$, otherwise $\gamma < 0$, which avoids falling into local optimum.

3.2. The Steps of Path Optimization

Step 1. Initialize phase. m: the number of ants, NC_max: the maximum number of iterations, the number of iterations NC = 0, best_len: the current optimum path length, and set concentration of pheromones on each node at initial time (0.001), all the ants at the rescue centre a_0, W indicates the vehicle load, and weigh[ ] stands for the demand of affected nodes. Set each ant a tabu list array tabu[ ] to store the city that the ants have already visited; that is to say, the cities will not be visited in the next choice, the path length passed ant.[k].path_length, array ROUTES[k][n] stored the kth ant's route where the path has already been visited in the nth iteration, and add a_0 to each ant's tabu list ($k = 1, \ldots, m$, $n = 1, \ldots,$ NC_max).

Step 2. Set variable $i = 1$.

Step 3. Calculate the next affected node the ant can choose and exclude the affected node in tabu list or the obstacle affected node. Array to_visit[ ] and Len_to_visit recorded separately the affected node's mark the ants can visit and the number of the affected node, while Len_to_visit ≥ 1 and $W \geq$ weigh[ ] go to Step 4, otherwise go to Step 7.

Step 4. Use (3) and roulette method to select an affected node to_visit for each ant k (k = 1 to m). And move ant k to this node. Meanwhile, update the ant k's length of path, route: ROUTES[k][n] = add[ROUTES[k][n], to_visit], the ant moves the next affected node to_visit, put to_visit in ant k's tabu list array tabu[ ], and $W = W -$ weigh [to_visit].

Step 5. Perform a local updating rule according to (4).

Step 6. Record the current iteration of the shortest route, update best_length with the best ant's path and record

the optimal path $G\_best\_route[\,]$, while update the optimal solution in accordance with 3-opt method.

Step 7. Ant k comes back to affected centre, update best_length and route, ROUTES$[k][n]$ = add[ROUTES $[k][n], a_0]$. Set $i = i + 1$. If $i \leq d$, switch to Step 4; otherwise, go to Step 8.

Step 8. Update global concentration of pheromones of each node by optimal ant in accordance with (5), (6), and (7).

Step 9. Set NC ← NC + 1. If NC < NC_max and the entire ant colony has not converged to take the same path, then set all the ants at the rescue node a_0 again and go to Step 3. If NC < NC_max and the entire ant colony has converged to take the same path, or NC = NC_max, then finish the cycle and output the best route.

4. Experiment of Computer Simulation and Analysis of Results

In this part, we use the established model shown in Part 2 to do the simulation experiment of the optimal algorithm. The experiment runs on Windows XP Professional SP3 (symmetric multiprocessor system) and the compiler environment is MATLAB R2009a. Each parameter is set to NC_max = 5, ρ = 0.85, α = 1.0, β = 1.0, γ = 17, Q = 100.0, and m = 20. All of the parameters are under the optimal experiment in the literature [15]. The results of the simulation are shown in Figures 3–8.

In order to illustrate the superiority of the proposed method, we use the example above to compare the results of three methods, including traditional ant colony algorithm, the literature [15], and improved ant colony algorithm. Figures 3, 4, and 5 shows the results that each polygon means one route. From Figure 3, we can see the path planning of traditional ant colony algorithm, the best length is 58.52 m, the route is

$$0 \to 10 \to 4 \to 13 \to 12 \to 0;$$
$$0 \to 3 \to 5 \to 7 \to 9 \to 2 \to 0;$$
$$0 \to 8 \to 6 \to 14 \to 0;\ 0 \to 1 \to 15 \to 11 \to 0.$$

When using the weight matrix, we define that if the two roads are not connected, the weight value is 0, then the higher the weight value the better the clearer road. The search result is shown as Figure 4, the best length is 48.49 m and the optimal path is

$$0 \to 1 \to 15 \to 8 \to 0;$$
$$0 \to 13 \to 2 \to 3 \to 5 \to 7 \to 9 \to 12 \to 0;$$
$$0 \to 14 \to 6 \to 0;\ 0 \to 11 \to 4 \to 0;\ 0 \to 10 \to 0$$

This paper introduced angle factor in the transfer rule, because in an emergency we need to avoid obstacles to the road to gain more time. Thus, we choose the next node where the offset angle is small and closer to the direction of the transfer, which is more likely to avoid obstacles in the roads. The save matrix is to compensate for the optimal path length

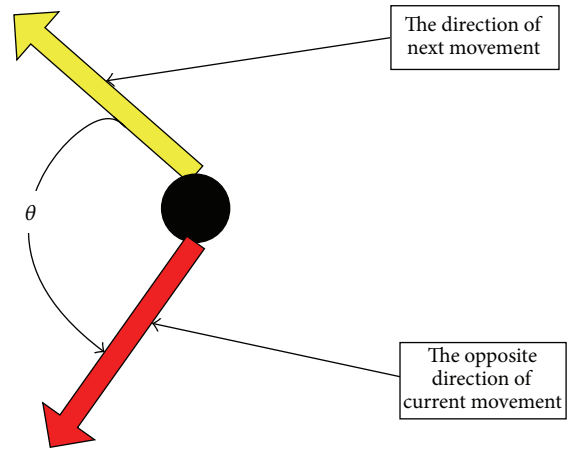


Figure 2: A view of added angle factor.

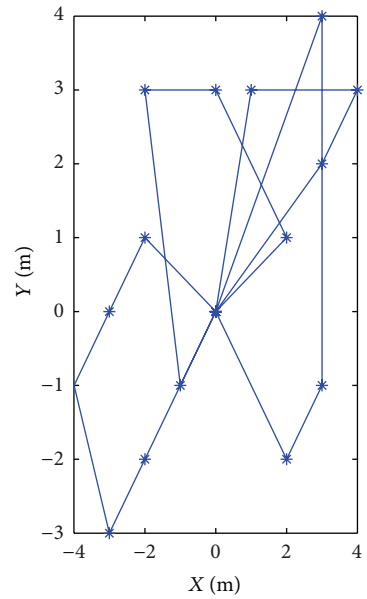


Figure 3: A result of ant colony algorithm path planning.

increasing brought by a perspective to factors. Meanwhile, after each cycle, update pheromone by reward function and use 3-opt algorithm to the optimal solution method to update the global optimal path length. Results shown in Figure 5, the best length is 41.61 m and the optimal path is

$$0 \to 13 \to 4 \to 11 \to 0;\ 0 \to 6 \to 14 \to 0;$$
$$0 \to 8 \to 15 \to 1 \to 10 \to 0;$$
$$0 \to 12 \to 3 \to 5 \to 7 \to 9 \to 2 \to 0.$$

In this instance, the literature [15] and this paper's proposed algorithm's optimal path length are best among the three methods. After adding the angle, the vehicle's route is more concise and clear, while the new transfer rules and pheromone update rule, this algorithm only sent four rescue vehicles to meet relief needs which is more reasonable.

Convergence property, where Figures 6, 7, and 8 show their convergence property. Blue line shows the optimal

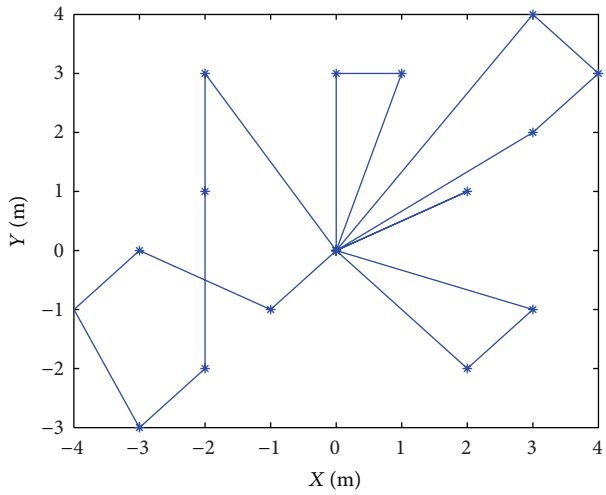


Figure 4: A result of the literature [15] path planning.

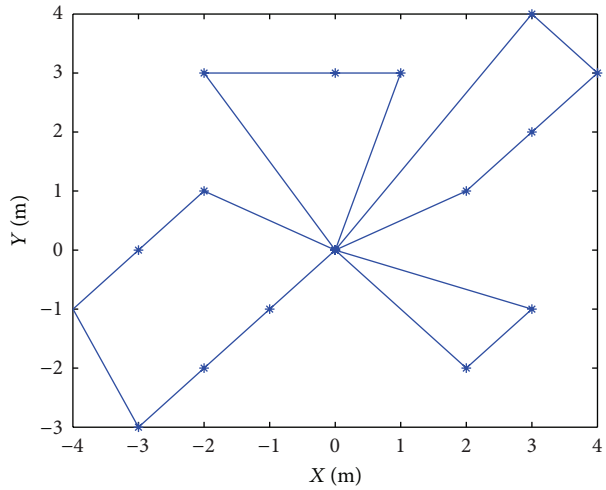


Figure 5: A result of improved ant colony algorithm path planning in this paper.

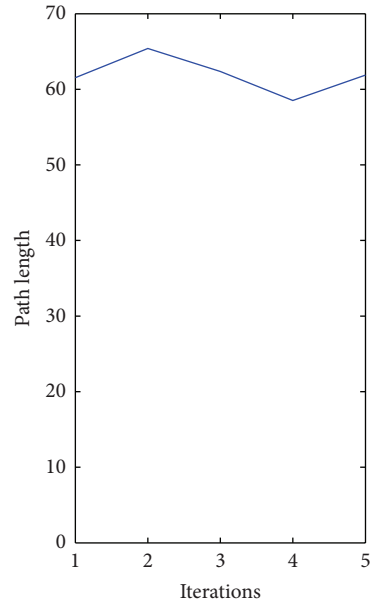


Figure 6: The convergence curve of ant colony algorithm path planning.

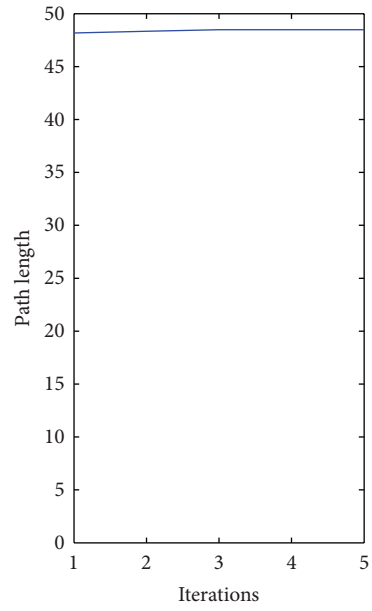


Figure 7: The convergence curve of the literature [15] path planning.

solution of the traditional ant colony algorithm convergence process, the black line is the literature [15] in the convergence process, and the red line is this paper's improved ant colony algorithm convergence process. As can be seen from the figure, in the same five iterations, the curve of traditional ant colony algorithm fluctuates, the optimal solution is 58.52 m, the literature [15] proposed optimization algorithm and the algorithm this paper proposed converges within 5 iterations. From the view of convergence, improved ant colony algorithm overcomes the shortcomings of slow convergence, and the optimal solution (41.61 m) that is superior to the literature [15] algorithm is the optimal solution (48.49 m).

In this part, the simulation instances from benchmark [10] were tested by the above algorithms. We labeled the example of examples that data points were non-decreasing order as I1, I2, . . . ,I14, where the conditions for instances I1–I5 are similar to the instances I9–I14, while instances I6–I10 and I11-I12 involve limited routing path length and service time. We conducted five experiments, the first group was the traditional GA, the second group was the traditional ACS, the third group was the literature [15], the fourth group was the literature [20], and the last group was for this paper proposed optimization algorithm. All parameters were under experimental environment of the optimum algorithm, shown in Table 3. Here parameters P_c and P_m are for Crossover probability and mutation probability in the GA algorithm respectively. Each experiment was performed 20 times, respectively, and compared with each optimal and

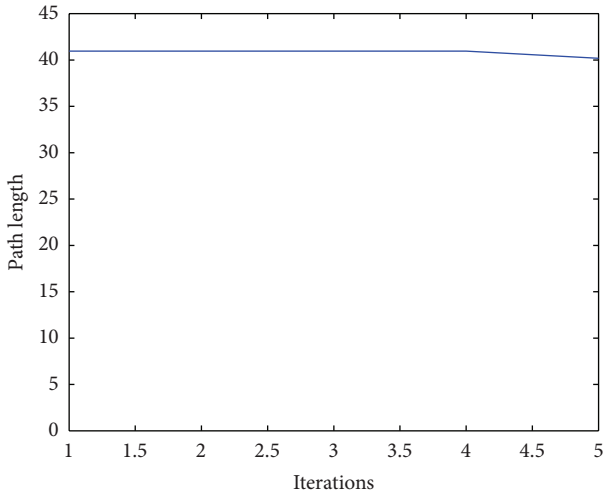


Figure 8: The convergence curve of improved ant colony algorithm path planning in this paper.

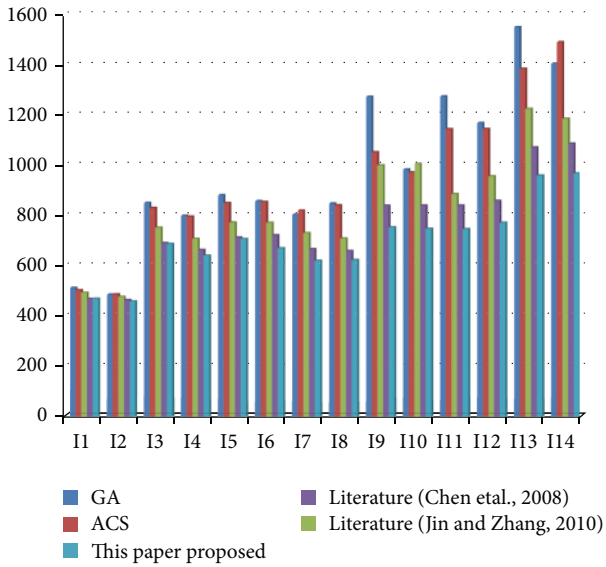


Figure 9: The simulation results of best length.

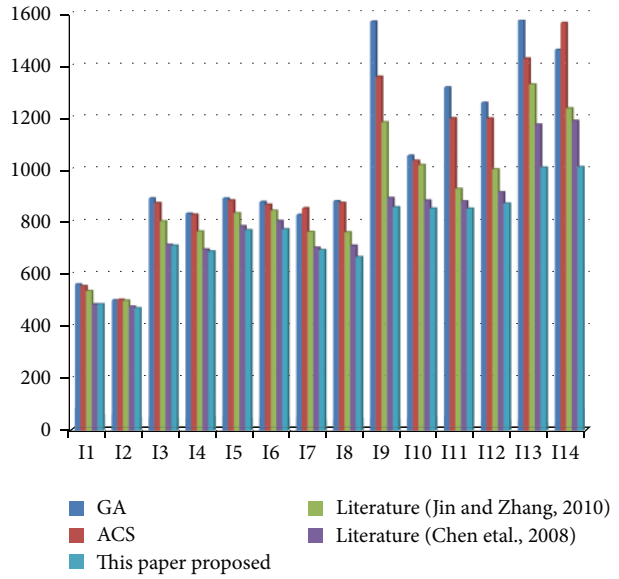


Figure 10: The simulation results of average length.

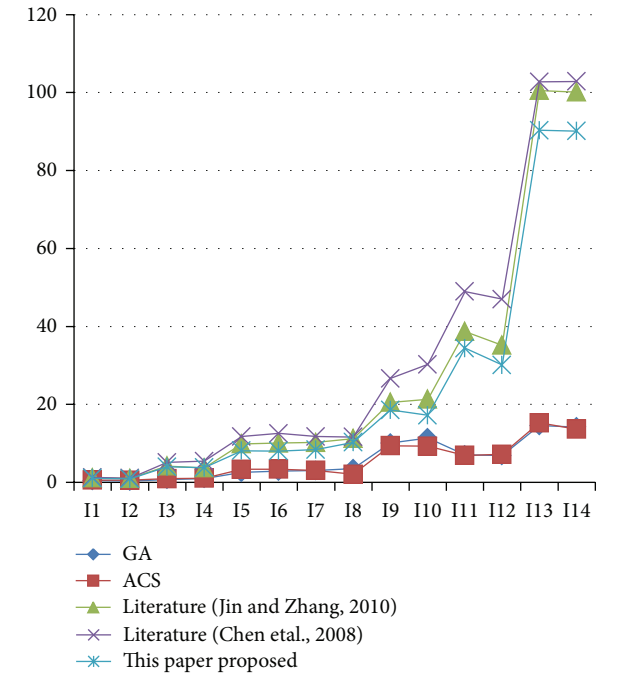


Figure 11: The simulation results of average time.

average path length. The experimental results are shown in Figures 9, and 10. To illustrate our proposed algorithm drawing better performance, we record CPU processing time under the experiment above, and then compare their average time, shown as Figure 11.

Figure 11 shows the average running time of the five algorithms in various numerical examples from I1 to I14. To compare against GA and ACS, our proposed algorithm has higher overhead time. However, it can still be able to obtain good solutions in a short time, and all the solving time is within 100 seconds. At the same time, our proposed algorithm performs better than [15] and [20]. From Figure 11, we can see that from I1 to I8, [15, 20] and our proposed algorithm have the almost same computation speed; however, from I9 to I14, our proposed algorithm takes less time than [15, 20] significantly.

As is seen from the experimental results, the optimal and the average results of the paper proposed algorithm are superior to other four methods clearly, comparison in terms of simulation time, our proposed algorithm also has advantage of average running time, as the problem size increases, that is to say, from I1 to I14 the size of examples increase; this advantage has become increasingly apparent. Taken all together, our proposed algorithm draws the best

Table 3: Parameter settings.

	α	β	γ	ρ(or P_c)	q_0 (or P_m)	NC_max	Q	m
GA	~	~	~	0.8	0.08	50	100	5
ACS	1	2	~	0.7	0.8	50	100	5
Literature [15]	0.1	2	15	0.1	0.9	50	100	5
Literature [20]	1	1	~	0.2	0.9	50	100	5
This paper	1	1	17	0.15	0.05	50	100	5

performance and in the future it can be applied in much more complex environments as well.

5. Conclusions

In this paper, we established a new model of disaster emergency response, which adopts new transfer rules, adds the path weight matrix, save matrix, angle factor functions, and new visibility functions, at the same time, updates pheromone model with reward function, overcoming the shortcomings that are of slow convergence and are easy to fall into local optimum and stagnation of ant colony algorithm in the vehicle applications, thus reducing the complexity of the search process. We use 3-opt method to update the optimal solution to shorten the length of rescue route. Finally, the computer simulation experiments confirm the correctness and validity of this method, and due to considering the actual road conditions and other factors such as the angle, which is more reasonable when applies to solve real emergency rescue.

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 a Grant from two Guangdong Province Production Education and Scientific Study Programs of China (no. 2011B090400056 and no. 2012B091100490).

References

[1] R.-M. Chen and C.-M. Wang, "Vehicle routing problem application in tour planning multimedia system with metaheuristic designed," *Advances in Information Sciences and Service Sciences*, vol. 3, no. 2, pp. 1–11, 2011.

[2] X. Li, P. Tian, and S. C. H. Leung, "Vehicle routing problems with time windows and stochastic travel and service times: models and algorithm," *International Journal of Production Economics*, vol. 125, no. 1, pp. 137–145, 2010.

[3] H. Malekly, B. Haddadi, and T.-M. Reza, "A fuzzy random vehicle routing problem: the case of Iran," in *Proceedings of the International Conference on Computers and Industrial Engineering (CIE '09)*, pp. 1070–1075, Troyes, France, July 2009.

[4] P. Toth and D. Vigo, *The Vehicle Routing Problem*, SIAM Monographs on Discrete Mathematics and Applications, Philadelphia, Pa, USA, 2002.

[5] B.-I. Kim, S. Kim, and S. Sahoo, "Waste collection vehicle routing problem with time windows," *Computers and Operations Research*, vol. 33, no. 12, pp. 3624–3642, 2006.

[6] F. Semet and E. Taillard, "Solving real-life vehicle routing problems efficiently using tabu search," *Annals of Operations Research*, vol. 41, no. 4, pp. 469–488, 1993.

[7] J. F. Cordeau and G. Laporte, "Tabu search heuristics for the vehicle routing problem," *Metaheuristic Optimization Via Memory and Evolution*, vol. 30, pp. 145–163, 2005.

[8] N. S. Niazy and A. Badr, "Complexity of capacitated vehicles routing problem using cellular genetic algorithms," *International Journal of Computer Science and Network Security*, vol. 12, no. 2, p. 5, 2012.

[9] E. F. Zulvia, R. J. Kuo, and T.-L. Hu, "Solving CVRP with time window, fuzzy travel time and demand via a hybrid ant colony optimization and genetic algorithm," in *Proceedings of the IEEE World Congress on Computational Intelligence (WCCI '2012)*, Brisbane, Australia.

[10] L. M. Gambardella, E. Taillard, and G. Agazzi, *MACS-VRPTW: A Multiple Ant Colony System For Vehicle Routing Problems With Time Windows in New ideAs in optimization*, 1999, Edited by D. Corne.

[11] J. E. Bell and P. R. McMullen, "Ant colony optimization techniques for the vehicle routing problem," *Advanced Engineering Informatics*, vol. 18, no. 1, pp. 41–48, 2004.

[12] B.-B. Jiang, H.-M. Chen, L.-N. Ma, and L. Deng, "Time-dependent pheromones and electric-field model: a new ACO algorithm for dynamic traffic routing," *International Journal of Modelling, Identification and Control*, vol. 12, no. 1-2, pp. 29–35, 2011.

[13] R. Zhou, H. P. Lee, and A. Y. C. Nee, "Applying Ant Colony Optimisation (ACO) algorithm to dynamic job shop scheduling problems," *International Journal of Manufacturing Research*, vol. 3, no. 3, pp. 301–320, 2008.

[14] M. López-Ibáñez and C. Blum, "Beam-ACO for the travelling salesman problem with time windows," *Computers & Operations Research*, vol. 37, no. 9, pp. 1570–1583, 2010.

[15] B. Jin and L. Zhang, "An improved ant colony algorithm for path optimization in emergency rescue," in *Proceedings of the 2nd International Workshop on Intelligent Systems and Applications (ISA '10)*, May 2010.

[16] X. C. Bian, C. Zhu, and Y. X. He, "Application of path weights ant colony algorithm in emergency rescue," *Industrial Safety and Environmental Protection*, vol. 36, no. 11, pp. 60–62, 2010.

[17] H. Wang and C. Cao, "Research on an intelligent ant algorithm in conjuction with local O3-opt optimization," *Computer Applications and Software*, vol. 10, pp. 89–91, 2010.

[18] R. Kamil and S. Reiji, "Accelerating 2-opt and 3-opt local search using GPU in the travelling salesman problem," *IEEE Conference Publications*, pp. 489–495, 2012.

[19] N. Christofides, A. Mingozzi, and P. Toth, "The vehicle routing problem," in *Combinatorial Optimization*, pp. 315–338, Wiley, Chichester, UK, 1979.

[20] P. Chen, H.-K. Huang, and X.-Y. Dong, "Hybrid heuristic algorithm for the vehicle routing problem with simultaneous delivery and pickup," *Chinese Journal of Computers*, vol. 31, no. 4, pp. 565–573, 2008.

[21] M. Y. Khoshbakht and M. Sedighpour, "An optimization algorithm for the capacitated vehicle routing problem based on ant colony system," *Australian Journal of Basic and Applied Sciences*, vol. 5, no. 12, pp. 2729–2737, 2011.

[22] B. Lyamine, H. Amir, and K. Abder, "A hybrid heuristic approach to solve the capacitated vehicle routing problem," *Journal of Artificial Intelligence*, vol. 1, no. 1, p. 31, 2010.