

Journal of Advanced Transportation

Sensors in Intelligent Transportation Systems

Lead Guest Editor: Krzysztof Okarma

Guest Editors: Darius Andriukaitis and Reza Malekian





Sensors in Intelligent Transportation Systems

Journal of Advanced Transportation

Sensors in Intelligent Transportation Systems

Lead Guest Editor: Krzysztof Okarma

Guest Editors: Darius Andriukaitis and Reza Malekian



Copyright © 2019 Hindawi. All rights reserved.

This is a special issue published in "Journal of Advanced Transportation." 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

Francesco Bella, Italy
Abdelaziz Bensrhair, France
Cesar Briso-Rodriguez, Spain
María Calderon, Spain
Juan C. Cano, Spain
Giulio E. Cantarella, Italy
Maria Castro, Spain
Oded Cats, Netherlands
Anthony Chen, USA
Nicolas Chiabaut, France
Steven I. Chien, USA
Antonio Comi, Italy
Emanuele Crisostomi, Italy
Luca D'Acierno, Italy
Andrea D'Ariano, Italy
Alexandre De Barros, Canada
Stefano de Luca, Italy
Rocío de Oña, Spain
Luigi Dell'Olio, Spain
Cédric Demonceaux, France
Sunder Lall Dhingra, India
Vinayak Dixit, Australia
Yuchuan Du, China
Nour-Eddin El-faouzi, France

Juan-Antonio Escareno, France
David F. Llorca, Spain
Peter Furth, USA
Francesco Galante, Italy
Md. Mazharul Haque, Australia
Jérôme Haïri, France
Samiul Hasan, USA
Serge Hoogendoorn, Netherlands
Hocine Imine, France
Lina Kattan, Canada
Victor L. Knoop, Netherlands
Alain Lambert, France
Ludovic Leclercq, France
Jaeyoung Lee, USA
Seungjae Lee, Republic of Korea
Zhi-Chun Li, China
Yue Liu, USA
Jose R. Martinez-De-Dios, Spain
Filomena Mauriello, Italy
Monica Menendez, UAE
Rakesh Mishra, UK
Andrea Monteriù, Italy
Giuseppe Musolino, Italy
Jose E. Naranjo, Spain

Aboelmaged Noureldin, Canada
Eneko Osaba, Spain
Eleonora Papadimitriou, Netherlands
Dongjoo Park, Republic of Korea
Paola Pellegrini, France
Luca Pugi, Italy
Nandana Rajatheva, Finland
Hesham Rakha, USA
Prianka N. Seneviratne, Philippines
Fulvio Simonelli, Italy
Richard S. Tay, Australia
Martin Trépanier, Canada
Pascal Vasseur, France
Antonino Vitetta, Italy
Francesco Viti, Luxembourg
S. Travis Waller, Australia
Yair Wiseman, Israel
Shamsunnahar Yasmin, Australia
Jacek Zak, Poland
Guohui Zhang, USA
Ning Zhang, USA
G. Homem de Almeida Correia, Netherlands

Contents

Sensors in Intelligent Transportation Systems

Krzysztof Okarma , Darius Andriukaitis , and Reza Malekian 
Editorial (2 pages), Article ID 7108126, Volume 2019 (2019)

A Particle Filter Localization Method Using 2D Laser Sensor Measurements and Road Features for Autonomous Vehicle

KyungJae Ahn  and Yeonsik Kang 
Research Article (11 pages), Article ID 3680181, Volume 2019 (2019)

Novel Registration and Fusion Algorithm for Multimodal Railway Images with Different Field of Views

Baoqing Guo , Xingfang Zhou, Yingzi Lin, Liqiang Zhu, and Zujun Yu
Research Article (16 pages), Article ID 7836169, Volume 2018 (2019)

Predicting Pedestrian Counts for Crossing Scenario Based on Fused Infrared-Visual Videos

Shize Huang , Wei Chen , Rongjie Yu , Xiaolu Yang , and Decun Dong 
Research Article (11 pages), Article ID 8703576, Volume 2018 (2019)

Neuromorphic Vision Based Multivehicle Detection and Tracking for Intelligent Transportation System

Guang Chen , Hu Cao, Muhammad Aafaque, Jieneng Chen, Canbo Ye, Florian Röhrbein, Jörg Conradt, Kai Chen, Zhenshan Bing, Xingbo Liu, Gereon Hinz, Walter Stechele , and Alois Knoll
Research Article (13 pages), Article ID 4815383, Volume 2018 (2019)

A Geostatistical Investigation into the Effective Spatiotemporal Coverage of Road Weather Information Systems in Alberta, Canada

Xu Wang , Lian Gu , Tae J. Kwon, and Tony Z. Qiu 
Research Article (11 pages), Article ID 5179694, Volume 2018 (2019)

Double Magnetic Loop and Methods for Calculating Its Inductance

Ferran Mocholí Belenguer , Antonio Mocholí Salcedo, Victor Milián Sánchez, and José Humberto Arroyo Núñez
Research Article (15 pages), Article ID 6517137, Volume 2018 (2019)

An Improved Robust Principal Component Analysis Model for Anomalies Detection of Subway Passenger Flow

Xuehui Wang , Yong Zhang , Hao Liu, Yang Wang, Lichun Wang , and Baocai Yin
Research Article (12 pages), Article ID 7191549, Volume 2018 (2019)

A Low-Cost Consistent Vehicle Localization Based on Interval Constraint Propagation

Zhan Wang and Alain Lambert 
Research Article (15 pages), Article ID 2713729, Volume 2018 (2019)

Editorial

Sensors in Intelligent Transportation Systems

Krzysztof Okarma ¹, **Dariusz Andriukaitis** ², and **Reza Malekian** ³

¹Department of Signal Processing and Multimedia Engineering, West Pomeranian University of Technology, Szczecin, Szczecin, 70-313, Poland

²Department of Electronics Engineering, Kaunas University of Technology, Kaunas, 51368, Lithuania

³Department of Computer Science and Media Technology/Internet of Things and People Research Center, Malmö University, Malmö, 20506, Sweden

Correspondence should be addressed to Krzysztof Okarma; okarma@zut.edu.pl

Received 23 December 2018; Accepted 24 December 2018; Published 25 February 2019

Copyright © 2019 Krzysztof Okarma 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.

Modern Intelligent Transportation Systems (ITS) are strictly related to the use of various types of sensors to allow gathering information about the parameters of vehicles and their motion. Rapid development of video technology and availability of relatively cheap high resolution cameras cause the increased interest in the use of vision sensors mounted inside vehicles as well as over the roads. Another relevant field of research is related to applications of various passive road sensors mounted in the road as well as traffic sensors such as piezoelectric or magnetic MEMS sensors. One of the most relevant recent challenges and trends in the ITS solutions is related to the fusion of data acquired by various types of sensors, including inertial sensors, GPS solutions, and model based estimators, which can be used as the data sources for connected vehicle technologies.

The purpose of this special issue is to bring together the ideas and recent achievements of various research groups specializing in various aspects of the ITS solutions. It contains papers related to newly developed methods and ideas combining the data obtained from various sensors, including cameras and other research projects and works in the area of Intelligent Transportation Systems with an emphasize on innovations based on sensor approach that address challenges in tracking, localization, road and weather prediction, sensor fusion algorithms, road safety, and anomalies detection.

The special issue consists of eight papers accepted after at least two rounds of reviews organized in the following topics and categories: (i) fusion algorithm for multimodal railway images; (ii) predicting pedestrian counts based on

fused infrared-visual videos; (iii) neuromorphic vision based multivehicle detection and tracking; (iv) geostatistical investigation for road weather information systems; (v) magnetic loops as the most used traffic sensors in Intelligent Transportation Systems; (vi) anomalies detection of subway passenger flow; and (vii) vehicle localization methods.

The first paper, written by B. Guo et al., presents the novel multimodal railway image registration method based on the matching of SURF features followed by Contourlet transform image fusion algorithm combined with total variation model and local region energy. The proposed method of image fusion, verified for daytime and night image pairs, allowed reduction of the strong disturbance light of torch at night.

The second article submitted by S. Huang et al. is related to fusion of infrared and visible light videos as well. The application of the proposed approach allows the estimation of the number of pedestrians in outdoor scenarios, in particular at the pedestrian waiting areas. It depends on the estimation of the number of pedestrians based on single-source video and the information fusion based on multisource detection results.

The topic of the third accepted submission is the multivehicle detection and tracking using the neuromorphic vision based approach. G. Chen et al. investigated the advantages of the transmission of only local pixel-level changes caused by the movement in a scene at the time they occur and proposed its application for the ITS. The verification of the detection and tracking accuracy is supported with the analysis of limitations of the proposed solution.

Another paper is devoted to the geostatistical approach to quantitative analysis of the spatiotemporal variations of the road weather and surface conditions. The reliability and applicability of the proposed approach have been analysed using the exemplary road weather information systems (RWIS) in Alberta, Canada, and may be useful for planning cost-effective RWIS networks.

The next article authored by F.M. Belenguer et al. covers the topic of inductance calculations of double magnetic loops used as traffic sensors in ITS. The authors discussed geometry, construction, operating mode, and three possible ways to calculate their inductance including their influence on the obtained precision.

The sixth paper deals with the subject of subway passenger flow analysis for anomalies detection using the robust PCA model. The improved anomaly detection method proposed by X. Wang et al., utilizing low-rank nature of the passenger flow data and the sparsity of anomaly data, has been verified experimentally using the real flow data of Beijing subway.

The two last papers correspond with vehicle localization methods using Interval Constraint Propagation and particle filters. The solution proposed by Z. Wang and A. Lambert, originating from the navigation of mobile robots, allows determination of consistently the vehicle's position and orientation by fusing low-cost sensors and map data obtained when a vehicle embedding expensive sensors navigates around the environment. The method has been validated in outdoor scenario for a vehicle equipped with odometers, gyro, and monocular camera.

The method described by K.J. Ahn and Y. Kang utilizes 2D laser range finder measurements and road features, such as curbs, lane lines, and other road markings, for the navigation of autonomous vehicles in urban areas. The authors proposed the application of particle filter for the estimation of the accurate position of a vehicle based on the map of the road features which may be further combined with data from other sensors, including cameras.

Conflicts of Interest

The guest editors declare that there are no conflicts of interest regarding the special issue.

Acknowledgments

The guest editors would like to thank the reviewers for the time and efforts they have devoted to provide detailed comments, advice, and suggestions for improvement of the technical and scientific level, as well as the presentation quality of the accepted papers.

*Krzysztof Okarma
Dariusz Andriukaitis
Reza Malekian*

Research Article

A Particle Filter Localization Method Using 2D Laser Sensor Measurements and Road Features for Autonomous Vehicle

KyungJae Ahn ¹ and Yeonsik Kang ²

¹Graduate School of Automotive Engineering, Kookmin University, Seoul 02707, Republic of Korea

²Department of Security Enhanced Smart Electric Vehicle, Kookmin University, Seoul 02707, Republic of Korea

Correspondence should be addressed to Yeonsik Kang; ykang@kookmin.ac.kr

Received 6 August 2018; Revised 11 November 2018; Accepted 17 December 2018; Published 14 February 2019

Academic Editor: Krzysztof Okarma

Copyright © 2019 KyungJae Ahn and Yeonsik Kang. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

This paper presents a method of particle filter localization for autonomous vehicles, based on two-dimensional (2D) laser sensor measurements and road features. To navigate an urban environment, an autonomous vehicle should be able to estimate its location with a reasonable accuracy. By detecting road features such as curbs and road markings, a grid-based feature map is constructed using 2D laser range finder measurements. Then, a particle filter is employed to accurately estimate the position of the autonomous vehicle. Finally, the performance of the proposed method is verified and compared to accurate Differential Global Positioning Systems (DGPS) data through real road driving experiments.

1. Introduction

The development of autonomous driving technology is one of the forefront fields of research, bound to have a significant influence on the automobile industry. Presently, autonomous driving technology is developed by both traditional automobile companies such as Daimler and Toyota, as well as information technology (IT) industries such as Uber and Apple. The four essential technologies currently being developed are environmental perception, localization, path planning, and vehicle control. Localization may be accomplished using various methods, including the utilization of the Global Positioning System (GPS), a vision system, and the road infrastructure, as well as the road environment maps. The efficient use of localization is one of the most essential features that influence the performance of a self-driving car. The most popular methods of localization use GPS sensors, which are advantageous because of their vast popularity on the market and their low price compared to the other sensors. However, the GPS-based methods do not perform as well in shaded areas (overpasses, tunnels, skyscrapers, etc.), where the satellite signal may not be received, which causes a large position error [1, 2]. The operation of autonomous vehicles in an urban environment requires a seamless localization

capability with an accuracy of several centimeters or less. To this end, researchers have recently been studying localization methods based on road environment maps, using on-board sensor measurements [3, 4].

A road environment map includes a variety of road features (curbs, road marking, structure, etc.) and helps a vehicle localize itself seamlessly and accurately. Many studies have been performed on map-based localization techniques that use a vision system or three-dimensional (3D) Light Detection and Ranging (LiDAR) sensors, which can acquire point clouds in a single scan [5]. However, the price of such a sensor is too high and the processing of such large amounts of data requires a lot of computational power. On the other hand, localization based on a vision system does not provide depth information as accurately as laser sensors and its performance is significantly affected by the available illumination.

In this paper we use two-dimensional (2D) laser sensor measurements to detect road features such as the curb position and the markings on the road surface. The curb, i.e., the boundary of the road, can be detected using the 2D laser sensor measurements and stored in the road environment map. Various studies have demonstrated the extraction of the curb position by an application of a line extraction algorithm,

including linear regression, the Hough transform, split and merge algorithms, the random sample consensus (RANSAC) algorithm, and the expectation–maximization (EM) algorithm. The Hough transform and the RANSAC algorithm are the most widely used procedures in the extraction of the curb position [6–8]. The Hough transform, the RANSAC algorithm, and the EM algorithm have been demonstrated to extract a straight line more accurately than other techniques; however, their accuracy comes at a large computational cost [9].

In general, the location of the vehicle is estimated by a Bayesian estimator, such as the Kalman filter, the extended Kalman filter, or a particle filter. The Kalman filter relies on the assumption that the system is linear and the posterior probability Gaussian, meaning that it achieves a high accuracy in environments that satisfy these assumptions; however, its performance may be degraded in nonlinear and non-Gaussian posterior probability environments. A particle filter can overcome the disadvantages of the Kalman filter in non-Gaussian environments by representing the posterior probability using samples. However, because the number of samples for this estimator increases, the computational complexity of the procedure increases, as well [10, 11].

We propose a map-based localization method that uses a 2D laser sensor to detect the road boundary and road markings for autonomous driving in an urban environment. Our main contributions are the following.

(i) We developed a seamless localization method that uses on-board sensor measurements with a reasonable accuracy and price. To detect road features, we used an affordable 2D laser sensor.

(ii) To detect the road environment features such as the lane and road markings, we used the reflectivity information and laser sensor measurements, instead of only relying on point measurements. We employed the classification method on the sensor measurements to extract the road feature information and construct a road feature map.

(iii) We employed a method based on a particle filter to calculate a posteriori distribution of the road features and estimate the position of the autonomous vehicle. A posteriori distribution is not Gaussian, meaning that the application of a localization method based on the Kalman filter would be inefficient.

2. Sensor System Configuration

To obtain the road boundaries and markings, we used a LMS-151 2D laser sensor manufactured by SICK. This sensor determines the distance from the object using time-of-flight measurements, which are used to calculate the propagation time of the infrared light that makes contact with an object and returns to the receiver at the sensor [12]. It is also possible to measure the reflectivity of the point on the road by determining the intensity of the reflected laser light. The 2D laser sensor was installed at a height of 1.6 m from the ground and set to scan the road surface 10 m in front of the vehicle, as shown in Figure 1.

The moving distance of the vehicle can be measured using the encoder sensor and the inertial measurement unit (IMU).

TABLE 1: Pseudocode of the Adaptive Breakpoint Detector algorithm.

| | |
|----|---|
| 1: | for j to N do |
| 2: | $D_j = r_{j-1} \frac{\sin(\Delta\phi)}{\sin(\lambda - \Delta\phi)} + 3\sigma_r$ |
| 3: | if $\ p_j - p_{j-1}\ > D_j$ then |
| 4: | $flag_j^b$ and $flag_{j-1}^b \leftarrow true$ |
| 5: | else |
| 6: | $flag_j^b \leftarrow false$ |
| 7: | end |
| 8: | end |

To predict the location of the vehicle in the particle filter method, it is necessary to obtain the moving distance of the vehicle during the sampling. The vehicle moving distance and velocity were calculated by Runge–Kutta integration for the dead reckoning obtained by inertial measurements [13].

3. Road Information Extraction by a 2D Laser Sensor

The road environment includes a variety of features, such as curbs, signal lights, lane markings, and guard rails. The locations of the curb and the lane were estimated using the distance and reflectivity measured by the 2D laser sensor. The road boundaries were classified by denoting the pattern of a continuous and long straight line followed by an abrupt change in its slope. The observed lane markings are painted with a special material that contains glass beads, which allows for a greater reflectivity compared to the remaining road surface. The retroreflection phenomenon of the glass beads is demonstrated in Figure 2(a), whereas Figure 2(b) shows a section of the experimental site used for the road feature extraction.

We estimate the road boundaries by identifying the curb position based on the measurements of the 2D laser sensor. The Adaptive Breakpoint Detector (ABD) and the Iterative Endpoint Fit (IEPF) algorithm were employed for their high computational speed for the extraction of straight lines [14]. The ABD algorithm is used to classify data into breakpoints when the interval of consecutive distance data is larger than the threshold circle, D_j . Figure 3 shows the threshold circle described by the following equation:

$$D_j = r_{j-1} \frac{\sin(\Delta\phi)}{\sin(\lambda - \Delta\phi)} + 3\sigma_r, \quad (1)$$

where the distance from the sensor to the point p_j on the road surface is denoted by r_j . The symbol $\Delta\phi$ represents the angle between the vectors r_j and r_{j-1} , whereas λ and σ_r are used as hyperparameters that control the size of the threshold circle. If the distance between the point data ($\text{dist} = \|p_j - p_{j-1}\|$) is larger than D_j , the points p_j and p_{j-1} are considered to be disconnected and classified as a breakpoint. Table 1 shows the pseudocode of the ABD algorithm, where j is the cardinal number of a point and N is the total number of points.

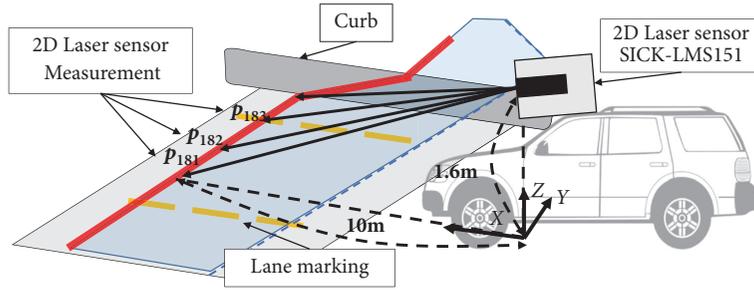


FIGURE 1: Vehicle configuration and setup of the two-dimensional (2D) laser sensor.

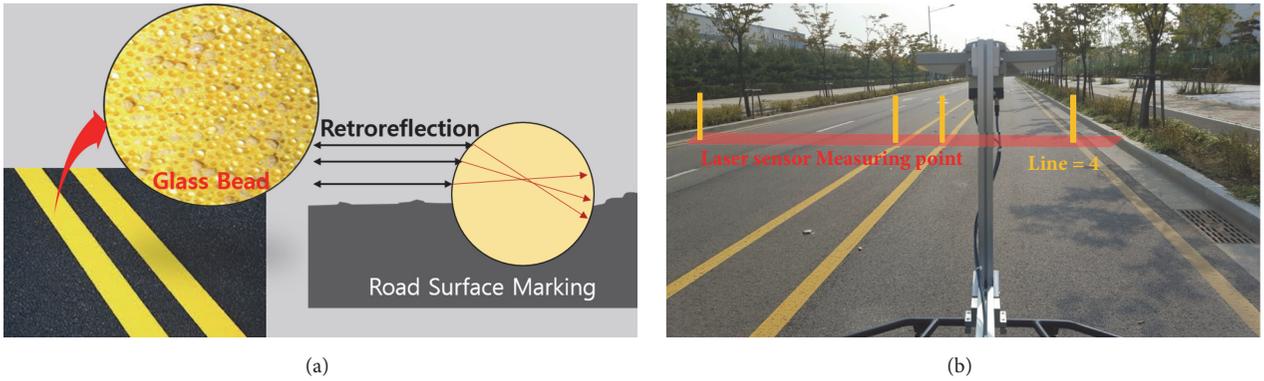


FIGURE 2: Glass bead in the lane marking and the experimental site used for road feature extraction: (a) the retroreflection phenomenon of the lane marking and (b) a section of the experimental site.

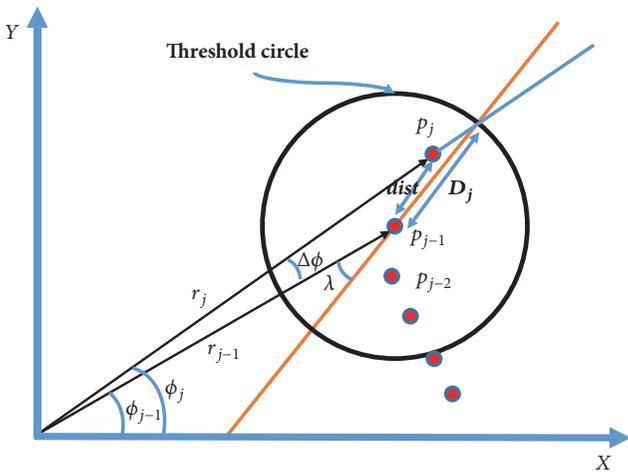


FIGURE 3: Schematic diagram of the threshold circle of the Adaptive Breakpoint Detector algorithm.

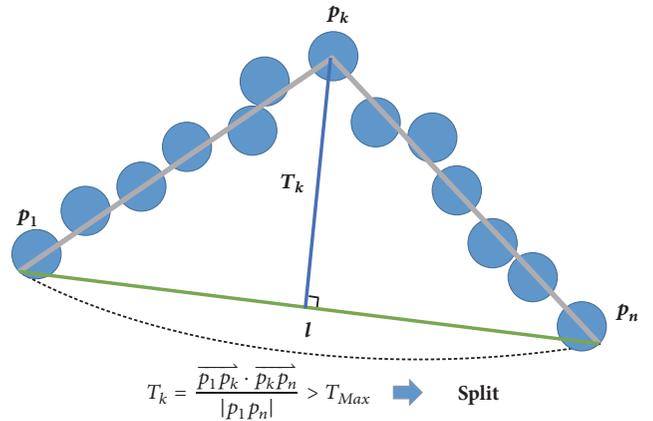


FIGURE 4: Schematic diagram of the Iterative Endpoint Fit (IEPF) algorithm.

We used the IEPF algorithm to extract a straight line for the M groups clustered by the ABD algorithm (S_1, \dots, S_M), where M is the total number of particles. The IEPF algorithm calculates the longest vertical distance at each of the points, T_k , as well as the line l connecting the endpoints of the first group, p_1 and p_n . If T_k is larger than the set parameter T_{max} , the algorithm splits S_1 based on the point p_k into two groups, S_1^1 and S_1^2 . Then, a new T_k and l are calculated

for one of the divided groups, S_1^1 . If the vertical distance is smaller than T_{max} , the process of splitting is started in the next group, S_1^2 . The IEPF algorithm is applied to the remaining groups (S_2, \dots, S_M), which enables us to represent the point data of the laser sensor as a group of straight lines. When the angular difference between the adjacent straight lines is small, they are merged into one group. Figure 4 presents a schematic diagram of the IEPF algorithm and Table 2 shows its pseudocode.

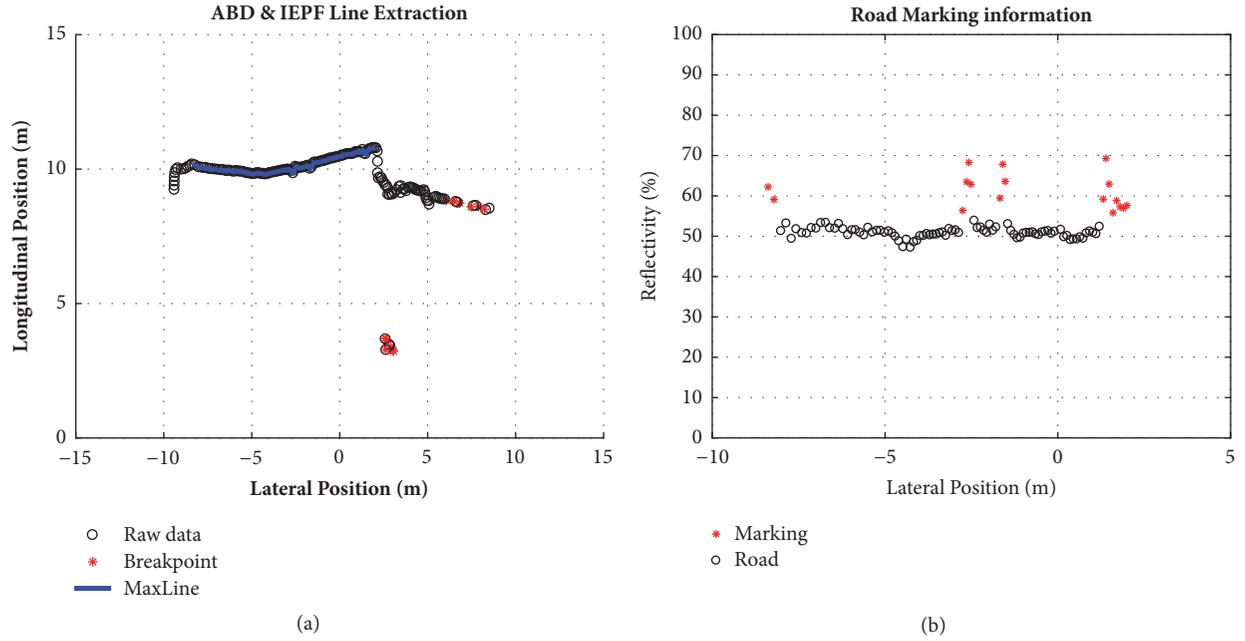


FIGURE 5: Result of the extraction of road markings and road region information: (a) the result of the road boundary extraction and (b) the result of the road marking extraction.

TABLE 2: Pseudocode of the Iterative Endpoint Fit algorithm.

| | |
|-----|---|
| 1: | choose cluster $S_i = \{p_1, \dots, p_n\}$ |
| 2: | while |
| 3: | <i>find</i> endpoint of line l ($l = \overline{p_1 p_n}$) |
| 4: | detect point p_k with maximum distance T_k |
| 5: | if $T_k > T_{max}$ then <i>split</i> |
| 6: | $S_1^1 = \{p_1, \dots, p_k\}$ and $S_1^2 = \{p_k, \dots, p_n\}$ |
| 7: | else |
| 8: | if last group <i>break</i> |
| 9: | go to next group |
| 10: | end |
| 11: | end |
| 12: | when all sets in S have been checked, merge collinear segments |

Figure 5(a) shows the extracted road area obtained by sequentially processing the distance measurement data using the ABD and the IEPF algorithms. The black circles in the figure represent the distance measurement data, the red stars are the breakpoints, and the blue line shows the road area extracted by applying the line extraction algorithm. As described, the discrimination of the lane markings on the road is enhanced by the use of a paint that contains glass beads [15]. The classification of the road and the road markings depend on the intensity of the reflected laser pulse [16]. The results of the lane marking classification are shown in Figure 5(b), where the vertical axis represents the reflectivity of the returned laser light. The reflectivity values larger than the threshold value are denoted by red stars, whereas those

lower than the threshold value are denoted by black circles. As the figure shows, four road surface markings with high reflectivity were classified and the location of the vehicle relative to the lane marking can be estimated.

4. Construction of a Road Feature Map

The road region and the lane markings were extracted using the method described in the previous section. Based on those results, a grid map was constructed with a resolution of $10 \text{ cm} \times 10 \text{ cm}$ for each grid cell. Figure 6(a) shows the classification of the grid states, according to the extracted road environmental conditions, whereas Figure 6(b) shows photos of the lane markings on the road and the lane marking data measured by the laser sensor. The information map of the grid state of the road contains data for the following four states:

- (i) Road surface: grid cells that represent a traversable road surface.
- (ii) Road boundary: grid cells of the road boundary distinguished by the curb.
- (iii) Lane marking: grid cell that contains lane markings.
- (iv) Unknown space: grid cells beyond the sensor measurements.

Figure 7 depicts the road environment information grid map created by measuring approximately 240 m of the road. The road surface, road boundary, and unknown space in the map are expressed in gray, black, and white, respectively. The cells that contain road markings are shown in red.

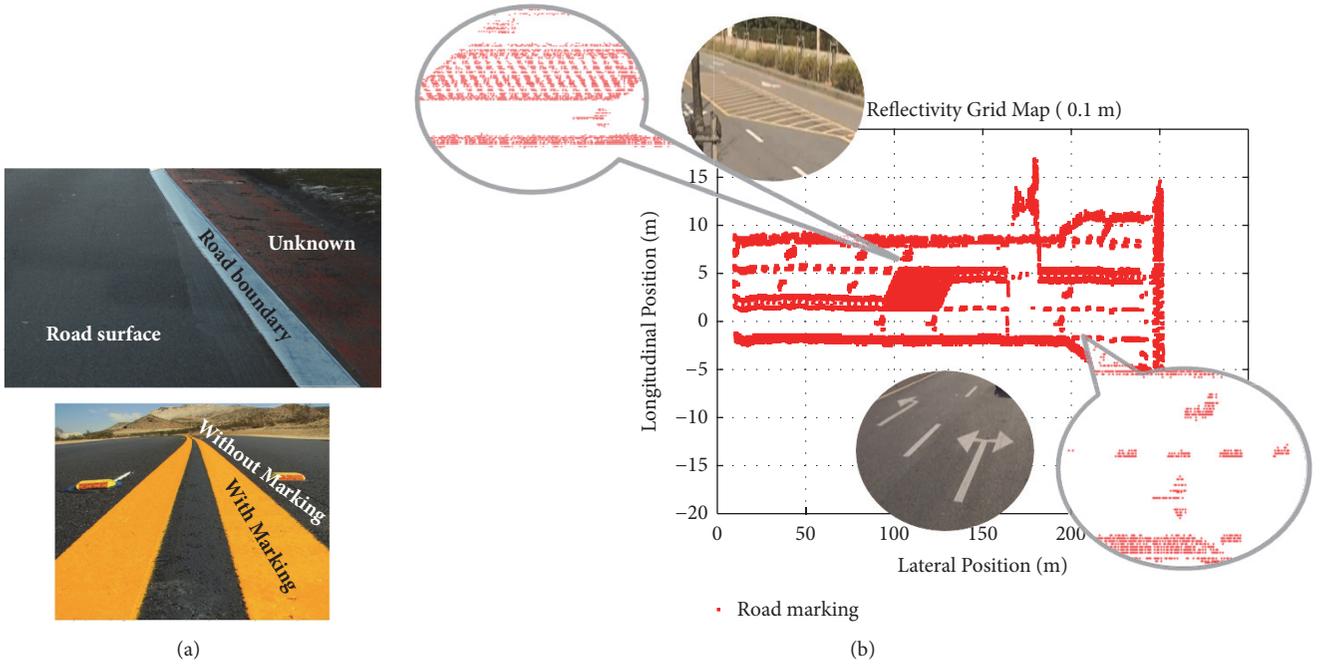


FIGURE 6: Four states of the road region and various lane markings used on the map: (a) the four states of the road region and (b) various lane markings shown on the information map.

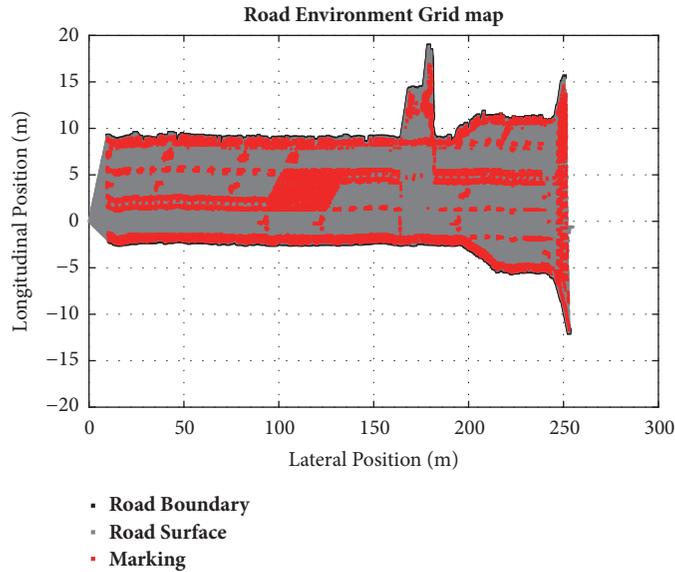


FIGURE 7: Road environment grid map. The road boundary is shown in black, the road surface in gray, the lane markings in red, and unknown space in white.

5. Particle Filter Localization

The probabilistic localization method we used employs a Bayesian filter [17], which relies on a belief, i.e., a probability distribution of the vehicle location x_{t-1} at a time t that can be expressed as follows:

$$bel(x_t) = p(x_t | z_{1:t}, u_{1:t}), \quad (2)$$

where the sensor measurements before the time t and the vehicle movement are denoted by $z_{1:t}$ and $u_{1:t}$, respectively. By using the previous location, x_{t-1} , and the movement, u_{t-1} , the location of the vehicle can be expressed in a recursive form:

$$p(x_t | x_{t-1}, u_{t-1}). \quad (3)$$

The probability distribution, $bel(x_{t-1})$, at the following time, t , can be estimated using the probability $\overline{bel}(x_t)$ at the moving

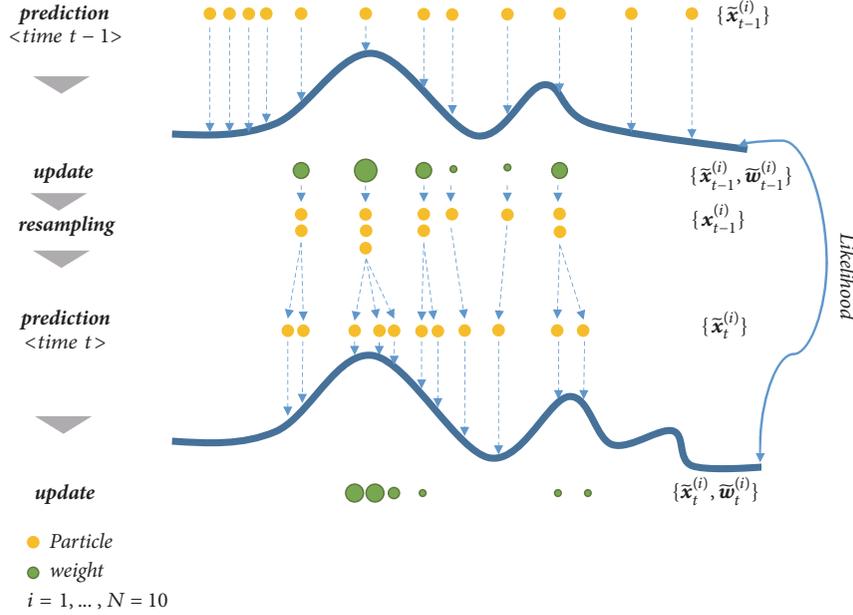


FIGURE 8: Importance sampling process for a particle filter.

position, and the previous position of the vehicle can be expressed using the theorem of total probability:

$$\overline{bel}(x_t) = \int p(x_t | x_{t-1}, u_{t-1}) \overline{bel}(x_{t-1}) dx_{t-1}. \quad (4)$$

In addition, the current positional probability, $bel(x_t)$, can be calculated by employing the Bayes rule with $\overline{bel}(x_t)$, the normalization constant η_t , and the conditional probability that uses the measured values, $p(z_t | x_t)$:

$$bel(x_t) = \eta_t p(z_t | x_t) \overline{bel}(x_t) \quad (5)$$

In this study we consider various features present in real road environments. In many cases, the posterior probability distributions are not Gaussian [18, 19], which is why we use a particle filter suitable for nonlinear systems to estimate the location of the vehicle. Our particle filter calculates the posterior probability without assuming the distribution to be Gaussian [10, 11, 19]. The posterior probability in the particle filter is expressed through the particles sampled using the non-Gaussian distributions:

$$\chi_t = x_t^{[1]}, x_t^{[2]}, x_t^{[3]}, \dots, x_t^{[M]} \quad (6)$$

where M represents the number of particles and each particle is denoted by $x_t^{[m]}$, with $1 \leq m \leq M$. As is the case for other Bayesian filters, a particle filter recursively obtains $bel(x_t)$ from $bel(x_{t-1})$. Since the belief of a particle filter is represented by the particles, χ_t is also recursively obtained from χ_{t-1} . The particle filter method comprises five steps: Initialization, Prediction, Update, Estimate, and Resampling; Figure 8 illustrates the importance sampling process for a particle filter [20].

(i) Initialization: the initial position and direction of a vehicle are not known in advance, so the M particles are

placed at arbitrary positions. The initialization is performed only at the beginning and the individual weights of the particles are all equal to $1/M$.

(ii) Prediction: the following particle locations, χ_t , are predicted by applying the vehicle movement, u_t , to the particle locations at the previous time step, χ_{t-1} . The probability $\overline{bel}(\chi_t)$ can be expressed using the distribution of the moved particles:

$$\chi_t \sim p(\chi_t | u_t, \chi_{t-1}). \quad (7)$$

(iii) Update: the importance factor of a particle $x_t^{[m]}$, i.e., its weight, is denoted by $w_t^{[m]}$. The weights are expressed using the information measured by the sensor and the measurement information available on the map. By comparing the road boundary information measured by the laser sensor and the road marking information in the road environment map constructed in advance, the weights can be calculated as follows:

$$w_t^{[m]} = p(z_t | x_t^{[m]}). \quad (8)$$

(iv) Estimation: by use of the normalization constant η_t , the weight $w_t^{[m]}$, and the probability $\overline{bel}(x_t)$, we can obtain the belief, i.e., the position of the vehicle:

$$\eta_t = \sum w_t^{[m]}, \quad (9)$$

$$bel(x_t) = \eta_t p(z_t | x_t^{[m]}) \overline{bel}(x_t).$$

(v) Resampling: the particles with high weights are cloned by creating a new set of particles that does not contain those with a low weight. This new set of particles is used for the next prediction, at the time $t + 1$.

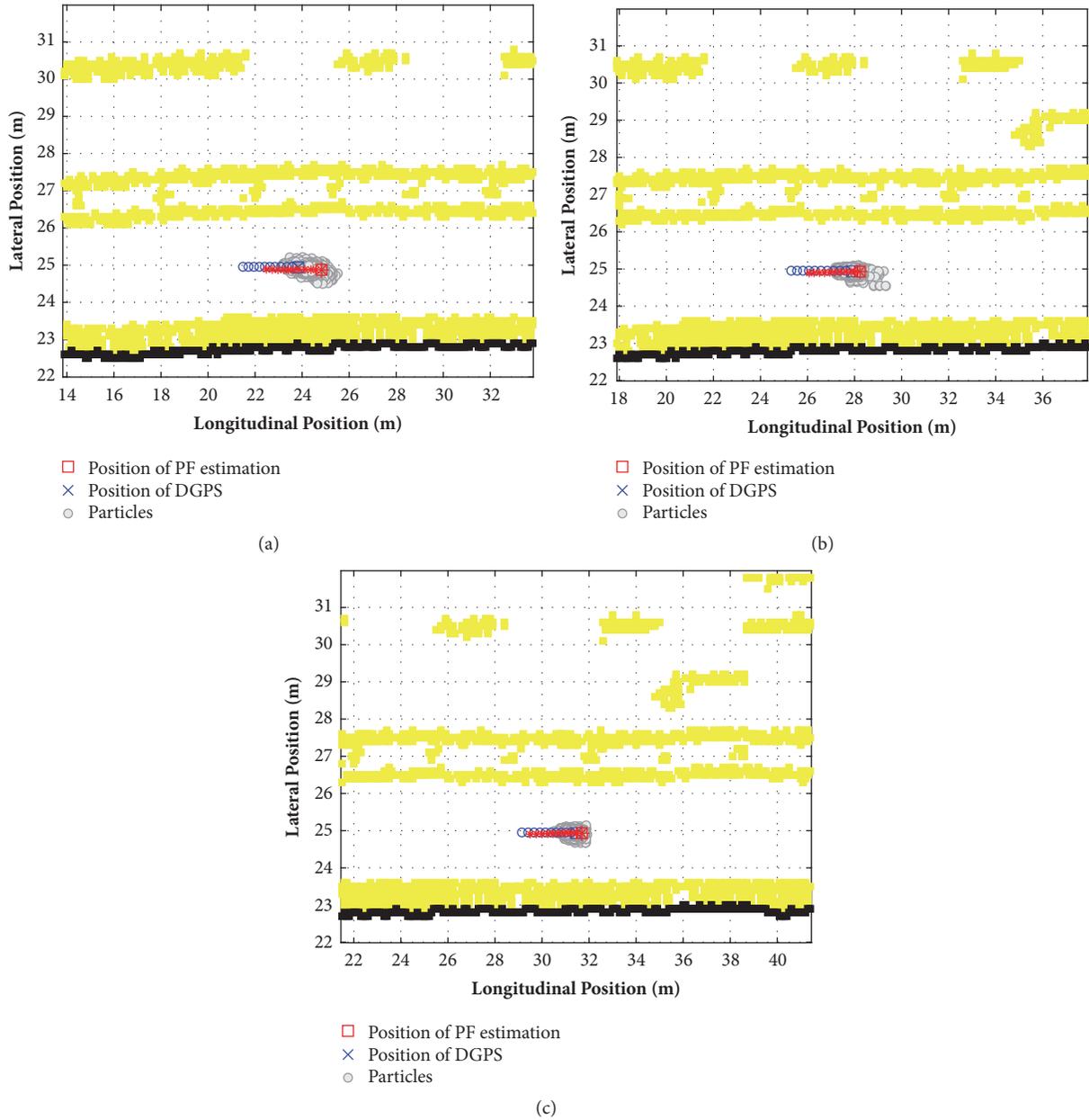


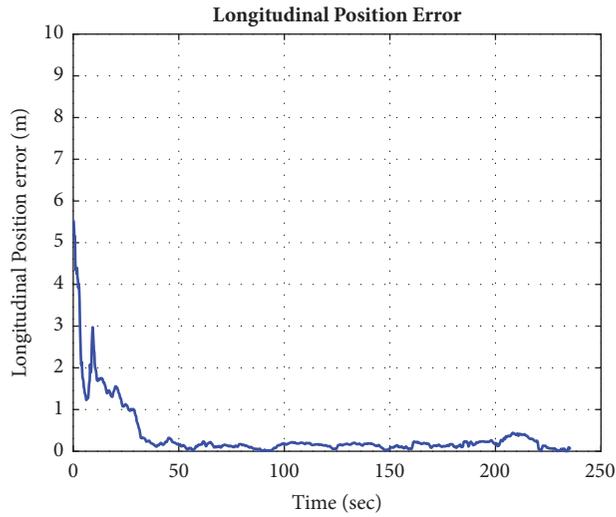
FIGURE 9: Convergence and divergence of particles according to road markings, obtained from three measurements.

6. Results and Discussion

We verified the performance of the proposed localization method based on a particle filter on a public road. Both sides of the public road contain curbs that separate the road region from the nonroad region. The road region has center lines, lanes that help distinguish the direction of travel, and various other road markings. The relative weight of the particles was obtained from the weights assigned by the road markings and the curb, which allowed them to diverge or converge. Figure 9 shows the convergence and divergence of particles according to the road surface information in some sections of the experiment. The yellow color in the figure represents the road markings and curbs, whereas the black color denotes

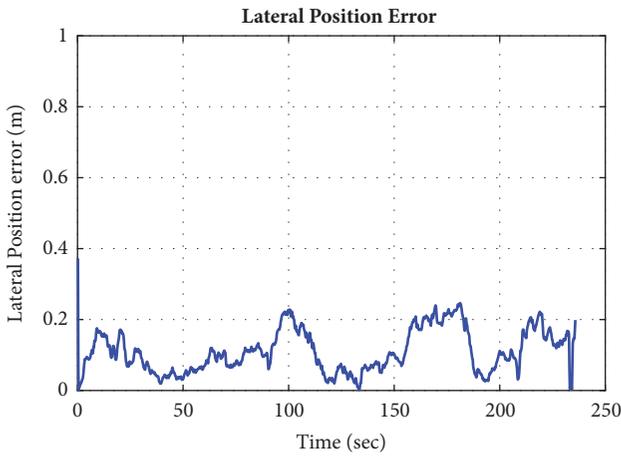
the road feature map information collected a priori, using a 2D laser sensor. The positions of each particle are shown in gray and the estimated positions from the particle filter are shown in red. The measurements obtained using the DGPS, showing the actual position of the vehicle with an accuracy of up to 10 cm, are presented as blue.

The lateral position error shown in Figure 9(a) is small, but the repeated lane markings and curbs diverged the particles longitudinally, and the longitudinal position error obtained from the DGPS and the particle filter position estimation is approximately equal to 1.1 m. Figure 9(b) shows a directional arrow marking between 34 m and 36 m of the longitudinal and between 28 m and 29 m of the lateral position. The road markings that indicate the



— DGPS vs PF

(a)



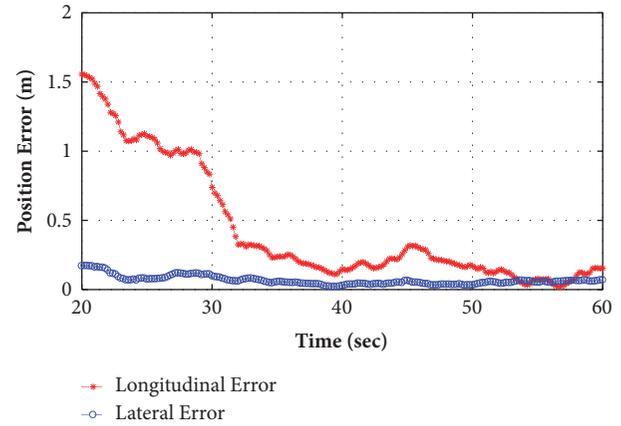
— DGPS vs PF

(b)

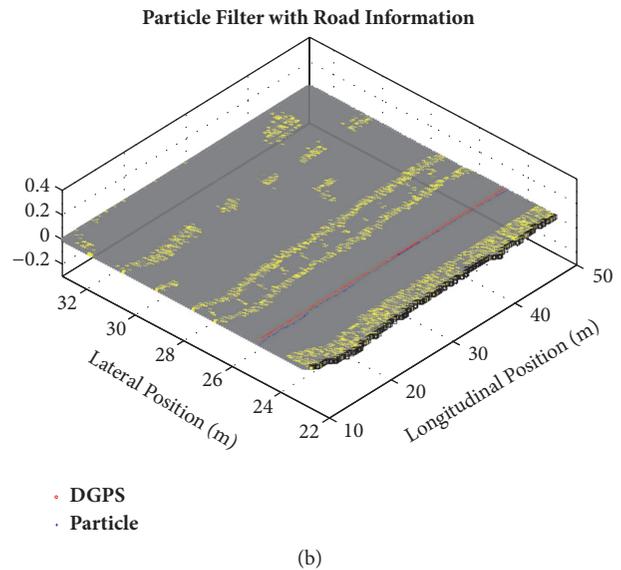
FIGURE 10: Root mean square error between the position estimated using our proposed method and that obtained by the DGPS: (a) the longitudinal position error and (b) the lateral position error.

road surface information (the directional arrow markings, the symbol markings, the stop lines, etc.) can be used to estimate the longitudinal position of the vehicle. Compared to Figure 9(a), such road markings show a gradual decrease in the longitudinal position error in Figure 9(b). Further road markings in Figure 9(c) indicate that particles were collected, which resulted in a significant reduction in the error between the DGPS and the particle filter position recognition.

The root mean square (RMS) error between the DGPS location measurements and the results obtained by the method presented in this paper is shown for the longitudinal and the lateral position in Figures 10(a) and 10(b), respectively. As Figure 10 shows, the longitudinal error is greater than the lateral error for the initial 40 seconds. The RMS value of the lateral axis error is 0.125 and it does not significantly deviate from the RMS value after the initial 40 seconds have



(a)

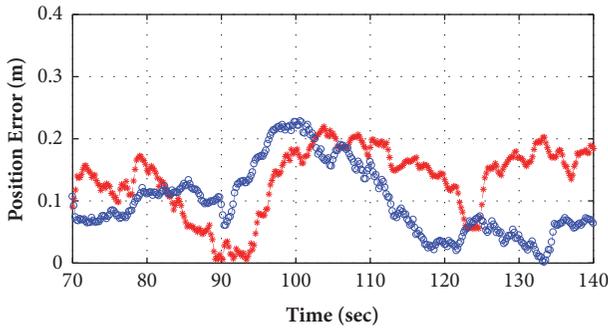


(b)

FIGURE 11: Particle filter localization at the first section (between 20 s and 60 s): (a) the position error and (b) the comparison of the trajectory obtained by the particle filter and the DGPS.

passed. However, the RMS of the longitudinal axis error equals 1.79 m immediately before 40 seconds had passed and 0.178 m after. Thus, as the image shows, it takes some time for the longitudinal axis estimation to converge to the true value. This effect stems from the 2D laser sensor scanning the road surface only in the lateral direction, thereby not providing information in the longitudinal direction. In addition, the road features for the travel direction are identical, which makes the identification of the longitudinal position difficult.

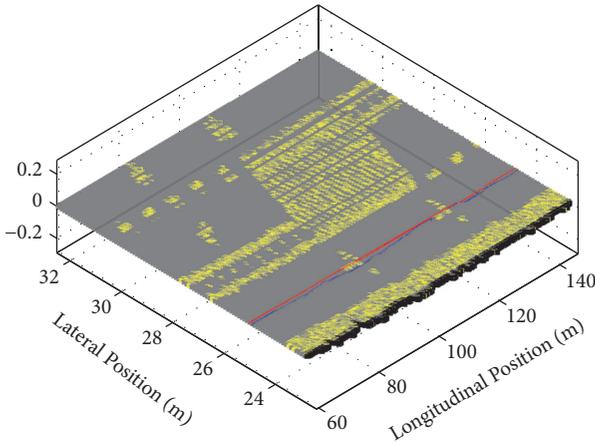
Our experimental data are divided into three sections: the first, second, and third sections comprise the data obtained during the temporal intervals from 20 s to 60 s, from 70 s to 140 s, and from 150 s to 200 s, respectively. Figure 11 shows the convergence of the particle filter estimates to the DGPS reference in the first section. The present longitudinal errors were significantly reduced by the detection of an arrow marking on the road between 30 s and 40 s. Then, the longitudinal error increased until 45 s, when a new feature was detected which allowed it to converge between 50 s



—+— Longitudinal Error
—o— Lateral Error

(a)

Particle Filter with Road Information



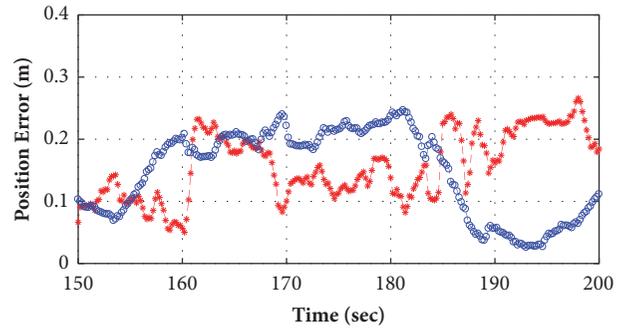
• DGPS
• Particle

(b)

FIGURE 12: Particle filter localization at the second section (between 70 s and 140 s): (a) the position error and (b) the comparison of the trajectory obtained by the particle filter and the DGPS.

and 60 s. The localization error during the second section is shown in Figure 12. Both the lateral and the longitudinal error maintained the value of approximately 0.1 m for the first 20 seconds of the second section. At that point, the road marking used to distinguish the center line began to widen, which made it difficult to estimate the exact lateral position of the vehicle. Once the vehicle passed the area where the center line widened, the lateral error converged to a small value, which did not happen for the longitudinal error, because no road feature in the longitudinal direction was detected.

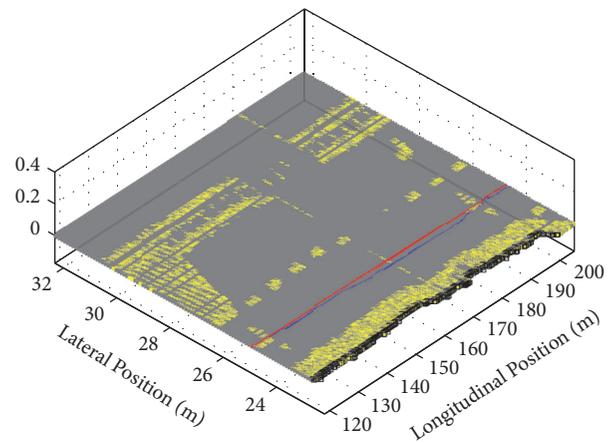
The third section, presented in Figure 13, mapped an intersection of the road. As can be seen in Figure 13(a), the error increased in both the lateral and the longitudinal direction between 160 s and 185 s, as the vehicle was approaching the intersection area, where the curb and the lane do not exist. The lack of features in the sensor field of view caused the localization algorithm to assign even weights to all particles, which resulted in the growth of the lateral error. The lateral error greatly diminished around 185 s, after the vehicle passed



—+— Longitudinal Error
—o— Lateral Error

(a)

Particle Filter with Road Information



• DGPS
• Particle

(b)

FIGURE 13: Particle filter localization at the third section (between 150 s and 200 s): (a) the position error and (b) the comparison of the trajectory obtained by the particle filter and the DGPS.

the intersection and the curb, when the lane markings could be detected again. The error in the longitudinal direction did not converge because no features were found to provide longitudinal information.

As has been shown, the proposed localization algorithm provided good estimates, with an accuracy of approximately 0.1 m, which is equal to the size of the grid cell, if the road feature exists in its sensor field of view. In general, lane markings and curbs can be found in most parts of a road, except for the intersections. Therefore, we expect the proposed algorithm to provide good localization estimates in the lateral direction. However, on parts of the road that contained no features to provide longitudinal information, it took time for the longitudinal position estimation to converge to the true value. Despite this, the accuracy of 0.1m in the lateral direction was good enough to enable the vehicle to follow the road successfully. Because accurate lateral estimates are more important than the longitudinal error for the urban navigation of an autonomous vehicle,

this method can be expected to successfully determine the location of the vehicle on roads in these environments.

7. Conclusions

We studied a localization method for autonomous vehicles, based on a road environment information map. Many map-based localization methods use expensive 3D LiDAR sensors, which also require a lot of computational power, whereas our proposed method uses a 2D laser sensor as a more affordable alternative. To overcome the difficulties stemming from the smaller number of measurements obtained using the 2D laser sensor, we utilized different types of road features, stored in the form of a grid map. By comparing the road features obtained during the vehicle motion to the information of the stored map and using our proposed particle filter method, we efficiently estimated the location information for the vehicle. The experiments were conducted on a public road, for which we could construct a map in advance, allowing for an evaluation of the localization performance. Another experiment was carried out with an accurate DGPS sensor, and the accuracy of the proposed method was evaluated by comparing the localization results with the accurate DGPS measurements.

Our analysis showed that the lateral error can reach the accuracy required for lateral vehicle control. However, the longitudinal error is greater than the lateral error due to the road features not being continuously found on the road. Despite this, as soon as a road feature relevant for the longitudinal direction appears, the error converges quickly and the longitudinal localization maintains a good accuracy. The used 2D laser sensor has a reliable return of the laser light from the road surface, which makes the reliability of the proposed method better than of those that use cheap GPS sensors. In conclusion, the proposed method can be expected to provide reliable localization data for autonomous vehicles in an urban environment, it comes at a reasonable price, and it can be combined with other sensing methods, such as camera vision or assistance from the road infrastructure.

Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this paper.

Acknowledgments

This work was supported by the Basic Science Research Program and the BK 21 Plus Program (Security Enhanced Smart Electric Vehicle Specialist Education Team) through the National Research Foundation of Korea funded by the Ministry of Education, Science and Technology (NRF-2018R1A2B6001888 and 31Z20130012993).

References

- [1] B. J. Clark and D. M. Bevlly, "GPS/INS integration with fault detection and exclusion in shadowed environments," in *Proceedings of the IEEE/ION Position, Location and Navigation Symposium, PLANS*, pp. 1–8, Monterey, CA, USA, May 2008.
- [2] S. Choi, G. Kim, Y. Kim, and J. Lee, "Outdoor positioning estimation of multi-GPS / INS integrated system by EKF / UPF filter conversion," *Journal of Institute of Control, Robotics and Systems*, vol. 20, no. 12, pp. 1284–1289, 2014.
- [3] M. Stubler, J. Wiest, and K. Dietmayer, "Feature-based mapping and self-localization for road vehicles using a single grayscale camera," in *Proceedings of the IEEE Intelligent Vehicles Symposium, IV 2015*, pp. 267–272, Seoul, South Korea, July 2015.
- [4] N. Sukanuma and D. Yamamoto, "Map based localization of autonomous vehicle and its public urban road driving evaluation," in *Proceedings of the IEEE International Symposium on System Integration*, pp. 467–471, Nagoya, Japan, December 2015.
- [5] J. Ziegler, H. Lategahn, M. Schreiber et al., "Video based localization for Bertha," in *Proceedings of the IEEE Intelligent Vehicles Symposium (IV)*, pp. 1231–1238, Dearborn, Michigan, USA, June 2014.
- [6] A. Harati and R. Siegwart, "A new approach to segmentation of 2D range scans into linear regions," in *Proceedings of the 2007 IEEE/RSJ International Conference on Intelligent Robots and Systems, IROS 2007*, pp. 2083–2088, San Diego, CA, USA, November 2007.
- [7] Y. Kang, C. Roh, S.-B. Suh, and B. Song, "A lidar-based decision-making method for road boundary detection using multiple Kalman filters," *IEEE Transactions on Industrial Electronics*, vol. 59, no. 11, pp. 4360–4368, 2012.
- [8] J.-K. Park and T.-H. Park, "Autonomous navigation system of mobile robot using laser scanner for corridor environment," *Journal of Institute of Control, Robotics and Systems*, vol. 21, no. 11, pp. 1044–1049, 2015.
- [9] V. Nguyen, A. Martinelli, N. Tomatis, and R. Siegwart, "A comparison of line extraction algorithms using 2D laser rangefinder for indoor mobile robotics," in *Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems*, pp. 1929–1934, Edmonton, Alta., Canada, August 2005.
- [10] M. S. Arulampalam, S. Maskell, N. Gordon, and T. Clapp, "A tutorial on particle filters for online nonlinear/non-Gaussian Bayesian tracking," *IEEE Transactions on Signal Processing*, vol. 50, no. 2, pp. 174–188, 2002.
- [11] J.-W. Park, C.-K. Yang, and D.-S. Shim, "Particle filter performance for ultra-tightly GPS/INS integration," *Journal of Institute of Control, Robotics and Systems*, vol. 14, no. 8, pp. 785–791, 2008.
- [12] B.-J. Jung, J.-H. Park, T.-Y. Kim, D.-Y. Kim, and H. Moon, "Lane marking detection of mobile robot with single laser rangefinder," *Journal of Institute of Control, Robotics and Systems*, vol. 17, no. 6, pp. 521–525, 2011.
- [13] S. W. Jang, J. S. Lee, K. J. Ahn, and Y. Kang, "A study on integration of particle filters and dead reckoning for efficient localization of automated guided vehicles," in *Proceedings of the IEEE International Symposium on Robotics and Intelligent Sensors (IRIS)*, pp. 81–86, Langkawi, Malaysia, October 2015.
- [14] G. A. Borges and M.-J. Aldon, "Line extraction in 2D range images for mobile robotics," *Journal of Intelligent & Robotic Systems*, vol. 40, no. 3, pp. 267–297, 2004.

- [15] F. Himm, N. Kaempchen, and D. Burschka, "Fusion of laser-scanner and video based lanemarking detection for robust lateral vehicle control and lane change maneuvers," in *Proceedings of the IEEE Intelligent Vehicles Symposium, IV'11*, pp. 969–974, Germany, June 2011.
- [16] S.-H. Im, J.-H. Im, S.-H. Yoo, and G.-I. Jee, "Pedestrian safety road marking detection using LRF range and reflectivity," *Journal of Institute of Control, Robotics and Systems*, vol. 18, no. 1, pp. 62–68, 2012.
- [17] S. Thrun, W. Burgard, and D. Fox, *Probabilistic Robotics*, vol. 026220162, MIT press, Cambridge, USA.
- [18] S. Thrun, "Particle filters in robotics," in *Proceedings of 17th Annual Conference on Uncertainty in AI*, pp. 511–518, Alberta, Canada, 2002.
- [19] M. W. Khan, N. Salman, A. Ali, A. M. Khan, and A. H. Kemp, "A comparative study of target tracking with Kalman filter, extended Kalman filter and particle filter using received signal strength measurements," in *Proceedings of the International Conference on Emerging Technologies (ICET)*, pp. 1–6, Peshawar, Pakistan, December 2015.
- [20] I. Rekleitis, G. Dudek, and E. Millios, "Probabilistic cooperative localization and mapping in practice," in *Proceedings of the IEEE International Conference on Robotics and Automation (ICRA '03)*, vol. 2, pp. 1907–1912, Taipei, Taiwan, September 2003.

Research Article

Novel Registration and Fusion Algorithm for Multimodal Railway Images with Different Field of Views

Baoqing Guo ^{1,2,3}, Xingfang Zhou,² Yingzi Lin,³ Liqiang Zhu,^{1,2} and Zujun Yu^{1,2}

¹School of Mechanical, Electronic and Control Engineering, Beijing Jiaotong University, Beijing 100044, China

²Key Laboratory of Vehicle Advanced Manufacturing, Measuring and Control Technology (Beijing Jiaotong University), Ministry of Education, Beijing 100044, China

³Department of Mechanical and Industrial Engineering, Northeastern University, Boston, MA 02115, USA

Correspondence should be addressed to Baoqing Guo; bqguo@bjtu.edu.cn

Received 16 May 2018; Revised 30 July 2018; Accepted 30 August 2018; Published 6 December 2018

Academic Editor: Krzysztof Okarma

Copyright © 2018 Baoqing Guo 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.

Objects intruding high-speed railway clearance do great threat to running trains. In order to improve accuracy of railway intrusion detection, an automatic multimodal registration and fusion algorithm for infrared and visible images with different field of views is presented. The ratio of the nearest to next nearest distance, geometric, similar triangle, and RANSAC constraints are used to refine the matching SURF feature points successively. Correct matching points are accumulated with multiframe to overcome the insufficient matching points in single image pair. After being registered, an improved Contourlet transform fusion algorithm combined with total variation and local region energy is proposed. Inverse Contourlet transform to low frequency subband coefficient fused with total variation model and high frequency subband coefficients fused with local region energy is used to reconstruct the fused image. The comparison to other 4 popular fusion methods shows that our algorithm has the best comprehensive performance for multimodal railway image fusion.

1. Introduction

High-speed railway is a competitive transportation all over the world. As of 2017, the length of commercial high-speed railway lines in China had exceeded 25,000 km. With the increasing of train speed, safety of railway operation has attracted more and more attention. Any object intruding railway clearance will pose a major threat to running trains. Effective intrusion detection methods are needed to prevent accidents.

Image analysis is an efficient and effective non-contact intrusion detection method widely used in target detecting and tracking [1–3]. Most of them use visible images to analyze intrusion activities. Visible images get the detailed information of colors and texture under good illumination. But their qualities are poor under low illumination and disturbance light circumstance. In contrast, infrared (IR) images are sensitive to thermal radiation of objects and tolerant to changing illumination and disturbance light. However, IR image typically has lower spatial resolution

and fewer details. Image fusion is widely used in remote sensing [4], military [5, 6], objects tracking and detecting [7, 8], etc. IR and visible image fusion will also bring great benefit to railway clearance intrusion detection. A typical image fusion process is usually divided into two steps: image registration and image fusion. Image registration is the basis of fusion. The quality of fusion is bounded not only by the quality of fusion algorithm, but also by the outcome of prior registration algorithm. Due to this dependency, images are always assumed to be pre-aligned [9]. However, in fact, visible and IR images acquired separately are not pre-aligned and even have different field of views and depth. These bring great difficulties to image fusion. We need both excellent registration and fusion algorithms in railway applications.

Image registration is the process of precisely overlaying two images with the same area through geometrically aligning common features identified in image pairs. It is usually divided into four steps: (1) feature detection and extraction, (2) feature matching, (3) transformation function fitting, and (4) image transformation and resample [10]. For multimodal

images, the first two steps are more difficult challenges since different modalities do not exhibit the same features. Z. Xiong [11] has provided a broad overview of registration methods based on feature extraction and matching. S. A. Raza [12] proposed a registration algorithm for thermal and visible diseased plants images based on silhouette extraction. J. Ma [13] presented a non-rigid IR and visible face images registration method by aligning edge maps with a robust regularized Gaussian fields criterion. G. A. Bilodeau [14] proposed a RANSAC-Based registration method using a novel function of moving objects' trajectories. These researches focused on simple scenes registration where the interested objects occupied most part of source images. For large scene images, M. Gong [15] and J. Woo [16] used mutual information for registration. M. Ghantous [9] presented an area-based registration and object-based fusion combined approach based on Dual-Tree Complex Wavelet Transform to reduce complexity. Z. Song [10] proposed an effective registration approach based on histogram of triangle area representation sample consensus for remote sensing images. J. Han [17] proposed a registration method combined line-based global transform approximation with point-based local transform adaptation. All above methods are registration for multimodal images with the same field of views. However, these methods are not available because railway IR and visible images usually have different field of views and depth.

Image fusion is a technique of combining or integrating complementary information from source images into the fused image. Fusion methods based on saliency detection [18, 19], guided image filters [7], and total variation model [20, 21] play dramatic effect in multimodal image fusion. Recently, multiscale decomposition fusion schemes have been a popular research area. Pyramid, wavelet, and Contourlet based decomposition are included in the schemes. Contourlet transform was proposed based on wavelet transform in 2002. Besides attributes of wavelets, it also has advantages of multiresolution, multidirection, and anisotropy. It can decompose images in arbitrary direction at any scale and is good at describing contours and directional texture of images [22]. Many researches focused on Contourlet transform combined with sparse representation [23, 24], low-level visual features [25], object region detection [26], and other methods [27, 28]. In order to solve the problem of unclear edge and textures, we proposed an improved Contourlet transform method combined with total variation (TV) model and local region energy.

In this paper, we present a novel registration and an improved Contourlet-based fusion algorithm for railway IR and visible images with different field of views. The initial candidate matching points are firstly generated with the ratio of the nearest to next nearest distance of SURF feature points. Then they are refined with consecutive geometric constraint, similar triangle, and RANSAC constraints. In order to overcome the shortcoming of insufficient matching points in single frame pair, matching points are accumulated to calculate the transformation matrix from multiframe sequence. After registration, an improved Contourlet transform combined with total variation and local region energy is proposed for image fusion. Qualitative and quantitative comparison

of proposed and other state-of-the-art fusion methods is discussed for fusion effect evaluation. The contributions of this paper are as follows:

- (1) A novel multimodal image registration method is presented based on SURF feature points refining with ratio of the nearest to next nearest distance, geometric, similar triangle, and RANSAC constraints consecutively.
- (2) An improved multimodal Contourlet transform fusion algorithm combined with total variation model and local region energy is proposed. The total variation model is used for low frequency subband coefficients fusing. The local region energy is used for high frequency subband coefficients fusing.
- (3) Qualitative and quantitative comparison of fusion effect on railway IR and visible images is presented to demonstrate the validity of our algorithm.

The remainder of this paper is organized as follows. A novel multimodal image registration method based on SURF feature points matching with a series of refining procedures is outlined in Section 2. Section 3 presents an improved Contourlet transform image fusion algorithm combined with total variation model and local region energy. The fusion evaluation of proposed and other state-of-the-art algorithms is discussed in Section 4. The conclusion is in Section 5.

2. Novel Image Registration with SURF Refining

2.1. Railway Images Registration Overview. There are some differences between familiar security monitoring and railway clearance intruding detection. In security monitoring application, visible and IR images always have the same field of views and depth. However, railway scene is a long and narrow field. Since the view of one camera is limited, visible cameras are arranged at an interval of 100 meters along rails to obtain seamless overlay images. Because of the much higher cost of IR camera, we usually use IR camera with shorter focal length to monitor a larger scene shown in Figure 1(a). The visible image in Figure 1(b) is a small part of IR image shown as the red rectangular in Figure 1(a). In IR image, the rail and person are clearer than other objects because of their higher thermal radiation. In visible image, the person and laying pole are very clear and easy to identify because of the great illumination at daytime. It contains more details such as color, texture, and edges than IR image. But it is not good enough at night. Image fusion will provide more information than each alone. However, in railway application, the typical multimodal fusion methods cannot be used directly due to their different field of views and depth.

Image registration is the basis of fusion. Our novel multimodal registration process for IR and visible images includes 3 steps:

- (1) Feature extraction and description of IR and visible images using SURF descriptor: the initial matching point pairs are generated with ratio of the nearest to next nearest neighbor distance.

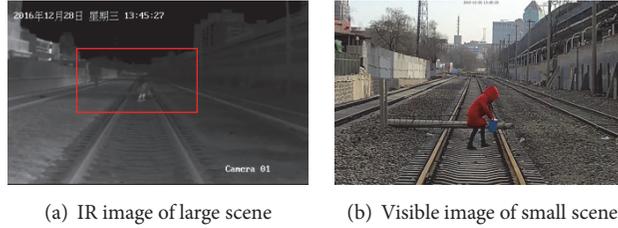


FIGURE 1: Railway IR and visible images.

(2) Matching point pairs refining with successively geometric constraints: similar triangle and RANSAC constraints in single frame pair.

(3) Transformation matrix calculation with accumulative matching point pairs obtained from multiframe sequences: this can overcome the shortcoming of insufficient matching points and matching deviation introduced by concentrated matching points in single image pair.

2.2. Feature Extraction and Comparison. Feature extraction is the first step of image matching and registration. In order to highlight the object, the source IR image is first negated. The commonly used feature descriptors can be divided into two types: histogram-based descriptors and binary descriptors. The *Scale Invariant Feature Transform* (SIFT), *Maximally Stable Extremal Region* (MSER), and *Speed Up Robust Features* (SURF) are typical histogram-based descriptors. Instead of expensive gradients operations, binary descriptors make use of simple pixel comparison, which result in binary strings of typically short length (commonly 256 bits) and lead to a major improvement in both running time and memory footprint [29]. The common binary descriptors include *Oriented Fast and Rotated BRIEF* (ORB), *Binary Robust Invariant Scalable Keypoints* (BRISK), and *Accelerated KAZE* (AKAZE). Each descriptor has its advantages and disadvantages. We will give a comparison of their results and select the best one as feature extraction method.

2.2.1. Histogram-Based Descriptors. (a) *SIFT* is the most renowned feature detection description method introduced by D.G. Lowe in 2004 [30]. It is based on Difference-of-Gaussians (DoG) operator which is an approximation of Laplacian-of-Gaussian (LoG). Feature points are detected by searching local maxim value using DoG at various scales of the subject images. SIFT is robustly invariant to image rotations, scale, and limited affine variations but its main drawback is high computational cost. Equation (1) shows the convolution of difference of two Gaussians (computed at different scales) with image “ $I(x, y)$ ”:

$$D(x, y, \sigma) = (G(x, y, k\sigma) - G(x, y, \sigma)) * I(x, y) \quad (1)$$

where “ G ” is the Gaussian function and σ is the scale.

(b) *SURF* also relies on Gaussian scale space analysis of images. It is based on *determinant of Hessian Matrix* and exploits integral images to improve feature-detection speed [31]. SURF features are also invariant to rotation and scale but

they have little affine invariance. The main advantage of SURF over SIFT is low computational cost. Equation (2) represents the Hessian Matrix in point “ $\mathbf{x} = (x, y)$ ” at scale “ σ ”:

$$H(x, \sigma) = \begin{bmatrix} L_{xx}(x, \sigma) & L_{xy}(x, \sigma) \\ L_{xy}(x, \sigma) & L_{yy}(x, \sigma) \end{bmatrix} \quad (2)$$

where $L_{xx}(x, \sigma)$ is the convolution of Gaussian second-order derivatives with the image in point \mathbf{x} and L_{xy} and L_{yy} are similar to L_{xx} .

When nonmaximal suppression is achieved in the $3 \times 3 \times 3$ stereo neighborhood detected by the Hessian Matrix, only the points larger than all 26 nearest neighbor’s responses are identified as feature points and then interpolated in the scale space to obtain a stable feature point location and scale.

(c) *MSER* generally attempts to detect a set of connected regions from an image. It is defined as

$$g(t) = \frac{(d/dt)|Q(t)|}{|Q(t)|} \quad (3)$$

where $|Q(t)|$ indicates the region area. MSER is computationally inexpensive and can be calculated in linear time. It was introduced by extremal to perform scene recognition under arbitrary viewpoints and illumination conditions. MSER performs very well when detected on flat surfaces and changing illumination. The only weakness for MSER is that it is sensitive to image blur. It can perform better in colorful visible image than in blurring IR image.

2.2.2. Binary Descriptors. (a) *ORB descriptor* was proposed by Reblee et al. [32]. It builds upon BRIEF by adding a rotation invariance mechanism that is based on measuring the patch orientation using first-order moments. BRIEF compares the same set of smoothed pixel pairs for each local patch that it describes. For each pair, if the first smoothed pixel intensity is larger than that of the second, BRIEF writes 1 in the final descriptor string and 0 otherwise. ORB also uses unsupervised learning to learn the sampling pairs.

(b) *BRISK descriptor* was proposed by Leutenegger in 2011 [33]. It uses a hand-crafted sampling pattern that is composed of set of concentric rings. BRISK uses the long-distance sampling pairs to estimate the patch orientation and the short distance sampling pairs to construct the descriptor itself through pixel intensity comparisons.

(c) *AKAZE* uses the Accelerated-KAZE detector estimation of orientation for rotating the LDB grid to achieve

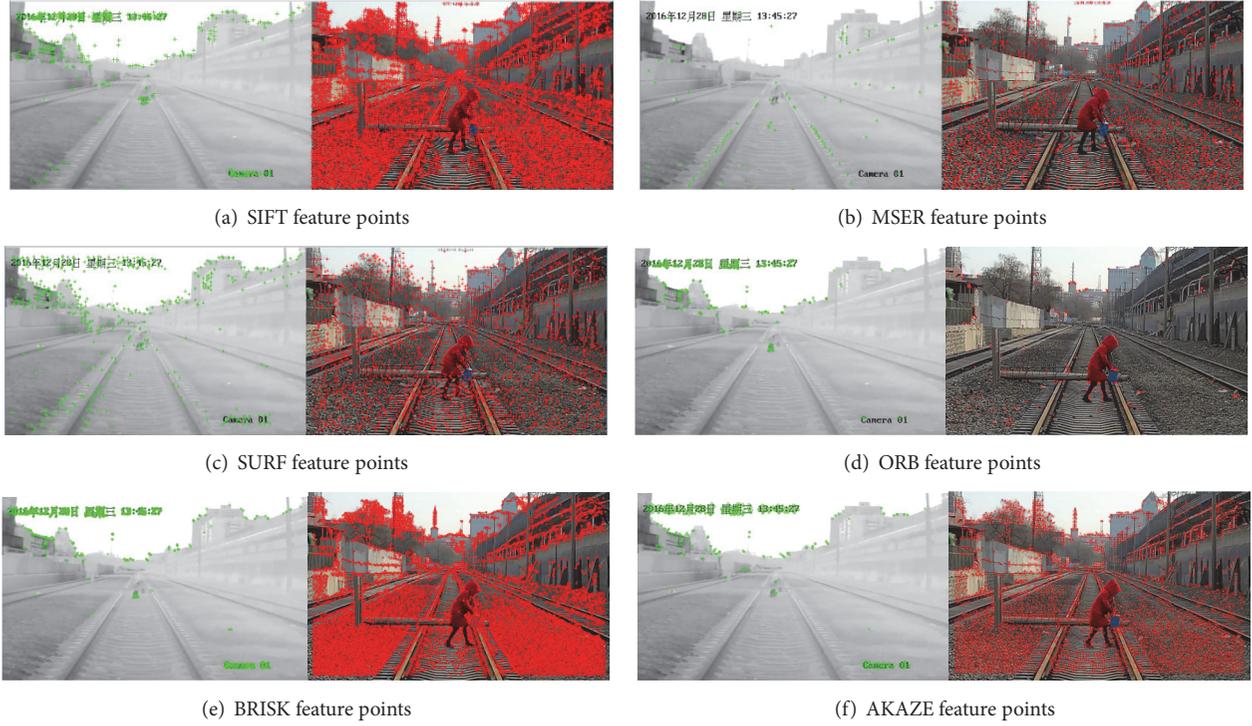


FIGURE 2: Comparison of different descriptors.

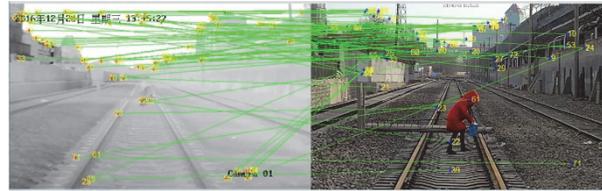


FIGURE 3: Initial candidate matching points.

rotation invariance [34]. In addition, AKAZE also uses the A-KAZE detector's estimation of patch scale to subsample the grid in steps that are a function of the patch scale.

2.2.3. Comparison of Different Descriptors. Figure 2 shows the results of SIFT, MSER, SURF, ORB, BRISK, and AKAZE descriptors. In general, binary descriptors are faster than histogram-based descriptors. ORB obtains the fewest feature points in both visible and IR images. For visible image, SIFT and BRISK obtain much more feature points than MSER, SURF, and AKAZE. For IR image, SIFT and SURF obtain much more feature points than other descriptors. But SURF feature points are more evenly distributed in whole IR image than all other methods. That is conducive to the following registration. We have tried to change the parameters of these detectors but found that parameters made little contribution to the improvement of uneven distribution. So we choose SURF as feature extractors for subsequent matching process.

2.3. Initial Candidate Matching Point Pairs Generation. After SURF features extraction, Euclidean distance is used to

measure the similarity of feature points in two images. For a feature point in IR image, it looks for two nearest points measured with Euclidean distance in visible image. If the ratio of the nearest distance d_{ND} to next nearest distance d_{NND} is less than threshold ϵ in formula (4), the points with the nearest distance are considered as a candidate matching point pair.

$$\frac{d_{ND}}{d_{NND}} < \epsilon \quad (4)$$

where ϵ controls the number of matching points. The number will decrease as ϵ decreases. After massive experiments, the default value is 0.8. The initial candidate matching points in IR and visible images are shown in Figure 3. There are many mismatching points that need to be refined.

2.4. Candidate Matching Points Refining

2.4.1. Geometric Constraints Refining. The centers of IR and visible image can be considered approximately identical when the images are obtained at a long distance. IR image

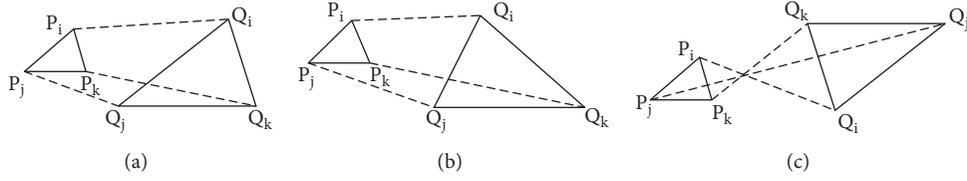


FIGURE 4: Similar triangle constraints.

center is noted as (x_{or}, y_{or}) , visible image center is noted as (x_{ov}, y_{ov}) , and any corresponding matching points (x_{ir}, y_{ir}) and (x_{iv}, y_{iv}) should obey the following geometric constraints:

(1) The dip angles of corresponding matching points to their own image centers are approximately equal:

$$\left| \arctan\left(\frac{(x_{ir} - x_{or})}{(y_{ir} - y_{or})}\right) - \arctan\left(\frac{(x_{iv} - x_{ov})}{(y_{iv} - y_{ov})}\right) \right| < T \quad (5)$$

where T is a threshold for dip angle difference; the matching points should be deleted when the dip difference is larger than T .

(2) The corresponding matching points should be within the same quadrants in both IR and visible images coordinate system. In other words, the vertical and horizontal coordinate differences of corresponding matching points and their respective image centers should have the same sign shown as formula (6):

$$\begin{aligned} (x_{or} - x_{ir}) * (x_{ov} - x_{iv}) &> 0 \\ (y_{or} - y_{ir}) * (y_{ov} - y_{iv}) &> 0 \end{aligned} \quad (6)$$

(3) Since the IR scene is larger than visible scene, the pixel distance in IR image should be less than that in visible image:

$$\begin{aligned} \sqrt{(x_{or} - x_{ir})^2 + (y_{or} - y_{ir})^2} \\ < \sqrt{(x_{ov} - x_{iv})^2 + (y_{ov} - y_{iv})^2} \end{aligned} \quad (7)$$

The candidate matching points should be deleted if any constraints above are not satisfied.

2.4.2. Similar Triangle Constraint Refining. After geometric constraints refining, some mismatching points are deleted. But there are still “many-to-one” and crossing mismatching points. We will refine them further with similar triangle constraint.

Assume that P and Q are matching point sets after geometric constraint. Any 3 feature points (noncollinear) in IR image can form a triangle $\Delta P_i P_j P_k$, ($i < j < k, P_i, P_j, P_k \in P$). The corresponding triangle in visible image is $\Delta Q_i Q_j Q_k$ ($i < j < k, Q_i, Q_j, Q_k \in Q$). These two triangles should be similar, shown in Figure 4(a).

According to the characters of similar triangles, there exists

$$\frac{P_i P_j}{Q_i Q_j} = \frac{P_j P_k}{Q_j Q_k} = \frac{P_k P_i}{Q_k Q_i} \quad (8)$$

We define dd_1 and dd_2 as

$$\begin{aligned} dd_1 &= \frac{|P_i P_j| / |Q_i Q_j|}{|P_j P_k| / |Q_j Q_k|} \\ dd_2 &= \frac{|P_i P_k| / |Q_i Q_k|}{|P_j P_k| / |Q_j Q_k|} \end{aligned} \quad (9)$$

From formula (8), we know that dd_1 and dd_2 in formula (9) should equal 1. But for noise interruption, they should approximately equal 1:

$$\begin{aligned} |dd_1 - 1| &< T_0 \\ |dd_2 - 1| &< T_0 \end{aligned} \quad (10)$$

T_0 is a threshold near to zero. In this paper, T_0 is 0.1.

In Figure 4(b), $|P_i P_j| / |Q_i Q_j| > |P_j P_k| / |Q_j Q_k|$, $dd_1 > 1$, and $dd_2 < 1$; they are not similar triangles even if their side lengths satisfy formula (10). In this situation, we should firstly sort their side lengths; only the similar triangle with the same sort can be kept. After that, there will also be another situation inverted as Figure 4(c). It can be removed with formula (11):

$$\left| \frac{\overrightarrow{P_i P_j}}{|P_i P_j|} - \frac{\overrightarrow{Q_i Q_j}}{|Q_i Q_j|} \right| < T_1 \quad (11)$$

Formula (11) means that unit vector of any pair of vertices (such as $\overrightarrow{P_i P_j}$ and $\overrightarrow{Q_i Q_j}$) in similar triangle should be approximated. T_1 is also a threshold near zero. In this paper, T_1 is 0.1.

In order to improve the matching accuracy, a traversal similar triangle test method is adopted. For the matching point set P and Q with n pairs of matching points, we randomly select 3 matching points to consist C_n^3 pairs of triangles. Then, we introduce an accumulator for each matching point. When any pairs of triangle satisfy the above similar triangle constraints, the accumulator increases by 1. After traversing all these triangles, the larger the accumulator is, the higher the probability of correct matching is for this point pair. When the accumulator value is larger than threshold T_2 , the corresponding point pair is considered as a correct matching point pair. The point pairs corresponding to the first k largest accumulators are considered as k correct candidate matching point pairs.

2.4.3. Fine Matching with RANSAC. After refining with geometric and similar triangle constraints, the reserved matching

TABLE 1: Comparison of different matching algorithms.

| Method | feature point numbers in IR/Visible image | Candidate matching points | Fine matching points | Correct points |
|-------------|---|---------------------------|----------------------|----------------|
| SIFT+RANSAC | 664/11,433 | 664 | 6 | 0 |
| SURF+RANSAC | 523/2,766 | 523 | 4 | 0 |
| Our method | 553/2,766 | 72 | 3 | 3 |

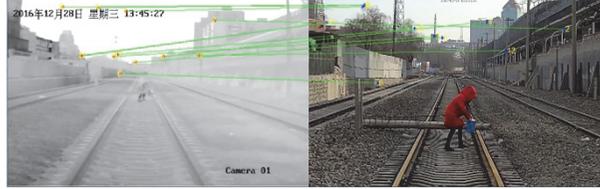


FIGURE 5: Refining result of geometric constraint.

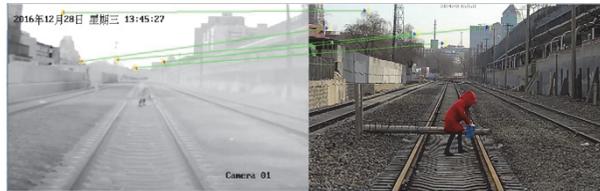


FIGURE 6: Refining result of similar triangle constraint.

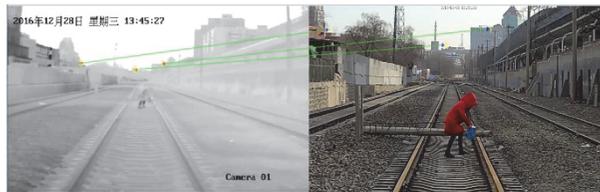


FIGURE 7: Result of RANSAC fine matching.

points are almost correct except several separate points. In order to improve the correct matching further, RANSAC fine matching algorithm is finally used for the prior reserved matching points.

The refining process is shown in Figures 5–7. After refining with geometric constraints there are 12 pairs of matching points in Figure 5. After similar triangle refining, 5 pairs are left in Figure 6. After RANSAC matching, only 3 pairs of correct matching points are reserved in Figure 7.

After consecutive procedure of geometric constraints, similar triangle, and RANSAC matching, the reserved point pairs are all correctly matched. In order to evaluate the effect of proposed algorithm, a comparison to SIFT+RANSAC and SURF+RANSAC is shown in Table 1. It is shown that the traditional SIFT+RANSAC and SURF+RANSAC matching methods are not applicable to multimodal images. Our proposed method with successive geometric, similar triangle, and RANSAC constraints can obtain the correct matching point pairs in IR and visible images.

2.5. Transformation Matrix Calculation. In our application, both IR and visible cameras are equipped with microdistortion lens. The distortion is so small that it can be ignored in

image registration process. The IR and visible cameras have approximate parallel optical axis and follow the perspective projection model. The coordinates between two images are shown as formula:

$$\begin{bmatrix} x' \\ y' \\ 1 \end{bmatrix} = \begin{bmatrix} m_0 & m_1 & m_2 \\ m_3 & m_4 & m_5 \\ m_6 & m_7 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = M \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} \quad (12)$$

where $(x, y), (x', y')$ are coordinates of corresponding matching points in IR and visible images. M is transformation matrix between the two images. Eight parameters require at least 4 pairs of matching points.

As shown in Figure 7, we cannot ensure that there are more than 4 pairs of correct matching points in a single image pair with the prior matching procedures. If the number of matching point pairs is less than 4, it is insufficient to calculate the 8 parameters in matrix M . On the other side, even though we got 4 pairs in one single image pair, the transformation will have deviation on another position if the matching pairs are too concentrated. It will still lead to bias if we use this matrix for registration.

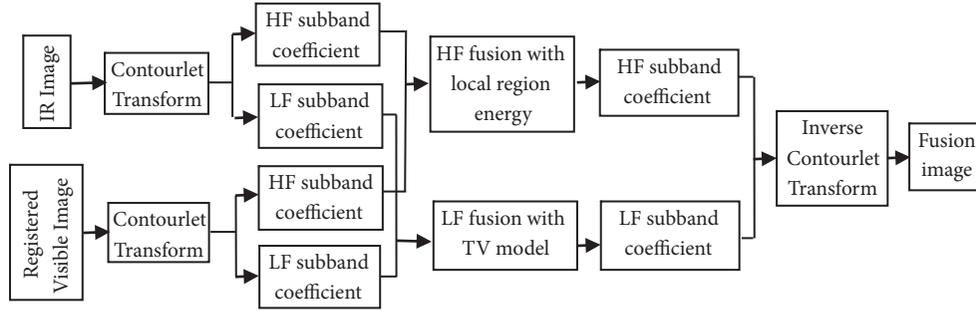


FIGURE 8: Flow chart of image fusion algorithm.

In order to improve the precision of transformation matrix, we create an accumulative set $Tpts$ from respective matching points sets $pts(i)$ of multiframe images where objects are in different position:

$$Tpts = \{pts(1), \dots, pts(i), \dots, pts(n)\} \quad (13)$$

where n is the number of selected frames, $pts(i)$ is the matching points set in the i th frame image pairs, and $Tpts$ is the accumulative point sets with the selected n frames of image pairs. In practice, we accumulate at least 12 correct matching point pairs for parameters calculation.

Substitute $Tpts$ coordinates into formula (12); we can get the transformation matrix parameters with the Least Squares method. The transformation matrix obtained with this method has higher accuracy and applicability.

3. Improved Contourlet Transform Fusion Combined with Total Variation and Local Region Energy

3.1. Improved Contourlet Transform. Contourlet transform is a multiscale geometric image analysis method. It completes multiscale analysis and direction analysis separately. Firstly, an image is multiscale decomposed with Laplacian Pyramid (LP) method to capture the singular points. With LP process, the source signal is decomposed into a low frequency and a band-pass signal of difference between source and predicted signal. Then, the band-pass signals of each pyramid level are direction-filtered. The singular points distributed in the same direction are combined into a coefficient by a *directional filter bank* (DFB). Both LP and DFB can be fully reconstructed. The discrete Contourlet transform of their combination can also be reconstructed perfectly. After registration, the resolution of both visible and IR images is 960×576 . There are 3 LP decomposition levels in Contourlet transform. The direction numbers of DFB are 2-3-4.

The procedure of IR and visible images fusion with improved Contourlet transform is shown in Figure 8. There are 3 steps:

(1) Contourlet transform to both IR and visible images, respectively, to get their coefficients of high and low frequency subbands under different scales.

(2) Different fusion rules fused low and high frequency coefficients. For low frequency subband, the coefficients

are fused with total variation model. For high frequency subbands, the coefficients are fused with local region energy.

(3) The fusion image is reconstructed by inverse Contourlet transform.

3.2. Low Frequency Fusion with Total Variation Model. Due to different principles of IR and visible imaging, different images include their own information. An IR image indicates the thermal distribution of objects. However, a visible image contains detailed information such as color and contour of objects. The low frequency subband keeps the approximate characteristics and contains most energy of source images. The familiar weighted fusion for low frequency subband will reduce the scales of both thermal radiation in IR image and colors in visible image, which will reduce the contrast of fusion. In this paper, the low frequency subband fusion is transformed into the minimization of optimization problem of total variation (TV) model.

Assume that the IR, registered visible, and fused image are all gray images of $m \times n$. Their column vectors are $u, v, x \in R^{mn \times 1}$. Since the targets in IR image are usually much more distinct than in visible image, we expected that the difference between fused and IR image in formula (14) is as small as possible:

$$\varepsilon_1(x) = \|x - u\|_1 \quad (14)$$

On the other hand, we also need details of objects. A simple method is to make the fused image have similar pixel intensity to the visible image. But the detailed appearance of object depends more on gradient than on pixel intensity. So fused image should have similar gradient rather than pixel intensity to visible image. The difference in formula (15) should also be as small as possible.

$$\varepsilon_2(x) = \|\nabla_x - \nabla_v\|_1 \quad (15)$$

where ∇ is gradient operator.

The coefficient solution of low frequency subband fusion can be expressed as a minimization problem of an objective function in formula (16).

$$\varepsilon(x) = \varepsilon_1(x) + \lambda \varepsilon_2(x) = \|x - u\|_1 + \lambda \|\nabla_x - \nabla_v\|_1 \quad (16)$$

where $\varepsilon_1(x)$ indicates that fused image x has similar pixel intensities to IR image u ; $\varepsilon_2(x)$ indicates that fused image x

has similar gradients to visible image v ; and λ is a positive constant that controls trade-off between the two terms. The objective function (16) aims to transfer the gradients/edges in visible image onto the corresponding positions in IR image. So the low frequency fusion image should look like an IR image but with more appearance details.

Obviously, the objective function (16) is convex and has global optimal solution. The first term $\varepsilon_1(x)$ is smooth, and the second term $\varepsilon_2(x)$ is nonsmooth. Assuming $z = x - v$, then (16) can be rewritten as

$$z^* = \arg \min \left\{ \sum_{i=1}^{mn} |z_i - (u_i - v_i)| + \lambda J(z) \right\} \quad (17)$$

where

$$J(z) = \sum_{i=1}^{mn} |\nabla_i z| = \sum_{i=1}^{mn} \sqrt{(\nabla_i^h z)^2 + (\nabla_i^v z)^2} \quad (18)$$

The new objective function (17) is a standard total variation (TV) minimization problem. It is the model for low frequency subband fusion. z^* can be calculated with λ regularization of the TV minimization problem. Then the global optimal solution of fusion image is $x^* = z^* + v$. The

column vector of $x^* \in R^{mn \times 1}$ is the low frequency fusion coefficients.

3.3. High Frequency Fusion with Local Region Energy. High frequency subbands stand for details of images. The local region energy is used for high frequency subband coefficients fusion. Not only the pixel itself but also its local neighbor pixels are involved in the calculation. The procedure is as follows:

(1) For two images A and B , calculate their own local region energies $E_{l,A}, E_{l,B}$ on their corresponding pyramid level centered (n, m) separately:

$$E_l(n, m) = \sum_{n' \in J, m' \in K} \omega'(n', m') [LP_l(n + n', m + m')]^2 \quad (19)$$

where $E_l(n, m)$ is local region energy centered (n, m) on the l th level of Laplacian Pyramid; LP_l is the l th level image; $\omega'(n', m')$ is the weight coefficient corresponding to LP_l ; J, K define the range of local region in images A and B , and n', m' are inner points in J and K .

(2) Calculate the matching degree between corresponding local regions in two images:

$$M_{l,AB}(n, m) = \frac{2 \sum_{n' \in J, m' \in K} \{\omega'(n', m') LP_{l,A}(n + n', m + m') LP_{l,B}(n + n', m + m')\}}{E_{l,A}(n, m) + E_{l,B}(n, m)} \quad (20)$$

where $E_{l,A}, E_{l,B}$ are calculated with formula (19).

(3) According to the matching degree, different fusion rules are adopted as follows:

If $M_{l,AB}(n, m) < \alpha$ ($\alpha = 0.85$), it indicates that the correlation between two source images is low. The coefficient with larger energy is used as fused coefficients shown as formula (21).

$$\begin{aligned} & \text{if } E_{l,A}(n, m) \geq E_{l,B}(n, m), \\ & LP_{l,F}(n, m) = LP_{l,A}(n, m) \\ & \text{if } E_{l,A}(n, m) < E_{l,B}(n, m), \\ & LP_{l,F}(n, m) = LP_{l,B}(n, m) \end{aligned} \quad (21)$$

If $M_{l,AB}(n, m) \geq \alpha$, it indicates that the correlation degree between two source images is very large, and the weighted average fusion method shown in (22) is better.

$$\begin{aligned} & \text{if } E_{l,A}(n, m) \geq E_{l,B}(n, m), \\ & LP_{l,F}(n, m) = W_{l,\max}(n, m) LP_{l,A}(n, m) \\ & \quad + W_{l,\min}(n, m) LP_{l,B}(n, m) \end{aligned}$$

if $E_{l,A}(n, m) < E_{l,B}(n, m)$,

$$\begin{aligned} LP_{l,F}(n, m) &= W_{l,\min}(n, m) LP_{l,A}(n, m) \\ & \quad + W_{l,\max}(n, m) LP_{l,B}(n, m) \end{aligned} \quad (22)$$

where

$$W_{l,\min}(n, m) = \frac{1}{2} - \frac{1}{2} \left(\frac{1 - M_{l,AB}(n, m)}{1 - \alpha} \right) \quad (23)$$

$$W_{l,\max}(n, m) = 1 - W_{l,\min}(n, m) \quad (24)$$

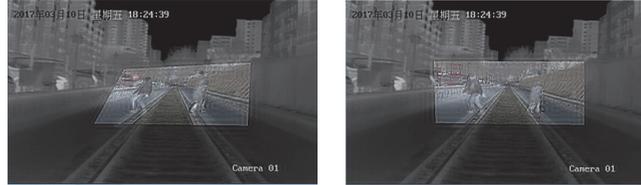
4. Experiment Results and Analysis

The performance of proposed registration and fusion algorithm is discussed in this section. The registration results of single frame and multiframe image pairs are introduced in Section 4.1. In Section 4.2, the performance comparison between source images and 5 state-of-the-art fusion results is presented and discussed. In Section 4.3, 15 frames of image pairs sampled from daytime and night videos separately will be used to evaluate the stability of proposed algorithm.

4.1. Results of Single and Multiframe Registration. Since single frame pair cannot ensure getting enough matching point pairs for transform matrix calculation, the multiframe



FIGURE 9: Infrared and visible images.



(a) Single frame registration

(b) Multiframe registration

FIGURE 10: Fusion result of single and multiframe registration.

accumulative features method not only solves the problem of insufficient matching points in single frame pair but also improves the applicability of transform model with evenly distributed feature points. The IR and visible source images of railway scene are shown in Figure 9. Their fused image with single frame registration (using 4 pairs of feature points in single image pair) is shown in Figure 10(a). The left rail in the fused image is not aligned very well because of the concentrated feature points for matrix calculation. The fused image with multiframe registration is shown in Figure 10(b). The coincidence of all objects is much better than that in Figure 10(a).

4.2. Fusion Performance Evaluation. Figure 11 shows the source images and fusion results of 5 state-of-the-art methods (including ours) for daytime and night images. From top to bottom, they are source IR image, registered visible image, fusion result with gray weighted average, Contourlet transform, wavelet transform, total variation fusion, and our proposed method. The left column shows daytime images; and the right column shows night images. In Figure 11, the fusion results of gray weighted average, wavelet transform, and total variation are much smoother and more blurred. The results of Contourlet transform and ours are much sharper and closer to source IR image. The result of our method at night is sharpest and of highest gray value. It not only kept the thermal information of IR image, but also greatly weakened the strong disturbance light at night.

Due to the subjective assessment varying from person to person, we also introduce objective evaluation with quantitative indicators. Average gray, standard deviation, average gradient, information entropy, and edge intensity are used for quantitative analysis. Their definitions are as follows:

(a) Average gray is the mean value of all pixels in the image, defined as

$$Ave = \frac{1}{n \times m} \sum_{i=0}^{n-1} \sum_{j=0}^{m-1} P(i, j) \quad (25)$$

where $P(i, j)$ is the gray value at pixel (i, j) ; n, m are horizontal and vertical size of image; and Ave is the average gray of the image.

(b) Standard deviation describes the dispersion between all pixels and average gray, defined as

$$\sigma = \sqrt{\frac{\sum_{i=0}^{m-1} \sum_{j=0}^{n-1} [P(i, j) - Ave]^2}{m \times n}} \quad (26)$$

(c) Average gradient, also known as sharpness, reflects image sharpness and variation in texture details, defined as

$$Ave_v = \frac{\sum_{i=0}^{n-1} \sum_{j=0}^{m-2} |F(i+1, j) - F(i, j)|}{(n-1) \times m} \quad (27)$$

$$Ave_H = \frac{\sum_{i=0}^{n-2} \sum_{j=0}^{m-1} |F(i, j+1) - F(i, j)|}{(n-1) \times m} \quad (28)$$

$$AG = \sqrt{Ave_v^2 + Ave_H^2} \quad (29)$$

where Ave_v , Ave_H are average gradients in vertical and horizontal direction, respectively and AG is the average gradient of the fused image F .

(d) Information entropy is used to evaluate the richness of image information. We assume that the gray values of each pixel in an image are independent samples; then the gray distribution is $P = \{P_0, P_1, \dots, P_i, \dots, P_n\}$; P_i is the probability of the gray value i in image, that is, the ratio of N_i (the pixels number with gray value i) to N (the total pixels number in the image). L is the total number of gray levels in the image. According to Shannon's theorem, information entropy of an image is defined as

$$H = - \sum_{i=0}^{L-1} P_i \log P_i \quad (30)$$

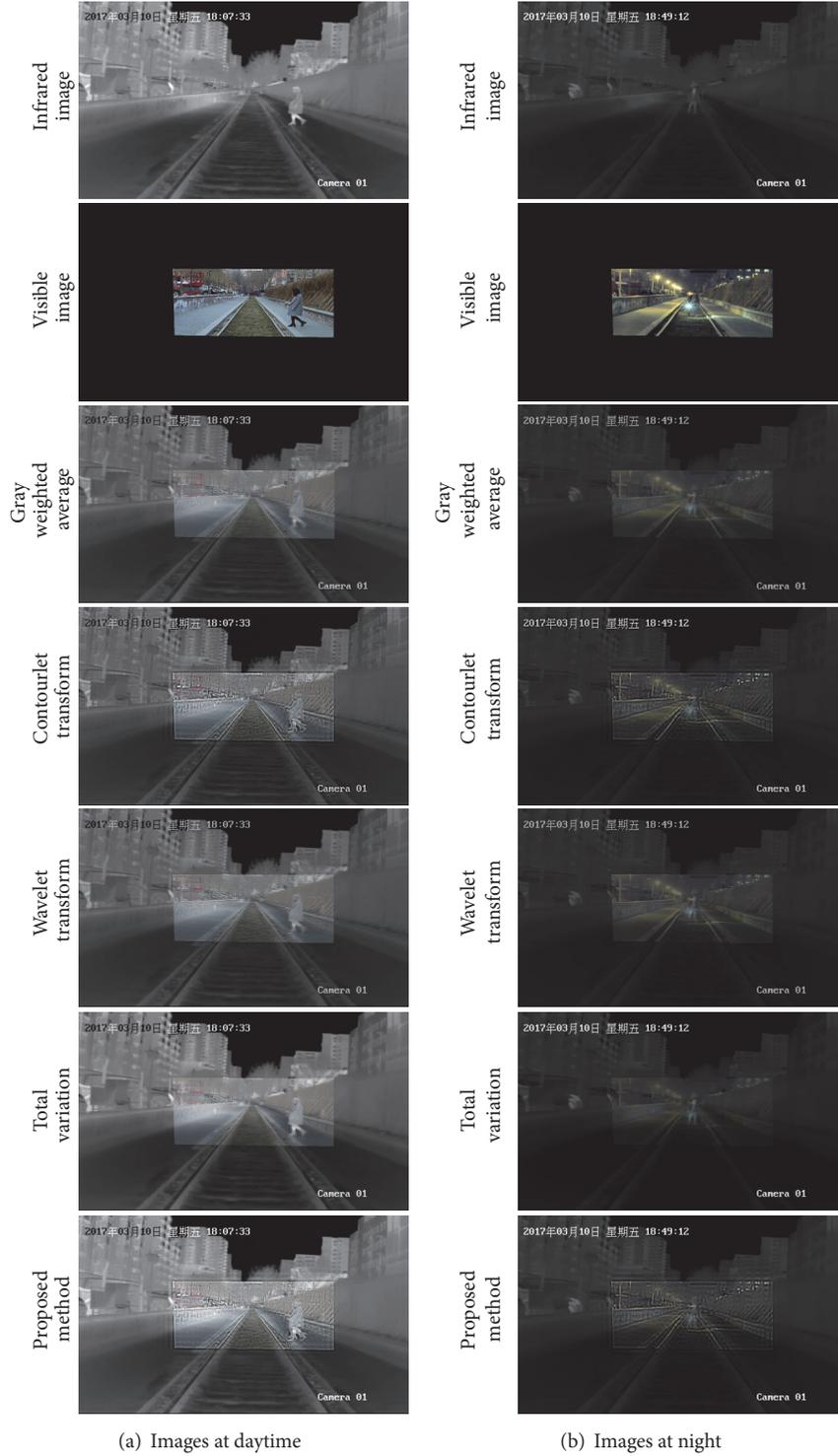


FIGURE 11: Comparison of visual quality of source and fused images of various methods at daytime and night.

(e) Edge intensity is essentially the amplitude of edge gradient. The first-order difference in directions x and y is defined as

$$\nabla P_x(i, j) = P(i, j) - P(i - 1, j) \quad (31)$$

$$\nabla P_y(i, j) = P(i, j) - P(i, j - 1) \quad (32)$$

The amplitude of gradients is

$$G(i, j) = \sqrt{\nabla P_x(i, j)^2 + \nabla P_y(i, j)^2} \quad (33)$$

TABLE 2: Quantitative evaluation of source image and fusion results at daytime.

| evaluation indicators | Source image | | | Fusion results | | | |
|-----------------------|--------------|--------------------|-----------------------|----------------------|-------------------|-----------------|--------------|
| | IR | Registered visible | gray weighted average | Contourlet Transform | Wavelet transform | Total variation | proposed |
| average gray | 88.37 | 17.28 | 67.09 | 67.37 | 67.07 | 76.66 | 91.60 |
| standard deviation | 47.73 | 43.20 | 37.07 | 40.13 | 37.85 | 46.46 | 51.63 |
| average gradient | 2.60 | 2.55 | 2.41 | 4.90 | 3.37 | 2.64 | 4.89 |
| Information entropy | 6.70 | 1.87 | 6.46 | 6.87 | 6.61 | 6.95 | 7.07 |
| edge intensity | 27.74 | 22.15 | 24.21 | 47.04 | 30.33 | 27.18 | 46.98 |

TABLE 3: Quantitative evaluation of source image and fusion results at night.

| evaluation indicators | Source image | | | Fusion results | | | |
|-----------------------|--------------|--------------------|-----------------------|----------------------|-------------------|-----------------|--------------|
| | IR | Registered visible | gray weighted average | Contourlet Transform | Wavelet transform | Total variation | proposed |
| average gray | 32.40 | 13.97 | 26.89 | 27.10 | 26.90 | 21.22 | 33.06 |
| standard deviation | 27.49 | 35.89 | 22.74 | 26.06 | 22.91 | 26.48 | 29.27 |
| average gradient | 1.67 | 1.07 | 1.44 | 2.61 | 1.65 | 1.55 | 2.65 |
| Information entropy | 5.82 | 1.83 | 5.76 | 5.93 | 5.82 | 5.09 | 6.06 |
| edge intensity | 17.70 | 11.53 | 15.23 | 27.69 | 17.07 | 16.56 | 28.05 |

For all the five indicators, the higher the value is, the better its performance is. The quantitative evaluation results for Figure 11 are shown in Tables 2 and 3.

For daytime image fusion results in Table 2, relative to the high value of source images, our proposed fusion method improved 3.66% at average gray, 8.17% at standard deviation, and 5.52% at information entropy. Because of the differences between IR and visible images, the improvement at average gradient and edge intensity is even more, 88.08% and 69.36%, respectively. Compared to other 4 methods, our method performs the best in average gray, standard deviation, and information entropy and the second best in average gradient and edge intensity. In general, compared to source images and other 4 fused results, our proposed method has the best performance.

For night image fusion results in Table 3, relative to the high value of source images, our proposed fusion method improved 2.03% at average gray, 4.12% at information entropy, and 56.68% and 58.47% at average gradient and edge intensity, respectively. But it has lower standard deviation than visible source image because of great illumination variation of lamps and shining of torch in dark environment. The illumination variety of large standard deviation leads to mass misdetection in visible images analysis. That is what we want to eliminate at night. In conclusion, our proposed fusion algorithm improved all other evaluation indicators except standard deviation. It has the best performance among the 5 fusion methods.

4.3. Stability Verification of Fusion Algorithm. Section 4.2 only shows the effectiveness of our algorithm on one typical frame pair at daytime and night separately. In order to verify the stability and reliability of our algorithm, 15 frames of image pairs sampled from different daytime and night videos, respectively, are used for quantitative evaluation. In this section, besides three indicators of standard deviation,

information entropy, and sharpness defined in Section 4.2, we also introduced another two indicators of mutual information and cross entropy defined as follows:

(a) Mutual information represents a measure of correlation between multiple variables. The mutual information $MI((A, B) : F)$ between images A , B , and F is defined as

$$MI((A, B) : F) = \sum_{i=0}^{L-1} \sum_{j=0}^{L-1} \sum_{k=0}^{L-1} P_{abf}(i, j, k) \log \frac{P_{abf}(i, j, k)}{P_{ab}(i, j) P_f(k)} \quad (34)$$

Similarly, the relation chart between i and $P_{ab}(i, j)$ can be thought as a normalized joint grayscale histogram of images A and B ; the relation chart between i and $P_{abf}(i, j, k)$ is the normalized joint grayscale histogram of images A , B , and F .

The larger the mutual information is, the more information the fusion image gets from source images.

(b) Cross entropy is used to measure the information difference of gray distribution between two images. Assuming reference image R and fusion image F , the cross entropy of R and F is defined as

$$CE_{R,F} = \sum_{i=0}^{L-1} P_{R_i} \log \frac{P_{R_i}}{P_{F_i}} \quad (35)$$

where P_{R_i} and P_{F_i} are gray distribution of reference image R and fusion image F , respectively.

The smaller the cross entropy value is, the more information the fusion image gets from source images and the better performance the fusion algorithm has. Because of the different quality of daytime and night images, we evaluate them separately as follows:

(1) *Fusion Evaluation of Daytime Images.* We sampled 15 frames of image pairs from daytime videos and compared 5 evaluation indicators of 4 state-of-the-art methods and our

TABLE 4: Average evaluation value of different fusion methods at daytime.

| | Standard deviation | Mutual information | Information entropy | Cross entropy | Sharpness |
|------------------------------|--------------------|--------------------|---------------------|---------------|--------------|
| Gray weighted average fusion | 36.686 | 1.571 | 6.454 | 1.084 | 3.099 |
| Contourlet transform fusion | 39.767 | 1.771 | 6.860 | 0.802 | 7.202 |
| Wavelet transform fusion | 37.463 | 1.644 | 6.602 | 0.860 | 5.224 |
| Total variation fusion | 44.223 | 1.687 | 6.889 | 0.590 | 3.099 |
| Proposed | 51.078 | 1.717 | 7.061 | 0.279 | 6.923 |

TABLE 5: Average evaluation value of different fusion methods at night.

| | Standard deviation | Mutual information | Information entropy | Cross entropy | Sharpness |
|------------------------------|--------------------|--------------------|---------------------|---------------|--------------|
| Gray weighted average fusion | 22.270 | 1.555 | 5.674 | 0.623 | 1.540 |
| Contourlet transform fusion | 25.425 | 1.722 | 5.841 | 0.628 | 2.852 |
| Wavelet transform fusion | 22.413 | 1.597 | 5.729 | 0.587 | 1.936 |
| Total variation fusion | 26.322 | 1.681 | 5.066 | 0.849 | 1.618 |
| Proposed | 28.846 | 1.690 | 6.001 | 0.134 | 2.828 |

proposed method. The 4 compared methods include gray weighted average, Contourlet transform, wavelet transform, and total variation methods. The 5 indicators including standard deviation, mutual information, information entropy, cross entropy, and sharpness of 5 methods are shown in Figures 12(a)–12(e). In Figure 12, except the large deviation in standard deviation, information entropy, and cross entropy of total variation method, all the other indicators of 5 methods have little deviation in the 15 frames. So, we can draw the conclusion that, except total variation method, the other 4 fusion methods have stability for different day images.

The average values of 15 daytime frame pairs are shown in Table 4. In the 5 evaluation indicators, for cross entropy, the less its value is, the more information of source images is translated into the fused image and the better the fusion effect is. For the other 4 indicators, the larger they are, the better the fusion effect is. For the 5 algorithms, our proposed method is the best in standard deviation, information entropy, and cross entropy and the second best in mutual information and sharpness. It improved 15.5% at standard deviation, 2.5% at information entropy, and 52.7% at cross entropy compared to the second best method. And as the second best one in mutual information and sharpness, the difference to the best one is very little. In summary, our proposed method is much better than other methods and can get best performance for day images fusion.

(2) *Fusion Evaluation of Night Images.* The visible images usually do not have good quality at night because of bad illumination and disturbance light. So we should also evaluate the fusion effect for night images. We still sampled 15 frames of railway night image pairs from different night videos. The 5 evaluation indicators of different methods are shown in Figures 13(a)–13(e). In Figure 13, different methods have relatively stable performance. The deviations of all indicators are much greater than daytime because of the poor quality of night visible images. The deviations of 5 methods on the same frame have similar trends. This

phenomenon indicates that deviations depended more on image qualities rather than on different fusion methods. In fact, frame 6 has the biggest deviation because of strong disturbance lights.

The average values of 15 night frames pairs are shown in Table 5. In Table 5, our proposed method has the best performance in standard deviation, information entropy, and cross entropy and improved 9.6%, 18.5%, and 77.2%, respectively, compared to the second best one. It is also the second best method in mutual information and sharpness. In summary, our proposed method has the most outstanding performance for night images fusion. It reduced the strong disturbance of torchlight and improved the contour and thermal information of objects at night.

5. Conclusion

In this paper, we propose a novel registration and fusion algorithm for multimodal railway images with different field of views. One of the main novelties is the multimodal matching SURF features refining procedure with geometric, similar triangle, and RANSAC constraints in registration process. Another novelty is the improved Contourlet transform image fusion algorithm combined with total variation model and local region energy. Experiment results of railway images validate the effectiveness of our proposed registration and fusion approach. Compared to other 4 state-of-the-art methods, our method performs best and improves 15.5%, 2.5%, and 52.7% for day images and 9.6%, 18.5%, and 77.2% for night images compared to the second best one in standard deviation, information entropy, and cross entropy. And it is the second best one in mutual information and sharpness. Our method greatly reduced the strong disturbance light of torch at night. This paper gives impetus to the research on objects detection in complex circumstance. As for future work, we plan to use the fusion results to study the railway clearance intrusion detection methods.

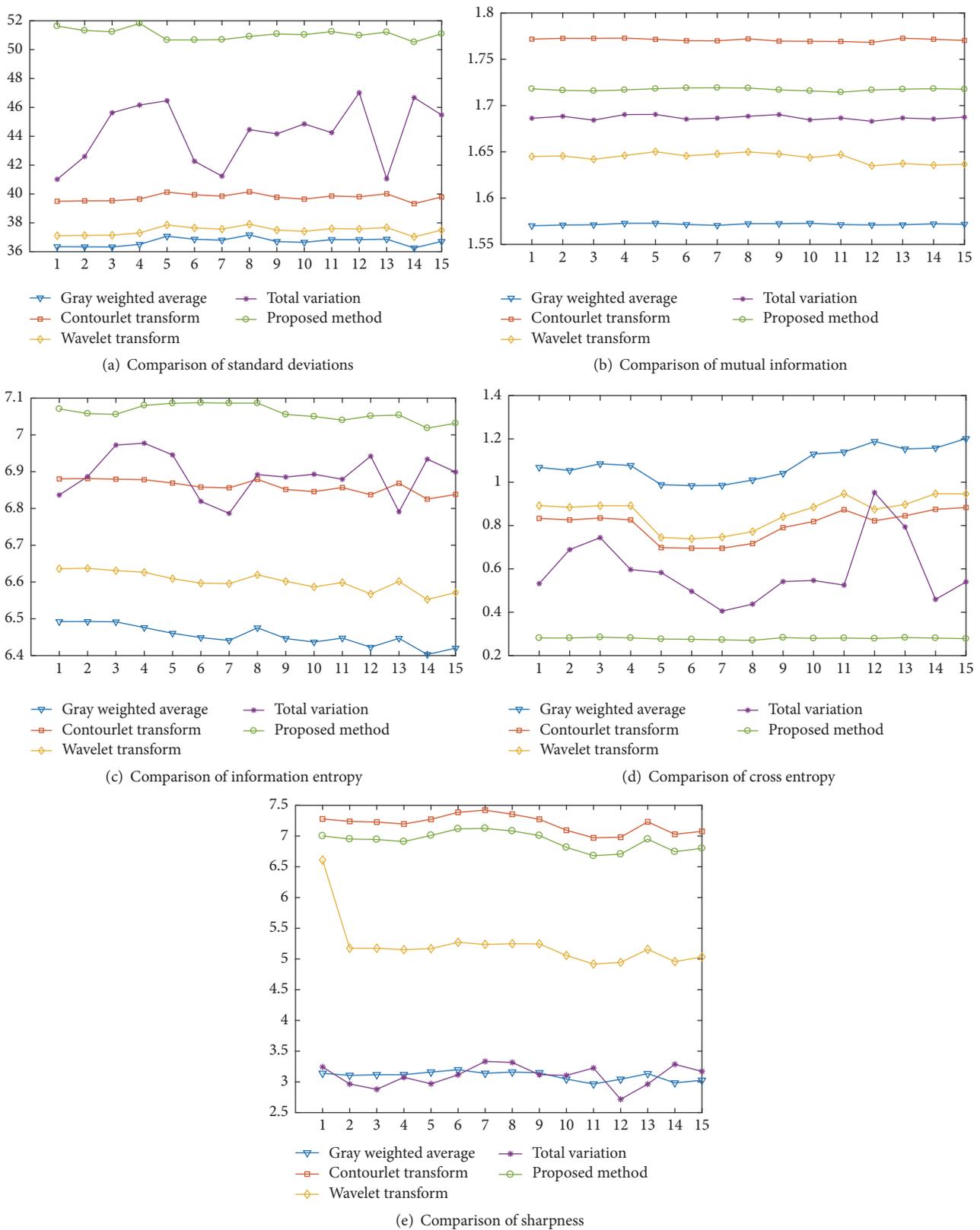
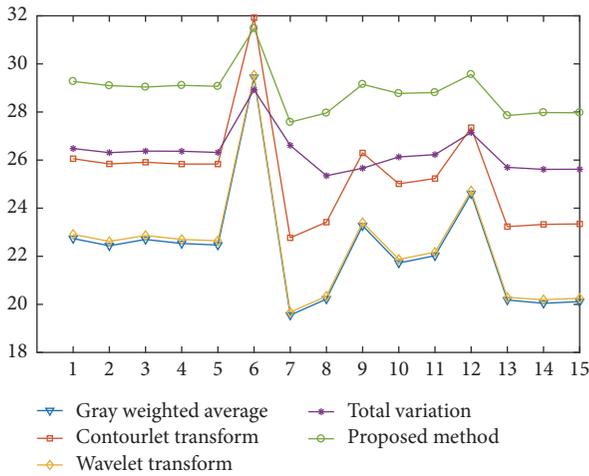
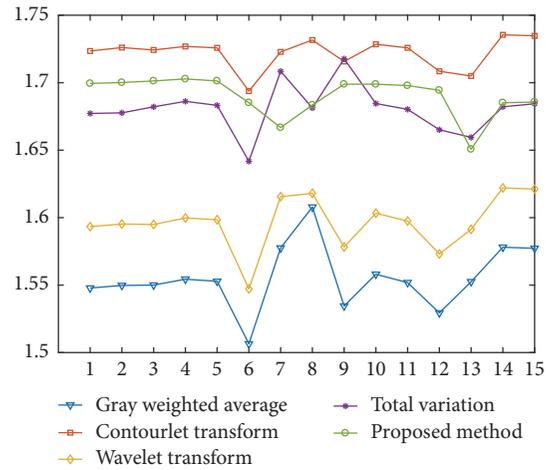


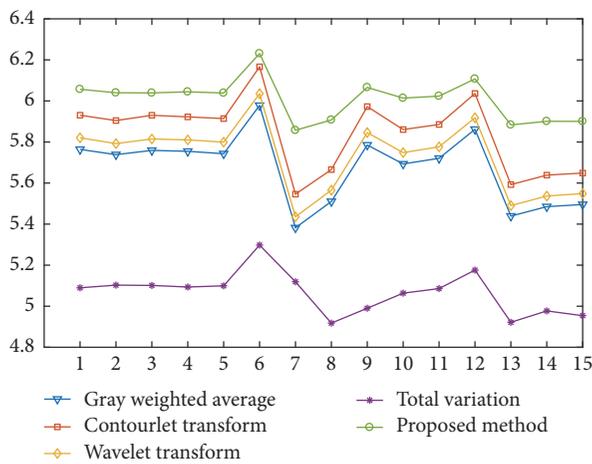
FIGURE 12: Quantitative evaluation of day images fusion.



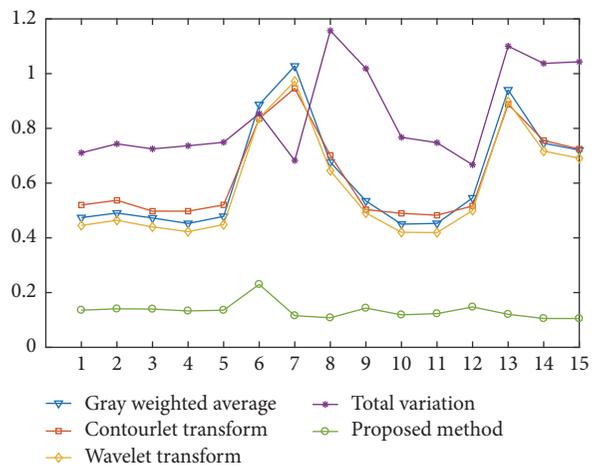
(a) Comparison of standard deviations



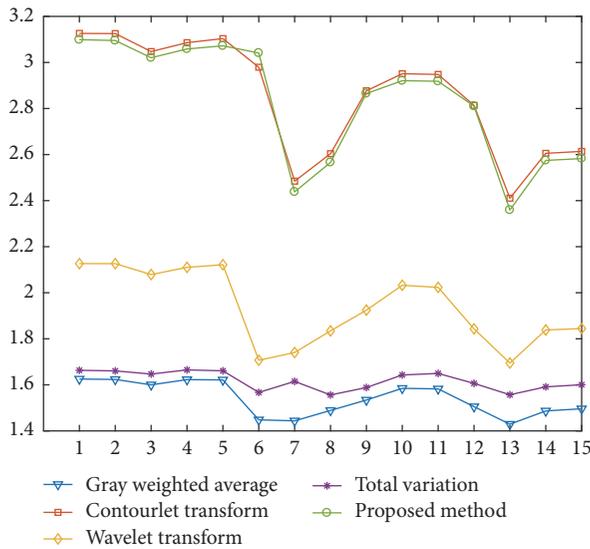
(b) Comparison of mutual information



(c) Comparison of information entropy



(d) Comparison of cross entropy



(e) Comparison of sharpness

FIGURE 13: Quantitative evaluation of night images fusion.

Data Availability

The data (including the images data) used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this paper.

Acknowledgments

This work is partially supported by the National Key Research and Development Program of China (2016YFB1200402), the Research and Development Plan of Chinese Railway Company (2017T001-B), and Chinese Scholarship Council (20170709507).

References

- [1] W. Zhong, H. Lu, and M.-H. Yang, "Robust object tracking via sparse collaborative appearance model," *IEEE Transactions on Image Processing*, vol. 23, no. 5, pp. 2356–2368, 2014.
- [2] Y. Xu, G. Yu, Y. Wang, X. Wu, and Y. Ma, "Car Detection from Low-Altitude UAV Imagery with the Faster R-CNN," *Journal of Advanced Transportation*, vol. 2017, Article ID 2823617, 10 pages, 2017.
- [3] S. Dugad, V. Puliyadi, H. Palod, N. Johnson, S. Rajput, and S. Johnny, "Ship Intrusion Detection Security System Using HoG and SVM," *International Journal of Advanced Research in Computer Engineering & Technology*, vol. 5, no. 10, pp. 2504–2507, 2016.
- [4] L. M. Dong, Q. X. Yang, H. Y. Wu, H. Xiao, and M. Xu, "High quality multi-spectral and panchromatic image fusion technologies based on Curvelet transform," *Neurocomputing*, vol. 159, pp. 268–274, 2015.
- [5] Y. Li, C. Tao, Y. Tan, K. Shang, and J. Tian, "Unsupervised Multi-layer Feature Learning for Satellite Image Scene Classification," *IEEE Geoscience and Remote Sensing Letters*, vol. 13, no. 2, pp. 157–161, 2016.
- [6] J. Zhao, Q. Zhou, Y. Chen, H. Feng, Z. Xu, and Q. Li, "Fusion of visible and infrared images using saliency analysis and detail preserving based image decomposition," *Infrared Physics & Technology*, vol. 56, pp. 93–99, 2013.
- [7] X. Qian, Y. Wang, and L. Han, "An object tracking method based on guided filter for night fusion image," *Infrared Physics & Technology*, vol. 74, pp. 38–43, 2016.
- [8] Z. Yu, L. Yan, N. Han, and J. Liu, "Image fusion algorithm based on contourlet transform and PCNN for detecting obstacles in forests," *Cybernetics and Information Technologies*, vol. 15, no. 1, pp. 116–125, 2015.
- [9] M. Ghantous and M. Bayoumi, "MIRF: A multimodal image registration and fusion module based on DT-CWT," *Journal of Signal Processing Systems*, vol. 71, no. 1, pp. 41–55, 2013.
- [10] Z. Song, S. Zhou, and J. Guan, "A novel image registration algorithm for remote sensing under affine transformation," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 52, no. 8, pp. 4895–4912, 2014.
- [11] Z. Xiong and Y. Zhang, "A critical review of image registration methods," *International Journal of Image and Data Fusion*, vol. 1, no. 2, pp. 137–158, 2010.
- [12] S.-E. Raza, V. Sanchez, G. Prince, J. P. Clarkson, and N. M. Rajpoot, "Registration of thermal and visible light images of diseased plants using silhouette extraction in the wavelet domain," *Pattern Recognition*, vol. 48, no. 7, pp. 2119–2128, 2015.
- [13] J. Y. Ma, J. Zhao, Y. Ma, and J. W. Tian, "Non-rigid visible and infrared face registration via regularized gaussian fields criterion," *Pattern Recognition*, vol. 48, no. 3, pp. 772–784, 2015.
- [14] G. A. Bilodeau, A. Torabi, and F. Morin, "Visible and infrared image registration using trajectories and composite foreground images," *Image and Vision Computing*, vol. 29, no. 1, pp. 41–50, 2011.
- [15] M. Gong, S. Zhao, L. Jiao, D. Tian, and S. Wang, "A novel coarse-to-fine scheme for automatic image registration based on SIFT and mutual information," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 52, no. 7, pp. 4328–4338, 2014.
- [16] J. Woo, M. Stone, and J. L. Prince, "Multimodal registration via mutual information incorporating geometric and spatial context," *IEEE Transactions on Image Processing*, vol. 24, no. 2, pp. 757–759, 2015.
- [17] J. Han, E. J. Pauwels, and P. De Zeeuw, "Visible and infrared image registration in man-made environments employing hybrid visual features," *Pattern Recognition Letters*, vol. 34, no. 1, pp. 42–51, 2013.
- [18] D. P. Bavirisetti and R. Dhuli, "Two-scale image fusion of visible and infrared images using saliency detection," *Infrared Physics & Technology*, vol. 76, pp. 52–64, 2016.
- [19] J. Ma, Z. Zhou, B. Wang, and H. Zong, "Infrared and visible image fusion based on visual saliency map and weighted least square optimization," *Infrared Physics & Technology*, vol. 82, pp. 8–17, 2017.
- [20] J. Ma, C. Chen, C. Li, and J. Huang, "Infrared and visible image fusion via gradient transfer and total variation minimization," *Information Fusion*, vol. 31, pp. 100–109, 2016.
- [21] Y. Ma, J. Chen, C. Chen, F. Fan, and J. Ma, "Infrared and visible image fusion using total variation model," *Neurocomputing*, vol. 202, pp. 12–19, 2016.
- [22] H. Li, L. Liu, W. Huang, and C. Yue, "An improved fusion algorithm for infrared and visible images based on multi-scale transform," *Infrared Physics & Technology*, vol. 74, pp. 28–37, 2016.
- [23] G. He, D. Dong, Z. Xia, S. Xing, and Y. Wei, "An Infrared and Visible Image Fusion Method Based on Non-Subsampled Contourlet Transform and Joint Sparse Representation," in *Proceedings of the 2016 IEEE International Conference on Internet of Things (iThings) and IEEE Green Computing and Communications (GreenCom) and IEEE Cyber, Physical and Social Computing (CPSCom) and IEEE Smart Data (SmartData)*, pp. 492–497, Chengdu, China, December 2016.
- [24] J. Cai, Q. Cheng, M. Peng, and Y. Song, "Fusion of infrared and visible images based on nonsubsampled contourlet transform and sparse K-SVD dictionary learning," *Infrared Physics & Technology*, vol. 82, pp. 85–95, 2017.
- [25] H. Li, H. Qiu, Z. Yu, and Y. Zhang, "Infrared and visible image fusion scheme based on NSCT and low-level visual features," *Infrared Physics & Technology*, vol. 76, pp. 174–184, 2016.
- [26] F. Meng, M. Song, B. Guo, R. Shi, and D. Shan, "Image fusion based on object region detection and Non-Subsampled Contourlet Transform," *Computers and Electrical Engineering*, vol. 62, pp. 375–383, 2017.

- [27] Z. Fu, X. Wang, J. Xu, N. Zhou, and Y. Zhao, "Infrared and visible images fusion based on RPCA and NSCT," *Infrared Physics & Technology*, vol. 77, pp. 114–123, 2016.
- [28] K. He, D. Zhou, X. Zhang, R. Nie, Q. Wang, and X. Jin, "Infrared and visible image fusion based on target extraction in the nonsubsampling contourlet transform domain," *Journal of Applied Remote Sensing*, vol. 11, no. 1, 2017.
- [29] X. Yang and K.-T. T. Cheng, "Local difference binary for ultrafast and distinctive feature description," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 36, no. 1, pp. 188–194, 2014.
- [30] D. G. Lowe, "Distinctive image features from scale-invariant keypoints," *International Journal of Computer Vision*, vol. 60, no. 2, pp. 91–110, 2004.
- [31] 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.
- [32] E. Rublee, V. Rabaud, K. Konolige, and G. Bradski, "ORB: an efficient alternative to SIFT or SURF" in *Proceedings of the IEEE International Conference on Computer Vision (ICCV '11)*, pp. 2564–2571, Barcelona, Spain, November 2011.
- [33] S. Leutenegger, M. Chli, and R. Y. Siegwart, "BRISK: binary robust invariant scalable keypoints," in *Proceedings of the IEEE International Conference on Computer Vision (ICCV '11)*, pp. 2548–2555, Barcelona, Spain, November 2011.
- [34] P. F. Alcantarilla, A. Bartoli, and A. J. Davison, "KAZE Features," in *Computer Vision – ECCV 2012*, vol. 7577 of *Lecture Notes in Computer Science*, pp. 214–227, Springer Berlin Heidelberg, Berlin, Heidelberg, 2012.

Research Article

Predicting Pedestrian Counts for Crossing Scenario Based on Fused Infrared-Visual Videos

Shize Huang , Wei Chen , Rongjie Yu , Xiaolu Yang , and Decun Dong 

Key Laboratory of Road and Traffic Engineering of Ministry of Education, Tongji University, 201804, China

Correspondence should be addressed to Rongjie Yu; yurongjie@tongji.edu.cn

Received 10 August 2018; Revised 26 October 2018; Accepted 27 November 2018; Published 4 December 2018

Academic Editor: Krzysztof Okarma

Copyright © 2018 Shize Huang 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.

Estimating the number of pedestrians based upon surveillance videos and images has been a critical task while implementing intelligent signal controls at intersections. However, this has been a difficult task considering the pedestrian waiting area is an outdoor scenario with complex and time-varying surrounding environment. In this study, a method for estimating pedestrian counts based on multisource video data has been proposed. First, the partial least squares regression (PLSR) model is developed to estimate the number of pedestrians from single-source video (either visible light video or infrared video). Meanwhile, the temporal feature of the scenario (daytime or nighttime) is identified based on visible light video. According to the recognized time periods, pedestrian count detection results from the visible light and infrared video data can be obtained with preset corresponding confidence levels. The empirical experiments showed that this fusion method based on environment perception holds the benefits of 24-hour monitoring for outdoor scenarios at the pedestrian waiting area and substantially improved accuracy of pedestrian counting.

1. Introduction

Estimating the number of pedestrians is critical within the intelligent transportation system. The pedestrian counts have been a vital input for intersection signal control [1], the guidance of passenger flow, and early warning of large-scale crowd gathering [2, 3]. However, the approach of estimating pedestrian counts under outdoor scenarios, such as the pedestrian waiting area, is still an unsolved challenge.

Generally, there are two main approaches to estimate the number of pedestrians. One kind is based on reliable tracking of individual pedestrians, which achieves the purpose of counting pedestrians through identifying each individual pedestrian based on image data [4–7]. However, this method is suitable for the case where the pedestrian density is low. If the pedestrian density is high and there is severe pedestrian overlapping, the performance of the method will be deteriorated. The other approach extracts feature from image data and applies regression analysis techniques to estimate the pedestrian counts rather than trying to identify each pedestrian in the image. This method is concluded to be more flexible since there is no need to track each pedestrian in the image.

The surveillance video data have been frequently adopted to estimate the number of pedestrians, which can be further divided into visible light video and infrared video. Infrared video is mostly used to determine whether there are people at scene or whether the target is a human being [8–11]. But, they were barely used for estimating the pedestrian counts. On the contrary, tremendous efforts have been investigated on the estimations of pedestrian counts using visible light video. For instance, Davies et al. [12] used geometric features such as areas and perimeters to estimate the number of pedestrians in the image. He [13] proposed a two-region learning algorithm, applying improved aggregate channel feature detection and Gaussian process regression to estimate the number of pedestrians. Chan [14] segmented the image, extracted the features of each segmentation region, and then used Gaussian process regression to learn the correspondence between the features and the number of pedestrians in each segment. Zhang [15] applied dimensionality reduction techniques to process high latitude features of images and performed regression analysis. Li [16] proposed a feature description operator combining wavelet transform and gray level cooccurrence matrix and used SVM to obtain the pedestrian density model.

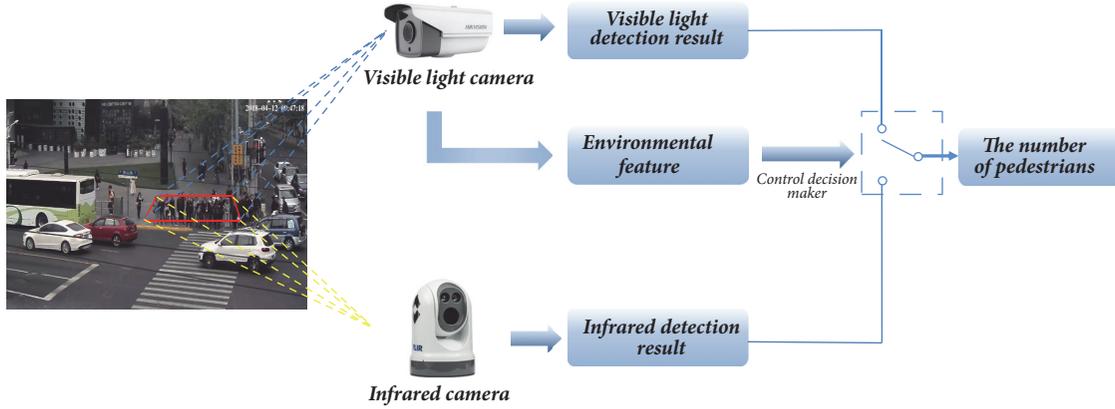


FIGURE 1: Pedestrian-counting framework.

Yan [17] used the simile classifier to optimize the subimage and then used the regression analysis model to establish the relationship between subimage blocks and the number of pedestrians. However, the abovementioned studies are based on visible light video which is sensitive to lighting conditions and cannot be implemented for monitoring the pedestrian waiting area for the whole day.

In this study, we propose a pedestrian number estimation method which is dependent on fusion of visible light video and infrared video based on environment perception, in order to realize 24-hour pedestrian counts detection for the pedestrian waiting area. First, partial least squares regression (PLSR) was employed to obtain the number of pedestrians from the image based upon visible light video and infrared video, respectively. Then, based on the environmental feature obtained from the visible light video, an information fusion model is established to obtain the number of pedestrians in the image. The specific schematic diagram is shown in Figure 1.

The remaining of this paper is organized as follows: in Section 2, we describe the image processing and how to extract features from images. Then, Section 3 describes the establishment of the pedestrian count estimation model and how to fuse the result of visible light detection with the result of infrared detection. And the report and analyses of the experimental results are given in Section 4 while Section 5 summarizes the work and discusses future directions.

2. Image Processing

In this section, visible light image processing and infrared image processing procedures are introduced correspondingly.

2.1. Visible Light Image Processing. The most important task in image processing is to extract the foreground of the motion from the image. For visible light images, background difference method was adopted to obtain the motion foreground in the image.

Since the background image would gradually change along the time in the actual scene, the background image

needs to be updated in real time. Kalman filter was used to update the background here. To be specific, the background image at the time t is determined by the background image at time $t - 1$ and the real-time image at time t , which includes both prediction and update. The forecast formula is as follows:

$$\begin{aligned}\widehat{B}(x, y, t) &= B(x, y, t - 1) \\ \widehat{P}(x, y, t) &= P(x, y, t - 1) + Q\end{aligned}\quad (1)$$

where $B(x, y, t - 1)$ is the background optimal value at time $t - 1$, $\widehat{B}(x, y, t)$ is the background prediction value at time t , $P(x, y, t - 1)$ is the covariance at time $t - 1$, $\widehat{P}(x, y, t)$ is the prediction of covariance at time t , and Q is the systematic process error.

The background update formula for time t is as follows:

$$\begin{aligned}k(x, y, t) &= \frac{\widehat{P}(x, y, t)}{(\widehat{P}(x, y, t) + R)} \\ B(x, y, t) &= \widehat{B}(x, y, t) \\ &\quad + k(x, y, t) [I_{vi}(x, y, t) - \widehat{B}(x, y, t)]\end{aligned}\quad (2)$$

$$P(x, y, t) = (1 - k(x, y, t)) \widehat{P}(x, y, t)$$

where R is the system measurement error, $k(x, y, t)$ is the system gain, $I_{vi}(x, y, t)$ is the gray image acquired by the visible light camera at time t , and $B(x, y, t)$ is the time t background. The optimal value $P(x, y, t)$ is the covariance at time t .

The visible light image $I_{vi}(x, y, t)$ is differentiated from the corresponding background image $B(x, y, t)$. The background difference result $F_{vi}(x, y, t)$ is

$$F_{vi}(x, y, t) = |I_{vi}(x, y, t) - B(x, y, t)|\quad (3)$$

Then, a binary region-of-interest (ROI) mask proposed by Chan [18] was applied to $F_{vi}(x, y, t)$, which not only reduces the amount of subsequent calculations, but also prevents some interference in noninterest areas. After applying the

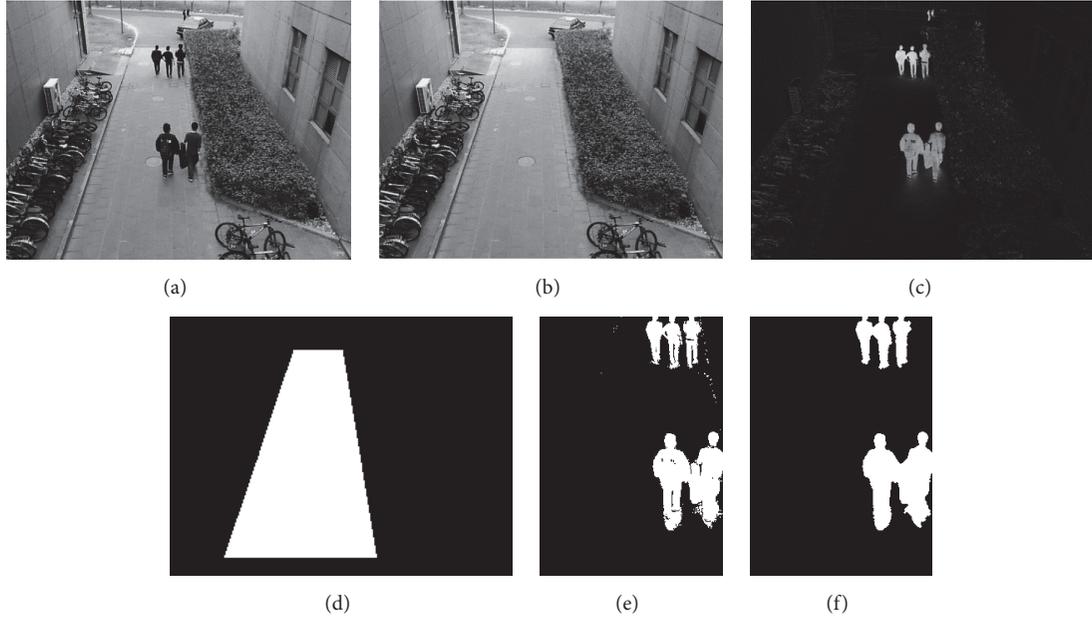


FIGURE 2: Visible light image processing (a) Original image. (b) Background image. (c) Background difference result. (d) ROI mask. (e) Foreground image. (f) Final result.

ROI mask, the binary foreground of visible light images $F'_{vi}(x, y, t)$ is calculated by

$$F'_{vi}(x, y, t) = \begin{cases} 1, & F_{vi}(x, y, t) > Th_{vi} \\ 0, & F_{vi}(x, y, t) \leq Th_{vi} \end{cases} \quad (4)$$

where Th_{vi} is the threshold used for binary processing. In our experiments, we set $Th_{vi} = 45$.

For the image $F'_{vi}(x, y, t)$, the closed operation (dilation followed by erosion operation) is to fill the small holes in the connected domain, connect adjacent objects, and smooth the boundary [19]. Then, it analyzes the connected domain and eliminates the connected domain with smaller area to remove noise [20]. The final result of visible light image processing is the image $M_{vi}(x, y, t)$. Then the set of blobs in $M_{vi}(x, y, t)$ is

$$B_{vi} = \{bvi_1, \dots, bvi_k, \dots, bvi_n\} \quad (5)$$

where bvi_k is the k -th blob in the image $M_{vi}(x, y, t)$ and n is the total number of blobs in the image $M_{vi}(x, y, t)$.

For example, Figures 2(a) and 2(b) are the original and background images, respectively. Figure 2(c) shows the background difference result $F_{vi}(x, y, t)$. Figure 2(d) is the ROI mask. Figure 2(f) is the final result $M_{vi}(x, y, t)$.

2.2. Infrared Image Processing. The infrared video data are imaged by thermal radiation, which is not sensitive to ambient light. Since the pedestrian generally appears as a highlighted area in the infrared image, we extract the foreground of the image by the gray value of the image. First, the projection images of the infrared images on the R, G, and B color channels are analyzed to find the projection image which has the greatest difference between pedestrians and

the surrounding environment. Figure 3 illustrates that, in the projection image on the G color channel, the characteristics of the pedestrian are the most prominent and easier to distinguish. This projection image is defined as the grayscale image $I_{in}(x, y, t)$.

With the application of the ROI mask, the binary foreground of infrared images $F'_{in}(x, y, t)$ is calculated by

$$F'_{in}(x, y, t) = \begin{cases} 1, & I_{in}(x, y, t) > Th_{in} \\ 0, & I_{in}(x, y, t) \leq Th_{in} \end{cases} \quad (6)$$

where Th_{in} is the threshold used for binary processing. In our experiments, we set $Th_{in} = 120$.

For the image $F'_{in}(x, y, t)$, the closed operation and connected domain analysis are also performed to remove the noise and ensure the integrity of the pedestrian. The final result of the infrared image is $M_{in}(x, y, t)$. Then the set of blobs B_{in} in $M_{in}(x, y, t)$ is

$$B_{in} = \{bin_1, \dots, bin_k, \dots, bin_n\} \quad (7)$$

where bin_k is the k -th blob in the image $M_{in}(x, y, t)$ and n is the total number of blobs in the image $M_{in}(x, y, t)$.

For example, Figure 4(a) is the image $I_{in}(x, y, t)$. Figure 4(b) is the ROI mask and Figure 4(d) is the final result $M_{in}(x, y, t)$.

2.3. Feature Extraction. Here the visible light image feature extraction procedure was taken as an example, while the feature extraction of infrared images is similar. The contained features of blobs and the inferred number of pedestrians were further extracted. Take the blob bvi_k as an example to calculate its geometric features and positional features using the following steps:

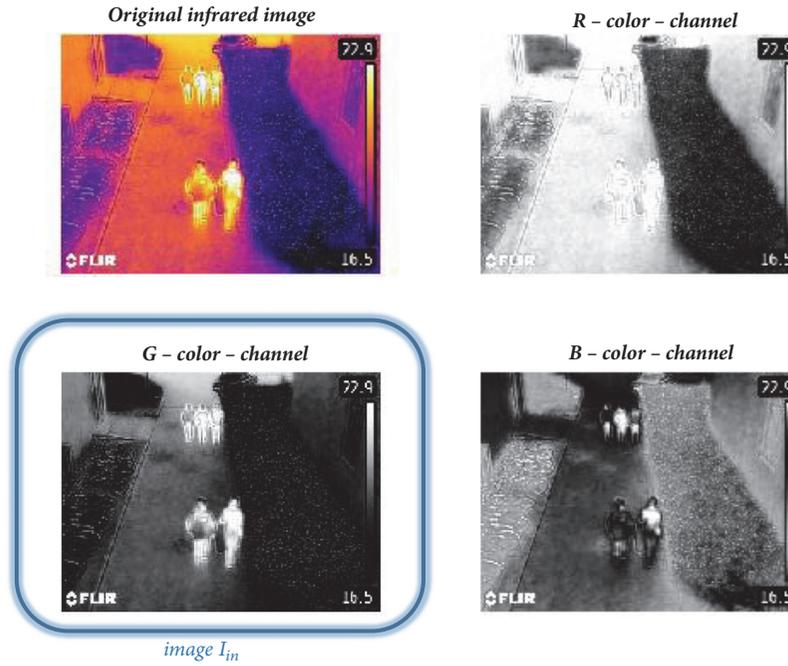


FIGURE 3: The RGB analysis of infrared images.

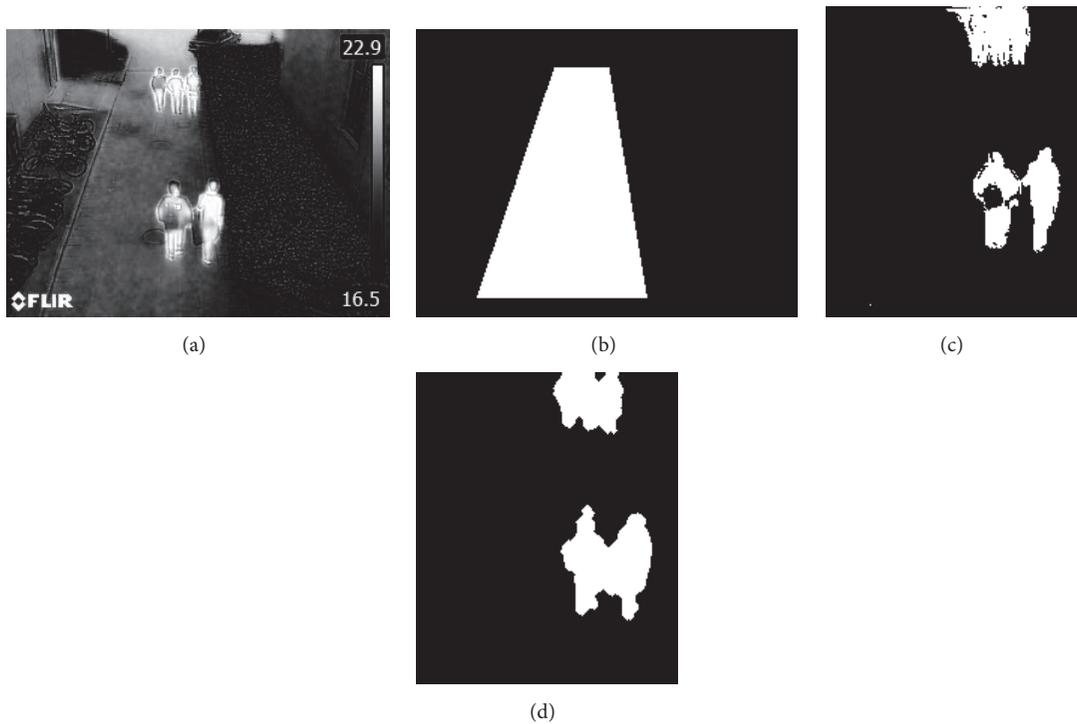


FIGURE 4: Infrared image processing (a) Original image. (b) ROI mask. (c) Foreground image. (d) Final result.

(1) Area A_k , which is the weighted sum of all pixels in the blob,

$$A_k = \sum_{x=1}^X \sum_{y=1}^Y bvi_k(x, y) \quad (8)$$

(2) Number of edge points EP_k , which is the weighted sum of pixels on the boundaries of the blob,

$$EP_k = \sum_{x=1}^X \sum_{y=1}^Y \varepsilon [bvi_k(x, y)] \quad (9)$$

where $\varepsilon[bvi_k(x, y)]$ denotes the edge image that is generated by the Sobel edge detector on the image bvi_k .

(3) Length of the spot L_k , which is the maximum number of pixels in the horizontal direction of the blob,

$$L_k = \max \left(\sum_{y=1}^Y bvi_k(1 : end, y) \right) \quad (10)$$

(4) Height of the spot H_k , which is the maximum number of pixels in the vertical direction of the blob,

$$H_k = \max \left(\sum_{x=1}^X bvi_k(x, 1 : end) \right) \quad (11)$$

(5) Horizontal position PX_k , which is the horizontal position of the center pixel of the blob in image M_{vi} (for infrared images it is M_{in}),

$$PX_k = \frac{\max(\sigma[bvi_k]) + \min(\sigma[bvi_k])}{2} \quad (12)$$

where $\sigma[bvi_k]$ denotes a horizontal position set of the pixels of the spot bvi_k in the image M_{vi} .

(6) Vertical position PY_k , which is the vertical position of the center pixel of the blob in image M_{vi} (for infrared images it is M_{in}),

$$PY_k = \frac{\max(\beta[bvi_k]) + \min(\beta[bvi_k])}{2} \quad (13)$$

where $\beta[bvi_k]$ denotes a vertical position set of the pixels of the spot bvi_k in the image M_{vi} .

Features A , EP , L , and H have strong correlations with pedestrian crowd density. In general, at the same position of the image, the larger the values of A , EP , L , and H , the more the number of pedestrians included in the blob. And the further away a pedestrian is from the camera lens, the smaller he is in the image. Therefore, we use position features PX and PY to record the positional relationship between pedestrians and the camera lens to ensure the accuracy of pedestrian counting.

Since the final decision result is based on visible light detection results and infrared detection results, an indicator was further introduced to selectively believe based on the distinct detection methods in different situations. In this study, the ambient brightness BR from the visible light image is the indicator.

$$BR_t = \frac{\sum_{x=1}^X \sum_{y=1}^Y I_{vi}(x, y, t)}{xy} \quad (14)$$

where BR_t denotes the ambient brightness at time t .

3. Model Establishment

This section focuses on how to infer the number of pedestrians from the extracted features. There are mainly two tasks being carried out: (1) a pedestrian count estimation model was developed based on the features of single-source video to establish; (2) then information fusion model was established based on the detection results of multisource video.

3.1. Pedestrian Count Estimation Model. In order to estimate the number of pedestrians in the blob and prevent the problem that overaggregated data might fail to reveal the true correlation between variables, we apply partial least squares regression (PLSR) [21, 22]. PLSR is a method for multivariate statistical analysis. It draws on the idea of extracting information from explanatory variables in principal component regression, and can effectively solve the multiple correlation problem between variables.

The independent variable $\{x_1, \dots, x_p\}$ contains p elements and the dependent variable $\{y_1, \dots, y_q\}$ contains q elements. In order to study the statistical relationships between the dependent variable and the independent variables, assume there are n sample observations, which constitute the independent variable set $X = [x_1, \dots, x_p]_{n \times p}$ and the dependent variable set $Y = [y_1, \dots, y_q]_{n \times q}$. The normalization results of X and Y are $E_0 = (E_{01}, \dots, E_{0p})_{n \times p}$ and $F_0 = (F_{01}, \dots, F_{0q})_{n \times q}$.

First, the main components are extracted in E_0 and F_0 . t_1 and u_1 are the first component of X and Y . Then t_1 and u_1 need to meet the following conditions:

$$\begin{aligned} t_1 &= E_0 w_1 & \|w_1\| \\ u_1 &= F_0 c_1 & \|c_1\| = 1 \\ \max_{w_1, c_1} \text{cov}(t_1, u_1) &= \text{cov}(E_0 w_1, F_0 c_1) & (15) \\ \text{s.t.} & \begin{cases} w_1' w_1 = 1 \\ c_1' c_1 = 1 \end{cases} & , = 1 \end{aligned}$$

where $\text{cov}(t_1, u_1)$ denotes the covariance between t_1 and u_1 .

After the first components t_1 and u_1 are extracted, the regression of X versus t_1 and the regression of Y versus u_1 are performed, respectively. If the regression equation has reached a satisfactory accuracy, the algorithm terminates; otherwise, the second round of component extraction will be performed using the residual information of X and Y . So reciprocate until a satisfactory accuracy is achieved. If we finally extract a total of m components t_1, \dots, t_m , PLSR will be implemented by implementing y_k regression of t_1, \dots, t_m and then expressed as y_k regression equations for the original variables x_1, \dots, x_p ($k = 1, 2, \dots, q$).

Take the visible light image $M_{vi}(x, y, t)$ as an example. Based on PLSR, we establish the pedestrian estimation model where the feature set $\{A_k, EP_k, L_k, H_k, PX_k, PY_k\}$ of the blob bvi_k is an input and the number of pedestrians PN_k included in the spot bvi_k is an output.

$$PN_k = f(A_k, EP_k, L_k, H_k, PX_k, PY_k) \quad (16)$$

Based on the above model, the number of pedestrians included in each blob in the image $M_{vi}(x, y, t)$ is calculated. The total number of pedestrians PN_{vi} in the image $M_{vi}(x, y, t)$ is

$$PN_{vi} = \sum_{k=1}^n [PN_k + 0.5] \quad (17)$$

where $[PN_k + 0.5]$ denotes rounding of PN_k .



FIGURE 5: Examples of experimental data.

PN_{vi} is the detection result of visible light. Using the same method, we can get the infrared detection result PN_{in} .

3.2. Information Fusion Model. The environment of outdoor scenarios like the pedestrian waiting area varies substantially along the daytime due to the lighting conditions, temperature, etc. In order to ensure the accuracy of the pedestrian count estimations, a method of combining the visible light detection result with the infrared detection result was proposed with its advantages of applying feasibility in different scenarios. First, the current scenario (day or night) is identified based on the ambient brightness BR_t obtained above. Then, according to the recognition result of the scenario, a corresponding confidence level is set for the detection result of the visible light and the detection result of the infrared. In the case of good daylight and good light, we believe the detection result of visible light; otherwise we believe the detection result of infrared. Therefore, the information fusion result PN_t at time t is

$$\begin{aligned}
 PN_t &= \alpha_{vi}PN_{vi} + \alpha_{in}PN_{in} \\
 \text{s.t. } \alpha_{vi} &= \begin{cases} 0, & BR_t < Th_{br} \\ 1, & \text{else,} \end{cases} \\
 \alpha_{in} &= \begin{cases} 1, & BR_t < Th_{br} \\ 0, & \text{else} \end{cases}
 \end{aligned} \quad (18)$$

where α_{vi} is the confidence level of visible light detection result and α_{in} is the confidence level of infrared detection result. Th_{br} is the environment segmentation threshold and we set it $Th_{br} = 75$.

4. Empirical Analysis

The empirical analysis was conducted at the campus of Tongji University. A total of 106 groups of daytime images and 18 groups of night images (as shown in Figure 5) were collected. The visible image is 640×480 pixels, and the infrared image

is 320×240 pixels. This section uses 8-fold cross validation to divide the image set into a training set and a test set and then to check the accuracy of the proposed method.

4.1. Daytime Scenario. For the subset of daytime images, the visible light detection results are shown in Table 1 and the infrared detection results are shown in Table 2. Figure 6 is a schematic diagram of information fusion in a daytime scenario. It can be seen from Figure 6 that the visible light image is clearer and the noise in the processing result of the visible light image is smaller. This is because the resolution of the visible light image is higher than that of the infrared image. Therefore, the result of information fusion gives credibility to the detection result of visible light, which is consistent with the actual situation.

4.2. Nighttime Scenario. For a group of night images, since there are no street lights near the experimental site, this would cause the visible light detection complete failure. Therefore, the visible light detection result is 0. The infrared detection results are shown in Table 3. Figure 7 is a schematic diagram of information fusion in a night scenario. Since the ambient brightness at this time is very low, the result of the information fusion is selected to believe the infrared detection result, which is consistent with the actual situation.

4.3. Influence of Thresholds Th_{vi} and Th_{in} . The thresholds Th_{vi} and Th_{in} are key parameters in this study, which were used to distinguish pedestrians from the background in the image. If Th_{vi} and Th_{in} are too large, a large number of pixels representing the pedestrians in the image will be misjudged as the background, which will result in incomplete motion foreground. As a consequence, the final pedestrian count result will be small. If Th_{vi} and Th_{in} are too small, a large number of pixels representing the background in the image will be misjudged as pedestrians, so that the foreground of the motion will contain a lot of noise. And the final pedestrian count result will be large.

Here different thresholds Th_{vi} were performed as an example, and the threshold Th_{in} is similar. For the same

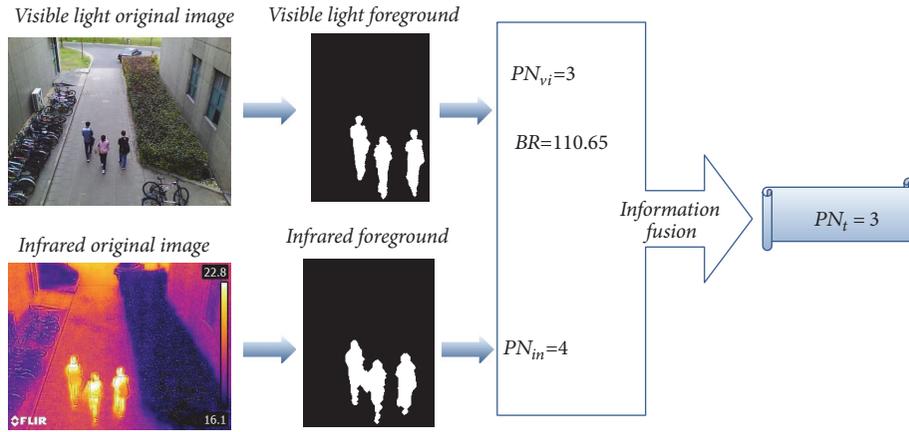


FIGURE 6: Schematic diagram of information fusion in a daytime scenario.

TABLE 1: Visible light detection results in a daytime scenario.

| Blob number k | Input | | | | | | Output | |
|--|-------|--------|-------|-------|--------|--------|---------|-----------------|
| | A_k | EP_k | L_k | H_k | PX_k | PY_k | PN_k | Rounded results |
| 1 | 2478 | 271 | 38 | 120 | 213.5 | 92.5 | 0.97255 | 1 |
| 2 | 2295 | 261 | 35 | 117 | 250 | 135 | 0.76629 | 1 |
| 3 | 2581 | 279 | 37 | 126 | 241.5 | 194 | 0.80096 | 1 |
| Visible light detection result PN_{vi} | | | | | | | | 3 |

TABLE 2: Infrared detection results in a daytime scenario.

| Blob number k | Input | | | | | | Output | |
|---|-------|--------|-------|-------|--------|--------|---------|-----------------|
| | A_k | EP_k | L_k | H_k | PX_k | PY_k | PN_k | Rounded results |
| 1 | 2504 | 343 | 55 | 103 | 152 | 84 | 2.5982 | 3 |
| 2 | 1345 | 173 | 28 | 76 | 154.5 | 136.5 | 0.78543 | 1 |
| Infrared detection result PN_{in} | | | | | | | | 4 |

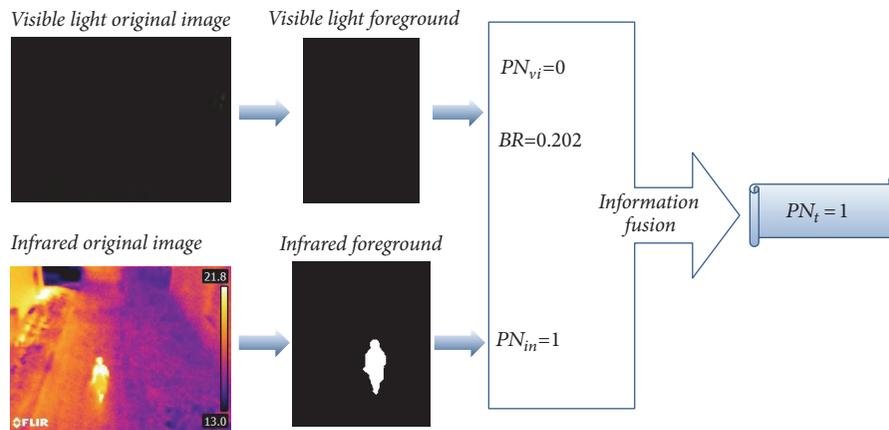


FIGURE 7: Schematic diagram of information fusion in a night scenario.

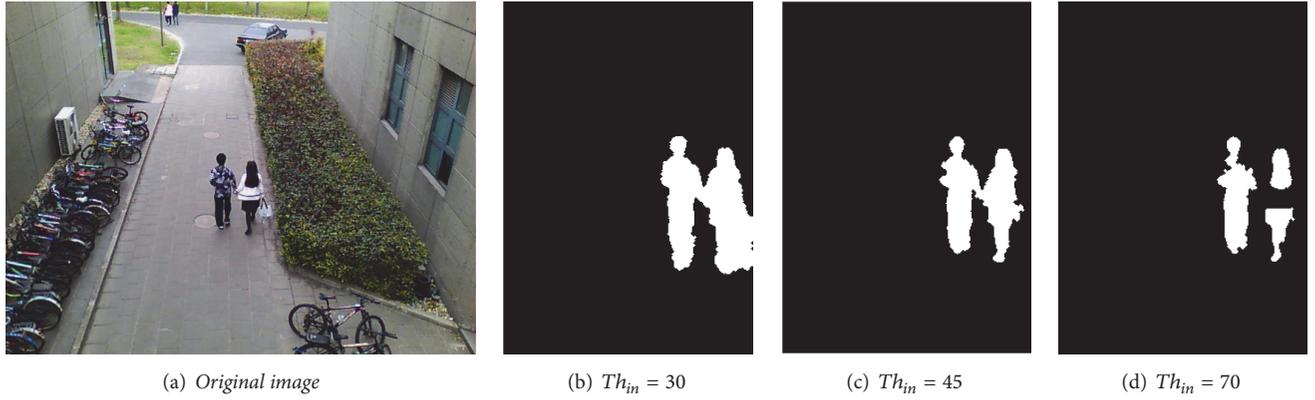
FIGURE 8: The results of motion foreground extraction with different threshold Th_{vi} .

TABLE 3: Infrared detection results in a nighttime scenario.

| Blob number k | Input | | | | | | | Output |
|---|-------|--------|-------|-------|--------|--------|---------|-----------------|
| | A_k | EP_k | L_k | H_k | PX_k | PY_k | PN_k | Rounded results |
| 1 | 1159 | 154 | 29 | 69 | 133 | 104 | 0.91752 | 1 |
| Infrared detection result PN_{in} | | | | | | | | 1 |

TABLE 4: The results of pedestrian counting with different threshold Th_{vi} .

| Th_{vi} | Blob number k | Input | | | | | | Output | Detection result | |
|-----------|-----------------|-------|--------|-------|-------|--------|--------|--------|------------------|-----------------|
| | | A_k | EP_k | L_k | H_k | PX_k | PY_k | PN_k | | Rounded results |
| 30 | 1 | 5840 | 534 | 82 | 122 | 181.5 | 180.5 | 2.73 | 3 | 3 |
| 45 | 2 | 4045 | 517 | 76 | 112 | 176.5 | 177.5 | 2.37 | 2 | 2 |
| 70 | 1 | 1864 | 263 | 36 | 104 | 172.5 | 158.5 | 1.04 | 1 | 1 |
| | 2 | 613 | 130 | 26 | 48 | 207.5 | 196.5 | 0.22 | 0 | |
| | 3 | 521 | 91 | 20 | 36 | 149.5 | 197.5 | 0.36 | 0 | |

visible image, we set the threshold Th_{vi} to 30, 45, and 70, respectively. The results of the extraction of the motion foreground are shown in Figure 8 and the results of the pedestrian detection are shown in Table 4. According to Figure 8 and Table 4, we can find that when the threshold Th_{vi} is too small ($Th_{vi}=30$), the motion foreground contains more noise, and the final pedestrian count result is too large. When threshold Th_{vi} is too large ($Th_{vi}=70$), the motion foreground is incomplete and the final pedestrian count results are small. Therefore, the thresholds Th_{vi} and Th_{in} need to be set according to the characteristics of the data and the actual situation.

4.4. Contribution of the Features. The method in this paper is based on six features (Section 2.3 Feature Extraction). In order to evaluate the contribution of these features to the final result, the average elastic coefficient is introduced. The bigger the average elastic coefficient of the feature, the greater contribution to the final result. And the average elastic coefficient \bar{E} is

$$\bar{E} = \frac{\Delta y / \bar{y}}{\Delta x / \bar{x}} \quad (19)$$

where \bar{x} is the average of the independent variables and \bar{y} is the average of the dependent variables.

In the visible light model and the infrared model, the average elastic coefficient of each feature is calculated separately. The calculation results are shown in Figure 9. We have found that the feature PX is the most influential feature of the final result in both the visible light model and the infrared model, because PX is the most important parameter to represent the distance from the pedestrian to the lens in the testing scenario of this paper. On the other hand, the features A , EP , and L are reasonable predictors of crowd density, which reflects the number of pedestrians from different angles. One possible explanation for the low contribution of features H and PY is that the camera's field of view is parallel to the road, not vertical or oblique in the testing scenario of this paper.

4.5. The Efficiency of Background Update. Visible light video detection is based on background differences to obtain motion foreground. Since the environment around the pedestrian waiting area varies greatly in a day, real-time background update is a must. Here, the efficiency of the background update method based on Kalman filter is tested. Three rounds of tests were performed on 102 images. The

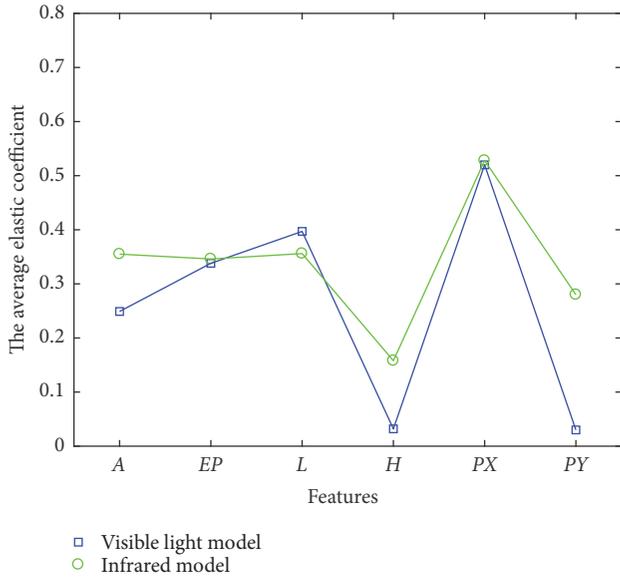


FIGURE 9: The calculation results of average elastic coefficient.

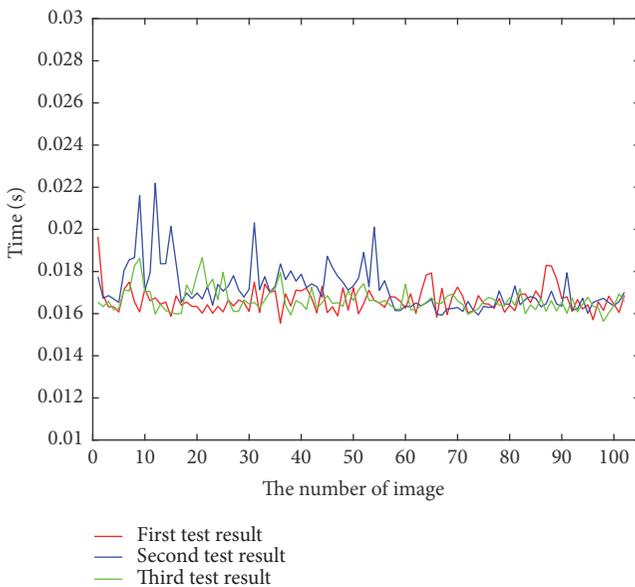


FIGURE 10: The test results of background update efficiency.

results are shown in Figure 10 and Table 5. (Note: this test was performed on a laptop and the test software is MATLAB 2017b.)

According to Table 5 and Figure 10, it can be found that the average time of background update is about 0.0168s. The result is ideal and can meet the needs of practical applications.

4.6. Accuracy Verification. The accuracy of the proposed method is verified based on 8-fold cross validation. 124 images are randomly divided into 8 groups. Each group is in turn used as a test set for the model, and the remaining 7 groups serve as a training set for the model. The experimental results are shown in Figure 11 and Table 6. As can be seen from Figure 11, the accuracy of the individual visible light detection

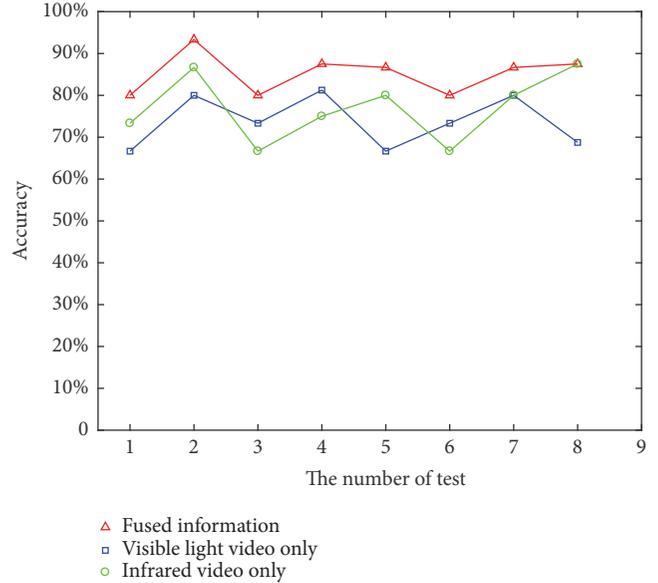


FIGURE 11: The results of 8-fold cross validation.

TABLE 5: The efficiency of background update.

| Number | Total time/s | Average time/s |
|--------|--------------|----------------|
| 1 | 1.6985 | 0.0167 |
| 2 | 1.7535 | 0.0172 |
| 3 | 1.6967 | 0.0166 |

is sometimes higher than that of the individual infrared detection, and sometimes lower. However, the accuracy of information fusion detection is always the highest. Combined with Table 6, the average accuracy of information fusion detection is higher than that of the individual visible light and infrared detections, while considering both daytime scenarios and nighttime scenarios.

Moreover, since there is no public dataset containing both infrared and visible light images, we test other methods on the dataset of this paper to show the advantage of the proposed method. Table 7 listed the prediction accuracy comparisons. It can be seen that the fusion method could provide better performance with lower MSE and higher accuracy as compared to the existing methods. Therefore, for 24-hour pedestrian counting in outdoor scenarios, the fusion method between visible light video and infrared video from the perspective of environment perception is more effective than the single video (visual videos or infrared videos).

5. Conclusion

In this study, a fused method between visible light video and infrared video based on environment perception for estimating the number of pedestrians has been proposed. And the method is intended to combine visual light information with infrared information to enable pedestrian counting techniques for complex outdoor scenarios. The proposed

TABLE 6: The results of 8-fold cross validation.

| The number of test | Visible light video only | Infrared video only | Fused information |
|-------------------------|--------------------------|---------------------|-------------------|
| 1 | 66.67% | 73.33% | 80.00% |
| 2 | 80.00% | 86.67% | 93.33% |
| 3 | 73.33% | 66.67% | 80.00% |
| 4 | 81.25% | 75.00% | 87.50% |
| 5 | 66.67% | 80.00% | 86.67% |
| 6 | 73.33% | 66.67% | 80.00% |
| 7 | 80.00% | 80.00% | 86.67% |
| 8 | 68.75% | 87.50% | 87.50% |
| Average accuracy | 73.75% | 76.98% | 85.21% |

TABLE 7: The comparison of MSE and accuracy.

| Paper | MSE | Accuracy |
|-------------------------------|-------|----------|
| [12] | 0.726 | 62.50% |
| [15] | 0.412 | 76.88% |
| Proposed method in this study | 0.246 | 85.21% |

approach is depending on two aspects: the estimation of the number of pedestrians based on single-source video and the information fusion based on multisource detection results. First, PLSR was applied to combine the dimensionality reduction analysis with the regression analysis to establish the pedestrian number estimation model based on single-source video. The method holds the advantages of reducing the redundancy of the data in the feature set and effectively solving the multiple correlations between variables. Meanwhile, the ambient brightness was employed to identify the scene of images and integrate the visible light detection result and the infrared detection result. The empirical analyses showed that, for 24-hour pedestrian counting in outdoor scenarios, the proposed method has better performance than the method using single information source, which expands the application scenario of pedestrian counting and provides reference for relevant research.

As for future analyses, one thing that needs to be expanded is the sample size of the empirical analyses and test the feasibility of utilizing deep learning networks to identify different scenarios (day, night, rain, fog, etc.). Besides, being under heavy fog or rain conditions will substantially increase the noise of video, and how to reduce the interference of these noises on pedestrian count would be a challenging issue in the future to be investigated. In addition, continued improvements of the information fusion model and the feasibility of employing new sensing equipment (such as laser scanners) to estimate the number of pedestrians will be tested.

Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

This research is supported by National Key R&D Program of China (2016YFB1200402), National Natural Science Foundation of China (61703308; 71771174), and the Fundamental Research Funds for the Central Universities.

References

- [1] K. H. Guan, *An adaptive optimization method of pedestrian crossing time based on video detection*, Jilin University, 2015.
- [2] J. Zhang, F.-Y. Wang, K. Wang, W.-H. Lin, X. Xu, and C. Chen, "Data-driven intelligent transportation systems: a survey," *IEEE Transactions on Intelligent Transportation Systems*, vol. 12, no. 4, pp. 1624–1639, 2011.
- [3] Y.-C. Chiu and P. B. Mirchandani, "Online behavior-robust feedback information routing strategy for mass evacuation," *IEEE Transactions on Intelligent Transportation Systems*, vol. 9, no. 2, pp. 264–274, 2008.
- [4] P. Viola and M. Jones, "Fast and Robust Classification using Asymmetric AdaBoost and a Detector Cascade," *Advances in Neural Information Processing Systems*, vol. 14, pp. 1311–1318, 2001.
- [5] 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.
- [6] P. Szabzmeydani and G. Mori, "Detecting pedestrians by learning shapelet features," in *Proceedings of the 2007 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, CVPR'07*, USA, June 2007.
- [7] 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.
- [8] Y. Fang, K. Yamada, Y. Ninomiya, B. K. P. Horn, and I. Masaki, "A shape-independent method for pedestrian detection with far-infrared images," *IEEE Transactions on Vehicular Technology*, vol. 53, no. 6, pp. 1679–1697, 2004.
- [9] M. Bertozzi, A. Broggi, P. Grisleri, T. Graf, and M. Meinecke, "Pedestrian detection in infrared images," in *Proceedings of the*

- 2003 *IEEE Intelligent Vehicles Symposium, IV 2003*, pp. 662–667, USA, June 2003.
- [10] R. O'Malley, E. Jones, and M. Glavin, "Detection of pedestrians in far-infrared automotive night vision using region-growing and clothing distortion compensation," *Infrared Physics & Technology*, vol. 53, no. 6, pp. 439–449, 2010.
- [11] S. Biswas K and P. Milanfar, "Linear Support Tensor Machine: Pedestrian Detection in Thermal Infrared Images," *IEEE Transactions on Image Processing*, p. 99, 2016.
- [12] A. C. Davies, J. H. Yin, and S. A. Velastin, "Crowd monitoring using image processing," *Electronics & Communication Engineering Journal*, vol. 7, no. 1, pp. 37–47, 1995.
- [13] G. He, Q. Chen, D. Jiang, X. Lu, and Y. Yuan, "A double-region learning algorithm for counting the number of pedestrians in subway surveillance videos," *Engineering Applications of Artificial Intelligence*, vol. 64, pp. 302–314, 2017.
- [14] A. B. Chan, Z. S. J. Liang, and N. Vasconcelos, "Privacy preserving crowd monitoring: Counting people without people models or tracking," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2008. CVPR 2008*, pp. 1–8, 2008.
- [15] J. Zhang, B. Tan, F. Sha, and L. He, "Predicting pedestrian counts in crowded scenes with rich and high-dimensional features," *IEEE Transactions on Intelligent Transportation Systems*, vol. 12, no. 4, pp. 1037–1046, 2011.
- [16] Y. Li, G. J. Wang, and H. G. Lin, "Crowd density estimation algorithm combining local and global features," *Journal of Tsinghua University (Science and Technology)*, vol. 4, pp. 542–545, 2013.
- [17] Q. Xinhui, W. Xiufei, and X. Zhou, "Counting people in various crowd density scenes using support vector regression," *Journal of Image and Graphics*, vol. 18, no. 4, pp. 392–398, 2013.
- [18] A. B. Chan, Z. S. J. Liang, and N. Vasconcelos, "Privacy preserving crowd monitoring: Counting people without people models or tracking," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 1–7, 2008.
- [19] Z. W. Kim, "Real time object tracking based on dynamic feature grouping with background subtraction," in *Proceeding of the 26th IEEE Conference on Computer Vision and Pattern Recognition (CVPR '08)*, pp. 1–8, Anchorage, AL, USA, June 2008.
- [20] K. Suzuki, I. Horiba, and N. Sugie, "Linear-time connected-component labeling based on sequential local operations," *Computer Vision and Image Understanding*, vol. 89, no. 1, pp. 1–23, 2003.
- [21] R. Yu, M. Quddus, X. Wang et al., "Impact of data aggregation approaches on the relationships between operating speed and traffic safety," *Accident Analysis & Prevention*, vol. 120, pp. 304–310, 2018.
- [22] S. D. Jong and A. Phatak, "Partial least squares regression," in *Proceedings of the International Workshop on Recent Advances in Total Least Squares Techniques and Errors-In-Variables Modeling*, pp. 25–36, Society for Industrial and Applied Mathematics, 1997.

Research Article

Neuromorphic Vision Based Multivehicle Detection and Tracking for Intelligent Transportation System

Guang Chen ^{1,2}, Hu Cao,³ Muhammad Aafaque,² Jieneng Chen,⁴ Canbo Ye,¹ Florian Röhrbein,² Jörg Conradt,⁵ Kai Chen,¹ Zhenshan Bing,² Xingbo Liu,¹ Gereon Hinz,² Walter Stechele ⁶ and Alois Knoll²

¹College of Automotive Engineering, Tongji University, China

²Robotics and Embedded Systems, Technische Universität München, Germany

³State Key Laboratory of Advanced Design and Manufacturing for Vehicle Body, Hunan University, China

⁴College of Electronics and Information Engineering, Tongji University, China

⁵Department of Computational Science and Technology, KTH Royal Institute of Technology, Sweden

⁶Integrated Systems, Technische Universität München, Germany

Correspondence should be addressed to Guang Chen; guang@in.tum.de

Received 10 August 2018; Revised 1 October 2018; Accepted 6 November 2018; Published 2 December 2018

Academic Editor: Krzysztof Okarma

Copyright © 2018 Guang Chen 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.

Neuromorphic vision sensor is a new passive sensing modality and a frameless sensor with a number of advantages over traditional cameras. Instead of wastefully sending entire images at fixed frame rate, neuromorphic vision sensor only transmits the local pixel-level changes caused by the movement in a scene *at the time they occur*. This results in advantageous characteristics, in terms of low energy consumption, high dynamic range, sparse event stream, and low response latency, which can be very useful in intelligent perception systems for modern intelligent transportation system (ITS) that requires efficient wireless data communication and low power embedded computing resources. In this paper, we propose the first neuromorphic vision based multivehicle detection and tracking system in ITS. The performance of the system is evaluated with a dataset recorded by a neuromorphic vision sensor mounted on a highway bridge. We performed a preliminary multivehicle tracking-by-clustering study using three classical clustering approaches and four tracking approaches. Our experiment results indicate that, by making full use of the low latency and sparse event stream, we could easily integrate an online tracking-by-clustering system running at a high frame rate, which far exceeds the real-time capabilities of traditional frame-based cameras. If the accuracy is prioritized, the tracking task can also be performed robustly at a relatively high rate with different combinations of algorithms. We also provide our dataset and evaluation approaches serving as the first neuromorphic benchmark in ITS and hopefully can motivate further research on neuromorphic vision sensors for ITS solutions.

1. Introduction

Neuromorphic vision sensors, inspired by biological vision, use an event-driven frameless approach to capture transients in visual scenes. In contrast to conventional cameras, neuromorphic vision sensors only transmit local pixel-level changes (called “events”) caused by movement in a scene *at the time of occurrence* and provide an information rich stream of events with a latency within tens of microseconds. To be specific, a single event is a tuple (t, x, y, p) , where x, y are the

pixel coordinates of the event in 2D space, t is the time-stamp of the event, and p is the polarity of the event, which is the sign of the brightness change (increasing or decreasing). Furthermore, the requirements for data storage and computational resources are drastically reduced due to the sparse nature of the event stream. Apart from the low latency and high storage efficiency, neuromorphic vision sensors also enjoy a high dynamic range of 120 dB. In combination, these properties of neuromorphic vision sensors inspire entirely new designs of intelligent transportation systems. In order

to elucidate the mechanism of neuromorphic sensors more clearly, a comparison between standard frame-based cameras and neuromorphic vision sensors is shown in Figure 1.

Traditionally, frame-based vision sensors serve as the main information sources for vision perception tasks of ITS, which results in well-known challenges, such as the limited real-time performance and substantial computational costs. The key problem lies in the fact that conventional cameras sample their environment with a fixed frequency and produce a series of frames, which actually contain enormous amounts of redundant information but lost all the information between two adjacent frames. Hence, traditional vision sensors waste memory access, energy, computational power, and time on the one hand and also discard significant information between continuous frames on the other hand. These properties bring about great limitations on its applications. For an intelligent transportation system equipped with conventional cameras, appearance feature extraction based on learning methods is the major strategy of environment perception tasks, which is acknowledged to be computationally demanding [3]. Moreover, in order to get good detection and tracking performance, large amounts of labeled data as well as dedicated and expensive hardware such as GPU are indispensable for the training and learning process.

In this paper, a novel approach for the tracking system of the intelligent transportation systems (ITS) is proposed based on the neuromorphic vision sensor. And we will publish our dataset and evaluation approaches as well, aiming to provide the first neuromorphic benchmark in ITS and motivate further research on neuromorphic vision sensors for ITS solutions. To fully demonstrate the feasibility and potential of the approach, different detection and tracking algorithms are presented and compared in this paper. In detection stage, we utilize and evaluate three classical clustering approaches: mean-shift clustering (MeanShift) [4], density based spatial clustering of applications with noise (DBSCAN) [5], and WaveCluster [6]. In terms of tracking stage, we carry out online multitarget tracking via four different algorithms: simple online and real-time tracking (SORT) [7], the Gaussian mixture probability hypothesis density filter (GM-PHD) [8], the cardinalized probability hypothesis density filter (GM-CPHD) [9], and probabilistic data association filter (PDAF) [10].

In combination, we propose the first neuromorphic vision based multivehicle detection and tracking system in ITS, with the unique properties of neuromorphic vision sensors mentioned above. The performance of the system is evaluated with a dataset recorded by a neuromorphic vision sensor mounted on a highway bridge. According to the experiment results, the tracking-by-clustering system can run at a rate of more than 110Hz, which far exceeds the real-time performance of traditional frame-based cameras. With priority given to accuracy, the tracking task can also be performed more robustly and precisely using different algorithm combinations. This work is extended from a conference paper which is published on the Joint German/Austrian Conference on Artificial Intelligence, 2017 [11]. We extended it from 4 aspects. First, we extend the testing data sequences

to 3 sequences for the experiment section. Second, we evaluate 3 detection-by-clustering approaches instead of 2 in [11]. Third, we extend to evaluate 4 tracking approaches instead of 1 in [11]. Finally, based on these differences, we analyze the results in different views with new outcomes [11].

The rest of this paper is organized as follows. In Section 2, we list the related work in the context of previous multivehicle detection and tracking methods. In Section 3, we introduce the variety of neuromorphic vision sensors and dataset. The algorithms utilized for detection and tracking are illustrated, respectively, in Section 4. The experiment results are analyzed and discussed in Section 5. In Section 6, we draw the conclusion and point out the possible further work.

2. Related Work

In the past decade, detecting and tracking multiple vehicles in traffic scenes for traffic surveillance, traffic control, and road traffic information systems is an emerging research area for intelligent transport systems [12–15]. Most of the existing vehicle tracking systems are based on the video cameras [16]. Previous approaches of vision based multiple vehicles detection and tracking could be subdivided into four categories: frames difference and motion based methods [17–19], background subtraction methods [15, 20], and feature based methods [21, 22]. Meanwhile, a few camera-based datasets for vehicle detection and tracking come to light in recent years [23–25], which promote the research for ITS.

All previous multivehicle detection and tracking methods leverage images acquired by traditional frame-based cameras. Conventional cameras may suffer from various motion-related issues (motion blur, rolling shutter, etc.) which may impact performance for high-speed vehicles detection and tracking. Neuromorphic vision sensors are widely applied to robotics [26–29] and vehicles [30–32]. A few relevant neuromorphic vision datasets [33, 34] have been released in recent years, which facilitate the neuromorphic vision application for object detection and tracking. Recent years also witness the various applications for detection and tracking tasks with neuromorphic vision sensor such as feature tracking [35, 36], line tracking [37], and microparticle tracking [38].

However, there is still a lack of neuromorphic datasets and relevant applications with neuromorphic vision sensors in intelligent transport system, albeit such sensors instinctively enjoy superiority in high-speed motion recording, which can correspondingly facilitates the high-speed multiple vehicle detection and tracking in ITS systems. Thus, it is meaningful to apply neuromorphic vision techniques to ITS systems.

3. Neuromorphic Vision Sensor and Dataset

3.1. Neuromorphic Vision Sensor. A short description of different versions of neuromorphic vision sensors is provided in this section which is also mentioned in [11]. The purpose is to encourage researchers who are not familiar with neuromorphic vision sensors to explore the potential applications

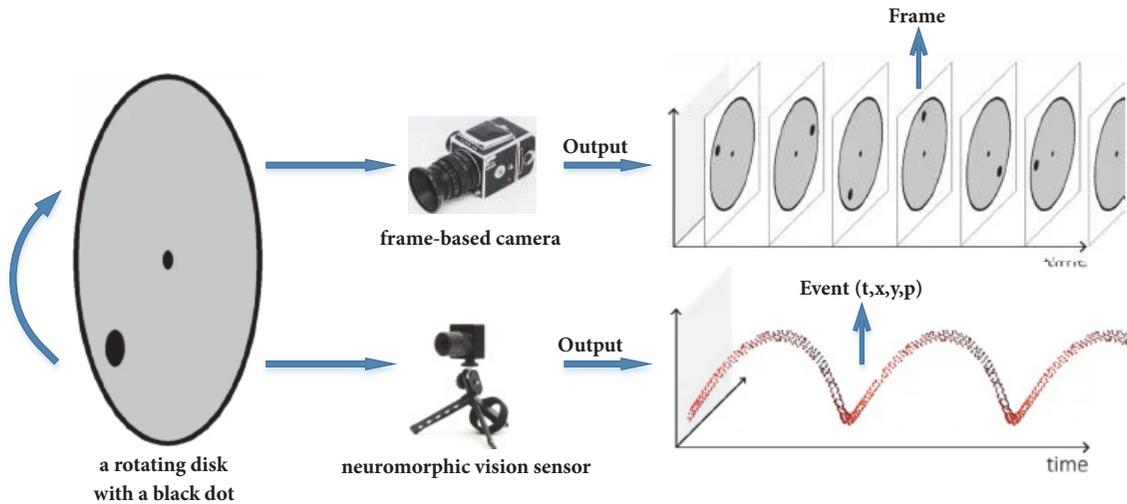


FIGURE 1: Visualization of the output from a neuromorphic vision sensor and a standard frame-based camera when facing a rotating disk with a black dot. Comparing to conventional frame-based camera which transmitted complete images at fixed latency, the neuromorphic vision sensor [1] emitted events individually and asynchronously at the time they occur. This figure is adopted from [2].

in the intelligent system. Figure 2 shows different versions of neuromorphic vision sensors.

Dynamic Vision Sensor (DVS). Dynamic Vision Sensors (DVS) are a new generation of cameras that are sensitive to intensity change, more specifically, to intensity logarithmic change. A DVS pixel typically generates one to four events (spikes) when an edge crosses it. DVS output consists of a continuous flow of events (spikes) in time, each with submicrosecond time resolution, representing the observed moving reality as it changes, without waiting to assemble or scan artificial time-constrained frames (images).

Embedded Dynamic Vision Sensor (eDVS). For embedded systems in mobile robotics such as unmanned aerial vehicle, an USB interface to transmit raw events is not desirable, nor is a desktop PC for event processing acceptable. For this purpose, a small embedded DVS (eDVS) is developed consisted of a DVS chip and a compact 64MHz 32bit microcontroller directly connected to the DVS chip.

Miniature Embedded Dynamic Vision Sensor (meDVS). The miniaturized version of the eDVS(meDVS) has minimum size (18cm×18cm) and lightest weight (2.2g) of DVS so far. The typical power consumption is 300mW. The strengths of meDVS make it desirable to any applications on the limited storage, bandwidth, and low latency of the on-board embedded system of the intelligent system.

Dynamic and Active Pixel Vision Sensor (DAVIS). In this paper we use a new neuromorphic vision sensor which is named the Dynamic and Active Pixel Vision Sensor (DAVIS) [39]. The model DAVIS240 camera has a higher resolution of 240x180, higher dynamic range, and lower power consumption and allows a concurrent readout of global shutter image frames, which are captured using the same photodiodes as for

the DVS event generation. In this work, we only use the event data.

3.2. Dataset and Benchmark. We present a labeled dataset for the evaluation of an online multivehicle detection and tracking system in ITS domain. The raw event data are collected by a neuromorphic vision sensor which is mounted on the bridge in a highway scenario. The neuromorphic vision sensor used in this paper is called dynamic and active pixel sensor (DAVIS) with a model No. DAVIS240C. We have labeled three event sequences in this work. The first event sequence (named EventSeq-Vehicle1) is of length 45.4s having 110.7Mevents and on average contains 2,438Keps (Kilo events per second). The second event sequences (named EventSeq-Vehicle2) is of length 32.4s with 79.4Mevents and on average contains 2,450Keps. The third event sequence (named EventSeq-Vehicle3) is of length 21.8s having 53.4Mevents and on average contains 2,450Keps. The vehicles are moving in both the directions, i.e., towards and away from the camera in multiple lanes. The vehicles in the dataset range from small cars to the trailers and trucks, which makes the dataset diverse and challenging in nature.

We manually annotated all the vehicles' positions and unique identity in three event sequences using the openly available video annotation tool called ViTBAT [40]. For annotation, the video was created from the binary events data file. We accumulated events data into video frames at three different time intervals: 10ms, 20ms, and 30ms. The description and data format of our dataset can be seen from Table 1.

4. Online Multitarget Detection and Tracking

We illustrate our multiobject tracking-by-clustering system in this section. In contrast to traditional object detection

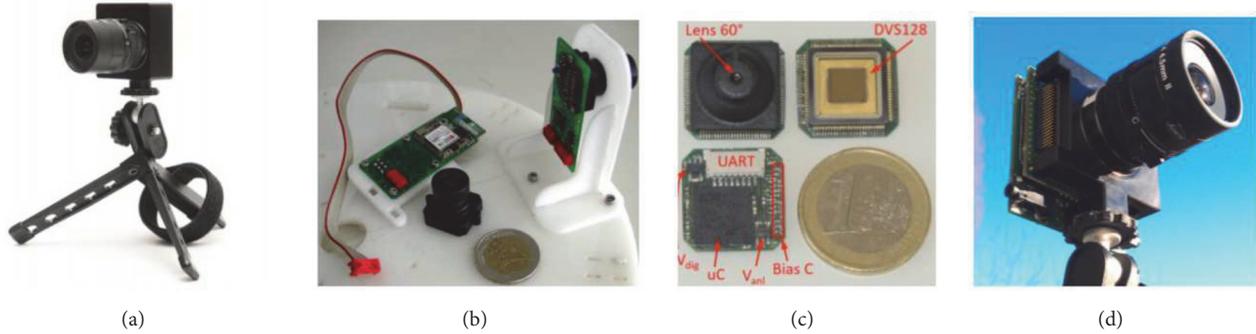


FIGURE 2: Different versions of neuromorphic vision sensors: (a) Dynamic Vision Sensor 128 (DVS128). (b) Embedded Dynamic Vision Sensor (eDVS). (c) Miniature Embedded Dynamic Vision Sensor (meDVS). (d) Dynamic and Active Pixel Vision Sensor (DAVIS). This figure is adopted from [11].

TABLE 1: The standard file format in our benchmark. In event.txt, timestamp recodes the timestamp of each raw event, x and y are the coordinates of the event's origin, and p is the event's polarity. In det.txt, box_x and box_y are x and y coordinates of the bounding box and box_width and box_height are the width and height of each bounding box. The events which fall into a bounding box from start timestamp to end timestamp are accumulated to a cluster. Each cluster has an ID. The items' definitions in gt.txt and track.txt are similar to det.txt file, except that the gt.txt shows the ground truth and track.txt shows the result of tracking.

| File name | Description | Format |
|-----------|---------------------------------------|---|
| event.txt | One event per line | (timestamp, x, y, p) |
| det.txt | One detection measurement per line | (starttimestamp, endtimestamp, box_x, box_y, box_width, box_height) |
| gt.txt | One ground-truth measurement per line | (starttimestamp, endtimestamp, ID, box_x, box_y, box_width, box_height) |
| track.txt | One tracking measurement per line | (starttimestamp, endtimestamp, ID, box_x, box_y, box_width, box_height) |

approaches, we generate our object hypothesis directly from the measurements with a classic clustering method. The advantage is that we can skip the background modeling step (dynamic foreground segmentation), as most events transmitted by the dynamic vision sensor are generated by dynamic objects. In order to estimate the states of the actual objects, we integrate an online multitarget tracking method into our system. It is our opinion that only highly effective and online tracking methodology can take full advantage of neuromorphic vision cameras.

4.1. Vehicle Detection by Clustering. As neuromorphic sensors only transmit relative light intensity changes for each pixel, methods using appearance features, such as color and texture as input, cannot be utilized. Clustering methods, on the contrary, are very suitable for this situation. Hence we present three different clustering algorithms in this section, which do not depend upon the prior knowledge of the number and shape of the clusters. In addition, only dynamic information in the form of sparse streams of asynchronous time-stamped events can be gained from neuromorphic vision sensors. In order to arrive at a meaningful interpretation and make the most of neuromorphic vision sensors' advantages, it is necessary to accumulate event streams before applying clustering algorithms. We accumulate event data for different time intervals (10ms, 20ms, and 30ms), making it synchronized and more informative, after which three

classic clustering approaches, MeanShift [4], DBSCAN [5] and WaveCluster [6], are carried out and compared. The following subsections illustrate these clustering approaches briefly.

Detection by MeanShift [11]. Estimation of the gradient of a density function via MeanShift and the iterative mode seeking procedure was developed by Fukunaga and Hostetler in [41]. The mean-shift algorithm has been exploited in low level computer vision tasks, including image segmentation, color space analysis, face tracking, etc., by reason of its properties of data compaction and dimensionality reduction [4]. Specifically, the mean-shift algorithm considers the input as a probability density function and the objective of the algorithm is to find the modes of this function. These modes represent the centers of the discovered clusters. The input points are fed to the kernel density estimation and then the gradient ascent method is applied to the density estimate. The density estimation kernel uses two inputs: the total amount of points and the bandwidth or the size of the window. The main disadvantage of the mean-shift algorithm lies in its iterative nature and difficulty of filtering out noise.

Detection by DBSCAN [11]. DBSCAN uses density based spatial clustering for applications with noise. For each point p , the associated density is calculated by counting the number of points in a search area of specified radius, Eps (the maximum radius of the neighbourhood from point), around

the point. The points with density higher than the specified threshold value, MinPts (the minimum number of points required to form a dense region), are classified as core points while the rest are classified as noncore points. A cluster is yielded if p is a core point; otherwise, if p is a border point, then no point is density-reachable from p and DBSCAN takes the next point from the database [5]. The main advantage of DBSCAN is that it can find the clusters of arbitrary shapes.

Detection by WaveCluster. The basic idea of WaveCluster is to quantize the feature space of the image firstly and then apply discrete wavelet transform on it, after which we can find the connected components (clusters) in the subbands of transformed feature space [6]. For best clustering result, the quantization scale as well as the component connection algorithm should be applied according to the raw data. In the context of this paper, the accumulated event data can be regarded as 2-dimensional data. With selected interval m in each dimension, we can now divide the event data into m^2 grids, and each grid i contains N_i data point. Considering the multiresolution property of wavelet transform, different grid sizes can be adopted at different scales of transform. In the second step of WaveCluster algorithm, discrete wavelet transform will be applied on the quantized feature space [6]. Afterwards, a new feature space T_k is acquired. We can also filter out the noise in T_k with a selected threshold. With the new set of units T_k , connected components in the transformed feature space can be detected as clusters. Details of the algorithm can refer to [6].

4.2. Online Multitarget Tracking. In order to make full use of the advantages of event data, we have chosen four classic tracking algorithms, which are relatively small in computation and highly effective. Our online multitarget tracking is a simple and standard method which is widely explored in traditional camera-based multiobject tracking [42]. As the event data have no texture information, we use the bounding box overlap as a simple association metric for the data association problem. All these tracking algorithms are briefly described in the following sections.

Tracking by SORT [11]. We utilize a single hypothesis tracking methodology with standard Kalman filter and data association using Hungarian method [7]. In order to assign detected clusters to existing targets, each target's geometry and image coordinates are estimated by predicting its new state in the current frame. The cost matrix for each detected cluster and each existing target is calculated as the intersection over union distance (IOU). The Hungarian algorithm is used to optimally solve the assignment problem. We also define a minimum IOU to reject assignments where the detected cluster to target cluster overlap is less than the threshold. When a new cluster enters into the camera field of view or when an existing target leaves the camera view, target identities get updated, either by adding new IDs or by according deletion. The same methodology for tracking has been used in this work as presented in [7]. Instead of solving for detection for tracking in a global assignment

problem, we choose an early deletion of lost targets policy, which prevents unbounded growth of the number of trackers.

Tracking by GM-PHD. GM-PHD filter is a recursive algorithm which jointly estimates the time-varying number of targets and their states from the observation sets in the presence of data association uncertainty, noise, false alarms, and detection uncertainty. The algorithm models the respective collection of targets and measurements as random finite sets and applies the probability hypothesis density (PHD) recursively for posterior intensity propagation, which is basically the first order-statistic of the random finite set in time. With linear and Gaussian assumptions, the target dynamics and birth process and the posterior intensity at any time step are considered to be Gaussian mixture. The recursions with number of Gaussian components management increase the efficiency. In tracking world, the intensity is also known as probability hypothesis density [8]. The further mathematical insights into the algorithm and its recursive linear Gaussian version can be studied in [8]. As stated in the previous section, the birth model for the targets is chosen to be linear in this work, which also stands for this and upcoming approaches for tracking.

Tracking by GM-CPHD. In probability hypothesis density (PHD) filter, the posterior intensity of the random finite set of targets is propagated, recursively. In cardinalized PHD (CPHD) filter, both the posterior intensity and posterior cardinality distribution are propagated jointly, hence making it a generalization of PHD recursion. The accuracy and stability are increased by incorporating the cardinality information [9]. This work is basically the implementation of closed-form solution to CPHD recursion under the assumption of linear Gaussian target dynamics and birth model. The algorithm can also be extended to nonlinear models using linearization and unscented transformation techniques. While comparing with standard PHD filter, CPHD filter not only side steps the need of data association task in conventional tracking methods but also improves the accuracy of the individual target state estimates and the variance of the estimated number of targets [9].

Tracking by PDAF. The probabilistic data association filter (PDAF) computes the probabilities for target being tracked for each valid measurement. This measurement origin uncertainty is accounted by this probabilistic or Bayesian information. As the linear models for the targets birth dynamics and measurement equations are assumed, therefore, the developed PDAF algorithm is based on Kalman filter. PDAF works on the validated measurements at the current time and for each measurement, an association probability is calculated for computing the weight of current measurement in a combined innovation. This combined innovation helps in updating the estimation of the state. And finally, the state covariances are updated for computing the measurement origin uncertainty [10]. The detailed mathematical insights into the PDAF algorithm with its extensions can be studied from [10].

TABLE 2: Evaluation metrics for vehicle detection, TP is the total number of true positives, FP is the total number of false positives, and GT indicates ground truth.

| Metrics | Better | Perfect | Description |
|-----------|--------|---------|---|
| Precision | higher | 100% | Ratio of TP / (TP+FP) |
| Recall | higher | 100% | Ratio of correct detections to total number of GT boxes |

5. Experiments and Results

We evaluate the performance of various tracking-by-clustering implementations on our dataset. The evaluation results are provided by following the standard MOT challenge metrics [43]. We analyze the performance and runtime of the three classical clustering algorithms, as well as the four tracking algorithms for multivehicle tracking-by-clustering task, where stream inputs are accumulated at different intervals (10ms, 20ms, and 30ms time intervals).

5.1. Metrics. For performance evaluation, we follow the current evaluation protocols for visual object detection and multiobject tracking. Although these protocols are designed for frame-based vision sensors, they are still suitable for quantitative evaluation of our tracking method. In this work, we accumulate events to frames in different time intervals. In this work we have two evaluation metrics (see Table 2) which are defined in [44].

Since our detection results from clustering methods have no probability score, we are not able to provide the mean precision to summarize the shape of the precision/recall (ROC) curve which is widely adopted in object detection evaluation in computer vision. The evaluation metrics for multivehicle tracking used in this work is defined in [43], well-known as the MOT challenge metrics. Evaluation scripts are available on MOT Challenge official website (<https://motchallenge.net>). More details are as follows:

- (i) MOTA(\uparrow): Multiple Object Tracking Accuracy. This measure combines three error sources: false positives, missed targets, and identity switches.
- (ii) MOTP(\uparrow): Multiple Object Tracking Precision. The misalignment between the annotated and the predicted bounding boxes.
- (iii) MT(\uparrow): mostly tracked targets. The ratio of ground-truth trajectories that are covered by a track hypothesis for at least 80% of their respective lifespan.
- (iv) PT(\uparrow): number of partially tracked trajectory.
- (v) ML(\downarrow): mostly lost targets. The ratio of ground-truth trajectories that are covered by a track hypothesis for at most 20% of their respective lifespan.
- (vi) FP(\downarrow): the total number of false positives.
- (vii) FN(\downarrow): the total number of false negatives (missed targets).
- (viii) IDs(\downarrow): the total number of identity switches.

- (ix) FM(\downarrow): the total number of times a trajectory is fragmented (i.e., interrupted during tracking).

5.2. Performance Evaluation. In this section, we report the performance and runtime of the selected approaches for multivehicle detection and tracking. Firstly, we compare the detection performance of the three clustering methods (DBSCAN, MeanShift, and WaveCluster). Then the impacts of different sampling time intervals on detection results are studied. Finally, the tracking performance and runtime for different tracking methods are evaluated.

5.2.1. Online Multivehicle Detection. In this work, event data are considered as pure 2D point data. The clustering technique is applied to generate object proposals. The event data for different time intervals (10ms, 20ms, and 30ms) are accumulated and can be seen in Figure 3. It is straightforward to see that clusters of event data reflect moving vehicles. The noise events surrounding each cluster are mainly generated by the environmental changes and sensor noise. Therefore, prior to generating object hypotheses, a background activity filtering step is performed to filter out the noise from the events. For each event, background activity filter checks whether one of the 8 (vertical and horizontal) neighbouring pixels has had an event within the last “us_Time” microseconds. If not, the event being checked will be considered as noise and removed. In other words, whether a new event is considered as “signal” or “noise” is determined by whether there is a neighbouring event generated within a set interval (us_Time). Figure 4 shows the accumulated events frame before and after the application of activity filter.

Figure 5(a) shows DBSCAN clustering results. For DBSCAN, the search radius, *Eps*, is chosen as 5 and the density, *MinPts*, is chosen as 10. The points with density higher than the specified threshold value, *MinPts*, are classified as core points while the rest are classified as noncore points. Those noncore points are also classified as noise points. Seven clusters including noise events have been detected. Figure 5(b) shows the MeanShift clustering results, with a chosen bandwidth of 20. The MeanShift algorithm successfully detected six clusters. And we can see many clusters were detected using WaveCluster from Figure 5(c). MeanShift divides many noises and objects into one cluster, and WaveCluster treats many noises as a single cluster. Their common shortcoming is that they cannot distinguish between object (here is car) and noise well.

The detection performance is assessed by clustering approach in terms of the metrics of recall and precision. The evaluation of DBSCAN, MeanShift, and WaveCluster on our neuromorphic data with different time intervals is shown in Table 3. We can see that the performance of clustering algorithms increases significantly from 10ms time interval to 20ms time interval, which shows that detection-by-clustering methods, used in this work, perform better with more events per time interval. But, from the performance of 30 ms interval, we can also know that, with the accumulation of events, more and more noise points appear, and the accuracy of the detection algorithm decreases. The results indicate

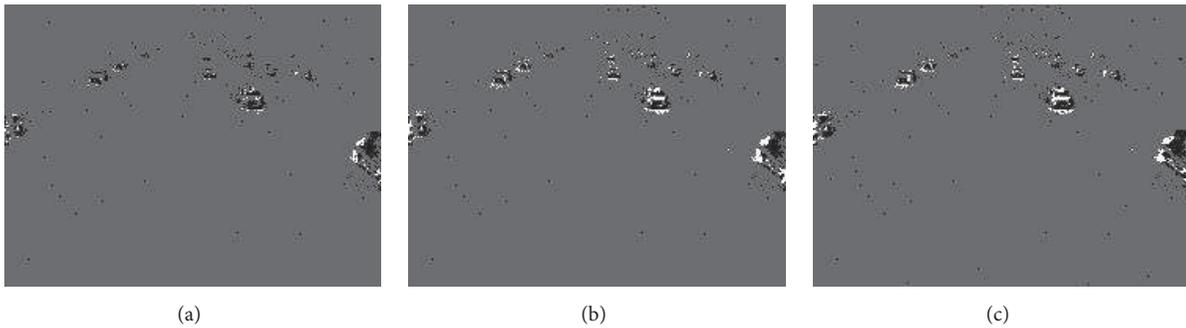


FIGURE 3: (a) Events accumulated in 10ms time interval. (b) Events accumulated in 20ms time interval. (c) Events accumulated in 30ms time interval.

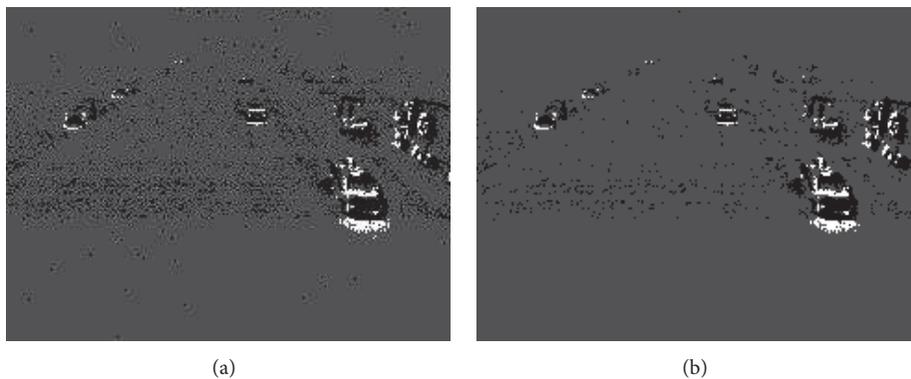


FIGURE 4: Background activity filter. (a) Events accumulated in 20ms time interval prior to filtering. (b) Events accumulated in 20ms time interval after filtering. This figure is adopted from [11].

that the detection performance is highly dependent on the number of events during the accumulated time. This points out an alternative way of accumulating a constant number of events instead of constant time intervals may increase the robustness of our detection-by-clustering approach. Among the three algorithms, MeanShift performs the worst. The reason behind it is that the density estimation of MeanShift is affected by the random noise from DAVIS. Secondly, since the MeanShift is aiming at globular clustering, it may merge some small targets when detecting as illustrated in Figure 5. Lastly, the kernel bandwidth and window size remain the same in detection, resulting in bad performance when detecting fast moving and size-changing vehicles in our scenario. From Table 3, the detection accuracy of WaveCluster is higher overall. But, the detection effect of WaveCluster at 10ms time interval is relatively poor, noise cannot be eliminated, and the detection performance is greatly affected by the number of events. In order to make the tracking algorithms get a better performance in different time intervals of the three datasets. We choose DBSCAN as detection algorithm used for comparing the tracking results.

5.2.2. Online Multivehicle Tracking. In this part, the four tracking algorithms have been implemented, i.e., simple

online and real-time tracking (SORT), GM-PHD filter, GM-CPHD filter, and the PDA filter. The tracking performance for the four trackers applied to the three vehicle sequences datasets is presented below.

Figure 6 shows the tracking results of SORT, GM-PHD filter, GM-CPHD filter, and the PDA filter using a series of input events for 20ms time interval. It can be seen from the continuous figures, such as Figures 6(a), 6(b), and 6(c), that our tracking algorithms perform better with moving vehicles when a new vehicle enters into the camera field of view or when an existing target leaves the camera view, target identities get updated, either by adding new IDs or by according deletion.

If any detected target in the current event frame had an overlap with an untracked detected target in previous frame, it would be registered with a new ID. As can be seen from Figure 6, most of the targets are well tracked. Especially, in the same continuous time interval, SORT tracks 29 targets, which is the largest number of targets tracked in the four algorithms. And the GM-PHD tracks 19 targets, followed by GM-CPHD with 15. However ID switching or target missing errors can also be witnessed from Figures 6(d)–6(l), PDAF performs the worst in terms of the problems of the number of targets, ID assignment and missed targets. And it can be

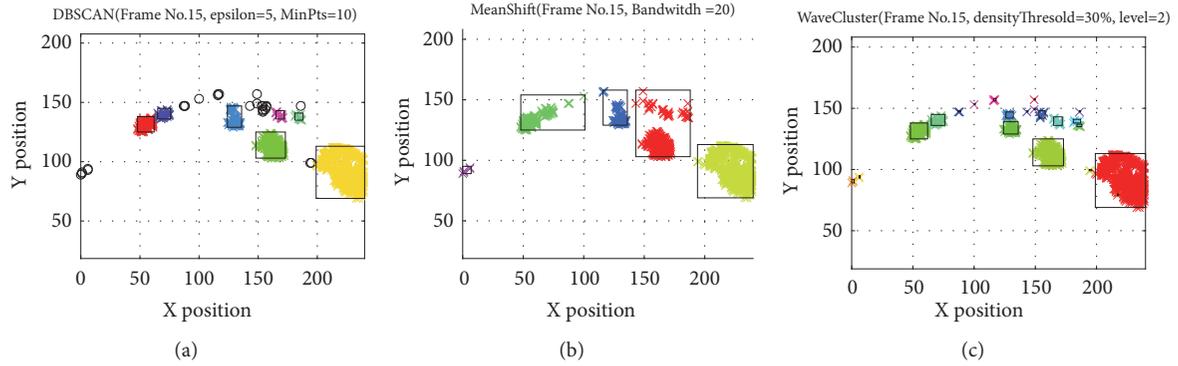


FIGURE 5: Detection-by-clustering results of (a) DBSCAN clustering applied to the events data accumulated at 20ms time interval. (b) MeanShift clustering applied to the events data accumulated at 20ms time interval. (c) WaveCluster clustering applied to the events data accumulated at 20ms time interval (best viewed in color).

TABLE 3: The performance of clustering algorithm.

| Dataset | Tis | DBSCAN | | MeanShift | | WaveCluster | |
|-------------------|------|--------|-----------|-----------|-----------|-------------|-----------|
| | | Recall | Precision | Recall | Precision | Recall | Precision |
| EventSeq-Vehicle1 | 10ms | 53.1% | 60.6% | 44.5% | 41.9% | 44.4% | 49.2% |
| EventSeq-Vehicle1 | 20ms | 62.8% | 64.5% | 46.6% | 40.7% | 63.1% | 64.4% |
| EventSeq-Vehicle1 | 30ms | 61.9% | 61.9% | 44.3% | 40.7% | 61.8% | 64.9% |
| EventSeq-Vehicle2 | 10ms | 46.4% | 51.9% | 39.8% | 38.0% | 38.7% | 41.4% |
| EventSeq-Vehicle2 | 20ms | 52.3% | 53.3% | 38.3% | 35.0% | 51.9% | 53.0% |
| EventSeq-Vehicle2 | 30ms | 46.2% | 47.2% | 33.2% | 33.5% | 47.2% | 52.1% |
| EventSeq-Vehicle3 | 10ms | 41.1% | 54.4% | 35.1% | 40.1% | 32.1% | 41.8% |
| EventSeq-Vehicle3 | 20ms | 49.7% | 59.4% | 35.5% | 36.8% | 49.6% | 58.0% |
| EventSeq-Vehicle3 | 30ms | 47.7% | 55.6% | 33.5% | 35.3% | 49.0% | 57.5% |

clearly seen from Figures 6(k) and 6(l) that the same target is given different ID at different times, indicating that target lost occurs. Hence, the performance of our algorithms shows the limitations of the tracking-by-clustering system to some extent.

Table 4 shows the tracking performance metrics, i.e., MOTA, MOTP, MT, PT, ML, FP, FN, IDs, and FM for all the four trackers, i.e., SORT, GM-PHD filter, GM-CPHD filter, and the PDA filter for each 10ms, 20ms, and 30ms time intervals fed with EventSeq-Vehicle1. As the tracking component is highly dependent on the detection results, the number of times an ID-switched (IDs) is pretty large due to the inconsistent detection results. From the overall tracking performance evaluation results in Table 4, the value of MOTA and MOTP for four tracking algorithms is relatively higher. After applying these frame-by-frame-based tracking approaches, it is not surprising that we get large number of false detection, missed detection, ID switch, and fragmentation (FM). One possible way to decrease the number of missed detection, ID switch, and fragmentation (FM) is replacing the simple association metric in this paper to a more informed metric including motion information; it is able to track objects through longer periods of occlusions and disappearances. Tables 5 and 6 present the tracking performance metrics for EventSeq-Vehicle2 and EventSeq-Vehicle3, which are not as good as that of EventSeq-Vehicle1. It is especially

obvious in EventSeq-Vehicle2 with 30ms time interval, where the evaluation metrics of MOTA for tracking algorithms are very low. The main reason behind it lies in the occasional flash of huge amount of noise as shown in Figure 7, which would seriously obscure the tracking targets, resulting in periodic fluctuation in the performance of the algorithm. This “noise flash” phenomenon can attributed to the unstable working state of the sensor and variable environmental conditions. It also indicates that our three datasets are very representative and challenging. Such limitation of the neuromorphic vision sensor will also be discussed in Section 6.2.

As the first work of multitarget tracking based on neuromorphic vision sensor, we are not able to compare to state-of-the-art tracking algorithms. Instead, we provide our evaluation results as a baseline tracker for future neuromorphic vision based multiobject tracking methods.

Runtime. The experiment is carried out on a laptop with Intel Core™i7-6700HQ CPU with 2.60GHz quad core processor and 8.00 GB of RAM. Table 7 shows that the average FPS of the DBSCAN algorithm is 36, 17, and 8 for 10ms, 20ms, and 30ms time intervals, respectively. The decreasing frame rate is due to the increased number of events in the density search area, resulting in a more iterative process. Of course, the runtime performance is related to the selection of algorithms; for example, WaveCluster has almost the same frame rate at

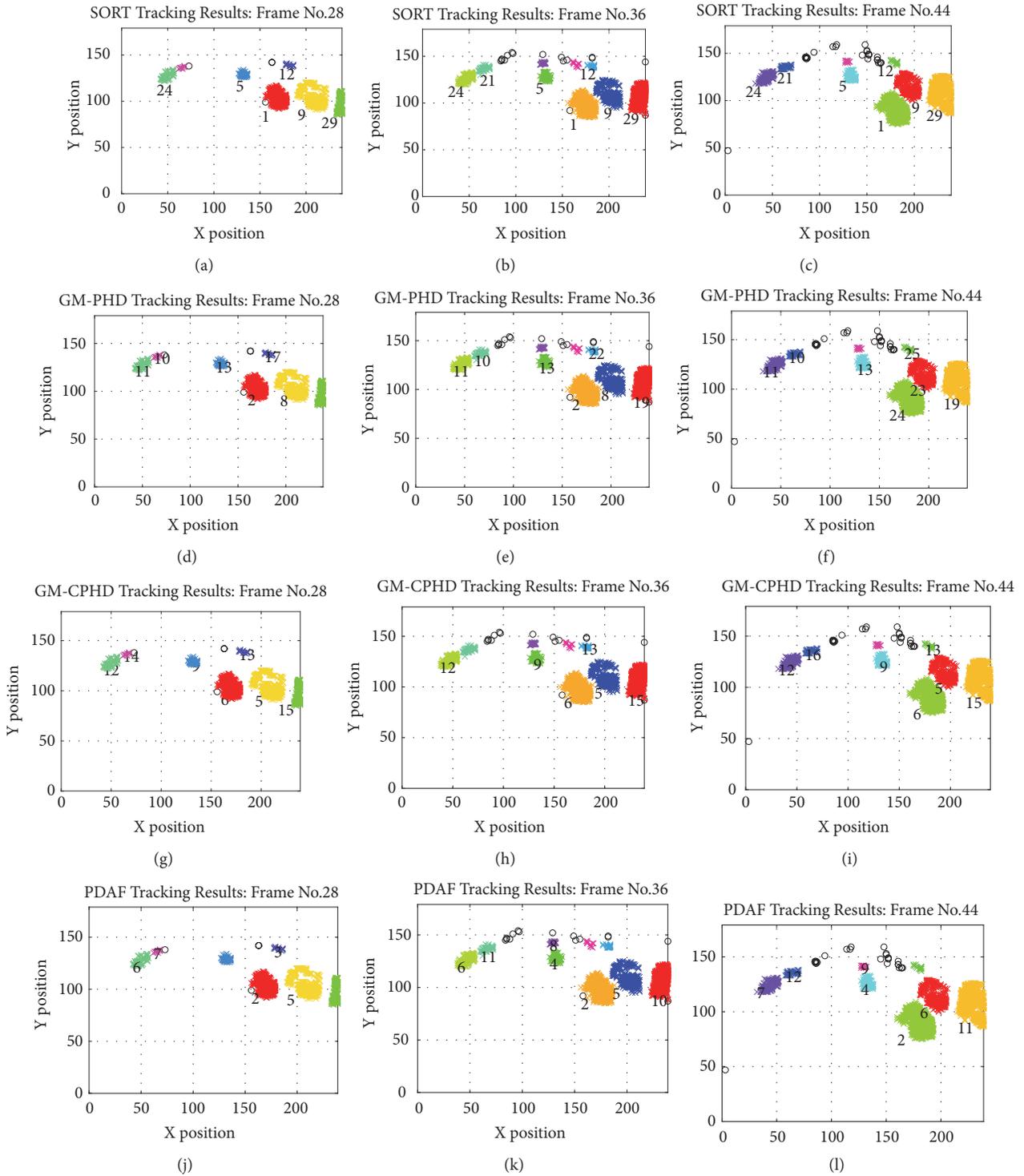


FIGURE 6: Tracking results from Frame 28 to Frame 44 for SORT, GM-PHD, GM-CPHD, and PDAF tracker. The numbers in black font near to cluster show the tracking ID (each object hypothesis is associated a tracker with a unique ID). The ID switching and lost IDs can also be witnessed from those sequences of input data (best viewed in color).

different time intervals. Additionally, in spite of the ordinary computer resources, MeanShift has a high running efficiency. For the tracking component, SORT is able to reach 552 FPS as shown in Table 8. Such a high frame rate indicates a promising

application of the sensors. According to the experimental results, our tracking-by-clustering system can run at a rate of more than 110 Hz when our tracking algorithms is combined with efficient detection algorithms, such as MeanShift. In

TABLE 4: The tracking performance for EventSeq-Vehicle1 using different tracking methods (SORT, GM-PHD, GM-CPHD, and PDAF) with detection through DBSCAN clustering. The impact of time intervals from 10ms to 30ms was studied.

| Tis | Tracker | MOTA↑ | MOTP↑ | MT↑ | PT↑ | ML↓ | FP ↓ | FN↓ | IDs↓ | FM↓ |
|------|---------|-------|-------|-----|-----|-----|------|-------|------|------|
| 10ms | SORT | 36.2% | 69.2% | 8 | 79 | 20 | 2891 | 16369 | 146 | 1302 |
| 10ms | GM-PHD | 24.0% | 69.1% | 1 | 85 | 21 | 4924 | 16649 | 1541 | 4097 |
| 10ms | GM-CPHD | 21.1% | 69.2% | 3 | 89 | 15 | 7480 | 15621 | 900 | 3616 |
| 10ms | PDAF | 20.9% | 69.1% | 0 | 86 | 21 | 5653 | 18228 | 158 | 4678 |
| 20ms | SORT | 35.0% | 70.2% | 18 | 71 | 18 | 2905 | 6893 | 92 | 444 |
| 20ms | GM-PHD | 35.1% | 70.6% | 18 | 70 | 19 | 2523 | 7019 | 323 | 770 |
| 20ms | GM-CPHD | 25.7% | 70.5% | 12 | 75 | 20 | 3974 | 7180 | 152 | 716 |
| 20ms | PDAF | 24.5% | 70.4% | 4 | 80 | 23 | 3576 | 7815 | 95 | 1371 |
| 30ms | SORT | 28.5% | 70.4% | 12 | 69 | 26 | 1950 | 5190 | 94 | 265 |
| 30ms | GM-PHD | 23.6% | 70.8% | 14 | 67 | 26 | 2323 | 5224 | 190 | 478 |
| 30ms | GM-CPHD | 18.3% | 70.7% | 8 | 76 | 23 | 2870 | 5259 | 135 | 481 |
| 30ms | PDAF | 19.3% | 70.5% | 1 | 76 | 30 | 2402 | 5701 | 66 | 900 |

TABLE 5: The tracking performance for EventSeq-Vehicle2 using different tracking methods (SORT, GM-PHD, GM-CPHD, and PDAF) with detection through DBSCAN clustering. The impact of time intervals from 10ms to 30ms was studied.

| Tis | Tracker | MOTA↑ | MOTP↑ | MT↑ | PT↑ | ML↓ | FP ↓ | FN↓ | IDs↓ | FM↓ |
|------|---------|-------|-------|-----|-----|-----|------|-------|------|------|
| 10ms | SORT | 24.4% | 70.2% | 3 | 53 | 29 | 2170 | 14929 | 183 | 942 |
| 10ms | GM-PHD | 13.4% | 69.4% | 0 | 60 | 25 | 3807 | 14994 | 1000 | 2524 |
| 10ms | GM-CPHD | 7.8% | 69.7% | 2 | 69 | 14 | 7452 | 13084 | 528 | 2524 |
| 10ms | PDAF | 13.8% | 69.8% | 0 | 57 | 28 | 4403 | 15206 | 100 | 3050 |
| 20ms | SORT | 5.7% | 68.1% | 7 | 49 | 28 | 3304 | 7393 | 81 | 331 |
| 20ms | GM-PHD | 15.6% | 70.6% | 11 | 52 | 21 | 2839 | 6524 | 290 | 729 |
| 20ms | GM-CPHD | 11.3% | 70.6% | 10 | 57 | 17 | 3824 | 6196 | 118 | 655 |
| 20ms | PDAF | 11.5% | 70.5% | 4 | 58 | 22 | 3091 | 6948 | 76 | 995 |
| 30ms | SORT | 0% | 67.3% | 3 | 45 | 37 | 2069 | 5496 | 53 | 183 |
| 30ms | GM-PHD | 7.6% | 70.3% | 5 | 50 | 30 | 1888 | 5001 | 149 | 328 |
| 30ms | GM-CPHD | -0.7% | 70.1% | 4 | 57 | 24 | 2872 | 4694 | 100 | 389 |
| 30ms | PDAF | 5% | 69.9% | 3 | 55 | 27 | 1969 | 5216 | 49 | 542 |

comparison, DeepSort method [45] only reaches a runtime speed of 40 Hz despite the use of high performance GPU.

6. Conclusion and Discussion

6.1. Conclusion. In this paper, the first neuromorphic vision based multivehicle detection and tracking system in ITS is proposed. We provide our datasets and approaches as a baseline tracker for future neuromorphic vision based multiobject tracking methods. A variety of algorithms to perform the tracking task are presented, of which different combinations can be chosen for different accuracy and rate requirements. Hopefully, our preliminary study can motivate further research in this field, considering that the sparse stream of event data from the sensor captures only motion and salient information, which is perfect for the intelligent infrastructure systems. The proposed event-based online multiple target tracking-by-clustering system utilizes strikingly simple algorithms while it achieves good detection and tracking performance with respect to runtime requirement.

Specifically, three clustering algorithms, i.e., DBSCAN, MeanShift, and WaveCluster, were explored to deal with

the sparse data from neuromorphic sensor. After studying the detection results, the DBSCAN was selected for further detection stage due to its more robust and accurate outcome. Based on the detection results from DBSCAN, four different trackers were studied and their results were compared. The selected trackers were SORT, GM-PHD filter, GM-CPHD filter, and the PDA filter. From the experimental results, the tracking algorithm combined with DBSCAN can achieve higher accuracy, while combined with MeanShift can achieve higher frame rate of more than 110Hz. Different combinations of algorithms can be applied depending on different requirements of accuracy and real-time performance.

6.2. Discussion. To the best of our knowledge, the presented system is the first application of neuromorphic vision sensor on ITS which makes it well suited as a baseline, allowing for new researcher to work on intersection of the neuroscience and intelligent system. In our future work, different event encoding methods will be tried, adaptive algorithms will be explored, and the benchmark will be extended to pedestrian detection and tracking. Different filters, other than the basic

TABLE 6: The tracking performance for EventSeq-Vehicle3 using different tracking methods (SORT, GM-PHD, GM-CPHD, and PDAF) with detection through DBSCAN clustering. The impact of time intervals from 10ms to 30ms was studied.

| Tis | Tracker | MOTA \uparrow | MOTP \uparrow | MT \uparrow | PT \uparrow | ML \downarrow | FP \downarrow | FN \downarrow | IDs \downarrow | FM \downarrow |
|------|---------|-----------------|-----------------|---------------|---------------|-----------------|-----------------|-----------------|------------------|-----------------|
| 10ms | SORT | 24.6% | 69.5% | 1 | 34 | 24 | 1258 | 10838 | 109 | 611 |
| 10ms | GM-PHD | 12.9% | 68.9% | 1 | 42 | 16 | 2484 | 10956 | 666 | 1903 |
| 10ms | GM-CPHD | 12.5% | 69.1% | 2 | 41 | 16 | 3844 | 9957 | 363 | 1646 |
| 10ms | PDAF | 13.8% | 69.0% | 0 | 38 | 21 | 2710 | 11173 | 70 | 2166 |
| 20ms | SORT | 10.1% | 69.3% | 5 | 33 | 21 | 1990 | 5225 | 64 | 235 |
| 20ms | GM-PHD | 21.4% | 70.1% | 5 | 38 | 16 | 1475 | 4699 | 188 | 490 |
| 20ms | GM-CPHD | 13.4% | 70.4% | 5 | 36 | 18 | 2218 | 4698 | 94 | 402 |
| 20ms | PDAF | 17.4% | 70.2% | 1 | 40 | 18 | 1682 | 4948 | 59 | 707 |
| 30ms | SORT | 4.0% | 69.0% | 1 | 32 | 26 | 1329 | 3815 | 41 | 141 |
| 30ms | GM-PHD | 14.3% | 70.3% | 7 | 35 | 17 | 1199 | 3311 | 120 | 285 |
| 30ms | GM-CPHD | 6.1% | 70.4% | 4 | 34 | 21 | 1712 | 3300 | 58 | 240 |
| 30ms | PDAF | 12.9% | 70.6% | 3 | 34 | 22 | 1112 | 3548 | 46 | 406 |

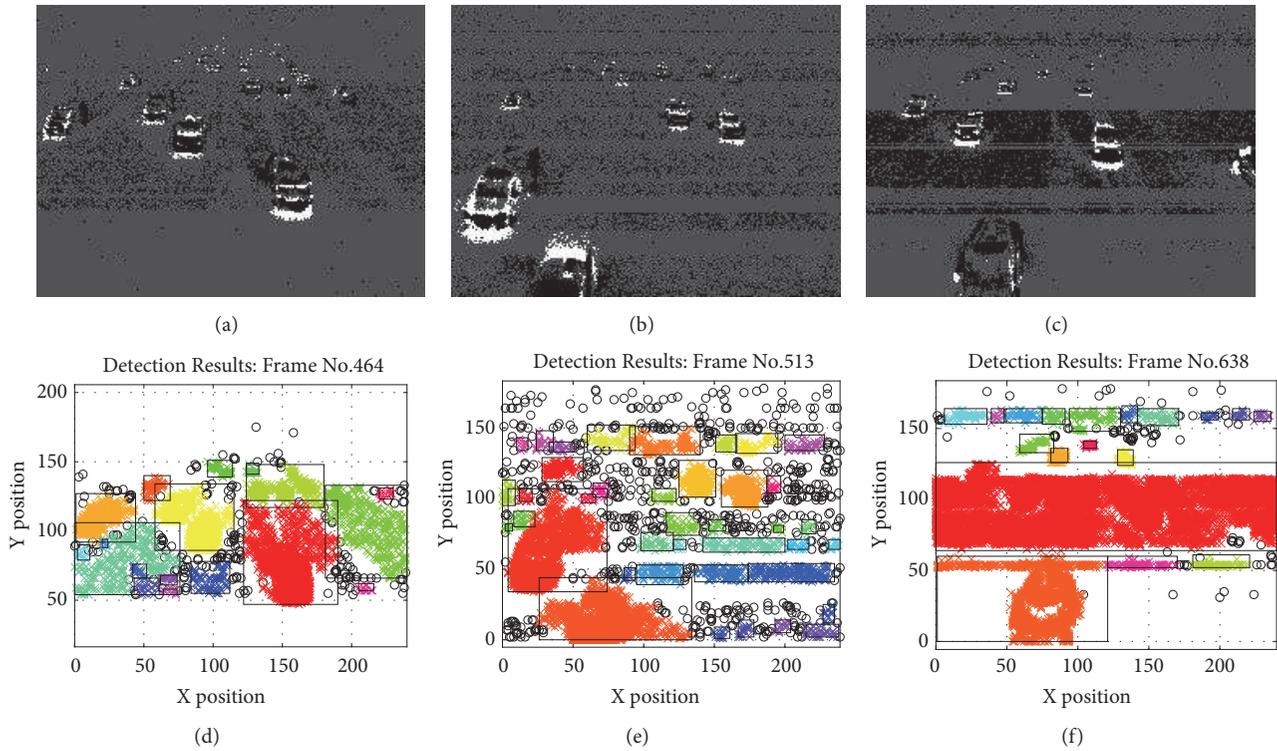


FIGURE 7: (a), (b), and (c) are the original images with the noise; (d), (e), and (f) are the results of the detection (best viewed in color).

activity filter, can be exploited to filter out the noise from the input data received from neuromorphic vision sensor. As a baseline, new approaches including recent deep learning based methods are supposed to improve the detection and tracking performance, especially the ability to identify vehicle types, such as trucks and cars and different pedestrians such as the elderly, children, etc.

Limitation. Admittedly, our algorithm still has some shortcomings. As can be seen in Figure 7, with noise becoming severer, the tracking system will make errors, such

as missed detection, multiple targets detection as one, the false detection of noise points as the target, and so on. The reason mainly lies in the immaturity of neuromorphic sensor technology. To be specific, the inherent defects of the current neuromorphic sensor lead to instability in collecting event information, which affects the quality of data and thus degrades the performance of the algorithms. Hence, a development of the sensor is indispensable before wide application of it in intelligent transportation system (ITS). It is also significant to note that, in order to take full advantage of event data, completely new neuromorphic vision algorithms are

TABLE 7: The FPS for EventSeq-Vehicle1 using different detection methods (DBSCAN, MeanShift, and WaveCluster). The impact of time intervals from 10ms to 30ms was studied.

| Detector | T _{is} | FPS |
|-------------|-----------------|-----|
| DBSCAN | 10ms | 36 |
| MeanShift | 10ms | 160 |
| WaveCluster | 10ms | 17 |
| DBSCAN | 20ms | 18 |
| MeanShift | 20ms | 107 |
| WaveCluster | 20ms | 19 |
| DBSCAN | 30ms | 8 |
| MeanShift | 30ms | 71 |
| WaveCluster | 30ms | 15 |

TABLE 8: The FPS for EventSeq-Vehicle1 using different tracking methods (SORT, GM-PHD, GM-CPHD, and PDAF).

| Tracker | FPS |
|---------|-----|
| SORT | 552 |
| GMPHD | 3 |
| GMCPHD | 4 |
| PDAF | 46 |

required instead of extending existing methods of computer vision, taking into account the brand new information stream and the extremely high frame rate of the neuromorphic vision sensor.

Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

This work was supported by the German Research Foundation (DFG) and the Technical University of Munich within the Open Access Publishing Funding Program. Part of the research has received funding from the European Unions Horizon 2020 Research and Innovation Program under Grant Agreement no. 785907(HBP SGA2) and from the Shanghai Automotive Industry Sci-Tech Development Program under Grant Agreement no. 1838. The authors would like to thank Zhongnan Qu for technical assistance and the help of data acquisition.

References

- [1] P. Lichtsteiner, C. Posch, and T. Delbruck, "A 128×128 120 dB 15 μ s latency asynchronous temporal contrast vision sensor," *IEEE Journal of Solid-State Circuits*, vol. 43, no. 2, pp. 566–576, 2008.
- [2] E. Mueggler, C. Bartolozzi, and D. Scaramuzza, "Fast event-based corner detection," in *Proceedings of the British Machine Vision Conference (BMVC)*, 2017.
- [3] X. Lagorce, C. Meyer, S.-H. Jeng, D. Filliat, and R. Benosman, "Asynchronous event-based multikernel algorithm for high-speed visual features tracking," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 26, no. 8, pp. 1710–1720, 2015.
- [4] D. Comaniciu and P. Meer, "Mean shift: a robust approach toward feature space analysis," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 24, no. 5, pp. 603–619, 2002.
- [5] M. Ester, H.-P. Kriegel, J. Sander, and X. Xu, "A density-based algorithm for discovering clusters a density-based algorithm for discovering clusters in large spatial databases with noise," in *Proceedings of the Second International Conference on Knowledge Discovery and Data Mining*, pp. 226–231, AAAI Press, 1996, <http://dl.acm.org/citation.cfm>.
- [6] G. Sheikholeslami, S. Chatterjee, and A. Zhang, "Wavecluster: A multi-resolution clustering approach for very large spatial databases in," in *Proceedings of the 24rd International Conference on Very Large Data Bases*, vol. 98, pp. 428–439, 1998.
- [7] A. Bewley, Z. Ge, L. Ott, F. Ramos, and B. Upcroft, "Simple online and realtime tracking," in *Proceedings of the 2016 IEEE International Conference on Image Processing (ICIP)*, pp. 3464–3468, Phoenix, AZ, USA, September 2016.
- [8] B. Vo and W. Ma, "The Gaussian mixture probability hypothesis density filter," *IEEE Transactions on Signal Processing*, vol. 54, no. 11, pp. 4091–4104, 2006.
- [9] B. Vo, B. Vo, and A. Cantoni, "Analytic implementations of the cardinalized probability hypothesis density filter," *IEEE Transactions on Signal Processing*, vol. 55, no. 7, part 2, pp. 3553–3567, 2007.
- [10] Y. Bar-Shalom, F. Daum, and J. Huang, "The probabilistic data association filter: estimation in the presence of measurement origin uncertainty," *IEEE Control Systems Magazine*, vol. 29, no. 6, pp. 82–100, 2009.
- [11] G. Hinz, G. Chen, M. Aafaque et al., "Online Multi-object Tracking-by-Clustering for Intelligent Transportation System with Neuromorphic Vision Sensor," *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics): Preface*, vol. 10505, pp. 142–154, 2017.
- [12] S. R. E. Datondji, Y. Dupuis, P. Subirats, and P. Vasseur, "A Survey of Vision-Based Traffic Monitoring of Road Intersections," *IEEE Transactions on Intelligent Transportation Systems*, vol. 17, no. 10, pp. 2681–2698, 2016.
- [13] J. C. Rubio, J. Serrat, A. M. López, and D. Ponsa, "Multiple-target tracking for intelligent headlights control," *IEEE Transactions on Intelligent Transportation Systems*, vol. 13, no. 2, pp. 594–605, 2012.
- [14] X. Zhang, S. Hu, H. Zhang, and X. Hu, "A real-time multiple vehicle tracking method for traffic congestion identification," *KSII Transactions on Internet and Information Systems*, vol. 10, no. 6, pp. 2483–2503, 2016.
- [15] J. Zhou, D. Gao, and D. Zhang, "Moving vehicle detection for automatic traffic monitoring," *IEEE Transactions on Vehicular Technology*, vol. 56, no. 1, pp. 51–59, 2007.
- [16] S. Sivaraman and M. M. Trivedi, "Looking at vehicles on the road: a survey of vision-based vehicle detection, tracking, and behavior analysis," *IEEE Transactions on Intelligent Transportation Systems*, vol. 14, no. 4, pp. 1773–1795, 2013.

- [17] R. Cucchiara, M. Piccardi, and P. Mello, "Image analysis and rule-based reasoning for a traffic monitoring system," in *Proceedings of the 1999 IEEE/IEEE/JSAI International Conference on Intelligent Transportation Systems*, pp. 758–763, October 1999.
- [18] M.-C. Huang and S.-H. Yen, "A real-time and color-based computer vision for traffic monitoring system," in *Proceedings of the 2004 IEEE International Conference on Multimedia and Expo (ICME)*, pp. 2119–2122, Taipei, Taiwan.
- [19] W. Zhang, Q. M. J. Wu, and H. B. Yin, "Moving vehicles detection based on adaptive motion histogram," *Digital Signal Processing*, vol. 20, no. 3, pp. 793–805, 2010.
- [20] B.-F. Wu, S.-P. Lin, and Y.-H. Chen, "A real-time multiple-vehicle detection and tracking system with prior occlusion detection and resolution," in *Proceedings of the 5th IEEE International Symposium on Signal Processing and Information Technology*, pp. 311–316, Greece, December 2005.
- [21] B. Aytekin and E. Altuğ, "Increasing driving safety with a multiple vehicle detection and tracking system using ongoing vehicle shadow information," in *Proceedings of the 2010 IEEE International Conference on Systems, Man and Cybernetics, SMC 2010*, pp. 3650–3656, Turkey, October 2010.
- [22] M. Betke, E. Haritaoglu, and L. S. Davis, "Real-time multiple vehicle detection and tracking from a moving vehicle," *Machine Vision and Applications*, vol. 12, no. 2, pp. 69–83, 2000.
- [23] A. Geiger, P. Lenz, and R. Urtasun, "Are we ready for autonomous driving? the KITTI vision benchmark suite," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR '12)*, pp. 3354–3361, Providence, RI, USA, June 2012.
- [24] L. Wen, D. Du, Z. Cai et al., "UA-DETRAC: A new benchmark and protocol for multi-object detection and tracking," *Computer Vision and Pattern Recognition*, 2015.
- [25] L. Leal-Taixé, A. Milan, I. Reid, S. Roth, and K. Schindler, "MOTChallenge 2015: Towards a benchmark for multi-target tracking," *Computer Vision and Pattern Recognition*, 2015.
- [26] D. P. Moeys, F. Corradi, E. Kerr et al., "Steering a predator robot using a mixed frame/event-driven convolutional neural network," in *Proceedings of the 2nd International Conference on Event-Based Control, Communication, and Signal Processing, EBCCSP 2016*, Poland, June 2016.
- [27] Z. Jiang, Z. Bing, K. Huang, G. Chen, L. Cheng, and A. Knoll, "Event-Based Target Tracking Control for a Snake Robot Using a Dynamic Vision Sensor," in *Neural Information Processing*, vol. 10639 of *Lecture Notes in Computer Science*, pp. 111–121, Springer International Publishing, Cham, Switzerland, 2017.
- [28] G. Chen, Z. Bing, F. Roehrbein et al., "Efficient Brain-inspired Learning with the Neuromorphic Snake-like Robot and the Neurobotic Platform," in *Proceedings of the IEEE Transactions on Cognitive and Developmental Systems*, p. 1, 2018.
- [29] H. Blum, A. Dietmüller, M. Milde, J. Conradt, G. Indiveri, and Y. Sandamirskaya, "A neuromorphic controller for a robotic vehicle equipped with a dynamic vision sensor," in *Proceedings of the Robotics: Science and Systems*, Cambridge, MA, USA, July 2017.
- [30] M. Litzengerger, A. N. Belbachir, and N. Donath, "Estimation of vehicle speed based on asynchronous data from a silicon retina optical sensor," in *Proceedings of the IEEE Intelligent Transportation Systems Conference (ITSC '06)*, pp. 653–658, Toronto, Canada, September 2006.
- [31] M. Litzengerger, C. Posch, D. Bauer et al., "Embedded vision system for real-time object tracking using an asynchronous transient vision sensor," in *Proceedings of the 2006 IEEE 12th Digital Signal Processing Workshop and 4th IEEE Signal Processing Education Workshop, DSPWS*, pp. 173–178, USA, September 2006.
- [32] A. I. Maqueda, A. Loquercio, G. Gallego, N. Garca, and D. Scaramuzza, "Event-based vision meets deep learning on steering prediction for self-driving cars in," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 5419–5427, 2018.
- [33] J. Binas, D. Neil, S.-C. Liu, and T. Delbruck, "Ddd17: End-to-end davis driving dataset, arXiv preprint," *Computer Vision and Pattern Recognition*, 2017.
- [34] Y. Hu, H. Liu, M. Pfeiffer, and T. Delbruck, "DVS benchmark datasets for object tracking, action recognition, and object recognition," *Frontiers in Neuroscience*, vol. 10, 2016.
- [35] D. Tedaldi, G. Gallego, E. Mueggler, and D. Scaramuzza, "Feature detection and tracking with the dynamic and active-pixel vision sensor (DAVIS)," in *Proceedings of the 2nd International Conference on Event-Based Control, Communication, and Signal Processing, EBCCSP 2016*, Poland, June 2016.
- [36] B. Kueng, E. Mueggler, G. Gallego, and D. Scaramuzza, "Low-latency visual odometry using event-based feature tracks," in *Proceedings of the 2016 IEEE/RISJ International Conference on Intelligent Robots and Systems, IROS 2016*, pp. 16–23, Republic of Korea, October 2016.
- [37] L. Everding and J. Conradt, "Low-latency line tracking using event-based dynamic vision sensors," *Frontiers in Neurobotics*, vol. 12, 2018.
- [38] Z. Ni, C. Pacoret, R. Benosman, S. Ieng, and S. Régnier, "Asynchronous event-based high speed vision for microparticle tracking," *Journal of Microscopy*, vol. 245, no. 3, pp. 236–244, 2012.
- [39] C. Brandli, R. Berner, M. Yang, S.-C. Liu, and T. Delbruck, "A 240×180 130 dB 3 μ s latency global shutter spatiotemporal vision sensor," *IEEE Journal of Solid-State Circuits*, vol. 49, no. 10, pp. 2333–2341, 2014.
- [40] T. A. Biresaw, T. Nawaz, J. Ferryman, and A. I. Dell, "ViTBAT: Video tracking and behavior annotation tool," in *Proceedings of the 13th IEEE International Conference on Advanced Video and Signal Based Surveillance, AVSS 2016*, pp. 295–301, USA, August 2016.
- [41] K. Fukunaga and L. D. Hostetler, "The estimation of the gradient of a density function, with applications in pattern recognition," *IEEE Transactions on Information Theory*, vol. 21, no. 1, pp. 32–40, 1975.
- [42] G. Chen, F. Zhang, D. Clarke, and A. Knoll, "Learning to track multi-target online by boosting and scene layout," in *Proceedings of the 2013 12th International Conference on Machine Learning and Applications, ICMLA 2013*, pp. 197–202, USA, December 2013.
- [43] L. Leal-Taixé, A. Milan, I. D. Reid, S. Roth, and K. Schindler, "Motchallenge 2015: Towards a benchmark for multi-target tracking," *Computer Vision and Pattern Recognition*, 2015.
- [44] M. Everingham, L. van Gool, C. K. I. Williams, J. Winn, and A. Zisserman, "The pascal visual object classes (VOC) challenge," *International Journal of Computer Vision*, vol. 88, no. 2, pp. 303–338, 2010.
- [45] N. Wojke, A. Bewley, and D. Paulus, "Simple online and realtime tracking with a deep association metric," in *Proceedings of the 2017 IEEE International Conference on Image Processing (ICIP)*, pp. 3645–3649, Beijing, China, September 2017.

Research Article

A Geostatistical Investigation into the Effective Spatiotemporal Coverage of Road Weather Information Systems in Alberta, Canada

Xu Wang ^{1,2}, Lian Gu ², Tae J. Kwon,² and Tony Z. Qiu ²

¹School of Qilu Transportation, Shandong University, Jinan 250061, China

²Department of Civil and Environmental Engineering, University of Alberta, Edmonton, Canada T6G2W2

Correspondence should be addressed to Xu Wang; xuwang@sdu.edu.cn

Received 2 August 2018; Revised 24 October 2018; Accepted 7 November 2018; Published 2 December 2018

Guest Editor: Darius Andriukaitis

Copyright © 2018 Xu 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.

Inclement weather acutely affects road surface and driving conditions and can negatively impact traffic mobility and safety. Highway authorities have long been using road weather information systems (RWISs) to mitigate the risk of adverse weather on traffic. The data gathered, processed, and disseminated by such systems can improve both the safety of the traveling public as well as the effectiveness of winter road maintenance operations. As the road authorities continue to invest in expanding their existing RWIS networks, there is a growing need to determine the optimal deployment strategies for RWISs. To meet such demand, this study presents an innovative geostatistical approach to quantitatively analyze the spatiotemporal variations of the road weather and surface conditions. With help of constructed semivariograms, this study quantifies and examines both the spatial and temporal coverage of RWIS data. A case study of Alberta, which is one of the leaders in Canada in the use of RWISs, was conducted to indicate the reliability and applicability of the method proposed herein. The findings of this research offer insight for constructing a detailed spatiotemporal RWIS database to manage and deploy different types of RWISs, optimize winter road maintenance resources, and provide timely information on inclement road weather conditions for the traveling public.

1. Introduction

Inclement weather acutely affects road and driving conditions. Approximately 22% of vehicle crashes and 25% of total travel time delays are reported to be adverse weather-related in the USA [1, 2]. To mitigate the risk of adverse weather on traffic, proactive strategies for transportation management and road maintenance operations have been developed based on road weather data [3–5]. As the predominant sources of road weather data, stationary road weather information systems (RWISs) [6] provide high temporal but limited spatial data coverage. In contrast, a mobile RWIS [7], which has vehicles collecting road and atmospheric condition information, provides spatially continuous but temporally discrete measurements. To reap the benefits of both systems, road weather data from stationary and mobile RWIS can be integrated to construct an RWIS database with both

high temporal and spatial coverage. However, installing and operating both stationary and mobile RWISs is costly. In practice, how to deploy stationary and mobile RWIS to achieve complete spatial and temporal coverage in a cost-effective way is yet unresolved. Spatiotemporal coverage is the maximum distance and time lag, beyond which the measurements are no longer representative. It is apparent that an RWIS network with better spatiotemporal coverage will generate more timely and reliable estimations. Timely and reliable road surface condition information is critical for transportation authorities to perform road maintenance and provide weather impact warnings to travelers [8–10]. Adverse road weather impact warnings, such as driving risk alert, advisory speed, and alternative route recommendations, can only be designed when real-time road weather condition data is of high spatiotemporal coverage and resolution. Thus, a cost-effective RWIS network design and deployment

require an understanding of spatiotemporal variability in road weather information. To realize such goals, the spatiotemporal coverage of RWIS must be quantified. This study quantifies that data.

Traditionally, RWISs measure atmospheric and pavement conditions in the field and send data to the traffic management centre (TMC) for road weather impact analysis. Tomás et al. [5] proposed an autonomous system that monitors, detects, and forecasts weather incidents with data from stationary RWISs. Gu et al. [11] proposed a regression Kriging method to interpolate road weather conditions using mobile RWIS data. Integrated with other data sources, RWISs can support advanced road weather management. Mahoney and Myers [12] introduced a winter road maintenance decision-support system, which integrates weather and road condition data along with maintenance operation rules of practice. Similarly, Cluett et al. [13] developed a framework concept that integrates weather and traffic operation data. In addition, the emerging connected vehicle (CV) technologies enable convenient communication between TMC and road users. In the state of Wyoming, USA, the road condition information is broadcasted from roadside units to CVs for spotting weather impact warnings [14].

Vehicles equipped with road weather sensors can also serve as mobile RWISs and enrich road weather-related data sources. Dey et al. [10] provided a comprehensive practice review and found that CVs can be used as mobile RWISs to enhance route-specific road weather data collection, condition estimation, and traffic management. Petty and Mahoney [15] revealed the potential of CVs in road weather data collection. In addition to the basic atmospheric observations, the data from onboard vehicle sensors, such as wiper state and antilock braking systems, can also be used in road weather condition estimation. With CV-enabled data in hand, Drobot et al. [16, 17] designed a quality and accuracy check process, which includes crosschecks with sensor specifications, climatological ranges, neighboring vehicle, and station measurements, for weather-related data from probe vehicles. Other data analysis shows that the quality of temperature and pressure data is affected by many factors, such as vehicle type, speed, and precipitation occurrence [16], and the temperature data from probe vehicles closely resembles data from stationary weather stations [17]. Most recently, Boyce et al. [18] developed a road weather condition assessment and forecast system, called Pikalert System. It integrates observations from connected vehicles with those from stationary RWIS, radar, and weather model analysis fields. Then it recommends snow and ice removal to road maintenance personnel and provides road weather and condition information to the public. The prototype of the tool has been tested in the Kansas City, USA, since the fall of 2017 [19].

RWISs, the stationary ones in particular, collect road weather condition information effectively but expensively. Additionally, a single RWIS data source can only take measurements that are either rich in space or rich in time—not both. Constructing a database that has both high spatial and temporal coverage data calls for an effective distribution of RWIS Environment Sensor Stations (ESSs)

and vehicles. To plan and distribute a cost-effective RWIS network, researchers have attempted to measure the monitoring capability of RWISs by experience-based and model-based methods. Manfredi et al. [20] suggested that an area with stable road weather and surface conditions can be regarded as regional representativeness in RWIS ESS allocation. Using experience-based methods, Mackinnon and Lo [21] conducted RWIS network expansion design by considering traffic loads, collision rates, climatic zones, and availability of meteorological information. For model-based methods, Singh et al. [22] assumed that the RWIS spatial coverage is a decreasing function of distance. Zhao et al. [23, 24] determined the spatial coverage based on spatial variability. They computed the standard deviation of weather severity in microzones. The spatial coverage of a station was then defined as the size of the critical buffers where the standard deviation changes most quickly with the buffer size. Kwon et al. [25, 26] introduced a spatial variogram approach to determine the autocorrelation range that measures the spatial variability of road weather conditions. The quantified monitoring capability of a stationary RWIS was then used as an input to determine optimal stationary RWIS density [25] and allocate RWIS in a road network [26].

All the studies above attempted to analyze the patterns of spatial dependence in road weather conditions. While time series analysis is a method commonly used to quantify correlation over time, to the authors' knowledge, there still lacks research that investigates spatiotemporal variability of road weather conditions. Incorporating spatial and temporal analysis, spatiotemporal analysis benefits from complete spatiotemporal information and the interactive effects of combining spatial and temporal data. The spatiotemporal variogram commonly used in geostatistics is a useful way to model spatial and temporal dependency and their interactions. It allows researchers to visualize the spatiotemporal variability and estimate the spatiotemporal autocorrelation. A variety of studies have applied the spatiotemporal variogram model to estimate spatiotemporal environmental, meteorological or climatological distributions [27–30].

To bridge the spatiotemporal research gap, this study applied the spatiotemporal variogram to measure the spatiotemporal coverage of RWISs. A case study of Alberta, Canada, was conducted to test the proposed method and look into seasonal trends of spatiotemporal data representativeness. There were three objectives of this study: (1) propose a method to quantify the spatiotemporal coverage of RWIS data; (2) evaluate the proposed method with real-world RWIS data; and (3) investigate the seasonal variations of spatiotemporal coverage. Determining the spatiotemporal coverage of RWIS data can provide a guideline that helps decision makers deploy stationary and mobile RWISs. Also, the spatiotemporal coverage works as an essential parameter in data integration and fusion among multiple data sources.

The remainder of this paper is organized into sections: The next section looks into the theory behind the spatiotemporal variogram; the Methodology section details the proposed method of this study, including the research procedure, data quality diagnostics, spatiotemporal variogram modelling, and cross validation; the Case Study section

deploys the proposed method in the study site to investigate the monthly spatiotemporal correlations; and, finally, the last section discusses the concluding remarks and suggests future work.

2. Spatiotemporal Variogram

Geostatistics is known as a class of numerical techniques that characterize spatial attributes and infers variability from the autocorrelation of insufficient data and measure the uncertainty associated with estimation [31]. However, data variability sometimes develops simultaneously in both space and time [32]. As such, the traditional pure spatial analysis has been incorporated with temporal analysis, namely, spatiotemporal geostatistics analysis.

Generally, the key element in spatiotemporal geostatistics is finding a valid and reliable spatiotemporal variogram that quantifies the data structure in space and time domains. A set of variables $z = \{z(s, t) \mid s \in S, t \in T\}$ varies within a spatial domain S and a temporal domain T . The spatiotemporal variogram is estimated as half of the mean squared difference between data separated by a given spatial and temporal lag (h_s, h_t) [33, 34]:

$$\begin{aligned} \gamma(h_s, h_t) \\ = \frac{1}{2n(h_s, h_t)} \sum_{i=1}^{n(h_s, h_t)} [z(s_i, t_i) - z(s_i + h_s, t_i + h_t)]^2 \end{aligned} \quad (1)$$

where $\gamma(h_s, h_t)$ is the semivariance of the empirical variogram; $z(s_k, t_k)$ is the measurement at the spatial location s_i and temporal location t_i ; and $n(h_s, h_t)$ is the number of pairs in the observation. Note that h_s is the Euclidean distance, which is a two- or three-dimensional spatial distance vector. Additionally, there should be no trend of systematic variation in the dataset and thus the estimated variogram is independent of the individual location. Figure 1 shows a three-dimensional spatiotemporal variogram (Figure 1(a)) together with its projection on the front or side planes (Figure 1(b)). The variogram in Figure 1(b) can be regarded as a pure spatial or temporal variogram.

A variogram model can be described by three parameters: the nugget effect, range, and sill. The nugget effect is the variance at the distance of zero, representing the microscale variation or measurement error. The range is the distance at which the data is no longer autocorrelated and the semivariance starts to level off. That is to say, the spatial and temporal ranges correspond to the spatially and temporally correlated portions in the variogram. Beyond the range, the semivariance becomes a constant value. The sill is the semivariance where the variogram levels off [35].

Typically, the empirical variogram is then smoothed by a mathematical model as the estimated model is commonly irregular, and the real data structure is likely unknown [36]. An appropriate model ensures the positive definiteness of the variogram and sufficient flexibility for the data autocorrelation structure. Thus, many models, including the separable covariance model, product-sum covariance model, and metric covariance model, have been proposed.

The separable covariance model assumes that the space-time covariance function is the product of separate spatial temporal components (i.e., $C(h_s, h_t) = C_s(h_s)C_t(h_t)$). This model separates the dependence on the two domains and simplifies the variogram modelling. In contrast, the product-sum covariance model (i.e., $C(h_s, h_t) = kC_s(h_s)C_t(h_t) + C_s(h_s) + C_t(h_t)$, with $k > 0$) was proposed to analyze the interaction between spatial and temporal domains. Also, the metric covariance model includes a space-time anisotropy ratio κ to incorporate distance in time and space (i.e., $C(h_s, h_t) = C_{joint} \sqrt{h_s^2 + (\kappa \cdot h_t)^2}$). Evaluation criteria, such as root-mean-square-error (RMSE) and mean squared error (MSE), can be applied to assess the goodness-of-fit. By measuring the difference between the observed values and the estimated values, the model that best reproduces the variability of the dataset can be selected.

3. Methodology

3.1. Research Procedure. This study analyzed the spatiotemporal patterns of RWIS data in three steps (see Figure 2). First, the raw RWIS data was extracted from the database and input into the data quality diagnostics module. The data quality diagnostics module included the data completeness test, reasonable range test, and model analysis test. Specifically, the data completeness test identified missing records; the reasonable range test recognized erroneous data that are out of a reasonable range; and the model analysis test searched for erroneous data using acceptable regions defined by models. Based on these tests, the data quality diagnostics module excluded invalid and erroneous data in the following steps. Then, with the processed data in hand, this study included the construction and fit of the variogram models by different types of spatiotemporal variograms. The fitted model parameters and the goodness-of-fit were obtained during this stage. Next, by comparing the goodness-of-fit, the best fitted variogram was selected. In this way, the spatiotemporal coverage of RWIS data was measured from the fitted spatiotemporal variogram.

3.2. Data Quality Diagnostics. In this study, missing and erroneous data was diagnosed by applying a diagnostics algorithm. RWISs typically take both road and weather-related measurements from various sensors. For data from a specific sensor, the diagnostics algorithm screened the raw data to check data completeness and quality in three tests: the data completeness test, reasonable range test, and model analysis test. First, if a data value from one sensor was void in one record, it was diagnosed as a missing measurement. Next, data that fell outside of the reasonable ranges was recognized as erroneous. The reasonable ranges were determined based on sensor specifications, location-specific climatological ranges, and historical data ranges [16, 17]. Then, in the model analysis test, data that fell beyond acceptable regions was reported as invalid. The algorithm predetermined acceptable regions according to the relationship among data from different sensors in one record. For example, the feasible road surface temperature (RST)

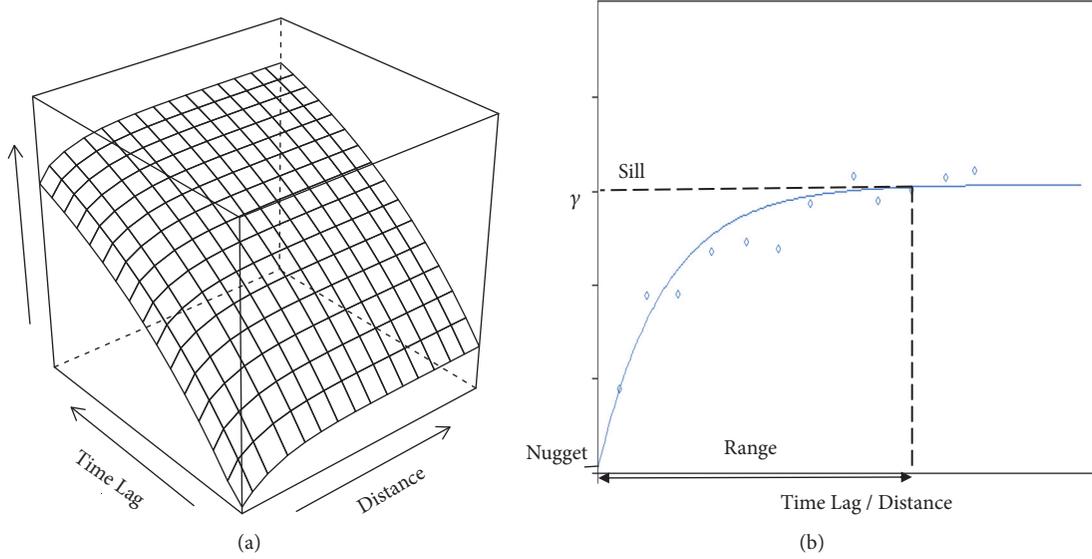


FIGURE 1: Spatiotemporal variogram. (a) Three-dimensional spatiotemporal variogram. (b) Two-dimensional variogram.

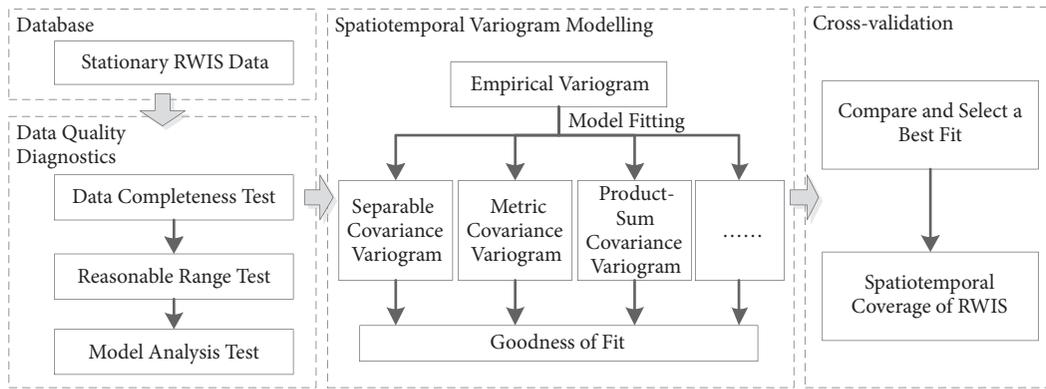


FIGURE 2: Spatiotemporal coverage quantification research procedure.

regions were defined by the regression of the generalized additive model (GAM) from historical data, which is detailed below. GAM was applied to build the relationship between RST and other data. Once RST measurements exceeded the model estimation by a certain percentage (e.g., $\pm 20\%$), the measurements were determined to be erroneous data. In this way, the diagnostics algorithm identified and removed missing and erroneous measurements from the dataset.

In the data quality diagnostics process, GAM became a generalized linear model with linear predictors. The linear predictors were linearly correlated with their unknown smooth functions [37]. The smooth functions expressed the nonparametric nonlinear associations with predictors. The following equations show the GAM formulation.

$$V = \beta_0 + f_1(x_1) + f_2(x_2) + \dots + f_I(x_I) \quad (2)$$

$$f_i(x_i) = \sum_{j=1}^J s(x_j) \quad (3)$$

In (2), V is the variable of interest; β_0 is the intercept; and $f_i(x_i)$ represents the smooth function of the predictor x_i , which is formulated as (3). The predictors can include spatial, temporal, and spatiotemporal parameters.

3.3. Spatiotemporal Variogram Modelling. The processed dataset was then applied to construct spatiotemporal variogram models. The basic steps to construct a spatiotemporal variogram are as follows.

First, the systematic trend was removed from the original data. The variogram estimator $\gamma(h_s, h_t)$ was valid only when there was no systematic variation. The measurements $z(s_k, t_k)$ were decomposed into a mean component $m(s_k, t_k)$ and a residual component $r(s_k, t_k)$ (i.e., $z(s_k, t_k) = m(s_k, t_k) + r(s_k, t_k)$). $m(s_k, t_k)$ represents the trend while $r(s_k, t_k)$ represents the fluctuations to be estimated by spatiotemporal variograms. The mean component can be estimated by various theoretical or numerical models. In this step, the residuals were extracted from the original measurements.

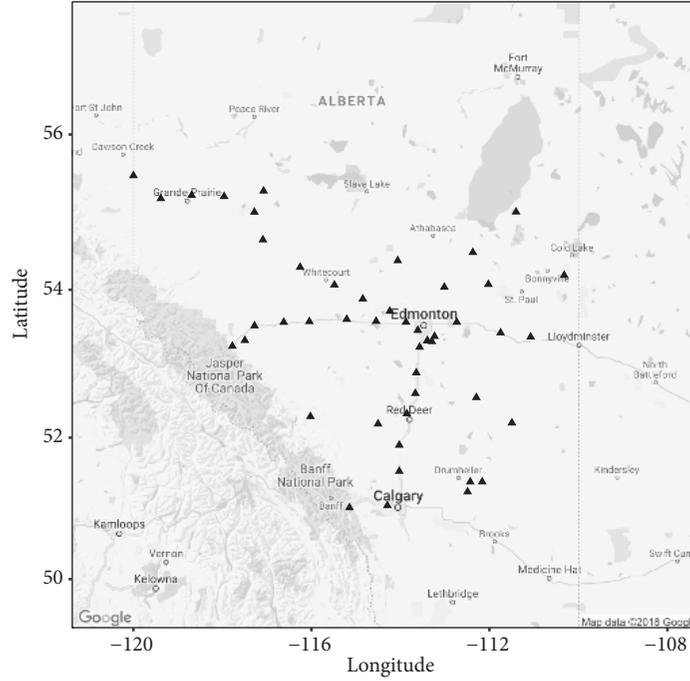


FIGURE 3: Data source: stationary RWIS ESSs.

Next, the variogram estimator $\gamma(h_s, h_t)$ was calculated as shown in (1). Note that the residuals $r(s_k, t_k)$ from the last step were regarded as measurements $z(s_k, t_k)$ in (1). The empirical variogram was obtained in this step.

Then, variogram models were fitted. The typical variograms of separable, product-sum, and metric covariance models are given by the following equations, respectively [38].

$$\gamma_{st}(h_s, h_t) = \text{sill} \cdot (\gamma_s(h_s) + \gamma_t(h_t) - \gamma_s(h_s) \cdot \gamma_t(h_t)) \quad (4)$$

where γ_{st} is the modelled spatiotemporal variogram, γ_s is the spatial variogram, γ_t is the temporal variogram, and sill is the overall sill.

$$\gamma_{st}(h_s, h_t) = (k \cdot \text{sill}_t + 1) \cdot \gamma_s(h_s) + (k \cdot \text{sill}_s + 1) \cdot \gamma_t(h_t) - k \cdot \gamma_s(h_s) \cdot \gamma_t(h_t) \quad (5)$$

where k is positive and sill_s and sill_t are sills in space and time.

$$\gamma_{st}(h_s, h_t) = \gamma_j \left(\sqrt{h_s^2 + (\kappa \cdot h_t)^2} \right) \quad (6)$$

where γ_j is any known variogram including a nugget effect and κ is the space-time anisotropy ratio.

When spatiotemporal variogram models were constructed, the model parameters, e.g., spatial and temporal ranges, were extracted from the models. In addition, for interpolation using Kriging methods, the residual component was then interpolated on a fine grid and added to the mean component to generate a fine grid map.

3.4. Cross Validation. To quantify the goodness-of-fit between the fitted models and the empirical variogram, the authors chose MSE as the measure of effectiveness. MSE allowed the authors to evaluate the overall difference between surfaces. MSE was calculated by the following equation. After comparing MSE among different models, the model with the best goodness-of-fit was selected.

$$\text{MSE} = \frac{1}{n \cdot m} \sum_{i=1}^n \sum_{j=1}^m (\gamma_{st}(h_{s,i}, h_{t,j}) - \gamma(h_{s,i}, h_{t,j}))^2 \quad (7)$$

where $\gamma_{st}(h_{s,i}, h_{t,j})$ and $\gamma(h_{s,i}, h_{t,j})$ are modelled and actual spatiotemporal variogram at spatial distance $h_{s,i}$ and temporal distance $h_{t,j}$, respectively.

4. Case Study

4.1. Study Site. The study area was located in the province of Alberta in Canada. Alberta has a humid continental climate, which often produces extremely cold winters [39]. The traffic problems brought by snowfalls in severe winters are always one of the major concerns of road authorities [21]. To acquire real-time road weather and surface condition data, Alberta is one of the leading provinces in Canada building an RWIS network. With innovative sensors and cameras, the stationary RWISs collect detailed road weather and surface conditions, which can be inputs for winter maintenance operations and weather impact warning applications. In this study, data from stationary RWIS ESSs in the vicinity of Edmonton were selected. Figure 3 schematically shows the locations of the 48 stationary RWIS ESSs selected. Stationary RWISs can record

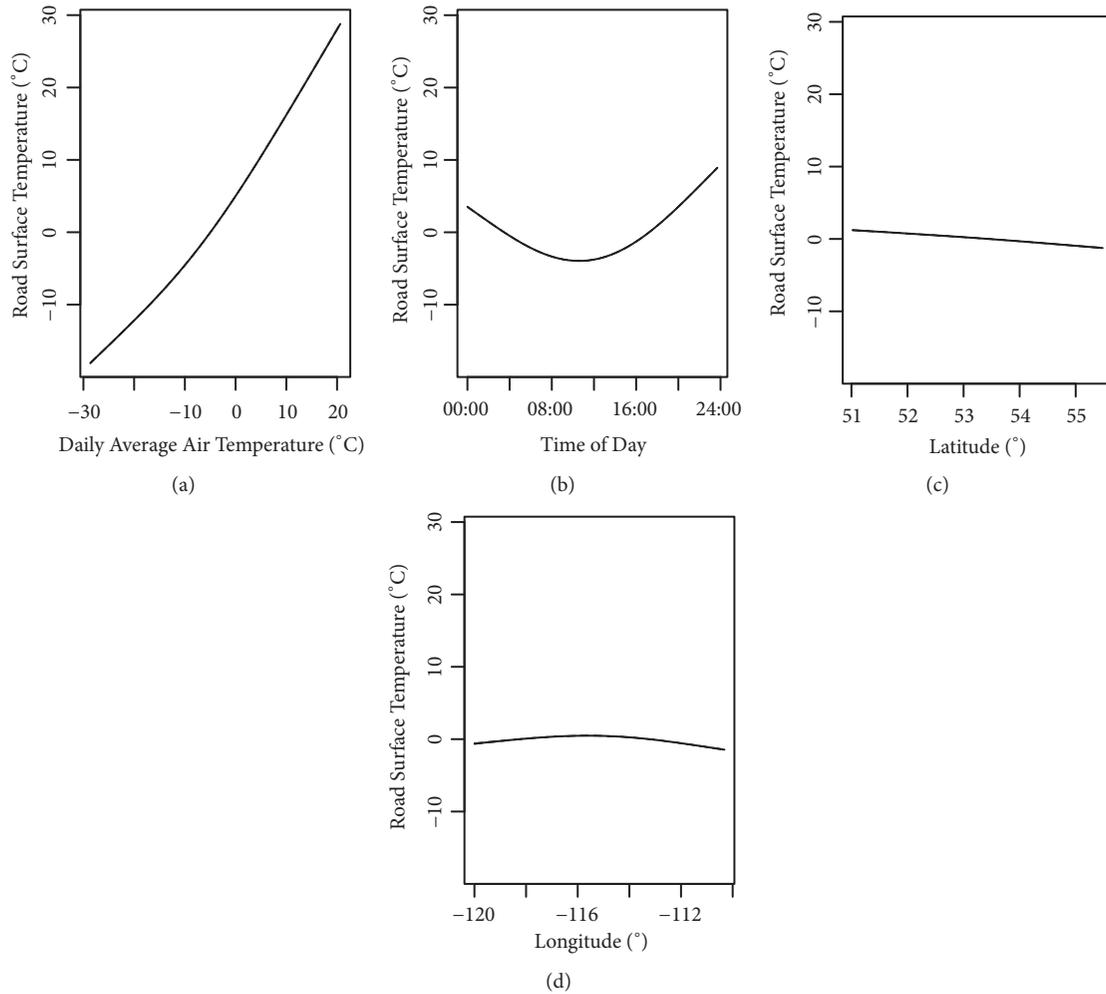


FIGURE 4: Nonlinear associations between RST and predictors. (a) Daily average air temperature. (b) Time of day. (c) Latitude. (d) Longitude.

both road- and weather-related data at prespecified temporal resolutions. The data they measure includes latitude, longitude, elevation, air temperature, humidity, precipitation, surface temperature, and wind speed. RST is one of the most important parameters for road condition prediction (e.g., black ice). Road conditions are sensitive to RST, so this research took it as a measure for estimation. RWIS datasets were available from 2014 to 2016. During this period, data from winter seasons (from October to March) was chosen.

4.2. Data Quality Diagnostics. In the dataset used for this study, the data missing ratio was 1.76% and the erroneous ratio was 6.59%. To identify erroneous RST measurements in the model analysis test, specific GAM model predictors were selected: daily average air temperature, time of day and the latitude and longitude of the stations. The “mgcv” package in the R software (version 3.4.3) [40] was used to fit the GAM model. The nonlinear relationships obtained from the GAM model to identify the acceptable regions are shown in Figure 4. The RST measurements that exceeded $\pm 20\%$ of the model estimation were defined as erroneous data.

The monthly sample size after data quality check is exhibited in Figure 5(a). Each month has 86,498 RST measurements on average. The descriptive statistics for monthly RST data are plotted in Figure 5(b). In Figure 5(b), the central mark on each box is the median value and the edges are the 25th and 75th percentiles. The individual data points beyond the whiskers are outliers. It can be observed that there is an obvious seasonal trend between shoulder months (October and March) and winter months (from November to February).

4.3. Spatiotemporal Coverage. The spatiotemporal variogram modelling was coded in the R software program [40]. The “gstat” package was used to construct and fit the spatiotemporal variogram models. The original data was detrended to obtain residuals by removing the local mean values. The empirical spatiotemporal variogram for each month was built from the variance of difference between any two spatiotemporal residual points. To fit the empirical spatiotemporal variogram, some initial guess of the model parameters was taken from pure spatial variograms and time

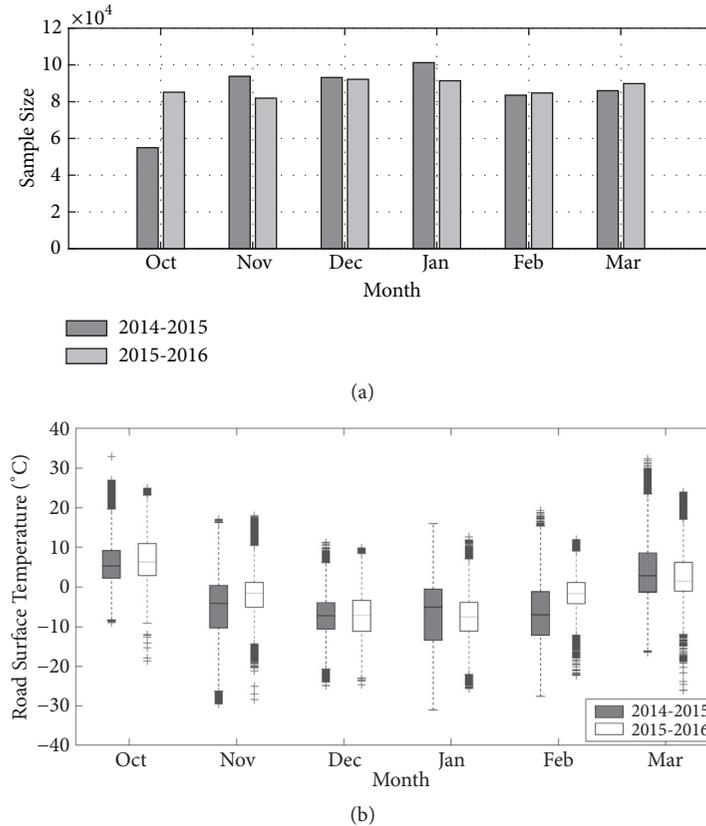


FIGURE 5: Data summary after data quality check. (a) Monthly sample size. (b) Boxplot of the processed dataset.

series analysis. Then the spatiotemporal variograms were fitted using the “fit.StVariogram” function. An algorithm in the family of quasi-Newton methods [41], called L-BFGS-B (Limited-memory Broyden–Fletcher–Goldfarb–Shanno-B) optimization algorithm, estimated the model parameters.

For the monthly empirical variogram, three spatiotemporal variograms were fitted: separable, product-sum, and metric variograms. The fitted model parameters are listed in Table 1. Among the models, the product-sum model gave the smallest MSE compared to others. Therefore, only the product-sum model was used to fit the variability structure of RST in the following analysis. Figure 6 shows the empirical and fitted product-sum variograms for November 2014 and January 2015. The empirical variograms visualize the spatiotemporal variability of RST data. The parameters from the fitted variograms quantify the spatiotemporal patterns and help to estimate RST in any unobserved space-time point.

The variograms in Figure 6 illustrate the correlation between any two pairs of spatiotemporal data points: the higher the variogram γ_{st} , the higher the correlation. The values for the range model parameter were extracted from the fitted variograms. The data was uncorrelated when the spatial distance or time lag between data points went beyond the ranges. Thus, the ranges were related to data representativeness and can be interpreted as the spatiotemporal coverage. In Figure 6, the lines on the front and side planes are the pure spatial variogram at the zero time lag and temporal variograms at the zero spatial distance, respectively.

As observed from Table 1 and Figure 6, the model parameters differ between months. Thus, the model parameters were extracted from monthly fitted spatiotemporal variograms to examine whether seasonal trends exist. Figure 7 illustrates the monthly variations of spatial and temporal ranges around the winter seasons from 2014 to 2016. The ranges indicate that RST measurements were correlated within a distance on the spatial scale and within a time lag on the temporal scale. The maximum and minimum ranges in space were 7.37 km (March 2015) and 35.55 km (January 2016), and those in time were 6.81 h (February 2016) and 16.90 h (January 2015). Although ranges varied across months and years, it is still obvious that the data from the months suffering from low temperatures had higher spatial and temporal coverage. RST measurements from RWIS in the winter months (December, January, and February) generally represented more in space and time than the shoulder months (October and March). This phenomenon is reasonable and related with meteorological variations. The severe winter months in Alberta, Canada, experience extremely cold temperatures during which the RST stays in a low level steadily. Thus, each RST measurement in winter months can cover more in space and time. On the contrary, shoulder months witness the transition between seasons. The meteorological variations in either space or time are larger than in other months. As a result, one RST measurement point in shoulder months (October and March) represents road surface conditions in a small coverage of space and time.

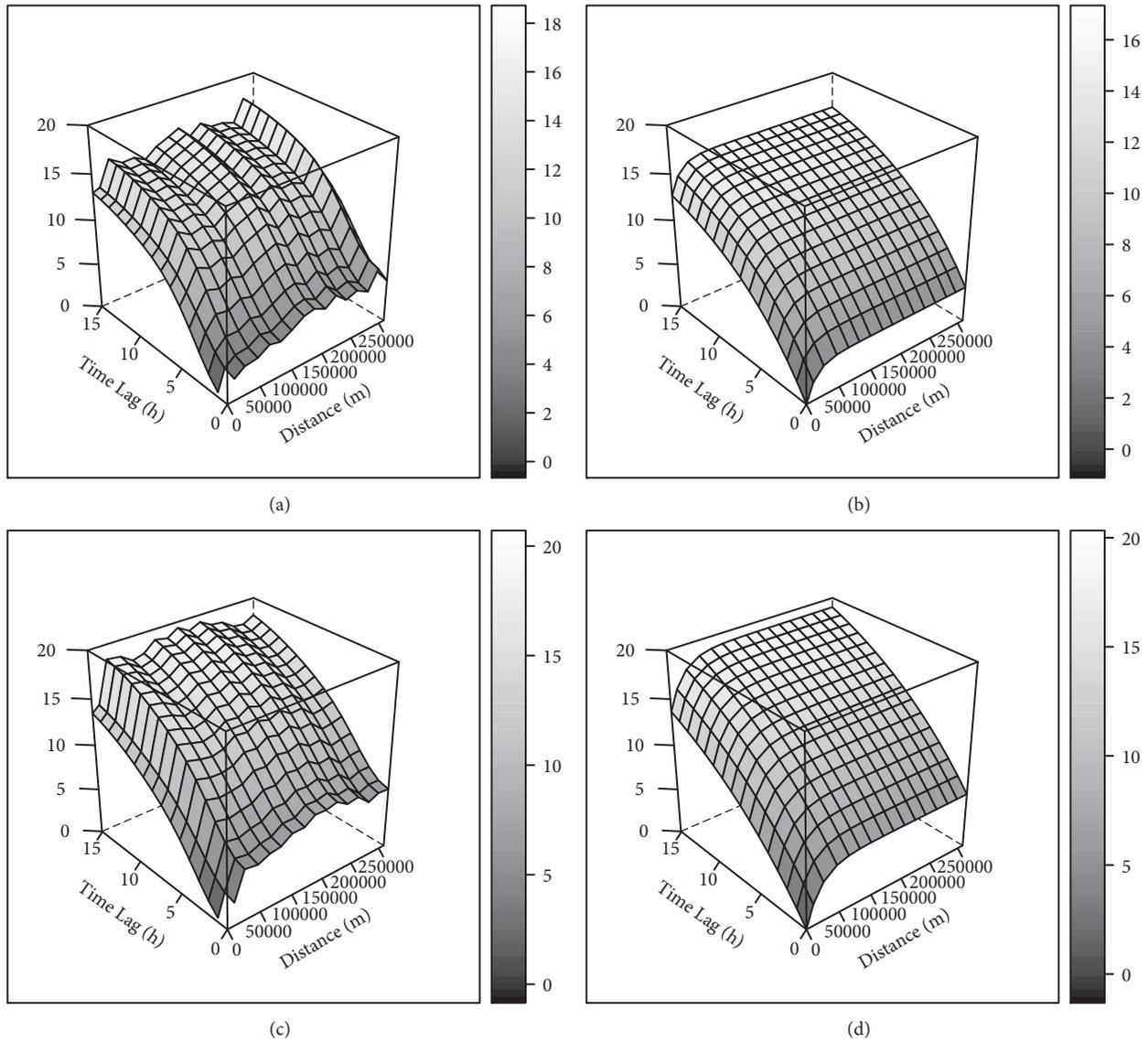


FIGURE 6: Spatiotemporal variograms. (a) Empirical variogram of Nov 2014. (b) Fitted variogram of Nov 2014. (c) Empirical variogram of Jan 2015. (d) Fitted variogram of Jan 2015.

The spatiotemporal coverage results provide a reference for decision makers to plan an RWIS network and hence undertake necessary road maintenance and traffic management. The spatiotemporal coverage from variograms is the maximum distance and time lag, beyond which the measurements are no longer representative. In practice, apart from the coverage values from variograms, lower coverage values can be defined to guarantee higher acceptable estimation accuracy.

As previously discussed, RST was chosen without loss of generality, but other road surface and weather indicators can also follow the same procedure to determine their respective spatiotemporal coverage. Once the spatiotemporal coverage is finalized, the density of stationary RWIS and the frequency of mobile RWIS can be specified. On the one hand, as data from stationary RWIS ESSs are discrete in the space

domain, the spatial coverage is an essential factor that needs to be considered in planning a stationary RWIS network. Since the spatiotemporal variogram gives a spatial range (SR), measurements from a stationary RWIS ESS can be regarded representative in the distance of SR . Then, along the target roadway, at least one RWIS is required to be allocated every $SR \times 2$. On the other hand, mobile RWIS provides discrete data in time, so the temporal coverage is a prerequisite in scheduling mobile RWIS vehicles. Using the temporal range (TR) from the variogram, measurements from a mobile RWIS can be regarded as representative in a time interval of TR . To construct a database with complete temporal information, one mobile RWIS vehicle should be scheduled to collect road weather data at least every $TR \times 2$. Moreover, to predict temporal changes in the near future, the time headway between two vehicles should be less than TR .

TABLE 1: Fitted model parameters for the spatiotemporal variogram. (a) Separable covariance model. (b) Product-sum covariance model. (c) Metric covariance model.

| (a) | | | | | | |
|---------------|-------------------------|---------------------------------|------------------------|---------------------------------|---------------------------------------|------------------------------|
| | Space | | Time | | Overall Sill [$^{\circ}\text{C}^2$] | MSE [$^{\circ}\text{C}^2$] |
| | Range [km^a] | Nugget [$^{\circ}\text{C}^2$] | Range [h^b] | Nugget [$^{\circ}\text{C}^2$] | | |
| November 2014 | 14.23 | 0.00 | 5.14 | 0.00 | 11.81 | 13.06 |
| January 2015 | 20.73 | 0.00 | 6.90 | 0.00 | 13.70 | 13.48 |

| (b) | | | | | | | | |
|---------------|------------|---------------------------------|-------------------------------|-----------|---------------------------------|-------------------------------|------|------------------------------|
| | Space | | | Time | | | k | MSE [$^{\circ}\text{C}^2$] |
| | Range [km] | Nugget [$^{\circ}\text{C}^2$] | Sill [$^{\circ}\text{C}^2$] | Range [h] | Nugget [$^{\circ}\text{C}^2$] | Sill [$^{\circ}\text{C}^2$] | | |
| November 2014 | 14.23 | 0.01 | 3.41 | 8.09 | 0.01 | 14.55 | 0.01 | 0.80 |
| January 2015 | 20.73 | 0.01 | 5.17 | 14.50 | 0.01 | 19.87 | 0.01 | 0.60 |

| (c) | | | | |
|---------------|--------|---------------------------------|-------------------------------|------------------------------|
| | Range | Nugget [$^{\circ}\text{C}^2$] | Sill [$^{\circ}\text{C}^2$] | MSE [$^{\circ}\text{C}^2$] |
| November 2014 | 532178 | 0.68 | 16.40 | 1.36 |
| January 2015 | 640430 | 2.69 | 18.84 | 2.74 |

^akm = kilometre.

^bh = hour.

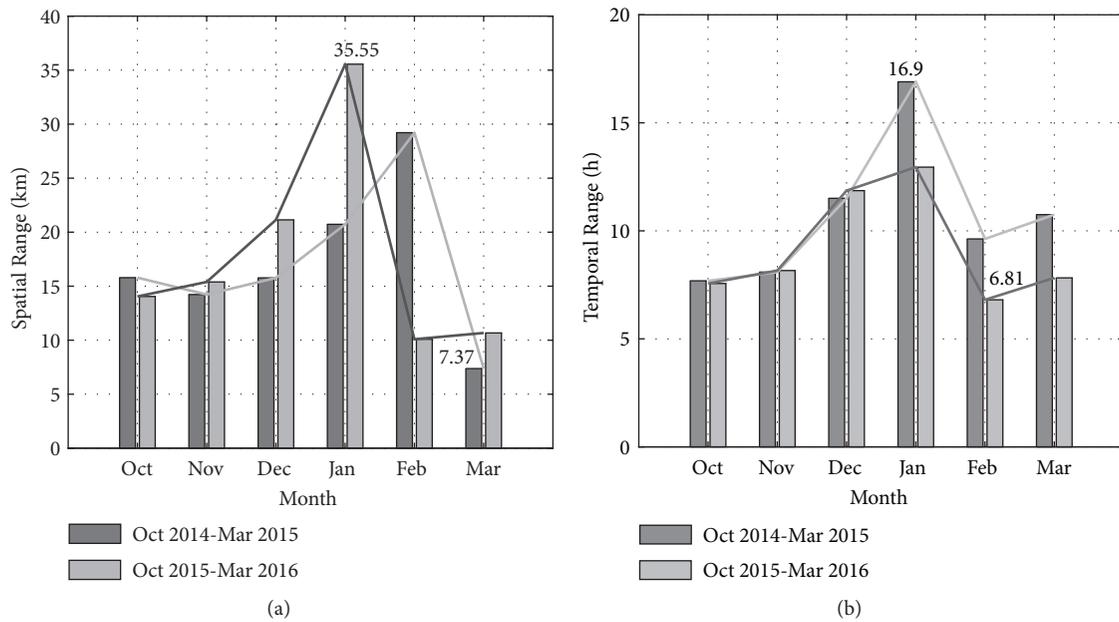


FIGURE 7: Monthly range variations from spatiotemporal variogram. (a) Spatial. (b) Temporal.

Similarly, in a CV environment, only when the penetration rate of CV-enabled vehicles fulfills a full temporal coverage, the CV data can then be interpolated to generate a fine map that shows spatiotemporal variability of road surface and weather conditions.

5. Conclusions and Future Research

Reliable road maintenance and traffic management call for road surface and weather data collection with high coverage;

however, RWISs collect road and weather data at a high cost. To help in planning a cost-effective RWIS network, a quantification method was developed to determine spatiotemporal coverage of RWIS data. The proposed method was assessed using authentic RWIS data. There are several key results from this research:

(a) Due to technical problems, the measurements that the studied RWIS ESSs collected contained a certain percentage of missing and erroneous data. Data quality diagnostics and imputation should be given great consideration.

(b) Among the three spatiotemporal variogram models, the product-sum model outperforms the others. The spatial and temporal range values from the model identify the spatiotemporal coverage of RWIS data. The spatiotemporal coverage can be used to decide the RWIS distributions to achieve a fine understanding of the spatiotemporal variability.

(c) A seasonal trend was found behind the monthly spatiotemporal coverage. Data from winter months covers more in space and time than data from shoulder months. The seasonal trend should be considered in the RWIS network design and operation.

(d) Analysis of spatiotemporal semivariograms suggested a stationary RWIS ESS should be located at least every 71.1 km, while a mobile RWIS vehicle should be deployed at least every 33.8 h in the Alberta case.

Based on the observations from this study, the proposed method can be applied in assigning an influence radius of RWIS ESS observations in various road weather-related applications. The current study can be extended in several directions. First, spatial stratified heterogeneity of the demand surface will be tested and attributed accordingly [42, 43]. In our present study, semivariograms are constructed by assuming a stationary process, which means the underlying spatiotemporal structures are translation-invariant over space and time. This rather strong assumption might not hold true when used in large regions with high spatial (i.e., landscape) and temporal (i.e., season) heterogeneity. Second, since semivariogram parameters are sensitive to the property of population, the way of sampling, and the method of estimation, further investigation is required on the spatial sampling trinity to examine the conclusiveness of the findings presented in our study. This will be particularly important when the study is extended to the stage of optimal RWIS network planning. Third, the stationary and mobile RWIS data will be integrated and fused to construct a database with high coverage in both space and time. Data from different sources may not match even in the same time and location. It is challenging to fuse data among multiple data sources and extend their temporal and spatial coverage. A practical, accurate, and time-efficient data integration and fusion method is still required for efficient and reliable road surface and weather condition estimation. Furthermore, the spatiotemporal coverage defined in this study could possibly vary depending upon several external factors, such as geographical and topological characteristics of the site under investigation. Hence, more case studies are essential to obtain more conclusive results.

Data Availability

The RWIS data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this paper.

Acknowledgments

This research work was jointly supported by the National Natural Science Foundation of China (61703236), Shandong Provincial Natural Science Foundation, China (ZR2017QF014, ZR2018MF027), and China Postdoctoral Science Foundation Funded Project (2017M612275).

References

- [1] Federal Highway Administration, "How do weather events impact roads?" https://ops.fhwa.dot.gov/weather/ql_roadimpact.htm.
- [2] R. M. Alfelori and C. Y. D. Yang, *Managing Traffic Operations during Adverse Weather Events*, Federal Highway Administration, Washington, Wash, USA, 2011, <https://www.fhwa.dot.gov/publications/publicroads/11janfeb/01.cfm>.
- [3] G. Liu, Y. Ge, T. Z. Qiu, and H. R. Soleymani, "Optimization of snow plowing cost and time in an urban environment: A case study for the city of Edmonton," *Canadian Journal of Civil Engineering*, vol. 41, no. 7, pp. 667–675, 2014.
- [4] S. Choi and C. Oh, "Proactive strategy for variable speed limit operations on freeways under foggy weather conditions," *Transportation Research Record*, vol. 2551, pp. 29–36, 2016.
- [5] V. R. Tomas, M. Pla-Castells, J. J. Martinez, and J. Martinez, "Forecasting adverse weather situations in the road network," *IEEE Transactions on Intelligent Transportation Systems*, vol. 17, no. 8, pp. 2334–2343, 2016.
- [6] P. A. Pisano and G. G. Nelson, "Advanced decision support for winter road maintenance: FHWA documentation of requirements for intelligent transportation systems," *Transportation Research Record*, no. 1741, pp. 129–136, 2001.
- [7] S. Drobot, M. Chapman, P. A. Pisano, and B. B. McKeever, "Using vehicles as mobile weather platforms," *Advances in Intelligent and Soft Computing*, vol. 81, pp. 203–213, 2010.
- [8] T. Kwon, L. Fu, and C. Jiang, "Effect of winter weather and road surface conditions on macroscopic traffic parameters," *Transportation Research Record*, no. 2329, pp. 54–62, 2013.
- [9] M. A. Linton and L. Fu, "Connected vehicle solution for winter road surface condition monitoring," *Transportation Research Record*, vol. 2551, pp. 62–72, 2016.
- [10] K. C. Dey, A. Mishra, and M. Chowdhury, "Potential of intelligent transportation systems in mitigating adverse weather impacts on road mobility: A review," *IEEE Transactions on Intelligent Transportation Systems*, vol. 16, no. 3, pp. 1107–1119, 2015.
- [11] L. Gu, T. J. Kwon, and T. Z. Qiu, "A geostatistical approach to winter road surface condition estimation using mobile RWIS data," in *Proceedings of the 97th Annual Meeting of Transportation Research Board*, 2018.
- [12] W. P. Mahoney and W. L. Myers, "Predicting weather and road conditions: integrated decision-support tool for winter road-maintenance operations," *Transportation Research Record*, vol. 1824, pp. 98–105, 2003.
- [13] C. Cluett, F. M. Kitchener, L. Osborne, and S. Conger, "Weather integration in transportation management centers," *Transportation Research Record*, no. 1978, pp. 201–208, 2006.
- [14] US Department of Transportation, "Connected Vehicle Pilot Deployment Program: Wyoming," 2017, https://wydotcyp.wyroad.info/assets/promotion/WyomingCV_Pilot_Factsheet_v2_020817.pdf.

- [15] K. R. Petty and W. P. Mahoney, "Enhancing road weather information through vehicle infrastructure integration," *Transportation Research Record*, vol. 2015, pp. 132–140, 2007.
- [16] S. Drobot, M. Chapman, E. Schuler et al., "Improving road weather hazard products with vehicle probe data," *Transportation Research Record*, vol. 2169, no. 1, pp. 128–140, 2010.
- [17] M. Chapman, S. Drobot, T. Jensen et al., "Diagnosing road weather conditions with vehicle probe data," *Transportation Research Record*, vol. 2169, no. 1, pp. 116–127, 2010.
- [18] B. Boyce, G. Weiner, A. Anderson, and S. Linden, *Pikalert® system vehicle data translator (VDT) utilizing integrated mobile observations*, U.S. Department of Transportation, Washington, Wash, USA, 2017.
- [19] Federal Highway Administration, "FHWA road weather management program update," 2018, https://sicop.transportation.org/wp-content/uploads/sites/36/2018/07/RWMP_ProgramUpdate_July2018.pdf.
- [20] J. Manfredi et al., *Road weather information system environmental sensor station siting guidelines*, U.S. Department of Transportation, Washington, Wash, USA, 2005.
- [21] D. Mackinnon and A. Lo, *Alberta transportation road weather information system (RWIS) expansion study*, Alberta Transportation, Edmonton, Alberta, Canada, 2009.
- [22] A. K. Singh, J. Li, M. Murphy, and C. M. Walton, "Optimizing environmental sensor station locations for road weather management: overview and a generalized modeling framework," *Transportation Research Record*, vol. 2551, no. 1, pp. 18–28, 2018.
- [23] L. Zhao, S. I. Chien, X. Liu, and W. Liu, "Planning a road weather information system with GIS," *Journal of Modern Transportation*, vol. 23, no. 3, pp. 176–188, 2015.
- [24] L. Zhao, S. Chien, J. Meegoda, Z. Luo, and X. Liu, "Cost-benefit analysis and microclimate-based optimization of a RWIS network," *Journal of Infrastructure Systems*, vol. 22, no. 2, 2016.
- [25] T. J. Kwon and L. Fu, "Spatiotemporal variability of road weather conditions and optimal RWIS density — an empirical investigation," *Canadian Journal of Civil Engineering*, vol. 44, no. 9, pp. 691–699, 2017.
- [26] T. J. Kwon, L. Fu, and S. J. Melles, "Location optimization of road weather information system (RWIS) network considering the needs of winter road maintenance and the traveling public," *Computer-Aided Civil and Infrastructure Engineering*, vol. 32, no. 1, pp. 57–71, 2017.
- [27] Y. Yang, J. Wu, and G. Christakos, "Prediction of soil heavy metal distribution using Spatiotemporal Kriging with trend model," *Ecological Indicators*, vol. 56, pp. 125–133, 2015.
- [28] L. Li, J. Zhang, W. Qiu, J. Wang, and Y. Fang, "An ensemble spatiotemporal model for predicting PM_{2.5} concentrations," *International Journal of Environmental Research and Public Health*, vol. 14, no. 5, article 549, 2017.
- [29] D. Hu, H. Shu, H. Hu, and J. Xu, "Spatiotemporal regression Kriging to predict precipitation using time-series MODIS data," *Cluster Computing*, vol. 20, no. 1, pp. 347–357, 2017.
- [30] S. O. Ahmed, R. Mazloum, and H. Abou-Ali, "Spatiotemporal interpolation of air pollutants in the Greater Cairo and the Delta, Egypt," *Environmental Research*, vol. 160, pp. 27–34, 2018.
- [31] R. A. Olea, *Geostatistics for Engineers and Earth Scientists*, Springer, Boston, Massachusetts, Mass, USA, 1999.
- [32] P. C. Kyriakidis and A. G. Journel, "Geostatistical space-time models: a review," *Mathematical Geology*, vol. 31, no. 6, pp. 651–684, 1999.
- [33] P. W. Gething, P. M. Atkinson, A. M. Noor, P. W. Gikandi, S. I. Hay, and M. S. Nixon, "A local space-time kriging approach applied to a national outpatient malaria data set," *Computers & Geosciences*, vol. 33, no. 10, pp. 1337–1350, 2007.
- [34] S. Shekhar, H. Xiong, and X. Zhou, *Encyclopedia of GIS*, Springer, New York, NY, USA, 2008.
- [35] K. T. Chang, *Introduction to Geographic Information Systems*, McGraw-Hill, New York, NY, USA, 5th edition, 2010.
- [36] M. A. Oliver and R. Webster, "Kriging: A method of interpolation for geographical information systems," *International Journal of Geographical Information Science*, vol. 4, no. 3, pp. 313–332, 1990.
- [37] T. J. Hastie and R. J. Tibshirani, *Generalized Additive Models*, vol. 43, Chapman and Hall, London, UK, 1990.
- [38] E. Pebesma and G. Heuvelink, "Spatio-temporal interpolation using gstat," *RFID Journal*, vol. 8, no. 1, pp. 204–218, 2016.
- [39] Government of Alberta, *Agroclimatic Atlas of Alberta: Climate of Alberta*, 2003, [https://www.lagric.gov.ab.ca/\\$department/deptdocs.nsf/all/sag6299](https://www.lagric.gov.ab.ca/$department/deptdocs.nsf/all/sag6299).
- [40] R Development Core Team, *A Language and Environment for Statistical Computing*, R Foundation for Statistical Computing, Vienna, Austria, 2017.
- [41] C. Zhu, R. H. Byrd, P. Lu, and J. Nocedal, "Algorithm 778: L-BFGS-B: fortran subroutines for large-scale bound-constrained optimization," *ACM Transactions on Mathematical Software*, vol. 23, no. 4, pp. 550–560, 1997.
- [42] J. Wang, A. Stein, B. Gao, and Y. Ge, "A review of spatial sampling," *Spatial Statistics*, vol. 2, no. 1, pp. 1–14, 2012.
- [43] J. F. Wang, T. L. Zhang, and B. J. Fu, "A measure of spatial stratified heterogeneity," *Ecological Indicators*, vol. 67, pp. 250–256, 2016.

Research Article

Double Magnetic Loop and Methods for Calculating Its Inductance

Ferran Mocholí Belenguer ¹, Antonio Mocholí Salcedo,²
Victor Milián Sánchez,³ and José Humberto Arroyo Núñez⁴

¹Traffic Control Systems Group, ITACA Institute, Universitat Politècnica de València, Valencia 46022, Spain

²Department of Electronic Engineering, ITACA Institute, Universitat Politècnica de València, Valencia 46022, Spain

³Chemical and Nuclear Engineering Department, Institute of Industrial, Radiological and Environmental Safety, Universitat Politècnica de València, Valencia 46022, Spain

⁴Electronic and Telecommunications Engineering Department, Universidad Politécnica de Tulancingo, Tulancingo 43629, Mexico

Correspondence should be addressed to Ferran Mocholí Belenguer; fermocbe@upv.es

Received 12 June 2018; Accepted 6 September 2018; Published 25 September 2018

Guest Editor: Darius Andriukaitis

Copyright © 2018 Ferran Mocholí Belenguer 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.

Due to their simplicity and operating mode, magnetic loops are one of the most used traffic sensors in Intelligent Transportation Systems (ITS). However, at this moment, their potential is not being fully exploited, as neither the speed nor the length of the vehicles can be surely ascertained with the use of a single magnetic loop. In this way, the vast majority of them are only being used to count vehicles on urban and interurban roads. For this reason, in order to contribute to the development of new traffic sensors and make roads safer, this paper introduces a theoretical study to explain the design and peculiarities of the innovative double loops, how to calculate their magnetic field and three different methods to calculate their inductance. Finally, the different inductance values obtained by these three methods will be analyzed and compared with experimental measurements carried out by our research group in order to know which method is more accurate and if all of them are equally reliable.

1. Introduction

Magnetic loops are the most common sensors on roads around the world since they are an affordable and highly developed technology with a simple operation that is not affected by environmental conditions [1–7]. Although these ones imply to drill and work on the road for their installation and possible future repairs like the rest of intrusive sensors [8], in practice, magnetic loops still have a long future ahead. Even though they might seem outdated, these are actually a widely extended and well-known reliable technology that offers good performance at a low price. Proof of this is that today they continue to be installed on the roads and they are even fundamental elements in the new algorithms for traffic management [9–15].

Their operation is straightforward, since it is based on the impedance variation that is recorded in the magnetic loops

during the passage of vehicles over them, and as shown in Figure 1, an entire system usually consists of three parts [16]:

- (I) A magnetic loop formed by a wire with one or more turns superficially buried in the pavement.
- (II) A cable that links the magnetic loop with the control booth, which is also buried in the pavement.
- (III) An electronic unit located in the control booth that contains an oscillator and amplifiers to excite the inductive loop.

In order to have a better understanding of how they work, there are many publications and bibliography [1] since they are one of the most widespread sensors. However, a brief physical explanation is provided in the following points:

- (i) The electronic unit together with the magnetic loop forms an oscillator circuit.

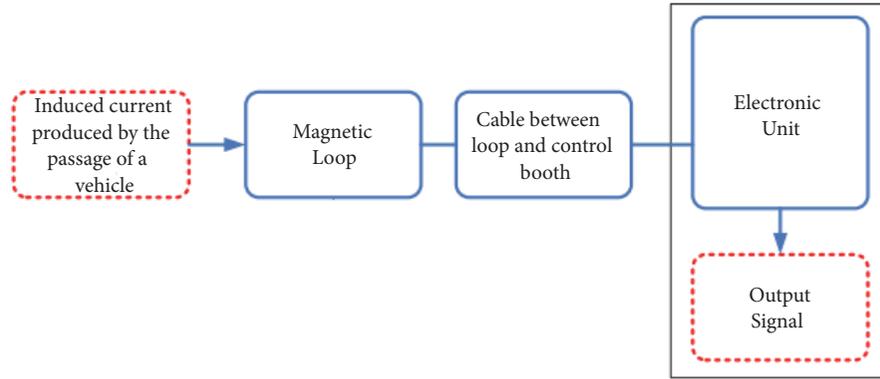


FIGURE 1: Magnetic loop system scheme.

- (ii) The current which passes through the loop produces a magnetic field \vec{H} around the cable as shown in (1), where N is the number of turns of the loop, I is the current expressed in Amperes, and l is the length of the loop expressed in meters.

$$\vec{H} = \frac{N \cdot I}{l} \quad (1)$$

- (iii) This magnetic field \vec{H} produces a magnetic flux \emptyset through the loop as shown in (2), where \vec{B} is the magnetic flux density, \vec{S} is the surface enclosed by the loop, μ_r is the relative magnetic permeability of the medium, and μ_o is a constant value ($4\pi \cdot 10^{-7} \text{ N / A}^2$).

$$\emptyset = \vec{B} \cdot \vec{S} = \mu_r \cdot \mu_o \cdot \vec{H} \cdot \vec{S} \quad (2)$$

- (iv) The result is that the inductance of a common single loop L expressed in Henrys is obtained as follows:

$$L = \frac{N \cdot \emptyset}{I} = \frac{N \cdot \vec{B} \cdot \vec{S}}{I} \quad (3)$$

In this way, when a vehicle or any object built with ferromagnetic materials passes through the magnetic field generated by a magnetic loop buried on the road with a surface area \vec{S} , a number of turns N , and a current intensity I as shown in Figure 2, there is a decrease in the global magnetic field because of the currents that are induced in the vehicle.

As seen in (3), the loop inductance is proportional to the magnetic flux, which causes that when passing a vehicle over it, the inductance also decreases. Moreover, like any oscillator circuit, the oscillation frequency of the whole system will be given by

$$f = \frac{k}{L} \quad (4)$$

where k is a constant that depends on the characteristics of the electronic components used in the construction of the

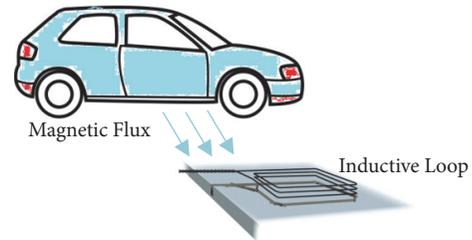


FIGURE 2: Magnetic loop operation mode.

oscillator circuit. Thus, when a vehicle passes over a loop, we can obtain what is commonly known as “the vehicle magnetic profile” or “the vehicle inductance signature” by analyzing the inductance or frequency variation recorded. This magnetic profile is different for each type of vehicle as seen in Figure 3, which allows classifying the different vehicles as motorcycles, cars, trucks and buses. However, while the vehicle magnetic profiles for the single loops are widely known, the magnetic profiles left by the passage of vehicles over the double loops have not yet been studied, although we can anticipate that these new magnetic profiles will offer much more information than the previous ones.

To estimate the vehicle speed and classify them, nowadays it is necessary to use two single loops, since a single one is not able to get all the necessary parameters to do it. After analyzing the magnetic profile, there are two unknown data (vehicle length and vehicle speed) with the variation of a single parameter (inductance or oscillation frequency).

For that reason, there are usually two loops per lane separated by a certain distance. In this manner, the passage of a vehicle over the first loop is recorded in the detector, and after a short interval of time, the vehicle passes again over the second loop where it is also recorded [8, 17]. Then, as the distance between both loops is known by design, the vehicle speed, the direction of traffic and the vehicle length can be finally estimated, as well as the vehicle axle detection [14].

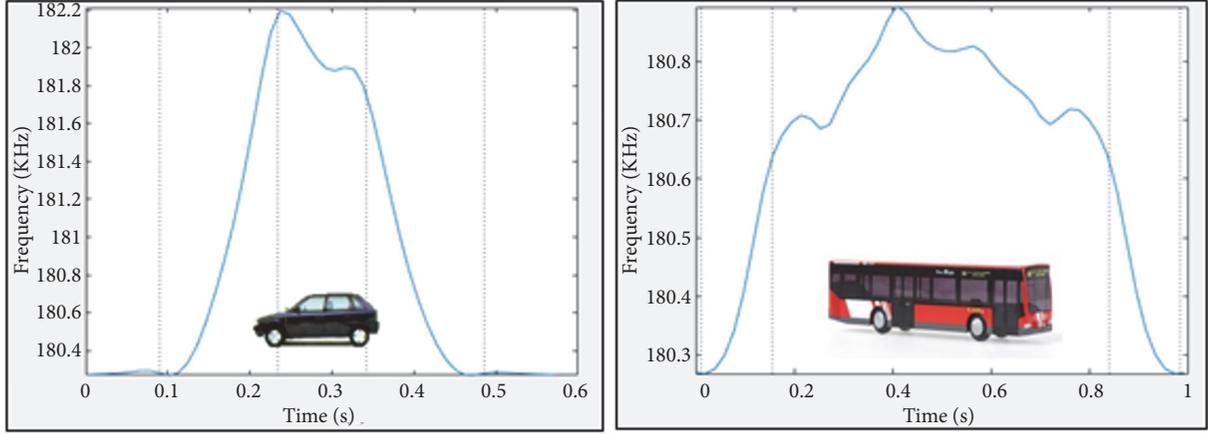


FIGURE 3: Magnetic profiles. (a) Car. (b) Bus.

Nevertheless, with the use of the double loops this problem would be solved and it would only be necessary to use one loop to find out all the previous data, since they have a simpler, more compact and more economical electronics. Moreover, having a single signal instead of two would facilitate the implementation of the measurement system.

Therefore, our work will aim to present and describe the characteristics of the double loop, to offer different methods to calculate its inductance, to verify which one provides better results and to improve the functional characteristics of the popular single loops, which despite being the most installed sensor on the roads, they are actually only dedicated to count vehicles. The presentation of the new vehicle magnetic profiles, the parameters that can be extracted with them and the advantages offered over the conventional loop will be the subject of the following paper.

2. Electromagnetic Analysis

The design, shape, and construction of a rectangular or circular single loop are well-known worldwide [3]. However, although every day there are more algorithms and new systems to regulate traffic based on magnetic loops [18, 19], the double loop capable of providing better performance than the current loops at lower cost is still unknown.

A double loop is no more than the union of two rectangular loops, which can have different dimensions and turns (not to be confused with two single loops spaced at a certain distance as described above). How to implement this type of loop can be quite varied, but a classic way to proceed would be to build the outer loop and then a smaller inner loop located at one extreme. However, another way to build it could be to construct a small loop and then put another small one next to the first one. In all cases, the direction of the current in each loop can be chosen with the aim of generating different types of configurations. At any rate, in our study we will present a general theoretical analysis capable of simulating any type of design. Therefore, in order to analyze this type of loop, the space will be divided into three sections as shown in Figure 4.

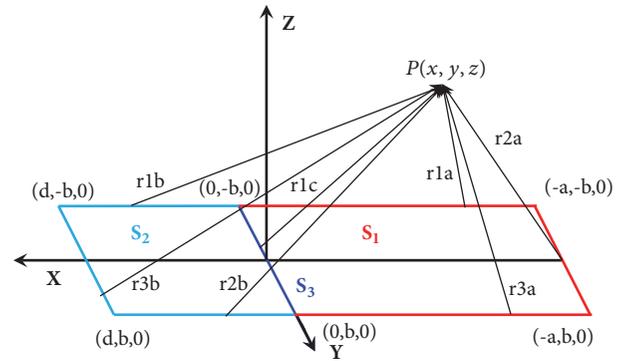


FIGURE 4: Double loop presented in three sections.

- (i) The first section, S_1 , corresponds to the three red segments located in the plane of the negative values of X . Two segments parallel to the X -axis with a length of a and one segment parallel to the Y -axis with a length of $2b$.
- (ii) The second section, S_2 , corresponds to the three turquoise segments located in the plane of the positive values of X . Two segments parallel to the X -axis with a length of d and one segment parallel to the Y -axis with a length of $2b$.
- (iii) The third section, S_3 , corresponds to the blue segment located on the Y -axis at $X = 0$ which has a length of $2b$.

In this way, point P where the magnetic field will be calculated will be determined by its coordinates (x, y, z) . Consequently and as shown in Figure 4, the distances from any point of the double loop (x', y', z') to the point of analysis of the magnetic field $P(x, y, z)$ will be defined as follows:

$$r1a = \sqrt{(x - x')^2 + (y + b)^2 + z^2}$$

$$r2a = \sqrt{(x + a)^2 + (y - y')^2 + z^2}$$

$$\begin{aligned}
r3a &= \sqrt{(x-x')^2 + (y-b)^2 + z^2} \\
r1b &= \sqrt{(x-x')^2 + (y+b)^2 + z^2} \\
r2b &= \sqrt{(x-x')^2 + (y-b)^2 + z^2} \\
r3b &= \sqrt{(x-d)^2 + (y-y')^2 + z^2} \\
r1c &= \sqrt{x^2 + (y-y')^2 + z^2}
\end{aligned} \tag{5}$$

In order to perform the electromagnetic analysis, the starting point will be the physical phenomenon of magnetic field generation due to the electrical currents flowing through a conductor. Maxwell's equations [20] revealed that the divergence of \vec{B} is zero:

$$\vec{\nabla} \cdot \vec{B} = 0 \tag{6}$$

This indicates that \vec{B} has the solenoidal property (with no divergence), which means that the magnetic field \vec{B} can be represented using an auxiliary vector function \vec{A} as follows:

$$\vec{B} = \vec{\nabla} \times \vec{A} \tag{7}$$

And thus

$$\vec{\nabla} \cdot (\vec{\nabla} \times \vec{A}) = 0 \tag{8}$$

The auxiliary vector \vec{A} is called magnetic potential vector [21] and it is related to the sources of the stable current density J which are responsible for generating the magnetic field. Then, in the case of a linear conductor, \vec{A} is given by

$$\vec{A} = \int_l \frac{\mu_0 I d\vec{l}}{4\pi r} \tag{9}$$

where I is the current in the linear conductor and r is the distance from the conductor to the analysis point. However, (9) represents a solution for the case of thin conductors, but the general solution must include a volumetric integral. In this way, the magnetic potential vector for each segment of our inductive double loop is obtained as follows:

Section S_1

$$\begin{aligned}
\vec{A}_{1s1x} &= \frac{\mu_0 I}{4\pi} \ln \frac{\sqrt{(x+a)^2 + (y+b)^2 + z^2} - a - x}{\sqrt{x^2 + (y+b)^2 + z^2} - x} \\
\vec{A}_{2s1y} &= \frac{\mu_0 I}{4\pi} \ln \frac{\sqrt{(x+a)^2 + (y-b)^2 + z^2} + b - y}{\sqrt{(x+a)^2 + (y+b)^2 + z^2} - b - y} \\
\vec{A}_{3s1x} &= \frac{\mu_0 I}{4\pi} \ln \frac{\sqrt{x^2 + (y-b)^2 + z^2} - x}{\sqrt{(x+a)^2 + (y-b)^2 + z^2} - a - x}
\end{aligned} \tag{10a}$$

Section S_2

$$\begin{aligned}
\vec{A}_{1s2x} &= \frac{\mu_0 I}{4\pi} \ln \frac{\sqrt{x^2 + (y+b)^2 + z^2} - x}{\sqrt{(x-d)^2 + (y+b)^2 + z^2} + d - x} \\
\vec{A}_{2s2x} &= \frac{\mu_0 I}{4\pi} \ln \frac{\sqrt{(x-d)^2 + (y-b)^2 + z^2} + d - x}{\sqrt{x^2 + (y-b)^2 + z^2} - x} \\
\vec{A}_{3s2y} &= \frac{\mu_0 I}{4\pi} \ln \frac{\sqrt{(x-d)^2 + (y+b)^2 + z^2} - b - y}{\sqrt{(x-d)^2 + (y-b)^2 + z^2} + b - y}
\end{aligned} \tag{10b}$$

Section S_3

$$\vec{A}_{1s3y} = \frac{\mu_0 I}{4\pi} \ln \frac{\sqrt{x^2 + (y-b)^2 + z^2} + b - y}{\sqrt{x^2 + (y+b)^2 + z^2} - b - y} \tag{10c}$$

In the previous expressions (A_{isjk}) it must be noted that, as shown in Figure 4, the subindex i refers to the considered segment of the section sj , which in turn corresponds to the k -axis. For example, \vec{A}_{3s2y} would mean the third segment of section S_2 located on the Y -axis, and the values a , b and d would correspond to the dimensions of the loop described above. However, before proceeding with the electromagnetic analysis of the double loop, we will define a series of terms to simplify the previous equations. These terms will be

$$\begin{aligned}
R_1 &= \sqrt{x^2 + (y+b)^2 + z^2} \\
R_2 &= \sqrt{(x-d)^2 + (y+b)^2 + z^2} \\
R_3 &= \sqrt{x^2 + (y-b)^2 + z^2} \\
R_4 &= \sqrt{(x-d)^2 + (y-b)^2 + z^2} \\
R_{1a} &= \sqrt{(x+a)^2 + (y+b)^2 + z^2} \\
R_{2a} &= R_1 \\
R_{3a} &= \sqrt{(x+a)^2 + (y-b)^2 + z^2} \\
R_{4a} &= R_3 \\
c_1 &= -x \\
c_2 &= d - x \\
d_1 &= -b - y \\
d_2 &= b - y \\
c_{1a} &= -a - x \\
c_{2a} &= c_1 \\
d_{1a} &= d_1 \\
d_{2a} &= d_2
\end{aligned} \tag{11}$$

In this way, combining (10a), (10b), and (10c) with (11), we can obtain the magnetic potentials vectors in a simpler way as seen in the following:

$$\begin{aligned}
\vec{A}_{1s1x} &= \frac{\mu_0 I}{4\pi} \ln \frac{R_{1a} + c_{1a}}{R_{2a} + c_{2a}} \\
\vec{A}_{2s1y} &= \frac{\mu_0 I}{4\pi} \ln \frac{R_{3a} + d_{2a}}{R_{1a} + d_{1a}} \\
\vec{A}_{3s1x} &= \frac{\mu_0 I}{4\pi} \ln \frac{R_{4a} + c_{2a}}{R_{3a} + c_{1a}} \\
\vec{A}_{1s2x} &= \frac{\mu_0 I}{4\pi} \ln \frac{R_1 + c_1}{R_2 + c_2} \\
\vec{A}_{2s2x} &= \frac{\mu_0 I}{4\pi} \ln \frac{R_4 + c_2}{R_3 + c_1} \\
\vec{A}_{3s2y} &= \frac{\mu_0 I}{4\pi} \ln \frac{R_2 + d_1}{R_4 + d_2} \\
\vec{A}_{1s3y} &= \frac{\mu_0 I}{4\pi} \ln \frac{R_{4a} + d_{2a}}{R_{2a} + d_{1a}}
\end{aligned} \tag{12}$$

For the section S_1 , located in the plane of the negative values of X , the magnetic potential vector would have two components, \hat{x} and \hat{y} :

$$\vec{A}_{s1} = (\vec{A}_{1s1x} + \vec{A}_{3s1x}) \hat{x} + (\vec{A}_{2s1y}) \hat{y} \tag{13}$$

For the section S_2 , located in the plane of the positive values of X , the magnetic potential vector would also have two components, \hat{x} and \hat{y} :

$$\vec{A}_{s2} = (\vec{A}_{1s2x} + \vec{A}_{2s2x}) \hat{x} + (\vec{A}_{3s2y}) \hat{y} \tag{14}$$

Nevertheless, in the case of the section S_3 , located on the axis Y at $X = 0$, it would only have component \hat{y} :

$$\vec{A}_{s3} = (\vec{A}_{1s3y}) \hat{y} \tag{15}$$

Thus, once the magnetic potential vector has been calculated, the magnetic field could already be obtained by applying the curl to the magnetic potential vector in the same way as (7). However, with the aim of obtaining the total magnetic field, the calculation must be performed for each loop section. In this manner, the equations of the three sections are presented in (16), (17), and (18) respectively.

Section S_1 –Negative Values of X

$$\begin{aligned}
\vec{B}_{s1x} &= \frac{\mu_0 I}{4\pi} \left[\frac{z}{R_{1a}(R_{1a} + d_{1a})} - \frac{z}{R_{3a}(R_{3a} + d_{2a})} \right] \\
\vec{B}_{s1y} &= \frac{\mu_0 I}{4\pi} \left[\frac{z}{R_{1a}(R_{1a} + c_{1a})} - \frac{z}{R_{2a}(R_{2a} + c_{2a})} \right. \\
&\quad \left. - \frac{z}{R_{3a}(R_{3a} + c_{1a})} + \frac{z}{R_{4a}(R_{4a} + c_{2a})} \right]
\end{aligned}$$

$$\begin{aligned}
\vec{B}_{s1z} &= \frac{\mu_0 I}{4\pi} \left[\left(\frac{-(x+a)}{R_{1a}(R_{1a} + d_{1a})} + \frac{(x+a)}{R_{3a}(R_{3a} + d_{2a})} \right) \right. \\
&\quad \left. - \left(\frac{(y+b)}{R_{1a}(R_{1a} + c_{1a})} - \frac{(y+b)}{R_{2a}(R_{2a} + c_{2a})} \right) \right. \\
&\quad \left. - \frac{(y-b)}{R_{3a}(R_{3a} + c_{1a})} + \frac{(y-b)}{R_{4a}(R_{4a} + c_{2a})} \right]
\end{aligned} \tag{16}$$

Section S_2 –Positive Values of X

$$\begin{aligned}
\vec{B}_{s2x} &= \frac{\mu_0 I}{4\pi} \left[\frac{-z}{R_2(R_2 + d_1)} - \frac{z}{R_4(R_4 + d_2)} \right] \\
\vec{B}_{s2y} &= \frac{\mu_0 I}{4\pi} \left[\frac{z}{R_1(R_1 + c_1)} - \frac{z}{R_2(R_2 + c_2)} \right. \\
&\quad \left. - \frac{z}{R_3(R_3 + c_1)} + \frac{z}{R_4(R_4 + c_2)} \right] \\
\vec{B}_{s2z} &= \frac{\mu_0 I}{4\pi} \left[\left(\frac{(x-d)}{R_2(R_2 + d_1)} - \frac{(x-d)}{R_4(R_4 + d_2)} \right) \right. \\
&\quad \left. - \left(\frac{(y+b)}{R_1(R_1 + c_1)} - \frac{(y+b)}{R_2(R_2 + c_2)} - \frac{(y-b)}{R_3(R_3 + c_1)} \right) \right. \\
&\quad \left. + \frac{(y-b)}{R_4(R_4 + c_2)} \right]
\end{aligned} \tag{17}$$

Section S_3 –Axis Y at $X = 0$

$$\begin{aligned}
\vec{B}_{s3x} &= \frac{\mu_0 I}{4\pi} \left[\frac{z}{R_{2a}(R_{2a} + d_{1a})} - \frac{z}{R_{4a}(R_{4a} + d_{2a})} \right] \\
\vec{B}_{s3z} &= \frac{\mu_0 I}{4\pi} \left[\frac{x}{R_{4a}(R_{4a} + d_{2a})} - \frac{x}{R_{2a}(R_{2a} + d_{1a})} \right]
\end{aligned} \tag{18}$$

Hence, all this double loop electromagnetic analysis have led us to the conclusion that the total magnetic field produced by a double loop of dimensions a , d and $2b$ at a certain point $P(x, y, z)$ as shown in Figure 4 will be the sum of the components obtained in (16), (17), and (18). The resulting expression is shown in the following:

$$\begin{aligned}
\vec{B} &= (\vec{B}_{s1x} + \vec{B}_{s2x} + \vec{B}_{s3x}) \hat{x} + (\vec{B}_{s1y} + \vec{B}_{s2y}) \hat{y} \\
&\quad + (\vec{B}_{s1z} + \vec{B}_{s2z} + \vec{B}_{s3z}) \hat{z}
\end{aligned} \tag{19}$$

3. Experimental Measurements

After obtaining the expression that describes the magnetic field generated by a double loop, the next step was to verify that the theoretical results coincided with the experimental ones. For this reason, a double loop was constructed in the laboratory by our research team (Group of Traffic Control

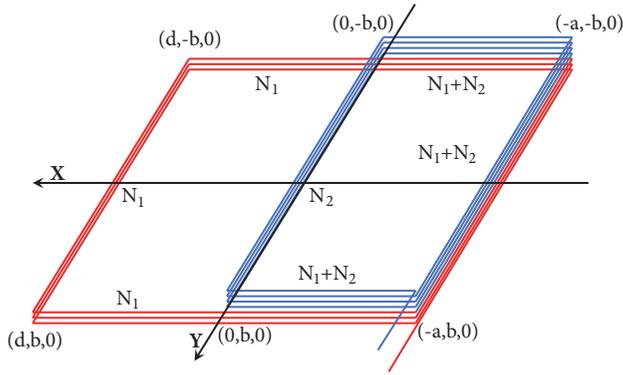


FIGURE 5: Outline of the double loop constructed.

System, ITACA Institute, Universitat Politècnica de València, Spain). It was implemented with an external loop of N_1 turns and, inside and over it, a smaller one of N_2 turns located in the negative half-plane, both with the same direction of circulation. The scheme is shown in Figure 5.

With these conditions, we only were interested in the component of the magnetic field perpendicular to the surface of the loop (\vec{B}_k), as the surface vector of the loop only has component along the Z-axis. Then, taking into account the turns of each of the loop sections, this value was easily obtained from (19) as follows:

$$\vec{B}_k = (N_1 + N_2) \cdot \vec{B}_{s1z} + N_1 \cdot \vec{B}_{s2z} - N_2 \cdot \vec{B}_{s3z} \quad (20)$$

3.1. Characteristics of the Double Loop Constructed. Conductor type is tinned copper wire conductors individually insulated with polyvinyl chloride with a cross section of 0.28 mm^2 .

Number of turns is 4 exterior turns (N_1) of $1.20 \times 0.46 \text{ m}$ and 5 interior turns (N_2) of $0.40 \times 0.46 \text{ m}$.

Dimensions is $1.20 \times 0.46 \text{ m}$.

With these values, the parameters of the loop according to the nomenclature used in Figures 4 and 5 would be as follows:

$$\begin{aligned} a &= 0.40\text{m} \\ b &= 0.23\text{m} \\ d &= 0.80\text{m} \end{aligned} \quad (21)$$

3.2. Characteristics of the Signal Used to Energize the Loop. Frequency of the signal applied to the loop is 139.2 kHz .

Signal type is rectangular.

Current intensity through the loop (RMS) is 51.80 mA .

3.3. Region Where the Readings of the Magnetic Field Were Carried Out. Height above the plane of the loop is 0.0825 m . Position is along the X-axis in the center of the loop ($Y = 0$).

With the above-mentioned characteristics, the theoretical calculation in the region of the measurements was carried out by applying and programming the expression obtained in



FIGURE 6: Exposure Level Tester ELT-400.



FIGURE 7: ELT-400 configuration.

(20) in Matlab ©, while the experimental measurements were performed in our laboratory in order to verify the goodness of the theoretical model developed for the calculation of the double loop magnetic field. The instrument used to measure this magnetic field was the Exposure Level Tester ELT-400 shown in Figure 6.

This device contains a series of turns with a diameter of 0.125 m and measures the magnetic field by means of this spherical sensor. It is able to detect frequencies from 1 Hz to 400 kHz , although this can be set as wished. To perform the measurements, the ELT-400 was configured as shown in Figure 7 and as follows:

Selected frequency range: $30 \text{ Hz} - 400 \text{ kHz}$.

Reading range: $320 \mu\text{T}$.

RMS signal value.

After collecting and processing all the information, the comparison between the calculated and measured values of the magnetic field as well as the tolerance of the measuring instrument ($\pm 35 \text{ nT}$) was made, which is shown in Figure 8.

It can be observed that the differences between the measured and calculated magnetic field values are, except in specific points, within the tolerance range of the instrument. The difference between the theoretical and measured values within the contour of the loop is below 20% of the reading and the mean value is below 8%. Therefore, it can be concluded that the theoretical model for double loops developed in this paper predicts with a good precision the behavior of the magnetic field. In addition, different types of tests were also carried out with other types of loops, both single and double,

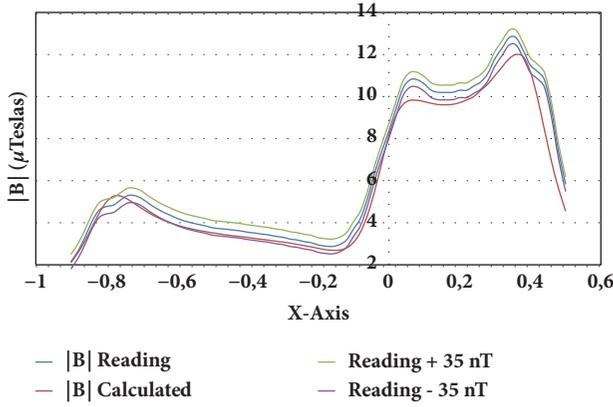


FIGURE 8: Calculated and measured values of the magnetic field $|B|$. The instrument tolerance is also considered.

varying the type and amplitude of the applied current, and very similar results were obtained.

4. Methods for Inductance Calculation

Throughout time, various methods to calculate the inductance of magnetic loops according to different geometric configurations have been proposed [22–28], such as the Mills' method based on Grover's equations [22]. Most of the mathematical expressions for these calculations appear in classical texts [23] and based on these expressions, the inductance values presented by other different models of loops can be easily deduced.

However, the development of computer systems has allowed to implement numerical methods that make use of the intrinsic definition of the physical process of magnetic induction. In this regard, it would be convenient to include our previously presented work [25], in which the methods to calculate the inductance of a rectangular loop were analyzed and compared with other calculation techniques and real measures.

In this way, this time we will present three methods to calculate the inductance of the innovative double loops, since there are no studies about them. Therefore, these methods will be deeply analyzed and compared in order to know which method is more effective and if all of them are equally good. These three methods are as follows:

- (A) Electromagnetic Analysis Method
- (B) Numerical Integration Method
- (C) Mills and Grover's Method

(A) *Electromagnetic Analysis Method.* The first method to calculate the inductance of a double loop is based on the electromagnetic analysis, which considers the use of a numerical method that employs the expression of the magnetic field generated by a double loop. This means to carry out a study applied to the case of a double loop with the same structure as Figure 4.

The procedure starts by making use of the flux calculation, since all the loops work in the same way. However, in this case, the magnetic field and the flux will present some peculiarities.

The magnetic flux will be obtained as the integral of the product of the magnetic field by the differential surface all along the entire surface of the loop:

$$\emptyset = \int_S \vec{B}_K \cdot d\vec{S} \quad (22)$$

The expression of \vec{B}_K used in (22), which only taken into account the field in the Z direction, will be the same as in (20), and the differential surface can be easily replaced by the product of the differential length according to the X-axis and Y-axis as shown in the following:

$$\emptyset = \int_S \vec{B} d\vec{S} = \int_{-a}^d \int_{-b}^b B_K dy dx \quad (23)$$

To solve this expression, the integrals will be replaced by summations. For this, it is necessary to obtain the magnetic field in a series of points in space and consider a differential surface $dydx$ around.

In this way, if we defined N_x as the number of points in which the magnetic field will be measured along the X-axis and N_y as the number of points in which the magnetic field will be measured along the Y-axis, the dx and dy differentials of (23) would be given by the following:

$$dx = \frac{d+a}{N_x} \quad (24)$$

$$dy = \frac{2b}{N_y}$$

Consequently, the total flux through the loop would be as shown in the following:

$$\emptyset = \sum_{n=1}^{N_x-1} \sum_{m=1}^{N_y-1} B_K(-a+ndx, -b+mdy, 0) f_x f_y dy dx \quad (25)$$

Summaries' limits have been designed to prevent the magnetic field measurements on the conductors, since at these points the value of B_K presents a singularity. The summation is extended to the rectangle defined by the points $(-a+dx, d-dx)$ in the X-axis and $(-b+dy, b-dy)$ in the Y-axis, but this causes an error in the measurement that has been tried to be solved by increasing the differential surface by 50% within the limits of the summation points. For this reason, f_x and f_y factors have been introduced [23]. The f_x factor takes a value 1 in all points except in those where $n = 1$ or $n = N_x - 1$ that has a value 1.5, and the f_y factor also takes a value 1 in all points except in those where $m = 1$ or $m = N_y - 1$ that has a value 1.5.

On the other hand, due to the abrupt change that appears in the B_K component of the magnetic field in the vicinity of the conductors, it is evident the importance of choosing the number of points N_x and N_y . In our previous article [25], it was shown that the optimal values of N_x and N_y to minimize

the error in the calculation of the inductance for a single rectangular loop were given by

$$\begin{aligned} N_x &= \frac{a+d}{3R_C} \\ N_y &= \frac{2b}{3R_C} \end{aligned} \quad (26)$$

where R_c is the radius of the conductor used to build the loop and a, d and $2b$, the dimensions of the loop shown in Figures 4 and 5. However, in this case, the space has been divided into three sections (two areas) as explained previously. For the positive region of the X -axis the flux is concatenated by N_1 turns, and for the negative region of the X -axis, the flux is concatenated by $N_1 \pm N_2$ turns as shown in Figure 5. The “+” sign would be used when the current goes through both loops in the same direction and the sign “-” when the current goes through the loops in opposite directions. Therefore, we could separate N_x into two sections as shown in Equation (27), which would represent the measured points, respectively, the negative and the positive ones at the X -axis.

$$\begin{aligned} N_{xa} &= \frac{a}{3R_C} \\ N_{xd} &= \frac{d}{3R_C} \end{aligned} \quad (27)$$

Finally, after all this mathematical and electromagnetic analysis, we could obtain the inductance of a double loop with dimensions a, d and $2b$ and with N_1 and $N_1 \pm N_2$ turns respectively by the following expression:

$$\begin{aligned} L_0 &= \frac{1}{I} \left[\sum_{n=1}^{N_{xa}-1} \sum_{m=1}^{N_y-1} (N_1 \pm N_2) B_K(-a+ndx, -b+mdy, 0) \right. \\ &\quad \cdot f_x f_y dy dx + \sum_{n=1}^{N_{xd}-1} \sum_{m=1}^{N_y-1} N_1 B_K(ndx, -b+mdy, 0) \\ &\quad \left. \cdot f_x f_y dy dx \right] \end{aligned} \quad (28)$$

In this expression, I represents the intensity used for the calculation of the magnetic field and the rest of values are known and have been described above. Nevertheless, it should be pointed out that the result of this first method will give the value of the inductance without considering the thickness of the loop, since this method is based only on the electromagnetic analysis of the magnetic field created by a double loop. Therefore, in this first method the spacing between the turns of the loop is considered null.

(B) *Numerical Integration Method.* After obtaining a first way to calculate the inductance of a double loop that does not take into account the spacing between turns, it seems reasonable that the second presented method does. These method, which apparently should provide more accurate

values, is the utilization of numerical integration techniques but considering the size and spacing between the turns of the loop (S_v). However, on this occasion the calculation will be more complex than the last one, but these values are supposed to be much more real, since in reality the cables have a thickness that although it is minimal, it should be contemplated.

For the purpose of calculation mentioned, it is assumed an assembly in which the loops are stacked vertically. First, the larger N_1 turns would be installed and after, over them, the smaller N_2 , assuming that are equi-spaced a distance S_v . In this way, for each turn of the loop, the flux through it would be produced by the current that flows through the turn itself plus the flux generated by each of the other turns of the loop and which is concatenated by the one that is being analyzed.

The different flux components mentioned above will be represented by the terms θ_{ijk} , where the subindex i indicates the type of loop which generates the field (1 Big, 2 Small), the subindex j indicates the type of loop which is crossed by the magnetic field (1 Big, 2 Small) and the subindex k indicates the distance between the loop that generates de field and the one which detects it. Hence, we can define the following terms:

- (i) θ_{110} is the flux generated by one of the turns of the largest loops and which goes through them.
- (ii) θ_{220} is the flux generated by one of the turns of the smallest loops and which goes through them.
- (iii) θ_{11i} is the flux generated by a big turn which goes through the other big ones separated a distance iS_v .
- (iv) θ_{12j} is the flux generated by a big loop which goes through the other small ones separated a distance of jS_v from the turn that has generated the flux.
- (v) θ_{22k} is the flux generated by a small turn which goes through the other small ones separated a distance of kS_v .
- (vi) θ_{21l} is the flux generated by a small loop which goes through the other big ones separated a distance lS_v from the turn that has generated the flux.
- (vii) N_{min} will be the minimum value between N_1 and N_2 and N_{max} will be the maximum value between N_1 and N_2 .

With this nomenclature, the inductance of the double loop would be given by the following:

$$\begin{aligned} L_{IIC} &= \frac{1}{I} \left[\sum_{i=0}^{N_1} (N_1 - i) \theta_{11i} \pm \sum_{i=0}^{N_2} (N_2 - i) \theta_{22i} \right. \\ &\quad + \sum_{i=1}^{N_{min}} i (\theta_{12i} \pm \theta_{21i}) + \sum_{i=N_{min}+1}^{N_{max}} N_{min} (\theta_{12i} \pm \theta_{21i}) \\ &\quad \left. + \sum_{i=N_{max}+1}^{N_1+N_2-1} (N_1 + N_2 - i) (\theta_{12i} \pm \theta_{21i}) \right] \end{aligned} \quad (29)$$

The “+” sign would be used when the current goes through both loops in the same direction while the sign “-” would

be used when the current goes through the loops in opposite directions. In any case, the fluxes described and shown in (29) are given by the following:

$$\begin{aligned}
\emptyset_{11i} &= \left[\sum_{n=1}^{N_{x1}-1} \sum_{m=1}^{N_{y1}-1} B_{KR1} \left(-\frac{d+a}{2} + ndx, -b + mdy, i \cdot S_v \right) \right. \\
&\quad \left. \cdot f_x f_y dy dx \right] \\
\emptyset_{22i} &= \left[\sum_{n=1}^{N'_{x1}-1} \sum_{m=1}^{N_{y1}-1} B_{KR2} \left(-\frac{a}{2} + ndx, -b + mdy, i \cdot S_v \right) \right. \\
&\quad \left. \cdot f_x f_y dy dx \right] \\
\emptyset_{12i} &= \left[\sum_{n=1}^{N'_{x1}-1} \sum_{m=1}^{N_{y1}-1} B_{KR1} \left(-\frac{a}{2} + ndx, -b + mdy, i \cdot S_v \right) \right. \\
&\quad \left. \cdot f_x f_y dy dx \right] \\
\emptyset_{21i} &= \left[\sum_{n=1}^{N_{x1}-1} \sum_{m=1}^{N_{y1}-1} B_{KR2} \left(-\frac{d+a}{2} + ndx, -b + mdy, i \cdot S_v \right) \right. \\
&\quad \left. \cdot f_x f_y dy dx \right]
\end{aligned} \tag{30}$$

In \emptyset_{11i} , $B_{KR1}(x, y, z)$ is the magnetic field component along the Z-axis generated by a large turn, extending from $-a$ to $+d$ along the X-axis and from $-b$ to $+b$ along the Y-axis at the point of space (x, y, z) .

In \emptyset_{12i} , $B_{KR1}(x, y, z)$ is also the magnetic field component along the Z-axis generated by a large turn, extending from $-a$ to $+d$ along the X-axis and from $-b$ to $+b$ along the Y-axis at the point of space (x, y, z) .

In \emptyset_{22i} , $B_{KR2}(x, y, z)$ is the magnetic field component along the Z-axis generated by a small turn, extending from $-a$ to 0 along the X-axis and from $-b$ to $+b$ along the Y-axis at the point of space (x, y, z) .

In \emptyset_{21i} , $B_{KR2}(x, y, z)$ is also the magnetic field component along the Z-axis generated by a small turn, extending from $-a$ to 0 along the X-axis and from $-b$ to $+b$ along the Y-axis at the point of space (x, y, z) .

The values used in (29) and (30) (N_{x1} , N_{y1} , dx , dy , f_x , f_y) are the same ones used to calculate the previous inductance and N'_{x1} takes the following value:

$$N'_{x1} = \frac{a}{3R_C} \tag{31}$$

As it can be seen, this method seems much more accurate than the previous one, but it is true that there is a high amount of operations and summations to perform, which will make it more difficult to implement and with higher computational

cost. Therefore, taking into account the fact that the first method did not consider spacing between the turns of the loop, which is not physically correct, and that the second method, although it does, is computationally complicated because of the large number of operations and summations, it leads us to think of a third method that also considers spacing but is much simpler to implement.

(C) *Mills and Grover's Method.* In order to calculate the inductance of a double loop, Mills and Grover's method will be proposed. As said above, in this third and last method it will be taken into account again that the size of the conductors prevents all turns from being in the same coordinate Z. To consider this phenomenon, studies such as Mills' emerged [22, 28], from which new expressions of inductance were deduced.

When working with double loops, the self-inductance of the N_1 turns of the loop and, over them, the N_2 turns, equally spaced a distance S_v , is represented by the following:

$$\begin{aligned}
L_T &= N_1 L_{10} \pm N_2 L_{20} + \sum_{n=1}^{N_1-1} (N_1 - n) M_{11}(nS_v) \\
&\quad \pm \sum_{n=1}^{N_2-1} (N_2 - n) M_{22}(nS_v) \\
&\quad + \sum_{n=1}^{N_1} \sum_{m=1}^{N_2} (M_{12}(S_v(N_1 + m - n)) \\
&\quad \pm M_{21}(S_v(N_1 + m - n)))
\end{aligned} \tag{32}$$

where

- (i) L_{10} is the self-inductance of a rectangular loop with only one large-size turn $(a + d) \times 2b$;
- (ii) L_{20} is the self-inductance of a rectangular loop with also only one small-size turn $(a \times 2b)$;
- (iii) $M_{11}(nS_v)$ is the mutual inductance between two turns of a rectangular large loop separated between them at a distance of nS_v ;
- (iv) $M_{22}(nS_v)$ is the mutual inductance between two turns of a rectangular small loop separated between them at a distance of nS_v ;
- (v) $M_{12}(S_v(N_1 + m - n))$ are the mutual inductances between a big-size turn and a small-size turn separated at a distance of $S_v(N_1 + m - n)$.

It must be noted that the above-mentioned parameters L_{10} and L_{20} are obtained as the sum of the internal and external inductance of the conductors that constitute the loop as follows:

$$\begin{aligned}
L_0 &= L_{0i} + L_{0e} \\
L_{10i} &= 2(a + d + 2b) L_i \\
L_{20i} &= 2(a + 2b) L_i
\end{aligned} \tag{33}$$

where L_i is the inductance per unit length, which must consider the relationship between the inductance at a certain

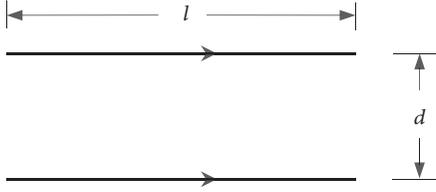


FIGURE 9: Features of two ideal parallel conductors of no straight section for measuring mutual induction.

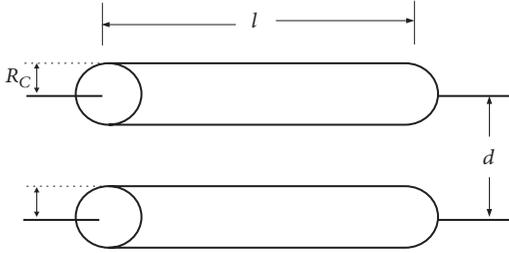


FIGURE 10: Features of two parallel conductors to measure the mutual induction.

frequency L_i and the inductance at low frequency L_{i0} . In order to carry out these operations, we will use Johnson's studies [23]. In this manner,

$$L_i = L_{i0} \cdot \frac{4}{q} \left[\frac{bei(q) \times bei'(q) + ber(q) \times ber'(q)}{(bei'(q))^2 + (ber'(q))^2} \right] \quad (34)$$

where $bei(q)$, $ber(q)$, $bei'(q)$, and $ber'(q)$ are the imaginary and real parts of the Bessel function of first order and their derivatives, which are necessary for the calculations when this method is applied. On the other hand, the inductance per unit length at a low frequency is given by (35), but for copper conductors it takes a value of $0.5 \cdot 10^{-7} H/m$:

$$L_{i0} = \frac{\mu_0 \mu_r}{8\pi} \quad (35)$$

To obtain L_{oe} [22–27], we focus on the mutual inductance of a pair of parallel conductors, whose expression is:

$$M(l, d) = \pm \frac{\mu_0 l}{2\pi} \left\{ \ln \left[\frac{l}{d} + \sqrt{1 + \left(\frac{l}{d}\right)^2} \right] - \sqrt{1 + \left(\frac{d}{l}\right)^2} + \frac{d}{l} \right\} \quad (36)$$

As seen in Figure 9, l is the length of the filaments and d is the separation between them (both quantities expressed in meters), which results in an inductance expressed in Henrys. This expression takes a positive sign when the current in both cables has the same direction and takes a negative sign when the current has opposite directions.

In addition, it is known that the external inductance of a pair of parallel conductors with the dimensions shown in Figure 10 is given by

$$L_p = L_1 - M_{12} + L_2 - M_{21} \quad (37)$$

L_1 and L_2 are the self-inductances of the simple conductors and M_{12} and M_{21} are the mutual inductance measures between the centers of the conductors, assuming a uniform current distribution throughout the cross section of the conductor. Mutual inductances have been considered to have a negative sign as it has been assumed that the current in the two conductors have opposite directions. Therefore, if both conductors have the same dimensions:

$$\begin{aligned} L &= L_1 = L_2 \text{ and} \\ M &= M_{12} = M_{21} \end{aligned} \quad (38)$$

Thus:

$$L_p = 2(L - M) \quad (39)$$

The external self-inductance of a conductor is obtained thanks to a method in which the conductor is replaced with two conductors with a null straight section separated by a distance equal to the radius of the conductor. In this way

$$L_p = 2(M(l, R_C) - M(l, d)) \quad (40)$$

In addition, the inductance of a rectangular loop with a single turn is given by the sum of the inductance of two pairs of parallel conductors as follows:

$$\begin{aligned} L_{0e} &= L_{p1} + L_{p2} \\ L_{0e} &= 2[M_1(l_1, R_C) - M_1(l_1, l_2) + M_2(l_2, R_C) \\ &\quad - M_2(l_2, l_1)] \end{aligned} \quad (41)$$

where $l_1 = a + d$ and $l_2 = 2b$ for the big loop and $l_1 = a$ and $l_2 = 2b$ for the small loop.

As it can be deduced from this expression, the external inductance of a rectangular loop with one turn is equal to the mutual inductance of two identical coaxial rectangular loops separated by a distance equal to the radius of the conductor. In this way, the mutual inductance of two parallel rectangular loops as shown in Figure 11 can be obtained from mutual inductances between parallel conductors.

Therefore, the mutual inductance between the two rectangular loops with the same dimensions as shown in Figure 11 may be expressed as follows:

$$\begin{aligned} M &= -2[M_{13}(A, \sqrt{H^2 + B^2}) - M_{11}(A, H) \\ &\quad + M_{24}(B, \sqrt{H^2 + A^2}) - M_{22}(B, H)] \end{aligned} \quad (42)$$

In the above expression, the terms M_{ij} represent the mutual inductance between the segment i of the bottom loop and the segment j of the top loop. The reason to multiply the expression by two is because the mutual inductances are all symmetric, that is, for all i and all j , $M_{ij} = M_{ji}$.

On the other hand, to calculate the mutual inductance between two parallel loops with different dimensions, Grover's equations must be used again, since they provide the mutual inductance between two parallel straight conductors as shown in Figure 12.

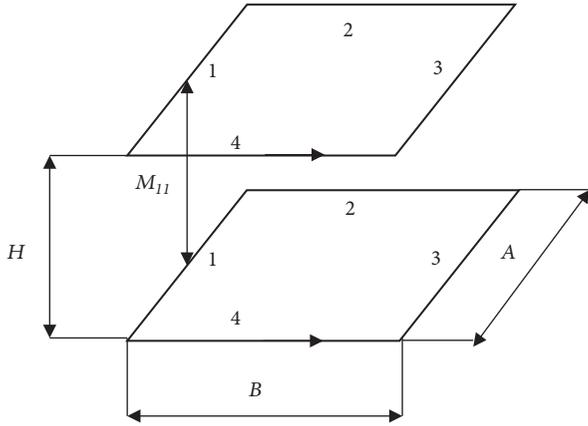


FIGURE 11: Geometry for calculating the mutual inductance between two parallel and coaxial rectangular loops with the same dimensions.

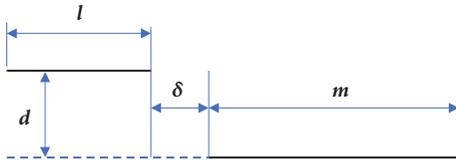


FIGURE 12: Disposition of two parallel straight conductors.

According to Grover's formula, the mutual inductance between two parallel conductors with sizes l and m spaced a distance d and displaced at a distance δ is given by

$$M_G(l, m, d, \delta) = \frac{\mu_0}{4\pi} \left[\alpha \sinh^{-1} \left(\frac{\alpha}{d} \right) - \beta \sinh^{-1} \left(\frac{\beta}{d} \right) - \gamma \sinh^{-1} \left(\frac{\gamma}{d} \right) + \delta \sinh^{-1} \left(\frac{\delta}{d} \right) - \sqrt{\alpha^2 + d^2} - \sqrt{\beta^2 + d^2} + \sqrt{\gamma^2 + d^2} - \sqrt{\delta^2 + d^2} \right] \quad (43)$$

where

$$\begin{aligned} \alpha &= l + m + \delta, \\ \beta &= l + \delta, \\ \gamma &= m + \delta \end{aligned} \quad (44)$$

It should be pointed out that if the two conductors overlap partially or totally, the parameter δ will have negative values.

In any case, from all these expressions it is finally possible to obtain the inductance between two parallel rectangular loops with different dimensions as shown in Figure 13, which is simply another way of seeing Figures 4 and 5.

Therefore, the mutual inductance between the two parallel loops is finally obtained as a sum of the mutual inductances of the parallel conductors where the terms M_{ij} represent mutual inductances between two parallel rectilinear conductors with the same dimensions, and the terms M_{Gij} represent the mutual inductances between two parallel rectilinear

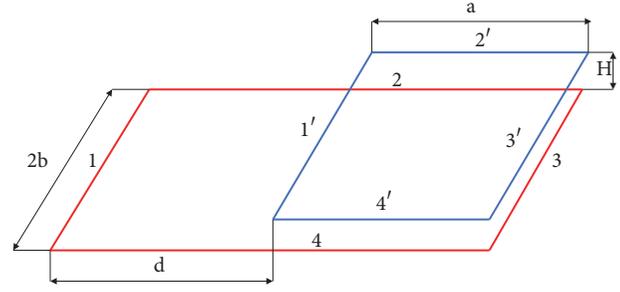


FIGURE 13: Disposition of two parallel loops of different dimensions.

conductors with different dimensions. This is shown in the following:

$$\begin{aligned} M_{EE}(d, a, 2b, H) &= M_{11} \left(2b, \sqrt{d^2 + H^2} \right) \\ &\quad - M_{13} \left(2b, \sqrt{(d+a)^2 + H^2} \right) \\ &\quad - M_{31} \left(2b, \sqrt{a^2 + H^2} \right) + M_{33} (2b, H) \\ &\quad + M_{G22} (d+a, a, h, -a) \\ &\quad - M_{G24} (d+a, a, \sqrt{4b^2 + H^2}, -a) \\ &\quad - M_{G42} (d+a, a, \sqrt{4b^2 + H^2}, -a) \\ &\quad + M_{G44} (d+a, a, h, -a) \\ &= M_{11} \left(2b, \sqrt{d^2 + H^2} \right) \\ &\quad - M_{13} \left(2b, \sqrt{(d+a)^2 + H^2} \right) \\ &\quad - M_{31} \left(2b, \sqrt{a^2 + H^2} \right) + M_{33} (2b, H) \\ &\quad + 2M_{G22} (d+a, a, h, -a) \\ &\quad - 2M_{G24} (d+a, a, \sqrt{4b^2 + H^2}, -a) \end{aligned} \quad (45)$$

Although this last equation may seem complicated and with a large number of operations, it should not be forgotten that in none of the previous expressions there are integrals or summations. The maximum operations that appear are square roots and trigonometric functions. Therefore, it apparently seems to be the method that best behaves computationally. Moreover, this also considered spacing between the turns of the loop as the second one.

Thus, once all methods have been presented, the inductance values obtained are going to be compared to each other to verify the similarity between them and know which the best methods are and which are not.

5. Results and Discussion

After making the description and performing the analysis of the three methods, the first thing that stands out is that there are two methods that consider the spacing between the turns (S_v) and one which does not. For that reason, what would be expected at first sight is that the accuracy and complexity of the two methods which do consider this are greater than the method which does not. Therefore, the inductance values obtained by using the electromagnetic analysis method should provide different values than the other ones, but a simple explanation for this phenomenon would be the fact of not having considered the flux lines that are lost between the loops because of the fact that they have a real thickness which is not zero.

To check all of the above, various test related to the inductance values of the different methods and types of loops were performed and are presented below. In fact, it was studied the effect of increasing the number of turns in a single and in a double loop and the effect of changing the dimensions of a double loop.

5.1. How the Number of Turns Affects the Value of the Inductance of a Single Loop. For the purpose of this study, a single 2×2 loop located and centered in the XY plane with a cable radius of 0.75mm and a turn spacing of 1.9mm was used. In this first test, the number of turns of the loop was gradually increased to see the effect that it has on the value of its inductance. With the nomenclature described in the paper, the characteristics of this loop were as follows:

- (i) $a = 1 \text{ meter}$
- (ii) $b = 1 \text{ meter}$
- (iii) $d = 1 \text{ meter}$
- (iv) $N_1 = \text{From } 1 \text{ to } 7$
- (v) $N_2 = 0$

As mentioned, in the ideal case it would be expected for the three methods to provide similar results, but it is clearly observed in Figure 14(a) that there are two methods that are practically identical, and one method whose results are far from the other ones. Therefore, the first point to emphasize is that the Mills and Grover's method and the numerical integration method, those that consider the separation between turns, provide good and identical results, while the method based on the electromagnetic analysis begins to fail when the turns increase.

This means that the more turns the loops have, the more different the results are, because the error of not considering separation between turns increases for each turn.

5.2. How the Number of Turns Affects the Value of the Inductance of a Double Loop. The next test was the same as the previous one but instead of working with a single loop, we will work with a double loop. This time, it was a double 2×2 loop formed by an external loop of N_1 turns located and centered in the XY plane and a smaller one of N_2 turns located in the half-plane of the negative values of X , both

with a cable radius of 0.75mm and a turn spacing of 1.9mm . In the previous analysis, only the number of turns N_1 was varied since it was a single loop and $N_2 = 0$. However, as we were working with a double loop, to check the effect of increasing the number of turns, the value of N_1 was kept fixed and N_2 was the value that we were gradually increasing. In this way, with the nomenclature described in the paper, the characteristics of this double loop were as follows:

- (i) $a = 1 \text{ meter}$
- (ii) $b = 1 \text{ meter}$
- (iii) $d = 1 \text{ meter}$
- (iv) $N_1 = 3$
- (v) $N_2 = \text{From } 1 \text{ to } 7$

The result obtained was really similar to the previous one. As it can be seen in Figure 14(b), the two methods that consider spacing provided practically identical values, while the electromagnetic analysis method, as it does not, provided different values, this time greater. The fact that the results were even greater is because as we had concluded before, the more turns the loops have, the more different the results are, and there are more turns in a double loop (N_1 in the first loop and $N_1 + N_2$ in the second one) than in a single one (only N_1). Thus, at this point we could already begin to draw conclusions, but in order to make a study as rigorous as possible, we will show more results from our analysis.

5.3. How the Dimensions of the Loop Affect the Value of the Inductance of a Double Loop. Once we had analyzed how the number of turns affects the inductance value, in the two remaining experiments we analyzed what happens when varying the dimensions of the loops. For this purpose, we worked with a double loop formed by N_1 large turns, centered with respect to the Y -axis, extending from $-a$ to $+d$ along the X -axis, and N_2 small turns, also centered with respect to the Y -axis, extending from $-a$ to 0 along the X -axis. Both had a cable radius of 0.75mm and a turn spacing of 1.9mm . In this way, the number of turns was kept fixed and the value that we varied was the length a , which means increasing the length of the smallest loop, which has N_2 turns and is located in the negative half-plane. In addition, it should be noted that this fact also increases the length of the biggest loop, since its dimensions are $(a + d) \times 2b$. Then, with the nomenclature described in the paper, the characteristics of this loop were as follows:

- (i) $a = \text{From } 1 \text{ to } 10 \text{ meters}$
- (ii) $b = 1 \text{ meter}$
- (iii) $d = 1 \text{ meter}$
- (iv) $N_1 = 3$
- (v) $N_2 = 2$

In Figure 14(c) it can be observed very clearly that although the trend of the three methods was the same, the results were different. Mills and Grover's method as well as the numerical integration one had a practically linear relationship between

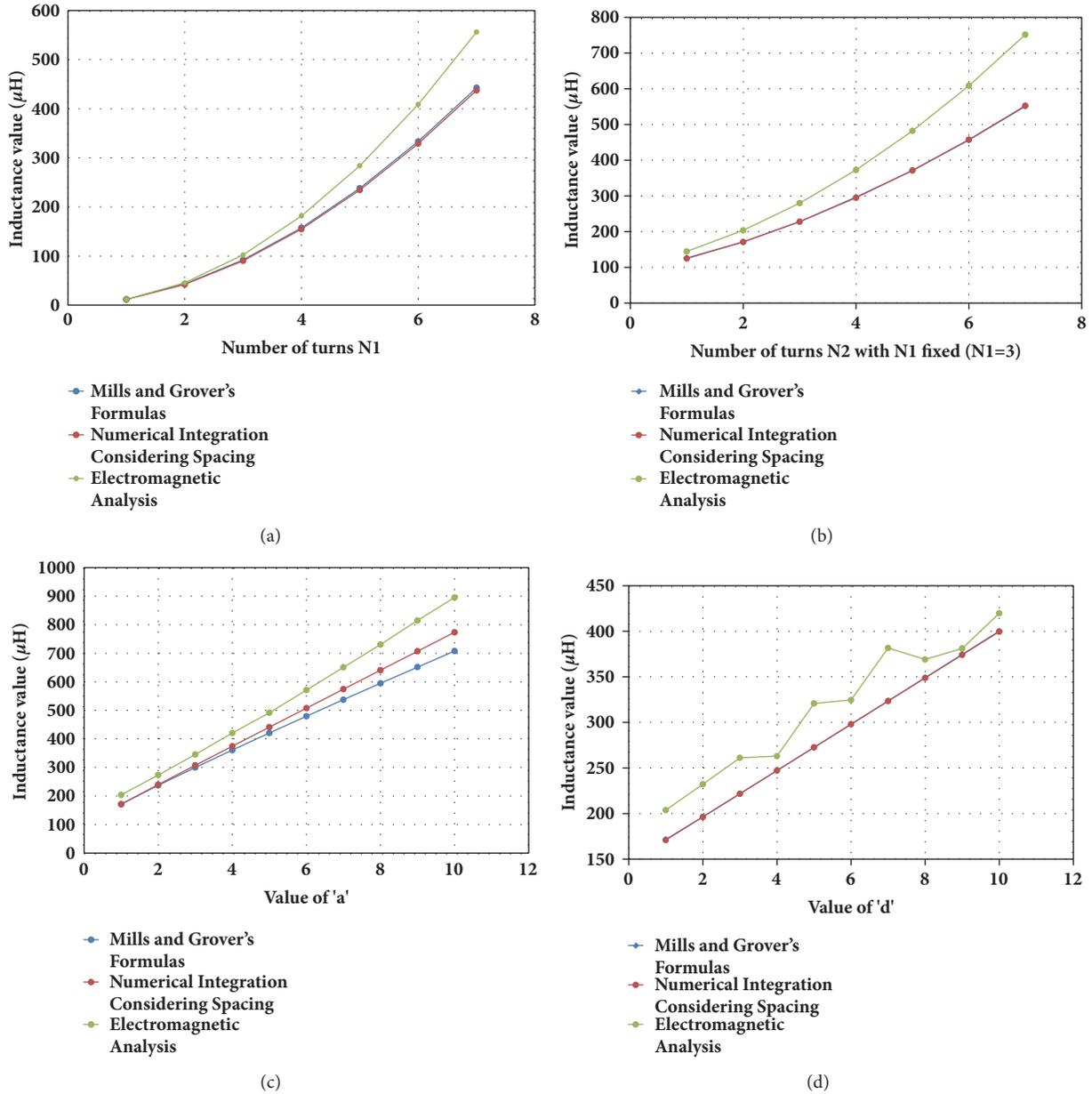


FIGURE 14: Inductance values. (a) When increasing the number of turns in a single loop. (b) When increasing the number of turns in a double loop. (c) When increasing the length of “a” in a double loop. (d) When increasing the length of “d” in a double loop.

the increase of a and the increase of the inductance. However, it is noted that the first method, which had extracted its equations from the analysis of the electromagnetic field without considering spacing, gave results distant from the real ones from the beginning, being this error greater as the length of a is bigger.

On the other hand, after analyzing what happens when the length of a is increased, it was analyzed what happens precisely if we increase the length of d , with whose variation we only be modify the length of the largest loop. The double loop was again formed by N_1 large turns, centered with respect to the Y -axis, extending from $-a$ to $+d$ along the X -axis, and N_2 small turns, also centered with respect to the

Y -axis and extending from $-a$ to 0 along the X -axis. Both had a cable radius of 0.75mm and a turn spacing of 1.9mm . The number of turns was kept fixed again and the value that we varied was d . In this manner, with the nomenclature described in the paper, the characteristics of this loop were as follows:

- (i) $a = 1$ meter
- (ii) $b = 1$ meter
- (iii) $d =$ From 1 to 10 meters
- (iv) $N_1 = 3$
- (v) $N_2 = 2$

In this test, it was seen how two methods followed the same trend again, providing almost the same results, but the other one did not behave properly. However, in Figure 14(d) it is clear that Mills and Grover's method and the method of numerical integration provide exactly the same values and the electromagnetic analysis method does not.

Therefore, after analyzing the previous results, we could conclude that of the three methods proposed, the electromagnetic method could be useful for very thin conductors with little separation between them and for complex geometries in which Mills and Grover can not be used, as it only serves for parallel and perpendicular conductors. When it was tried to increase the turns of the loop, it was the only one that differed from the rest. On the other hand, when the length of the loop was increased, it was also evident that this method did not behave correctly. For this reason, we can affirm that although it is a valid method and can give us an approximation if the separation between turns is big, it should be only used as a reliable source if that separation is minimal.

Regarding the two remaining methods, it must be noted that both offer good and similar results as it can be seen in Figures 14(a), 14(b), 14(c) and 14(d), but in the case of having to opt for one of them, we must emphasize that the numerical integration method carries a much higher computational cost than the Mills and Grover's one. In fact, it is clear by observing Figure 14 that this method had the best behavior in every of the cases. Then, we could conclude that because of its low computational cost and extreme accuracy, Mills and Grover's method would be the best way of calculating the inductance of a double magnetic loop, which is why it will be the one that we will use in our simulation programs.

6. Conclusions

This article is aimed to be a presentation of the double loop, where geometry, construction, operating mode and three possible ways to calculate its inductance have been explained. After presenting these three above-mentioned methods, an analysis has clarified that if precision is required, Mills and Grover's method or the numerical integration method must be used, as they both take into account the separation between turns, although we recommend to choose the first one because of its low computational cost.

In future papers, we will focus more closely on the advantages offered by using this type of loops, what will help to understand the need and importance of this paper. It will focus on the new vehicle magnetic profiles, the parameters that can be extracted from them and the benefits of using them in comparison with the conventional loops.

The reality is that magnetic loops, despite being from the eighties, are still the most used technology to capture data from traffic. For that reason, we must improve the existing infrastructure and provide this sensor with greater potential and reliability.

Data Availability

The data used to support the findings of this study are included within the article.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

References

- [1] R. L. Anderson, "Electromagnetic loop vehicle detectors," *IEEE Transactions on Vehicular Technology*, vol. VT-19, no. 1, pp. 23–30, 1970.
- [2] M. J. Caruso, T. Bratland, C. H. Smith, and R. Schneider, "A new perspective on magnetic field sensing," *Sensors (Peterborough, NH)*, vol. 15, no. 12, pp. 34–46, 1998.
- [3] L. A. Klein, D. R. P. Gibson, and M. K. Mills, *Traffic Detector Handbook*, vol. FHWAHRT-06-108, Federal Highway Administration, U.S. Department of Transportation, 2006.
- [4] N. L. Nihan, "Evaluation of forced flows on freeways with single-loop detectors," *Journal of Advanced Transportation*, vol. 34, no. 2, pp. 269–296, 2000.
- [5] H. Wang, Z. Li, D. Hurwitz, and J. Shi, "Parametric modeling of the heteroscedastic traffic speed variance from loop detector data," *Journal of Advanced Transportation*, vol. 49, no. 2, pp. 279–296, 2015.
- [6] M. Bugdol, Z. Segiet, M. Krecichwost, and P. Kasperek, "Vehicle detection system using magnetic sensors," *Transp. Problems*, vol. 9, no. 1, pp. 49–60, 2014.
- [7] F. C. Nemtanu, J. Schlingensiepen, D. Buretea, and V. Iordache, "Mobility as a service in smart cities," in *ICEIRD 2016 – Responsible Entrepreneurship Vision, Development and Ethics*, pp. 425–435, 2016.
- [8] T. V. Mathew, "Transportation systems engineering," in *IIT Bombay, Chapter 9– Intrusive Technologies*, School: Jessore Science & Technology University, Bangladesh, 2014.
- [9] S. Meta and M. G. Cinsdikici, "Vehicle-classification algorithm based on component analysis for single-loop inductive detector," *IEEE Transactions on Vehicular Technology*, vol. 59, no. 6, pp. 2795–2805, 2010.
- [10] D. Guilbert, S.-S. Ieng, C. L. Bastard, and Y. Wang, "Robust blind deconvolution process for vehicle reidentification by an inductive loop detector," *IEEE Sensors Journal*, vol. 14, no. 12, pp. 4315–4322, 2014.
- [11] L. Bhaskar, A. Sahai, D. Sinha, G. Varshney, and T. Jain, "Intelligent traffic light controller using inductive loops for vehicle detection," in *Proceedings of the 1st International Conference on Next Generation Computing Technologies, NGCT 2015*, pp. 518–522, India, September 2015.
- [12] J. Gadjia, R. Sroka, M. Stencil et al., *Measurements of Road Traffic Parameters*, Publishing House PWN, 2015.
- [13] L. A. Klein, *Sensors Technologies and Data Requirements for ITS*, Artech House, 2011.
- [14] Z. Marszalek, R. Sroka, and T. Zeglen, "Inductive loop for vehicle axle detection from first concepts to the system based on changes in the sensor impedance components," in *Proceedings of the 2015 20th International Conference on Methods and Models in Automation and Robotics (MMAR)*, pp. 765–769, Miedzyzdroje, Poland, August 2015.
- [15] Z. Marszalek, R. Sroka, and T. Zeglen, "Multi-frequency conditioning system of the inductive loop sensor — Simulation investigations," in *Proceedings of the 2017 22nd International Conference on Methods and Models in Automation and Robotics (MMAR)*, pp. 889–893, Miedzyzdroje, Poland, August 2017.

- [16] D. S. García, J. T. Arjona, and J. M. M. García, *Cuaderno Tecnológico de la PTC – Sistemas de Adquisición de Información de Tráfico: Estado Actual y Futuro*, Plataforma Tecnológica Española de la Carretera (PTC), Madrid, 2011.
- [17] A. L. J. Ortega, *OEP 2013. Especialidad: Gestión Técnica del Tráfico 2011*, Chapter 30, 2013.
- [18] S. S. M. Ali, B. George, L. Vanajakshi, and J. Venkatraman, “A multiple inductive loop vehicle detection system for heterogeneous and lane-less traffic,” in *Proceedings of the IEEE Transactions on Instrumentation and Measurement*, vol. 61, pp. 1353–1360, 2012.
- [19] J. H. Arroyo-Núñez, *Estudio del Comportamiento Magnético de Espiras Rectangulares para la Transmisión de Información de Corto Alcance en Sistemas Inteligentes de Transporte [Ph.D. thesis]*, Dept. of Electrical Engineering University Politécnica València, Valencia, 2016.
- [20] G. TurnBull, “Maxwell’s Equations,” in *Proceedings of the IEEE*, vol. 101, pp. 0018–9219, 2013.
- [21] S. Errede, “Lecture Notes 16. The magnetic vector potential,” *Physics 435 EM Fields & Sources I*, 2007.
- [22] M. Mills, “Self inductance formulas for multi-turn rectangular loops used with vehicle detectors,” in *Proceedings of the 33rd IEEE Vehicular Technology Conference*, pp. 65–73, May 1983.
- [23] E. B. Rosa and F. W. Grover, “Formulas and tables for the calculation of mutual and self-inductance (Revised),” *Bulletin of the Bureau of Standards*, vol. 8, no. 1, p. 1, 1912.
- [24] F. C. Nemptanu, I. M. Costea, and L. G. Obreja, “Model of intelligent traffic sensors — Application in hardware in the loop,” in *Proceedings of the 2017 40th International Spring Seminar on Electronics Technology (ISSE)*, pp. 1–5, Sofia, Bulgaria, May 2017.
- [25] A. Mocholí-Salcedo, J. H. Arroyo-Núñez, M. Víctor, J. Gumersindo, and A. Arroyo-Núñez, “Magnetic field generated by the loops used in traffic control systems,” in *Proceedings of the IEEE Transactions on Intelligent Transportation Systems*, 2017.
- [26] M. Misakian, “Equations for the magnetic field produced by one or more rectangular loops of wire in the same plane,” *Journal of research of the National Institute of Standards and Technology*, vol. 105, no. 4, pp. 557–564, 2000.
- [27] Y. Niwa, “A study of coils wound on rectangular frames with special reference to the calculation of inductance,” *Researches of the Electrotechnical Laboratory*, vol. 141, 1924.
- [28] F. W. Grover, *Inductance Calculations: Working Formulas and Tables*, New York, NY, USA, 1962.

Research Article

An Improved Robust Principal Component Analysis Model for Anomalies Detection of Subway Passenger Flow

Xuehui Wang ¹, Yong Zhang ¹, Hao Liu,² Yang Wang,¹
Lichun Wang ¹ and Baocai Yin¹

¹Beijing Advanced Innovation Center for Future Internet Technology, Beijing Key Laboratory of Multimedia and Intelligent Software Technology, Faculty of Information Technology, Beijing University of Technology, Beijing 100124, China

²Beijing Transportation Information Center, Beijing, China

Correspondence should be addressed to Yong Zhang; zhangyong2010@bjut.edu.cn

Received 16 May 2018; Accepted 31 July 2018; Published 14 August 2018

Academic Editor: Darius Andriukaitis

Copyright © 2018 Xuehui 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.

Subway is an important transportation means for residents, since it is always on schedule. However, some temporal management policies or unpredicted events may change passenger flow and then affect passengers requirement for punctuality. Thus, detecting anomaly event, mining its propagation law, and revealing its potential impact are important and helpful for improving management strategy; e.g., subway emergency management can predict flow change under the condition of knowing specific policy and estimate traffic impact brought by some big events such as vocal concerts and ball games. In this paper, we propose a novel anomalies detection method of subway passenger flow. In this method, an improved robust principal component analysis model is presented to detect anomalies; then ST-DBSCAN algorithm is used to group the station-level anomaly data on space-time dimensions to reveal the propagation law and potential impact of different anomaly events. The real flow data of Beijing subway are used for experiments. The experimental results show that the proposed method is effective for detecting anomalies of subway passenger flow in practices.

1. Introduction

Owing to the high efficiency and the comfort, subway has generally become first choice for citizens' daily travel, and it directly facilitates the city's economic development and people's quality of life. For example, as one of the busiest subway systems in the world, the Beijing subway has the world's largest annual ridership with 3.03 billion trips delivered in 2016, averaging 8.26 million per day, with peak single-day ridership reaching 10.52 million. The public transportation in Beijing accounts for 45% of total traffic, in which the ridership of subway dominates nearly 40%.

Although bringing great convenience for residents, the subway system becomes more vulnerable at the same time, as the subway system is a large and complicated network running in a restricted time schedule. For example, there are 22 lines and 370 stations in Beijing subway, and more than 500 trains are running on the network with the minimal peak

headway in 90 seconds. This will be more critical in the cases of encountering exceptional events, such as station accidents, major activities, and bad weathers. Once a station has an anomaly event, such as failure operation and chaos in station, the retention of passengers would happen, which would bring great loss with high security risks. Moreover, the bad situation would propagate through the urban subway system since it is a relatively closed and connected network. So the impact of anomaly event will not be restricted in a specific station, it may affect the traffic system in a large region, and the influence of abnormal events usually shows a certain space-time law. Thus, it is necessary to detect anomalies in urban subway transportation system and figure out its spreading rules, which can provide valuable proof for management to making strategy for dealing with abnormal events.

However, in the traditional road transportation system, many methods have been proposed for detecting transit anomalies, such as the Automatic Incident Detection (AID)

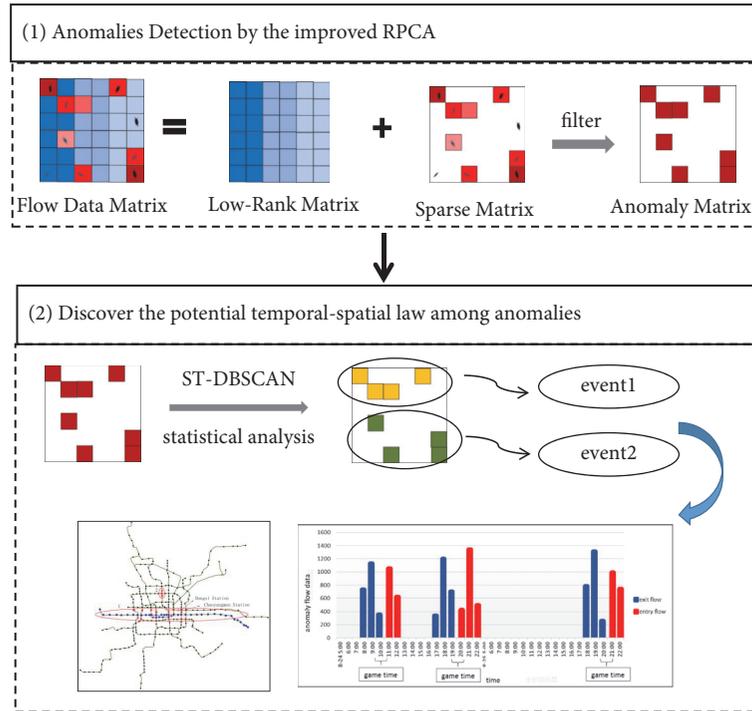


FIGURE 1: The whole procedure of anomalies detection by the improved RPCA. (1) Decompose the matrix of subway passenger flow into a low-rank matrix and a sparse matrix, filter the sparse matrix decomposed by the improved RPCA, and then acquire the anomaly matrix. (2) Mine the potential laws and discover the possible relations among anomalies through ST-DBSCAN algorithm and statistical analysis.

algorithms [1–3] based on comparison, statistics techniques, traffic flow model, and so on. These methods are mostly applied in freeway and urban roads, and they link the main regions of a city and try to find unexpected traffic flow between any two regions [4, 5]. As the subway is a different traffic system from the traditional road transportation, the above methods are difficult to introduce to subway system. At the present, few studies focus on the anomalies detection of passenger flow in subway transportation.

Since the subway anomaly events are always uncertain and sporadic, the anomalies show obvious sparse property among the whole subway traffic data. From this observation, based on the available subway passenger flow data collected by the AFC (Automatic Fare Collection) system, we propose a novel anomalies detection method based on Robust Principle Component Analysis (RPCA) model [6], which represents the temporal-spatial distribution of data and the sparsity of the anomalies by the low-rank and sparse regularization. Additionally, in order to reveal the propagation law and potential impact of different anomaly events, the ST-DBSCAN algorithm is adopted to group the station-level anomaly data on temporal-spatial dimensions. Thus the proposed method can not only detect anomaly of a single station but also find the relations among anomalies. Figure 1 shows the main structure of our model.

The main contributions of this paper are summarized as follows:

- (i) A novel anomaly detection method of subway passenger flow based on RPCA is proposed, which

utilizes low-rank nature of the passenger flow data and the sparsity of anomaly data. Experimental results demonstrate that our approach can achieve an accurate anomaly detection.

- (ii) The ST-DBSCAN clustering algorithm is adopted to explore the temporal-spatial propagation law of anomalies, and the obtained expected results are verified by tweet data. The distribution law of anomalous flow caused by different anomaly events can provide prior information to cope with possible anomalies.

The rest of this paper is organized as follows. Related works are summarized in Section 2. Section 3 gives the methodology. Section 4 reports experimental results on real data and their visualization analysis. Finally, we conclude this paper in Section 5.

2. Related Works

In this section, we review the commonly used methods for anomalies detection in traffic systems and introduce RPCA model related to our work.

2.1. Anomalies Detection Methods of Passenger Flow. Most of the existing anomalies detection methods of traffic flow are in highways or urban roads scenarios, and the traffic data from fixed detectors is usually used for analysis. The typical approaches include the statistical methods, the comparison methods, and the traffic flow model based methods. The

famous comparison algorithms are the California algorithm and its derivation [1], which discriminate the anomaly event by comparing traffic parameters between adjacent detectors. But they are not suitable for subway passenger flow because the relation between neighboring stations is not exactly similar with the relation between neighboring detectors. The statistical methods (like SND [2]) achieve traffic anomaly by judging change rate of the traffic parameters, and they adopt the threshold method [7] to identify unreasonable data values based on historical data. To these methods, the suitable thresholds are difficult to choose. The traffic flow model based methods (like McMaster algorithm [3]) define boundary between crowded traffic flow and noncrowded traffic flow to determine a speed threshold for distinguishing, which is not well for subway due to the difference of flow pattern while time and space scales change. In addition, the wavelet analysis [8] is used for detecting anomalous samples by separating the high frequency components and the low frequency components of traffic data.

For the transportation system of a city scale, the current studies on anomalies detection are mostly region-based. Pang and Linsey Xiaolin [4, 5] partition city into uniform grids and report anomalies if traffic volumes in neighboring cells are different, while Shekhar [9] focuses on detecting spatial outliers in graph structured datasets. Similarly, Liu [10] and Chawla [11] partition the city into disjoint regions linked by major roads and then find unexpected traffic flow between any two regions. However, the above methods are either road-based or region-based and the former cannot accurately identify location of events, and the latter may result in loss of information because of the coarse region partition.

There are few works concentrating on the anomalies detection of subway passenger flow. Some anomalies detections in subway system focus on the pedestrian abnormal activity inside the station [12, 13], and they generally adopt visual recognition techniques based on the video surveillance system in the station; thus the applied scale is valid only in the view of the cameras. Besides, the other studies on passenger flow data of subway mainly focus on passenger flow prediction and analysis [14, 15]. Differently, in this paper, we conduct anomalies detection of subway passenger flow and explore the temporal-spatial impact of anomaly events.

2.2. Robust Principal Component Analysis. Recently, due to the power of revealing the intrinsic structure or property underlying the data, the low-rank and sparse theory have been successfully applied in numerous areas such as image recovery and denoising [16], background modeling, and foreground object detection of video image [17]. RPCA is a typical model utilizing the low-rank and sparse matrix decomposition for data restoration and denoising. The basic idea is that the original data in form of a numerical matrix can be decomposed into a low-rank matrix and a sparse matrix as follows:

$$\begin{aligned} \min_{\mathbf{X}, \mathbf{A}} \quad & \text{rank}(\mathbf{X}) + \lambda_1 \|\mathbf{A}\|_0, \\ \text{s.t.} \quad & \mathbf{D} = \mathbf{X} + \mathbf{A} \end{aligned} \quad (1)$$

where $\mathbf{D} \in \mathbb{R}^{m \times n}$ is the raw data usually having noise, \mathbf{X} represents the expected clean data which is assumed having low-rank property, and \mathbf{A} represents the noise data or outlier which is considered being sparse. The target of RPCA in (1) is to estimate the unknown \mathbf{X} and \mathbf{A} given \mathbf{D} .

However, the optimization problem in (1) is a NP-hard problem [18] due to its nonconvexity and discontinuity. On one hand, the low-rank term should be processed properly. For this purpose, a widely used solving scheme is replacing rank(\mathbf{X}) by its convex envelope, nuclear norm $\|\cdot\|_*$ [6, 19], as nuclear norm minimization approaches can perform stably without knowing the target rank of the recovery matrix in advance. On the other hand, the nonconvexity and discreteness of the ℓ_0 penalty make it be not preferred. Considering that ℓ_1 is also good at modeling the sparse noise [6] and has high efficient solution, the ℓ_0 term in (1) is replaced with ℓ_1 . Thus, (1) can be written as

$$\begin{aligned} \min_{\mathbf{X}, \mathbf{A}} \quad & \|\mathbf{X}\|_* + \lambda_1 \|\mathbf{A}\|_1, \\ \text{s.t.} \quad & \mathbf{D} = \mathbf{X} + \mathbf{A} \end{aligned} \quad (2)$$

where $\|\mathbf{X}\|_* := \sum_i \sigma_i(\mathbf{X})$ denotes the nuclear norm, $\sigma_i(\mathbf{X})$ is the i th largest singular value of matrix \mathbf{X} , and $\|\mathbf{A}\|_1 := \sum_{i=1}^m \sum_{j=1}^n |a_{i,j}|$, $a_{i,j}$ is the element of \mathbf{A} .

In this paper, we introduce RPCA into the anomaly detection of subway passenger flow. Moreover, the passenger flow data matrix has low-rank structure because it shows regular cycles with respect to day, week, month, and year. In addition, the real-world data is usually polluted by noise or outliers, and the outliers are considered anomalies for detection. So we adopt RPCA to represent the subway passenger flow and detect the anomalies by the sparse outliers. Additionally, we consider the temporal correlation among the data and propose an improved RPCA. The next section will give the improved RPCA in detail.

3. Methodology

In this section, we first represent the subway passenger flow as a matrix and give it decomposition for anomalies detection. Then the improved RPCA is applied to obtain preliminary abnormal flow information. Finally, the detected anomalies are grouped into several clusters for revealing the temporal-spatial laws.

3.1. Subway Passenger Flow Representation and Decomposition. The raw subway passenger riding data are collected from the subway AFC system; they include the boarding or alighting time at a station, the boarding line ID or alighting, and the boarding station ID or alighting. Based on the raw riding data, the subway passenger flow data are calculated in one hour interval, and then we obtain the subway passenger flow matrix \mathbf{D} , which is constructed with the row and column corresponds to the date and the time interval of each day, respectively. Therefore, each element in the matrix represents the passenger flow of a station at a certain time interval of a certain day.

As the passenger flow of a subway station shows similar varying degrees taking year, month, week, hour, or minute as a cycle, the temporal patterns of the passenger flow matrix \mathbf{D} are typically a low-rank matrix [20]. Besides, passenger flow of adjacent stations also show certain similarity, which further supports the low-rank property of the passenger flow matrix. However when anomaly events happen, the low-rank property of the flow would be ruined by the outliers. So the matrix \mathbf{D} can be considered as a combination of normal and outliers. Let $x_{e,t}$ and $a_{e,t}$ represent the expected flow component and the outlier interference of a station at time t on date e , so the measured passenger flow at time t can be expressed as $d_{e,t} = x_{e,t} + a_{e,t}$. Collecting n measurements and introducing matrices $\mathbf{X} := [x_{e,t}]$, $\mathbf{A} := [a_{e,t}]$, the passenger flow matrix can be decomposed by

$$\mathbf{D} = \mathbf{X} + \mathbf{A} \quad (3)$$

By this decomposition, the subway passenger flow can be represented by two components: the expected flow \mathbf{X} and the anomalous part \mathbf{A} . The anomalous part \mathbf{A} is explained as special events or special activities around the station; it is sporadic over time and may last for short periods relative to the (possibly long) measurement period \mathbf{T} , which means that only a small fraction of the elements in observation traffic flow matrix \mathbf{D} is supposed to be anomalous. Therefore, the anomaly matrix \mathbf{A} would be sparse both in rows and columns.

From the above analysis, the subway passenger flow \mathbf{D} completely has the RPCA model in (2) with the low-rank and sparse terms. In the following, we further exploit the temporal constraint for the model and propose our improved RPCA model for the anomalies detection of subway passenger flow data.

3.2. The Improved RPCA. For the subway passenger flow matrix \mathbf{D} , the two adjacent rows of the same weekdays in different weeks are often approximately equal except some outliers, derived from the obvious day cycle of the passenger flow measurement. This property is conductively true for the corresponding expected flow \mathbf{X} , while the current RPCA model has no specific description for this important property. So we propose a constraint to keep the consistence among rows of \mathbf{X} by adding an item $\|\mathbf{H}\mathbf{X}\|_1$ to the current RPCA model. The matrix \mathbf{H} is defined as follows:

$$\mathbf{H} = \begin{pmatrix} 1 & -1 & 0 \cdots 0 & 0 \\ 0 & 1 & -1 \cdots 0 & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 \cdots 1 & -1 \\ 0 & 0 & 0 \cdots 0 & 1 \end{pmatrix} \quad (4)$$

(m-1) × m.

The above temporal differential matrix is $\mathbf{H} = \text{Toeplitz}(0, 1, -1)$, in which the central diagonal is defined as ones and the first upper diagonal is defined as negative ones. The temporal constraint matrix intuitively expresses the fact that nominal passenger flow matrices at same time

intervals for the same weekdays are usually similar. Actually, $\|\mathbf{H}\mathbf{X}\|_1$ captures consistence between two adjacent rows of \mathbf{X} . Moreover, compares with ℓ_2 norm, ℓ_1 norm is more inclusive and robust while considering temporal abrupt changes [6]. Thus we choose ℓ_1 norm to minimize $\mathbf{H}\mathbf{X}$, as it enforces the matrix \mathbf{X} temporally stable [21]. Hence, we revise the RPCA model in (5) and obtain the following improved RPCA model:

$$\begin{aligned} \min_{\mathbf{X}, \mathbf{A}} \quad & \|\mathbf{X}\|_* + \lambda_1 \|\mathbf{A}\|_1 + \lambda_2 \|\mathbf{H}\mathbf{X}\|_1, \\ \text{s.t.} \quad & \mathbf{D} = \mathbf{X} + \mathbf{A} \end{aligned} \quad (5)$$

where λ_2 controls weight of the term $\|\mathbf{H}\mathbf{X}\|_1$.

To solve the improved model, we adopt the Alternating Direction Method of Multiplier (ADMM) [22], which is a popular algorithm for solving convex optimization problems. For this purpose, three auxiliary variables $\mathbf{L} \in \mathbb{R}^{m \times r}$, $\mathbf{Q} \in \mathbb{R}^{r \times n}$, and $\mathbf{S} \in \mathbb{R}^{(m-1) \times n}$ are introduced; let $\mathbf{X} = \mathbf{L}\mathbf{Q}$ and $\mathbf{H}\mathbf{X} = \mathbf{S}$, where r is the decomposition rank of \mathbf{X} . Therefore (5) is rewritten as

$$\begin{aligned} \min_{\mathbf{X}, \mathbf{A}, \mathbf{S}, \mathbf{L}, \mathbf{Q}} \quad & \|\mathbf{L}\mathbf{Q}\|_* + \lambda_1 \|\mathbf{A}\|_1 + \lambda_2 \|\mathbf{S}\|_1, \\ \text{s.t.} \quad & \mathbf{D} = \mathbf{X} + \mathbf{A}, \\ & \mathbf{X} = \mathbf{L}\mathbf{Q}, \\ & \mathbf{H}\mathbf{X} = \mathbf{S} \end{aligned} \quad (6)$$

Remove the linear equality constraints in (6) with augmented Lagrangian method, and then we have the following objective function:

$$\begin{aligned} \mathbf{L}(\mathbf{X}, \mathbf{A}, \mathbf{S}, \mathbf{L}, \mathbf{Q}) = & \|\mathbf{L}\|_F^2 + \|\mathbf{Q}\|_F^2 + \lambda_1 \|\mathbf{A}\|_1 + \lambda_2 \|\mathbf{S}\|_1 \\ & + \langle \mathbf{Y}_1, \mathbf{D} - \mathbf{X} - \mathbf{A} \rangle \\ & + \frac{\mu}{2} \|\mathbf{D} - \mathbf{X} - \mathbf{A}\|_F^2 + \langle \mathbf{Y}_2, \mathbf{X} - \mathbf{L}\mathbf{Q} \rangle \\ & + \frac{\mu}{2} \|\mathbf{X} - \mathbf{L}\mathbf{Q}\|_F^2 + \langle \mathbf{Y}_3, \mathbf{H}\mathbf{X} - \mathbf{S} \rangle \\ & + \frac{\mu}{2} \|\mathbf{H}\mathbf{X} - \mathbf{S}\|_F^2 \end{aligned} \quad (7)$$

where \mathbf{Y}_1 , \mathbf{Y}_2 , and \mathbf{Y}_3 are Lagrange multipliers, $\mu > 0$ is adaptive penalty parameter, and $\langle \cdot, \cdot \rangle$ represents the standard trace inner product. We adopt an alternative iterations to solve this optimization as follows.

Update \mathbf{X} . When \mathbf{A} , \mathbf{S} , \mathbf{L} , and \mathbf{Q} are fixed, (7) degenerates into a function with respect to \mathbf{X} . So we solve \mathbf{X} by the following optimization problem:

$$\begin{aligned} \mathbf{X}^{(i+1)} = & \arg \min_{\mathbf{X}} \frac{\mu^{(i)}}{2} \left\| \mathbf{D} - \mathbf{X} - \mathbf{A}^{(i)} + \frac{\mathbf{Y}_1^{(i)}}{\mu^{(i)}} \right\|_F^2 \\ & + \frac{\mu^{(i)}}{2} \left\| \mathbf{X} - \mathbf{L}^{(i)}\mathbf{Q}^{(i)} + \frac{\mathbf{Y}_2^{(i)}}{\mu^{(i)}} \right\|_F^2 \end{aligned}$$

$$+ \frac{\mu^{(i)}}{2} \left\| \mathbf{H}\mathbf{X} - \mathbf{S}^{(i)} + \frac{\mathbf{Y}_3^{(i)}}{\mu^{(i)}} \right\|_F^2 \quad (8)$$

Taking derivative of the objective function in (8) and setting it to 0, the closed-form solution is given by

$$\begin{aligned} \mathbf{X}^{(i+1)} &= \left(2\mu^{(i)} \mathbf{I}^{n \times n} + \mu^{(i)} \mathbf{H}^T \mathbf{H} \right)^{-1} \\ &\cdot \left(\mu^{(i)} \left(\mathbf{D} - \mathbf{A}^{(i)} + \mathbf{L}^{(i)} \mathbf{Q}^{(i)} + \mathbf{H}^T \mathbf{S}^{(i)} \right) + \mathbf{Y}_1^{(i)} - \mathbf{Y}_2^{(i)} \right. \\ &\left. + \mathbf{H}^T \mathbf{Y}_3^{(i)} \right) \end{aligned} \quad (9)$$

Update A. When others are fixed, in order to update **A**, one needs to solve the following ℓ_1 minimization problem:

$$\begin{aligned} \mathbf{A}^{(i+1)} &= \arg \min_{\mathbf{A}} \lambda_1 \|\mathbf{A}\|_1 \\ &+ \frac{\mu^{(i)}}{2} \left\| \mathbf{A} - \left(\mathbf{D} - \mathbf{X}^{(i)} + \frac{\mathbf{Y}_1^{(i)}}{\mu^{(i)}} \right) \right\|_F^2 \end{aligned} \quad (10)$$

whose solution is given by [23]:

$$\mathbf{A}^{(i+1)} = \delta \left(\mathbf{D} - \mathbf{X}^{(i)} + \frac{\mathbf{Y}_1^{(i)}}{\mu^{(i)}}, \frac{\lambda_1}{\mu^{(i)}} \right) \quad (11)$$

where $\delta(a, b) = \text{sgn}(a)(|a| - b)$ for $|a| \geq b$ and is zero otherwise.

Update S. In a similar way with updating **A**, the closed-form solution of **S** is given by

$$\mathbf{S}^{(i+1)} = \delta \left(\mathbf{H}\mathbf{X}^{(i)} + \frac{\mathbf{Y}_3^{(i)}}{\mu^{(i)}}, \frac{\lambda_2}{\mu^{(i)}} \right) \quad (12)$$

Update L, Q. In a similar way with updating **X**, the closed-form solutions of **L**, **Q** are given by

$$\mathbf{L}^{(i+1)} = \left(\mu^{(i)} \mathbf{X}^{(i)} + \mathbf{Y}_2^{(i)} \right) \mathbf{Q}^{(i)T} \left(2\mathbf{I}^{r \times r} + \mu^{(i)} \mathbf{Q}^{(i)} \mathbf{Q}^{(i)T} \right)^{-1} \quad (13)$$

$$\begin{aligned} \mathbf{Q}^{(i+1)} \\ = \left(2\mathbf{I}^{r \times r} + \mu^{(i)} \mathbf{L}^{(i)T} \mathbf{L}^{(i)} \right)^{-1} \left(\mu^{(i)} \mathbf{L}^{(i)T} \mathbf{X}^{(i)} + \mathbf{L}^{(i)T} \mathbf{Y}_2^{(i)} \right) \end{aligned} \quad (14)$$

Update Y₁, Y₂, Y₃, and μ . The Lagrangian multipliers **Y₁**, **Y₂**, and **Y₃** and penalty parameter μ could be updated as follows:

$$\begin{aligned} \mathbf{Y}_1^{(i+1)} &= \mathbf{Y}_1^{(i)} + \mu^{(i)} \left(\mathbf{D} - \mathbf{X}^{(i)} - \mathbf{A}^{(i)} \right), \\ \mathbf{Y}_2^{(i+1)} &= \mathbf{Y}_2^{(i)} + \mu^{(i)} \left(\mathbf{X}^{(i)} - \mathbf{L}^{(i)} \mathbf{Q}^{(i)} \right), \\ \mathbf{Y}_3^{(i+1)} &= \mathbf{Y}_3^{(i)} + \mu^{(i)} \left(\mathbf{H}\mathbf{X}^{(i)} - \mathbf{S}^{(i)} \right), \\ \mu^{(i+1)} &= \min \left(\rho \mu^{(i)}, \mu^{max} \right) \end{aligned} \quad (15)$$

where $\rho > 1$ is a constant and μ^{max} is the upper bound of μ .

Input: Data matrix **D**, the parameters $\lambda_1 > 0, \lambda_2 > 0, r > 0$,
Initialize: $\mathbf{X}^{(0)} = \mathbf{A}^{(0)} = \mathbf{S}^{(0)} = \mathbf{1} \in \mathbb{R}^{m \times n}$,
 $\mathbf{Y}_1^{(0)} = \mathbf{Y}_2^{(0)} = \mathbf{1} \in \mathbb{R}^{m \times n}$, $\mathbf{Y}_3^{(0)} = \mathbf{1} \in \mathbb{R}^{m \times n}$,
 $\mathbf{L}^{(0)} = \mathbf{1} \in \mathbb{R}^{m \times r}$, $\mathbf{Q}^{(0)} = \mathbf{1} \in \mathbb{R}^{r \times n}$,
 $\mu^{(0)} = 10^{-6}, \rho = 1.1, \varepsilon = 10^{-6}, MaxIter = 1000, i = 0$.
1: **while** not converged and $i < MaxIter$ **do**
2: Update **X** : via (9)
3: Update **A** : via (11)
4: Update **S** : via (12)
5: Update **L** : via (13)
6: Update **Q** : via (14)
7: Update the multipliers: via (15)
8: $i = i + 1$.
Output: Expected matrix **X**, Sparse matrix **A**.

ALGORITHM 1: Solving for the improved RPCA.

Convergence Conditions. The stopping criterion is measured by the following problem:

$$\max \left\{ \begin{array}{l} \|\mathbf{X}^{(i+1)} - \mathbf{X}^{(i)}\|_{\infty}, \\ \|\mathbf{A}^{(i+1)} - \mathbf{A}^{(i)}\|_{\infty}, \\ \|\mathbf{S}^{(i+1)} - \mathbf{S}^{(i)}\|_{\infty}, \\ \|\mathbf{L}^{(i+1)} - \mathbf{L}^{(i)}\|_{\infty}, \\ \|\mathbf{Q}^{(i+1)} - \mathbf{Q}^{(i)}\|_{\infty} \end{array} \right\} \leq \varepsilon. \quad (16)$$

where ε is tolerance error. If the convergence condition is met, the iteration terminates. The overall algorithm is summarized in Algorithm 1.

Once solving the improved RPCA, we obtain the expected flow **X** and the anomalous part **A**. To eliminate the interference of the noise, we use the three-sigma rule of thumb [24] to filter elements of **A**. ω_j is the standard deviation of $x_{1,j}, x_{2,j}, \dots, x_{i,j}, \dots, x_{m,j}$; if $-3\omega_j \leq a_{i,j} \leq +3\omega_j$ is considered an allowable deviation and set as $a_{i,j} = 0$, then we get the anomaly flow matrix $\hat{\mathbf{A}}$. Each element of $\hat{\mathbf{A}}$ represents the abnormal amplitude of the space-time position; it may be positive or negative. The positive indicates that the passenger flow is higher than the expected flow and the negative indicates the passenger flow is lower than the expected flow.

3.3. Discovering the Potential Temporal-Spatial Laws among Anomalies. Based on the improved RPCA, the anomalies of subway passenger flow are detected. To explore the potential laws of anomalies propagation, we group the anomalies into several clusters to identify anomalies in the region and their propagation laws. Because the detected anomalies have similar temporal-spatial characteristics, we use ST-DBSCAN algorithm [25] to cluster the station-level anomalies to find the anomaly in the region. We regard $p(lon, lan, t)$ as the feature of an anomaly data object, lon and lan are the longitude and the latitude of a station, and t denotes the time interval. The ST-DBSCAN algorithm requires three parameters: space radius R , time window ΔT , and density threshold $MinPts$; the

first two parameters determine neighborhood on temporal-spatial dimension.

The algorithm starts with the earliest anomaly data object p and retrieves all neighbors of point p within spatiotemporal neighborhood. If the number of neighbors is greater than $MinPts$, a new cluster is created which has p as core of the cluster. Then, the algorithm iteratively collects neighbors beginning with another core point. The above procedure continues until all points have been processed.

4. Experiments

In this section, we evaluate the robustness of the improved RPCA by adding noise on a set of real subway passenger flow data, comparing with RPCA [6] the wavelet transform method [8] and the threshold method [7]. Then we apply our proposed framework on the real subway passenger flow data for anomalies detection and analysis; meanwhile the results are verified with traffic related tweet data.

4.1. Robustness Evaluation. The improved RPCA model is characterized by its robustness to noise, so we first validate the performance of our methods on noisy passenger flow data compared with the related methods.

First, we construct three real-world passenger flow datasets from three different geographical positions shown in Figure 2. By exploiting the strong weekly seasonality observed in the data, we convert hourly flow within one week into a row vector and stack 12 weeks vector to form the data matrix which contains much noise. To implement the verification experiment, it needs to know the ground truth. In the case of ground truth being unavailable, we have to estimate a relatively accurate ground truth. Here, we use 4 layers of wavelet to filter the small white noise and then take the average as ground truth value. As a result, we get three relatively clean and ideal ground truth datasets, denoted by \mathbf{G}_i ($i = 1, 2, 3$).

Next, we add sparse noise on the ground truth matrices to simulate the corresponding noise matrices \mathbf{A}_i ($i = 1, 2, 3$). The randomly corrupted proportion cp of these matrices varies from 0.06 to 0.50; the fluctuation range is $\pm 80\%$ of the average of \mathbf{G}_i ($i = 1, 2, 3$). So we obtain the noisy passenger flow matrices by mixing the ground truth matrices \mathbf{G}_i and the produced noise matrices \mathbf{A}_i by $\mathbf{D}_i = \mathbf{G}_i + \mathbf{A}_i$. These datasets will be used as the test datasets for anomalies detection and evaluating the robustness of the proposed method. The properties of the constructed datasets are summarized in Table 1.

Evaluation Criteria. To evaluate the performance of the improved RPCA algorithm, we use the precision rate pr in [21] to evaluate the recognition accuracy of anomalies, which are defined as follows:

$$pr = \frac{2 * precision * recall}{precision + recall} \quad (17)$$

where $precision = a_{true}/a_{all}$, $recall = a_{true}/a$, a_{all} and a_{true} denote the number of anomalies recognized by our model

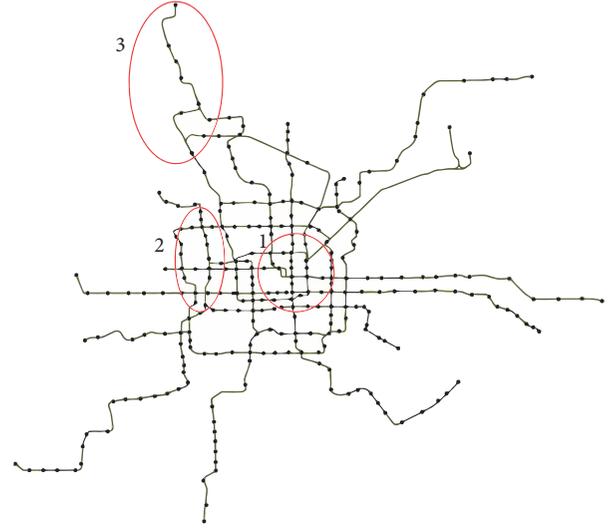


FIGURE 2: The selected stations in three areas of Beijing.

TABLE 1: Datasets description.

| Dataset | Numbers of stations | Size |
|---------|---------------------|-----------|
| D1 | 20 | 240 × 126 |
| D2 | 18 | 216 × 126 |
| D3 | 12 | 144 × 126 |

and the number of true anomalies among them, respectively, and a represents the actual number of anomalies. pr is calculated by averaging results over 10 runs.

Parameters Setting. The improved RPCA has three parameters λ_1 , λ_2 , and decomposition rank r , and they are important for the performance. Rank r needs to be as small as possible to minimize matrix sparsity and low-rank error. Here, we use singular value decomposition (SVD) [26] to estimate a superior rank for these three datasets. Figure 3 shows the distribution of the singular value of the three ground truth datasets. The x-axis presents the i th singular values and the y-axis presents the cumulative ratio of the first i singular values to the sum of all singular values. It can be found that the first 14 singular values almost dominate nearly 90% energy in all three datasets. To simplify, the rank r is set as 14 for all datasets.

For λ_1 and λ_2 , we first change one parameter while fixing the other parameter in the model, and the parameter is gradually taken as 10^{-3} , 10^{-2} , \dots , 10^3 . Then we achieve the relatively superior parameters. Next, we tune these parameters in a narrow range from 10^{-1} to 10 by step of 0.1. Finally we obtain the relatively optimal λ_1 and λ_2 . The setting of experimental parameters is shown in Table 2.

In our experiments, we apply 5 layers of discrete wavelet transform based on the wavelet of DB4. For the threshold method, the threshold value of flow at different time interval is different and we compute the mean value Av_j and standard deviation σ_j of $d_{1,j}$, $d_{2,j}$, \dots , $d_{i,j}$, \dots , $d_{m,j}$ and

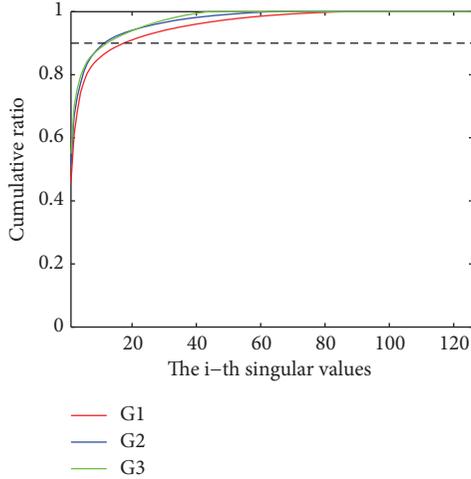


FIGURE 3: Cumulative ratio distribution of singular value in three ground truth datasets.

TABLE 2: Experiment parameters.

| Dataset | The improved RPCA | | RPCA |
|---------|-------------------|-------------|-----------|
| | λ_1 | λ_2 | λ |
| D1 | 1 | 0.4 | 0.1 |
| D2 | 1.4 | 0.4 | 0.2 |
| D3 | 1.1 | 0.5 | 0.1 |

set the confidence interval as $(Av_j - 3\sigma_j, Av_j + 3\sigma_j)$; it is judged to be an anomaly if $d_{i,j}$ is beyond the confidence interval.

Experiment Result Analysis. Figure 4 is the comparison results of the four methods. As can be seen, there is a downward trend of pr with the increasing of data corrupted proportion. The improved RPCA is superior to other methods, followed by RPCA. Notice that the threshold method has worse performance, because the noise data reduce the calculation accuracy of the threshold range. When the data corrupted proportion is low, the wavelet transform has a good detection result, but the local stationarity is destroyed in a high corrupted proportion, which results in the detection accuracy of a steady decline. The improved RPCA is more robust than RPCA even in high corrupted proportion, because the constraint item $\|\mathbf{HX}\|_1$ can capture the feature of abrupt changes of the time series data when the sparseness of anomalies becomes weak. Additionally, the improved RPCA performs well on different stations from different geographical positions. In a word, the improved RPCA is more suitable for anomalies detection of subway passenger flow.

4.2. Anomalies Detection and Verification. In order to demonstrate the practicability and the authenticity of the improved RPCA, we conduct anomalies detection experiments on real-world datasets and verify the results with tweet data. Figure 5 shows the decomposition results of the exit

flow of Xidan station. The low-rank expected flow matrix \mathbf{X} represents the weekly pattern and the anomaly matrix $\hat{\mathbf{A}}$ successfully captures multiple outliers.

To further analyze and verify the anomalies, we collect tweet data which contain a wide variety of information and retrieve events information through natural language processing method. There are four explanatory anomaly regions, highlighted by ellipses in Figure 5. They correspond with the following events 1 ~ 3 extracted from tweet data, as shown in Figure 6, and specific analysis as follows:

- (i) Event1: The ellipse region 1 in Figure 5 shows the increasing of the exit flow lasting about three hours in the evening. It is because many large shopping malls near Xidan station held sales for Chinese Valentine's Day, which attracted massive customers and led a rise in exit flow.
- (ii) Event2: In ellipse regions 2 and 3, the flow was declined. It is because Xidan station was closed for facilitating celebration parade for the 70th anniversary victory of the anti-Japanese war.
- (iii) Event3: In ellipse region 4, the exit flow was higher than usual. Because it was a commuter day due to legal exchanging holiday, therefore the flow was increasing and consistent with the flow of a working day.

The improved RPCA can not only identify anomalies at the station level but also accurately detect anomalies. These anomalies could be used for a reference for real-time alerting.

4.3. Discovering the Potential Temporal-Spatial Laws among Anomalies. An isolated anomaly may affect neighbored stations consecutively, so anomalies among some stations have strong temporal-spatial correlations. Grouping several anomalies along temporal-spatial dimensions may reveal the evolution or the impact of the isolated anomaly; hence we adapt ST-DBSCAN clustering algorithm to group the anomalies to analyze the propagation feature of the anomalies.

In experiments, space radius $R = 0.03$ (Euclidean distance of latitude and longitude between two adjacent stations), time window $\Delta T = 1$ hour interval, and density threshold $MinPts = 3$ work well. We cluster the anomalies of all stations in one week and name each cluster as an anomaly event.

In Figure 7, we use ellipses to highlight four anomaly events grouped by ST-DBSCAN, and these clustered results are easier to be verified by tweet data and visually analyzed. The ellipse region 1 in Figure 7(a) shows the entry flow decrease of the stations on the same line. It was lasting for five hours and induced by the closure strategy. Meanwhile, it also led to the flow increase of the nearby stations. In particular, the transfer stations such as Dongsi station and Chaoyangmen station had an obvious flow increase. The ellipse region 2 in Figure 7(a) shows the surges of exit flow as attendees traveled to Bird's Nest stadium for the opening ceremony of IAAF World Championships. In

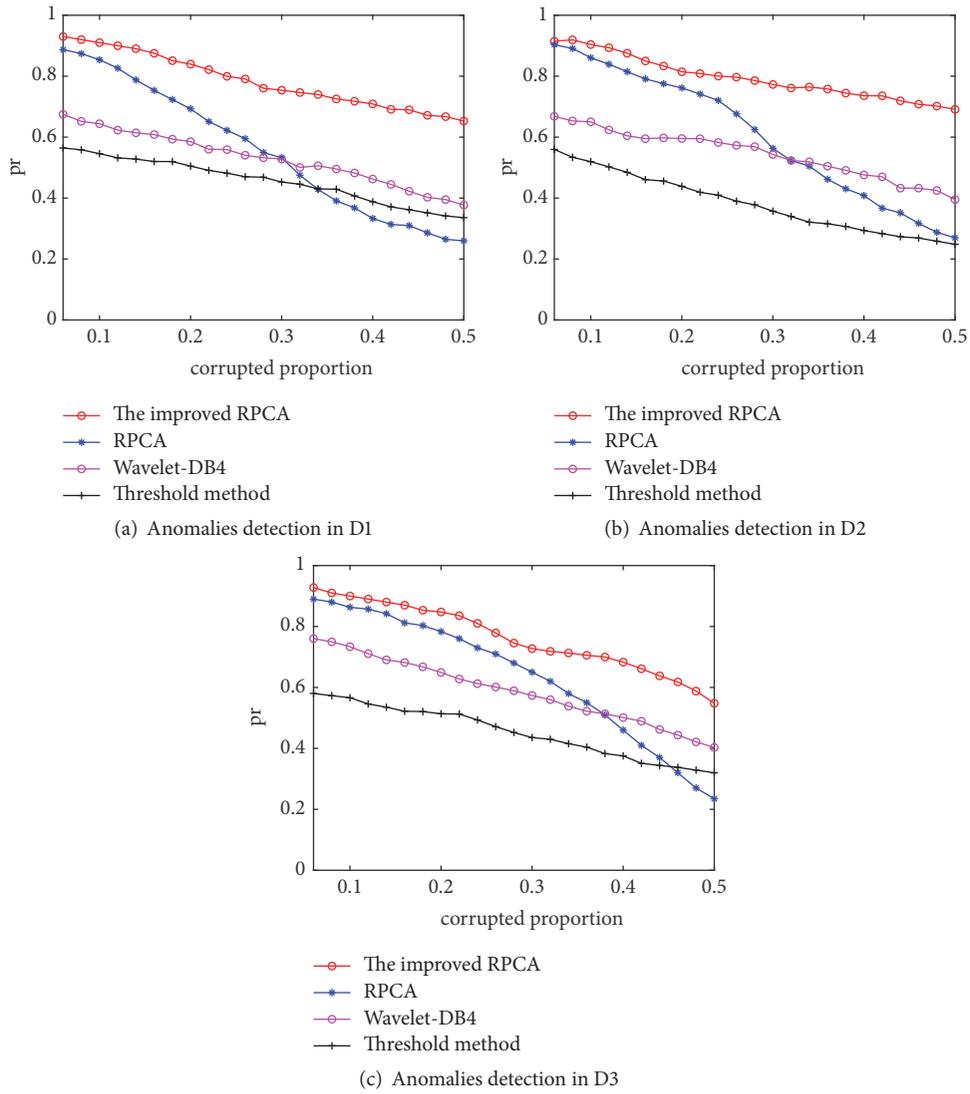


FIGURE 4: Performance of anomalies detection in three ground truth datasets.

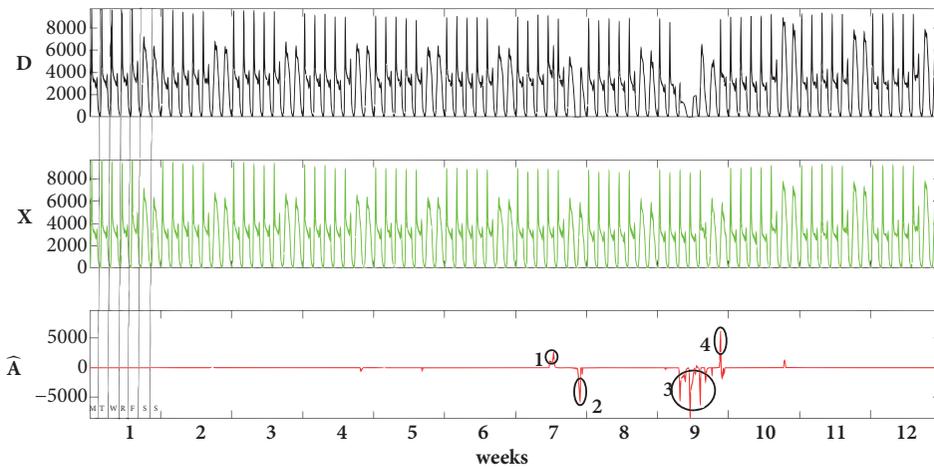
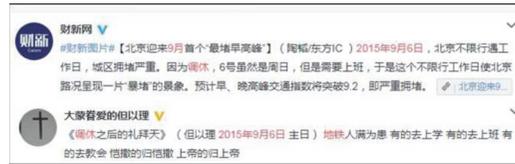


FIGURE 5: Component decomposition of exit flow of Xidan station by the improved RPCA. The horizontal axis is time interval covering 12 weeks, and the vertical axis is flow value.



(a) Event: August 20, 2015, is Chinese Valentine's Day, and many large shopping malls near Xidan station held sales activities

(b) Event2: To facilitate celebration parade for the 70th anniversary victory of the anti-Japanese war, the traffic authorities in Beijing imposed closure restriction measures on Xidan station



(c) Event3: September 6, 2015, is Sunday but exchanged with the working day because of legal exchanging holiday

FIGURE 6: The events information sent by tweet users.

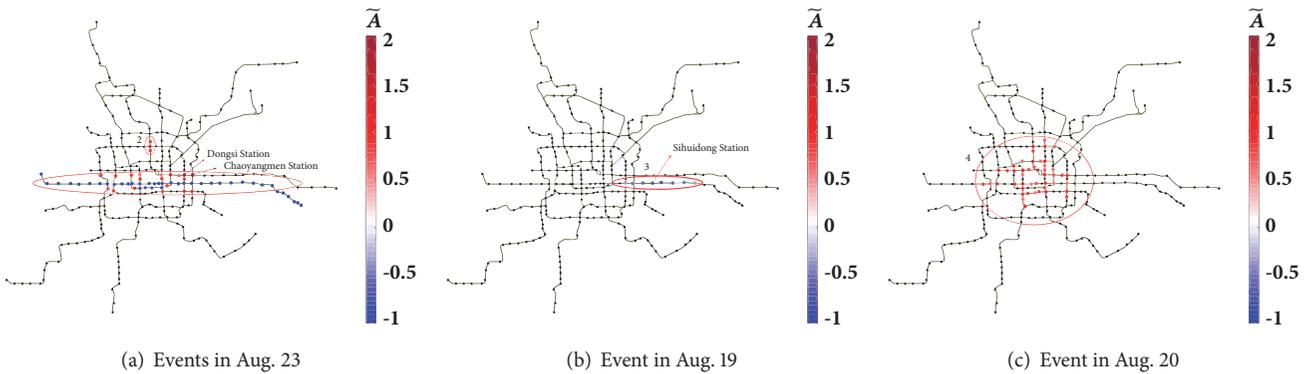


FIGURE 7: Anomaly events (labeled with ellipse). Red points mean traffic flow increasing and blue points mean traffic flow declining. We normalize anomaly A to get $\tilde{A} = A/X$ and values of \tilde{A} denotes deviation from the expectation.

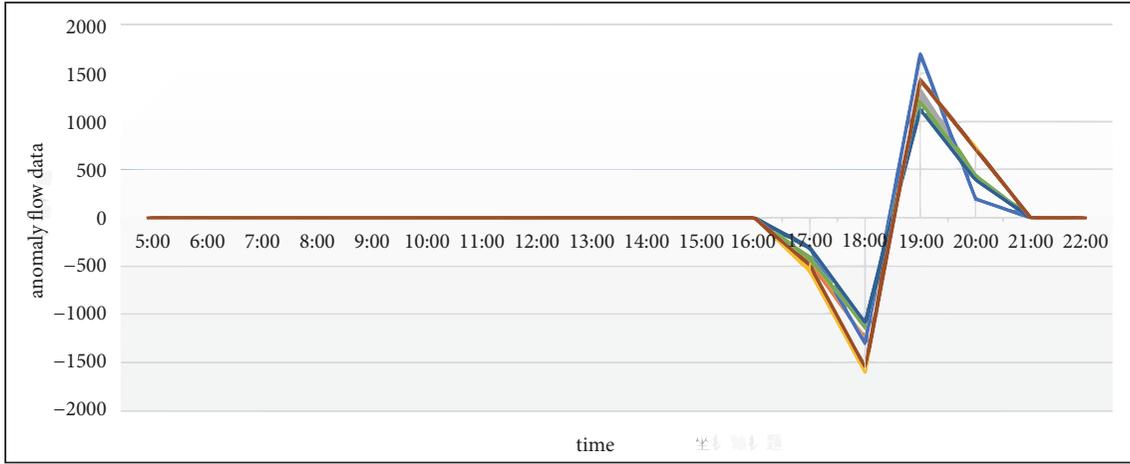
Figure 7(b), ellipse region 3 shows the spreading of the anomaly caused by an hour's breakdown of the train on Sihuidong station. In Figure 7(c), ellipse region 4 shows the entry flow increase of many stations for one day, since the traffic control of city roads led more people to choose the subway.

After the clustering and verification process, we can discover the potential temporal-spatial laws among anomalies from the following three aspects:

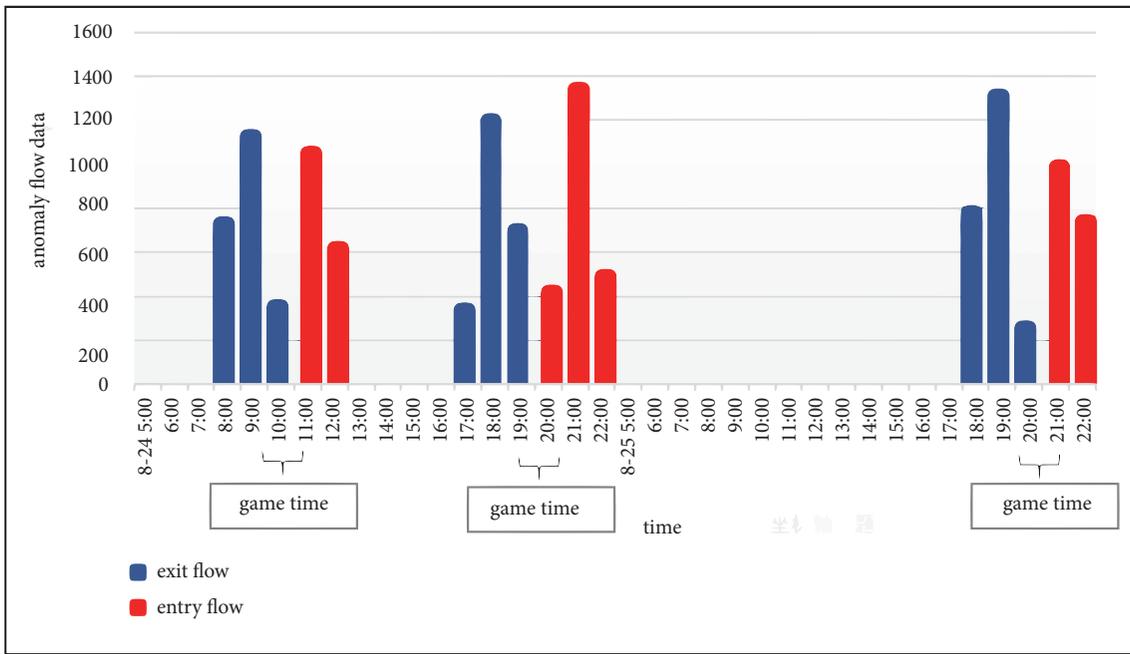
- (i) Distribution and spreading of anomalies along time and space: how many stations? And how long are they affected? The center of anomalies and the range of spreading from some destine anomaly. As shown in Figure 7(c), ellipse region 4 shows the stations affected by traffic control measures. Ellipse region 3 shows the anomaly in Sihuidong station spread to the adjacent stations.

- (ii) The serious degree of anomalies: Values in \tilde{A} reflect the serious degree of anomalies. In Figure 7, affected stations labeled with the red color having different levels reveal the degrees of impact. The heavier the color, the more severe the anomaly impact.
- (iii) The potential impact caused by events: Some anomaly events not only affect the corresponding stations, but also cause the potential impact on the surrounding stations. As shown in ellipse region 1 of Figure 7(a), the closure strategy resulted in a potential increasing flow of the surrounding stations.

Furthermore, we apply the statistical analysis to get some rules shown in Figure 8. The detected events are classified into two categories: some are predefined such as traffic control and vocal concerts and some are emergencies such as subway device failure and a sudden heavy rain. For both



(a)



(b)

FIGURE 8: In the two figures, the horizontal axis is time interval and the vertical axis is anomaly flow. (a) shows the exit anomaly flow of multiple stations on August 7; lines represent different stations. From 17:00 to 20:00, there are continuous negative and positive anomalies, because there was a sudden heavy rain at 17:00 that resulted in the passenger flow peak moving back. (b) shows the anomaly flow of Olympic Sports Center Subway Station near the Bird's Nest stadium. It is obvious that the exit flow surged before the beginning of game and the entry flow surged after the end of game.

two categories events, the above analysis with our method can provide a beneficial suggestion for subway managers:

- (i) For emergency events, our framework provides distribution laws of anomaly events, and these can be used for estimating anomalies' propagation and impact on adjacent stations. As shown in Figure 8(a), a sudden heavy rain caused a delay of the evening rush hour, so that managers can further push announcement timely to remind passengers and take emergency measures. This would prevent subway station from chaos and

hazard spreading and also save the travel time of passengers.

- (ii) For predefined events, our framework indicates detailed rules along spatial and temporal dimension, so that subway managers can obtain prior information and make sufficient preparations to cope with possible anomalies. As shown in Figure 8(b), the exit flow of Olympic Sports Center Station surged in the two hours before the beginning of one game, and the entry flow surged in the two hours after the end of

this game. These anomalies rules can help to estimate the impact of anomaly flow involved to major urban events and then take mitigation strategies in advance.

5. Conclusion

In this paper, the improved RPCA is suggested to detect station-level anomalies in subway, and ST-DBSCAN algorithm is used to group the detected station-level anomalies into clusters named as anomaly events. This framework can not only precisely locate anomalies in temporal dimension but also find the distribution and spreading in temporal and spatial dimension. With the detection results and impact analysis of events, subway managers can estimate traffic flow impact involved to predicted events and then take corresponding measures. Besides, they can push announcement timely for unpredicted events through decomposing the real-time data.

In future, we shall improve our work in three aspects. First, we shall extend our model to anomalies prediction as well as anomalies propagation process. Second, we shall consider temporal-spatial distribution by extracting comprehensive temporal and spatial information, e.g., OD flow data. Third, we shall propose more efficient ADMM algorithm for solving the proposed model and propose convergence analysis of the algorithm.

Data Availability

The data used to support the findings of this study are included within the article.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

This work is supported by Beijing Municipal Science & Technology Commission (Grant nos. Z151100002115040, Z171100000517003, and Z171100000517004), Project of Beijing Municipal Education Commission (Grant nos. KM201510005025, KM201610005033), and Funding Project for Academic Human Resources Development in Institutions of Higher Learning under the Jurisdiction of Beijing Municipality (Grant no. IDHT20150504).

References

- [1] H. J. Payne and S. C. Tignor, "Freeway Incident-Detection Algorithms Based on Decision Trees with States," *Transportation Research Record*, no. 682, pp. 30–37, 1978.
- [2] C. L. Dudek, C. J. Messer, and N. B. Nuckles, "Incident detection on urban freeways," *Transportation Research Record*, 1974.
- [3] B. N. Persaud and F. L. Hall, "Catastrophe theory and patterns in 30-second freeway traffic data- Implications for incident detection," *Transportation Research Part A: General*, vol. 23, no. 2, pp. 103–113, 1989.
- [4] L. X. Pang, S. Chawla, W. Liu et al., "On mining anomalous patterns in road traffic streams," in *Proceedings of the International Conference on Advanced Data Mining and Applications*, pp. 237–251, 2011.
- [5] L. X. Pang, S. Chawla, W. Liu, and Y. Zheng, "On detection of emerging anomalous traffic patterns using gps data," *Data & Knowledge Engineering*, vol. 87, no. 9, pp. 357–373, 2013.
- [6] E. J. Candes, X. Li, Y. Ma, and J. Wright, "Robust principal component analysis?" *Journal of the ACM*, vol. 58, no. 3, article 11, 2011.
- [7] S. M. Turner, "Guidelines for developing its data archiving systems," *Data Collection*, 2001.
- [8] L. Zhi Min, Y. Liang You, P. Xue et al., "Short-term traffic flow detection based on wavelet," *Application Research of Computers*, vol. 28, no. 5, pp. 1677–1678, 2011.
- [9] S. Shekhar, C.-T. Lu, and P. Zhang, "Detecting graph-based spatial outliers: algorithms and applications (a summary of results)," *Intelligent Data Analysis*, vol. 6, no. 5, pp. 451–468, 2001.
- [10] W. Liu, Y. Zheng, S. Chawla, J. Yuan, and X. Xie, "Discovering spatio-temporal causal interactions in traffic data streams," in *Proceedings of the 17th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pp. 1010–1018, August 2011.
- [11] S. Chawla, Y. Zheng, and J. Hu, "Inferring the root cause in road traffic anomalies," in *Proceedings of the 12th IEEE International Conference on Data Mining (ICDM '12)*, pp. 141–150, December 2012.
- [12] D. Jin and S. Zhu, "Spatio-temporal feature based anomaly event recognition," in *Proceedings of the 35th Chinese Control Conference, CCC 2016*, pp. 3818–3822, China, July 2016.
- [13] X. Sun, S. Zhu, and Y. Cheng, "Temporal-spatial coherence based abnormal behavior detection," in *Proceedings of the 29th Chinese Control and Decision Conference, CCDC 2017*, pp. 1997–2001, China, May 2017.
- [14] P. Wang, C. Wu, and X. Gao, "Research on subway passenger flow combination prediction model based on RBF neural networks and LSSVM," in *Proceedings of the 28th Chinese Control and Decision Conference, CCDC 2016*, pp. 6064–6068, China, May 2016.
- [15] D. Y. Zhang and H. N. Yang, "Passenger Flow Analysis in Subway Using a Kind of Neural Network," *Applied Mechanics and Materials*, vol. 713–715, pp. 2284–2287, 2015.
- [16] Z. L. Zhang and Y. F. Yan, "Image denoising based on recovering low dimensional manifolds in patch space," *Computer Engineering & Design*, vol. 34, no. 2, pp. 579–583, 2013.
- [17] B. Y. Wang and X. G. Wang, "Research on background modeling based on low-rank matrix recovery," *Applied Mechanics and Materials*, vol. 635–637, pp. 1056–1059, 2014.
- [18] B. K. Natarajan, "Sparse approximate solutions to linear systems," *SIAM Journal on Computing*, vol. 24, no. 2, pp. 227–234, 1995.
- [19] X. Zhou, C. Yang, and W. Yu, "Moving object detection by detecting contiguous outliers in the low-rank representation," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 35, no. 3, pp. 597–610, 2013.
- [20] A. Lakhina, M. Crovella, and C. Diot, "Diagnosing network-wide traffic anomalies," in *Proceedings of the Conference on Applications, Technologies, Architectures, and Protocols for Computer Communications*, pp. 219–230, 2004.
- [21] W. Ye, L. Chen, G. Yang, H. Dai, and F. Xiao, "Anomaly-tolerant traffic matrix estimation via prior information guided matrix completion," *IEEE Access*, vol. 5, pp. 3172–3182, 2017.

- [22] E. Wei and A. Ozdaglar, "Distributed alternating direction method of multipliers," *Decision and Control*, pp. 5445–5450, 2012.
- [23] A. M. Bruckstein, D. L. Donoho, and M. Elad, "From sparse solutions of systems of equations to sparse modeling of signals and images," *SIAM Review*, vol. 51, no. 1, pp. 34–81, 2009.
- [24] M. Zhang and Y. Hui, "The pata criterion and rejecting the abnormal value," *Journal of Zhengzhou University of Technology*, vol. 18, no. 1, pp. 84–88, 1997.
- [25] D. Birant and A. Kut, "St-dbscan: An algorithm for clustering spatial;ctemporal data," *Data & Knowledge Engineering*, vol. 60, no. 1, pp. 208–221, 2007.
- [26] D. Kalman, "A Singularly Valuable Decomposition: The SVD of a Matrix," *The College Mathematics Journal*, vol. 27, no. 1, pp. 2–23, 1996.

Research Article

A Low-Cost Consistent Vehicle Localization Based on Interval Constraint Propagation

Zhan Wang and Alain Lambert 

Laboratoire de Recherche en Informatique (LRI), CNRS, Univ Paris-Sud, Université Paris-Saclay, 91403 Orsay, France

Correspondence should be addressed to Alain Lambert; alain.lambert@u-psud.fr

Received 10 March 2018; Revised 3 May 2018; Accepted 13 May 2018; Published 24 June 2018

Academic Editor: Krzysztof Okarma

Copyright © 2018 Zhan Wang and Alain Lambert. 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.

Probabilistic techniques (such as Extended Kalman Filter and Particle Filter) have long been used to solve robotic localization and mapping problem. Despite their good performance in practical applications, they could suffer inconsistency problems. This paper proposes an interval analysis based method to estimate the vehicle pose (position and orientation) in a consistent way, by fusing low-cost sensors and map data. We cast the localization problem into an Interval Constraint Satisfaction Problem (ICSP), solved via Interval Constraint Propagation (ICP) techniques. An interval map is built when a vehicle embedding expensive sensors navigates around the environment. Then vehicles with low-cost sensors (dead reckoning and monocular camera) can use this map for ego-localization. Experimental results show the soundness of the proposed method in achieving consistent localization.

1. Introduction

Using sensory information to locate the robot in its environment is one of the most fundamental problem providing a mobile robot with autonomous capabilities. A vast number of works [1, 2] dedicated to such problems use both proprioceptive and exteroceptive sensors. Proprioceptive sensors such as odometers and gyrometers are used to calculate the elementary movements, which are used to estimate the robot pose (position and orientation). However, this generates cumulative errors as the robot moves [3, 4]. To cope with this problem, exteroceptive sensors (lasers, sonars, GPS, cameras, etc.) are used to gather information either from the surrounding environment or from a global reference, aiming at eliminating the accumulated errors and improving the estimation of the robot pose [5, 6].

Probabilistic techniques have long been used in solving robotic localization and mapping problem. The most commonly used methods are Kalman filtering and Particle filtering [7, 8]. These methods associate a probability to a position, but nothing ensures that the robot is indeed in the position with the highest probability. It has long been noticed by the research community that probabilistic methods suffer

the inconsistency problem [9]. It is due to the fact that they are based on the assumption that the sensor noises are described by probability density function. However, in real world, many sensors do not possess Gaussian noise model or their distributions are simply not available [10].

Interval analysis [11] is an alternative and less known method which allows us to solve nonlinear problems in a guaranteed way. Instead of making hypothesis on the probability distribution of sensor noises, interval methods take the assumption that all the noises are bounded within known limits. This seems to be a realistic representation as most sensor manufacturers provide the maximum and minimum measurement errors under suitable working conditions. These extreme values of error can then be regarded as the error bounds. Interval based algorithms can be used to recursively propagate such bounded errors by using consistency techniques and systematic search methods. Contrary to probability methods, interval methods provide guaranteed sets which enclose the real state, without losing any feasible value.

Interval methods have achieved promising results in parameter and state estimation tasks [12–14], and they found their way to the mobile robotic localization and mapping

area. Instead of considering Gaussian noise, all model and measurement noises are assumed to be bounded. The aim is to guarantee the configurations of the robot pose that are consistent with the given measurements and noise bounds. Kieffer [15] and Jaulin [16] provided simulation results for a robot equipped with a belt of on-board ultrasonic sensors. The feasibility of interval methods on localization applications had been shown. Followed-up researchers focused on real-time implementation of such localization technique. Kieffer [17] and Seignez [18] presented results for a mobile robot navigating in an indoor environment, with odometers mounted on each rear wheel to measure the motion and sonars located in different directions to perceive the environment. The ability of interval method to cope with erroneous data and to obtain accurate estimations of the robot pose was demonstrated. Lambert [19] extended such work by providing results for an outdoor vehicle equipped with odometers, gyro, and GPS receiver. Comparison was made with the Particle Filter localization, showing the better performance of interval method in terms of consistency. Gning [2] and Kueviakoe [20] dealt with the outdoor localization problem in the framework of Interval Constraint Satisfaction Problem (ICSP). Those works used constraint propagation techniques to fuse the redundant data of sensors. Drevelle [21] uses relaxed constraint propagation approach to deal with erroneous GPS measurements. Bonnifait [22] combined constraint propagation and set inversion techniques and presented a cooperative localization method with significant enhancement in terms of accuracy and confidence domains. Experimental results illustrate that constraint propagation techniques are well adapted to real-time implementation and provide consistent result for localization. Some other researchers proposed specific contractors [23] or separators [24] to solve the localization problem.

Indeed, when considering consistency issue, interval methods are advantageous in vehicle localization. However, they are prone to large output imprecision, which can make them unsuitable for high level task. We improve on it and present an extension of others' works dealing with the localization problem by using ICP. Instead of using ultrasonic sensors [16], sonar [18, 25], or GPS [19, 26] as exteroceptive sensors, we propose to use a monocular camera. The aim is to achieve both consistent and low imprecision localization result. We propose to drive a vehicle embedding precise but expensive sensors around the environment in order to build up a map represented in an interval way. Then the proposed method enables a set or a single vehicle to navigate consistently in such environment using this map and low-cost sensors. By constructing the problem as an ICSP, the proposed method is able to handle both mapping and localization in a unified form using Interval Constraint Propagation techniques.

The paper is organized as follows. Section 2 introduces the basics of interval analysis. Section 3 illustrates how the mapping and localization problem can be formulated as an Interval Constraint Satisfaction Problem. Sections 4 and 5 present the simulation and experimental results in terms of mapping and localization. Section 6 draws a conclusion of the paper.

2. Overview of Interval Analysis and Constraint Propagation

Interval analysis and constraint propagation are the theoretical tools used in interval based methods for robotic localization and mapping. In this section, we give a brief overview of these tools.

2.1. Interval Analysis. Interval analysis [11, 27] is a numerical method which allows us to solve nonlinear problems in a guaranteed way. One of the pioneers of interval analysis, Ramon E. Moore, proposed to represent a solution of a problem by an interval in which the real solution is guaranteed. An interval $[x]$ is a connected subset of \mathbb{R} , defined by its lower bound \underline{x} and upper bound \bar{x} :

$$[x] = [\underline{x}, \bar{x}] = \{x \in \mathbb{R} \mid \underline{x} \leq x \leq \bar{x}\} \quad (1)$$

The width of a nonempty interval $[x]$ is calculated by $w([x]) = \bar{x} - \underline{x}$ and the midpoint (or center) is $mid([x]) = (\bar{x} + \underline{x})/2$. If $[x]$ is an interval containing the x position of a vehicle, $mid([x])$ can be taken as an estimation of x and $w([x])/2$ can be regarded as the estimation error [28].

Let \mathbb{IR} be the set of the intervals of \mathbb{R} . A *box* (also called *interval vector*) $[\mathbf{x}]$ ($[\mathbf{x}] \in \mathbb{IR}^n$) is a generalization of the interval concept. It is a vector whose components are intervals:

$$[\mathbf{x}] = [x_1] \times [x_2] \times \cdots \times [x_n] \quad (2)$$

For instance, the configuration of a vehicle's pose usually contains 3 parameters: position x , position y , and heading angle θ . Consequently, for vehicle localization, the solution is a three-dimensional box: $[x] \times [y] \times [\theta]$.

The width of a nonempty box $[\mathbf{x}]$ can be computed by $w([\mathbf{x}]) = \max_{i=1, \dots, n} w([x_i])$. We define the volume of $[\mathbf{x}]$ as

$$vol([\mathbf{x}]) = \prod_{1 \leq i \leq n} w([x_i]) \quad (3)$$

The volume of the box is usually used to evaluate the estimation uncertainty of a state vector [28].

2.2. Operations of Interval Arithmetic. Rules have been defined to apply the basic arithmetical operations on intervals [27]. Let us consider two intervals $[x]$ and $[y]$ and a binary operator $\diamond \in \{+, -, \times, \div\}$; the smallest interval which contains all feasible values for $[x] \diamond [y]$ is defined as follows:

$$[x] \diamond [y] = \{x \diamond y \mid x \in [x], y \in [y]\} \quad (4)$$

Example 1. Addition: $[-4, 3] + [2, 5] = [-2, 8]$; subtraction: $[3, 5] - [1, 2] = [1, 4]$; multiplication: $[-1, 8] \times [-3, 1] = [-24, 8]$; division: $[-2, 8] \div [2, 4] = [-1, 4]$.

The set-theoretic operations can be applied to intervals. The *intersection* of two intervals is defined by

$$[x] \cap [y] = \{z \in \mathbb{R} \mid z \in [x] \text{ and } z \in [y]\} \quad (5)$$

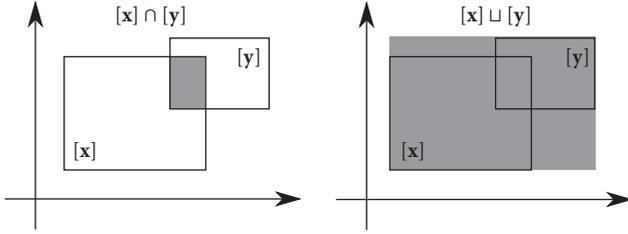
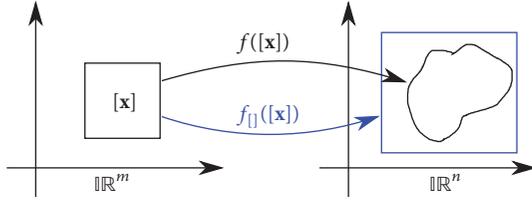


FIGURE 1: Intersection and interval hull of interval boxes.

FIGURE 2: Images of an interval box by function f and its inclusion function $f_{[]}$.

The result of *intersection* is always an interval, but it is not the case for their *union*

$$[x] \cup [y] = \{z \in \mathbb{R} \mid z \in [x] \text{ or } z \in [y]\} \quad (6)$$

Consequently, an *interval hull* is defined as the smallest interval that contains all the subsets of $[x] \cup [y]$, denoted by $[x] \sqcup [y]$:

$$[x] \sqcup [y] = [[x] \cup [y]] \quad (7)$$

For instance, the interval hull of $[2, 4] \cup [5, 8]$ is the interval $[2, 8]$. Figure 1 gives an example of *intersection* and *union* for 2 interval boxes; it is realized by applying the intersection and interval hull operator to each component of the box.

2.3. Inclusion Function. The image of a vector function $f : \mathbb{R}^m \rightarrow \mathbb{R}^n$ (defined by arithmetical operators and elementary functions) over an interval box $[x]$ can be evaluated by its inclusion function $f_{[]}$, whose output contains all possible values taken by $f(\cdot)$ over $[x]$ (see Figure 2):

$$\forall [x] \in \mathbb{R}^m, \quad f([x]) \subset f_{[]}([x]) \quad (8)$$

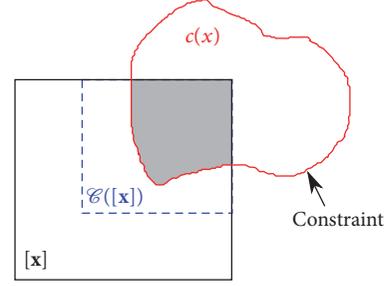
The simplest way to create an inclusion function is to replace all the variables and operators by their interval equivalents. The resulting interval function is called “natural inclusion function”. For instance,

$$\forall x \in \mathbb{R}, \quad \hat{\mathbf{N}}\mathbf{L}f(x) = 4x^4 - 7\ln(x) - 3 \quad (9)$$

has the natural inclusion function:

$$\forall [x] \in \mathbb{R}, \quad f_{[]}([x]) = 4[x]^4 - 7\ln_{[]}([x]) - 3 \quad (10)$$

Inclusion function can also be used to represent equations defined by interval variables. Such equations, also called constraints, are the core of Interval Constraint Satisfaction Problems.

FIGURE 3: Contraction of a box by a contractor \mathcal{C} .

2.4. Contractor. The concept of contractor is directly inspired from the ubiquitous concept of filtering algorithms in constraint programming. Given a constraint c relating a set of variables x , an algorithm \mathcal{C} is called a contractor (or a filtering algorithm) if \mathcal{C} returns a subdomain of the input domain $[x]$ and the resulting subdomain $\mathcal{C}([x])$ contains all the feasible points (the gray domain in Figure 3) with respect to the constraint c :

$$\mathcal{C}([x]) \subseteq [x] \quad \text{and} \quad \forall x \in [x], \quad c(x) \implies x \in \mathcal{C}([x]) \quad (11)$$

2.5. Interval Constraint Satisfaction Problem (ICSP). The concept of Interval Constraint Satisfaction Problem (ICSP) was introduced by Hyvönen [29] in 1992. It is typically defined as

- (i) a set V of variables $\{v_1, v_2, \dots, v_n\}$,
- (ii) a set D of domains $\{d_1, d_2, \dots, d_n\}$, such that for each variable v_i , a domain d_i with the possible values for that variable is given; d_i is an interval or union of intervals,
- (iii) a set C of p constraints $\{c_1, c_2, \dots, c_p\}$; each constraint c_i defines the relationships of a number of variables from V ; e.g., $c_1(v_1, v_2, v_3) = 0$ restricts the possible domains of v_1, v_2 , and v_3 .

An ICSP is a mathematical problem whose solution is the minimal domain of the variables satisfying all the constraints.

2.6. Interval Constraint Propagation (ICP). To solve an ICSP, an Interval Constraint Propagation (ICP) algorithm iterates domain reductions until no domain can be contracted. To satisfy the set of p constraints, it removes, from the domain of the variable, every value that is not compatible with the constraints and the other variables. ICP reduces the size of the domains in a consistent way by repeating this removal operation. This solving process is called *contraction* and the corresponding algorithm (e.g., Forward/Backward) used for contraction is regarded as a *contractor*.

Let us consider a simple ICSP instance which is defined by 3 variables and 2 constraints:

$$\begin{aligned} z &= x + \ln(y) \\ y &= z^2 \end{aligned} \quad (12)$$

The domains of x , y , and z are, respectively, $[x]$, $[y]$, and $[z]$. An inclusion function of (12) is

$$\begin{aligned} [z] &= [x] + \text{In}_{\square}([y]) \\ [y] &= [z]^2 \end{aligned} \quad (13)$$

The Forward/Backward Propagation (FBP), an ICP algorithm, can be applied to solve the ICSP. FBP executes the contraction process in two phases: firstly, the Forward propagation reduces the left terms of (13) via

$$[z] = [z] \cap ([x] + \text{In}_{\square}([y])) \quad (14a)$$

$$[y] = [y] \cap [z]^2 \quad (14b)$$

Secondly, Backward propagation reduces the right terms of (13) by

$$[x] = [x] \cap ([z] - \text{In}_{\square}([y])) \quad (15a)$$

$$[y] = [y] \cap \exp_{\square}([z] - [x]) \quad (15b)$$

$$[z] = [z] \cap \sqrt{\square}[y] \quad (15c)$$

Forward and Backward equations, obtained from the constraints, are computed one after the other (here we compute from (14a) to (15c)). When all the equations have been computed, the algorithm restarts from the beginning equation until the domains of $[x]$, $[y]$, and $[z]$ are no more contracted (or contracted less than a specified parameter).

3. Localization and Mapping

3.1. Problem Statement. Let us consider a vehicle navigating in a 2D environment. At time step k , the vehicle pose is represented by a three-dimensional vector $\mathcal{X}_k = (x_k, y_k, \theta_k)^T$ with x_k , y_k being the vehicle's position and θ_k the orientation. The n stationary landmarks are represented by $\mathcal{M} = (\mathcal{L}_1, \mathcal{L}_2, \dots, \mathcal{L}_n)^T$. \mathcal{L}_i is a state vector defining the landmark position (see Section 3.3). For the sake of simplicity, we assume that the pose evolution follows *Markov chain*, where the current pose \mathcal{X}_k depends only on the very last pose \mathcal{X}_{k-1} and the current control input vector \mathcal{U}_k . The input vector \mathcal{U}_k consists of linear and angular elementary movements, which were collected by proprioceptive sensors (e.g., odometer and gyro). If these odometric measurements are available, the state vector \mathcal{X}_k can be evaluated through the dead reckoning function (also called motion model) expressed as follows:

$$\mathcal{X}_k = f(\mathcal{X}_{k-1}, \mathcal{U}_k) + w_k \quad (16)$$

where w_k is the measurement noise affecting the odometric measurements.

The vehicle is equipped with exteroceptive sensors, recording information about the environment. At each time step k , the vehicle observes a number of landmarks which can be regarded as a subset of the entire map; the measurement \mathcal{X}_k depends only on the current vehicle pose \mathcal{X}_k and the map state \mathcal{M} :

$$\mathcal{X}_k = \mathcal{H}(\mathcal{X}_k, \mathcal{M}) + r_k \quad (17)$$

where $\mathcal{H}(\cdot)$ is generally a nonlinear function; it differs for different types of exteroceptive sensors. The output of $\mathcal{H}(\cdot)$ is a geometric parameterization of the landmark (e.g., distance, angle, or pixel coordinates), and r_k is the measurement noise of the exteroceptive sensors which introduce uncertainty to the landmark state.

Interval methods are based on the assumption that the support of sensor noises is unknown but bounded by real intervals. Mathematically, $\forall k > 0$, $w_k \in [\mathbf{w}]$ and $r_k \in [\mathbf{r}]$, where $[\mathbf{w}]$ and $[\mathbf{r}]$ are real intervals defined by lower and upper bounds. Then the localization problem can be regarded as an ICSP: the set of variables is defined by the vehicle pose and the landmark position; the motion and observation functions ((16) and (17)) define the constraints. ICP techniques allow us to propagate the interval uncertainties in a consistent way without any probability hypothesis or linearization process. As a result, all possible solutions are found and guaranteed.

3.2. Motion Model. The motion estimation process uses the measurements of proprioceptive sensors to predetermine the localization vector $[\mathcal{X}_k]$. EKF use a probabilistic motion model to fuse sensor data [30]. We make an extension of this model to obtain an interval based model defined by Seignez et al. [18]:

$$[\mathcal{X}_k] = f_{\square}([\mathcal{X}_{k-1}], [\mathcal{U}_k]) \quad (18)$$

$$\begin{aligned} &\begin{pmatrix} [x_k] \\ [y_k] \\ [\theta_k] \end{pmatrix} \\ &= \begin{pmatrix} [x_{k-1}] + [\delta s_k] \cos_{\square} \left([\theta_{k-1}] + \left[\frac{\delta \theta_k}{2} \right] \right) \\ [y_{k-1}] + [\delta s_k] \sin_{\square} \left([\theta_{k-1}] + \left[\frac{\delta \theta_k}{2} \right] \right) \\ [\theta_{k-1}] + [\delta \theta_k] \end{pmatrix} \end{aligned} \quad (19)$$

where $[\mathcal{U}_k] = ([\delta s_k], [\delta \theta_k])^T$ is the input vector. δs_k and $\delta \theta_k$ represent, respectively, the elementary displacement and rotation.

The elementary movements can be deduced from raw odometric data (odometer and gyro) using an interval based static fusion method:

$$\begin{aligned} [\delta s_k] &= \frac{[\pi]([\omega_l][\delta p_l] + [\omega_r][\delta p_r])}{[P]} \\ [\delta \theta_k] &= \frac{[\pi]([\omega_l][\delta p_l] - [\omega_r][\delta p_r])}{[e] \cdot [P]} \cap [\delta \theta_{gyro}] \end{aligned} \quad (20)$$

where δp_l and δp_r are the odometer measurements of the left and right wheels. ω_l and ω_r represent the radius of the wheels. P is the odometer resolution and e is the length of the rear axle. $\delta \theta_{gyro}$ is the gyro measurement. For further information about interval based odometric sensor integration, the reader might refer to [31]. Once we have constituted intervals from sensor data, the vehicle pose $[\mathcal{X}_k]$ can be estimated by applying FBP algorithm (Section 2.6) over constraints deduced from the motion model.

3.3. Landmark Parameterization. A landmark (or feature point) is a 3D point in the global world. Probabilistic methods generally represent landmark with an inverse depth parameterization [32]. They model the uncertainty of each parameter by Gaussian distributions. However, it turns out to be not so efficient to represent the linear depth uncertainty of monocular vision. Interval analysis provides an easy and efficient way to parameterize landmarks. Each landmark \mathcal{L}_i is defined as a six-dimensional state vector: $([x_o], [y_o], [z_o], [\alpha_i], [\varphi_i], [d_i])^T$, which models the estimated landmark position at

$$[\mathcal{L}_i] = ([x_o], [y_o], [z_o])^T + [d_i] \cdot [\mathbf{m}]([\alpha_i], [\varphi_i]) \quad (21)$$

where coordinates $[x_o]$, $[y_o]$, and $[z_o]$ represent the optical center of the camera when the landmark was seen for the first time, $[\alpha_i]$ and $[\varphi_i]$ represent azimuth and elevation angle for the ray which traces the landmark, and $[d_i]$ is the depth of the landmark. $[\mathbf{m}]([\alpha_i], [\varphi_i])$ is a unitary vector pointing from the camera to the landmark $[\mathcal{L}_i]$:

$$[\mathbf{m}]([\alpha_i], [\varphi_i]) = \begin{pmatrix} \cos_{\square}([\alpha_i]) \cos_{\square}([\varphi_i]) \\ -\sin_{\square}([\alpha_i]) \cos_{\square}([\varphi_i]) \\ \sin_{\square}([\varphi_i]) \end{pmatrix} \quad (22)$$

Since all the parameters are represented by intervals, and $[d_i]$ can be initialized as $[0, +\infty]$, the landmark's uncertainty is modeled as an infinite cone which combines the vehicle pose uncertainty and the observation uncertainty. It is a realistic representation for the monocular vision uncertainty. The major advantage is that the initialization of $[d_i]$ is undelayed, guaranteed, and efficient for landmarks over a wide range as it always includes all possible value without any prior information.

3.4. Observation Model. The observation of a landmark is the coordinates of the pixel where the landmark is projected on the image. In order to detect and extract landmark information from an image, we used Speed Up Robust Features (SURF) [33]. The work [34] recommended it after a comparison between several algorithms, showing that SURF provide results with a very low error rate. The projection of the i^{th} landmark to the image frame can be formulated as follows:

From (21), the landmark's position in the world coordinate \mathbf{O}_w can be expressed:

$$[\mathcal{L}_i^w] = ([x_o], [y_o], [z_o])^T + [d_i] \cdot [\mathbf{m}]([\alpha_i], [\varphi_i]) \quad (23)$$

In the camera coordinate \mathbf{O}_c , the landmark $[\mathcal{L}_i^w]$ will be seen in the position $[\mathcal{L}_i^c]$:

$$[\mathcal{L}_i^c] = [R^{wc}]([\mathcal{L}_i^w] - [R^{rw}][\mathcal{X}_k]) \quad (24)$$

where $[\mathcal{X}_k]$ is the current vehicle pose. R^{wc} and R^{rw} are, respectively, the transformational matrix between the world and the camera coordinate and the robot and world coordinate, respectively.

The camera does not observe $[\mathcal{L}_i^c]$ directly but its projection in the image frame $(u, v)^T$. According to the pin-hole

camera model [35], the observation \mathcal{Z}_k^i of $[\mathcal{L}_i^c]$ on the image frame can be predicted by

$$[\mathcal{Z}_k^i] = \mathcal{H}_{\square}([\mathcal{X}_k], [\mathcal{L}_i^c]) \quad (25)$$

$$\begin{pmatrix} [u_{pre}^i] \\ [v_{pre}^i] \end{pmatrix} = \begin{pmatrix} [c_u] + [f][k_u] \frac{[\mathcal{L}_i^c(x)]}{[\mathcal{L}_i^c(z)]} \\ [c_v] + [f][k_v] \frac{[\mathcal{L}_i^c(y)]}{[\mathcal{L}_i^c(z)]} \end{pmatrix} \quad (26)$$

where f , k_u , k_v , c_u , c_v are the camera intrinsic parameters. Equation (26) is also used as the projection function.

3.5. Data Association. Data association focuses on finding out the correspondence between the map and observations. Different from probabilistic method, which deals with uncertainty ellipsoid in projection procedure, interval method proceeds with bounding box. The matching process is carried out as follows:

- (i) Project the landmarks to the current image plane; discard those which lay out of the image bounds. Each landmark defines a rectangle searching area.
- (ii) In the searching area, search for the candidate feature that is matched to the landmark. We adopt Dan's method [36] which performs a two-stage matching process (considering both the Euclidean distance and dominant orientation information of feature descriptors) to provide a robust and accuracy matching result. For each pair of matched features, a Zero Normalized Cross Correlation (ZNCC) [37] score is computed in the mapping stage to help to evaluate the robustness of the landmarks.
- (iii) Filtering step is necessary to keep a desirable amount of matched points. The landmark uncertainty and ZNCC score can be used for the selection process. We make the matched results distributed uniformly over the image, providing enough parallax for the localization.

3.6. ICSP Formulation. The data association process generates a set of 3D to 2D correspondences, which we can use to build the ICSP constraints. Let us consider the i^{th} landmark $[\mathcal{L}_i]$ that has been matched. The observation of $[\mathcal{L}_i]$ is $([u_{obs}^i], [v_{obs}^i])^T$. The predicted position on the image plane, denoted by $([u_{pre}^i], [v_{pre}^i])^T$, is computed via (23), (24), and (26). The observation can be used to correct the system state through the predicted observation:

$$\begin{aligned} [u_{pre}^i] &= [u_{obs}^i] \\ [v_{pre}^i] &= [v_{obs}^i] \end{aligned} \quad (27)$$

Equation (27) creates a link between the prediction and observation data of the i^{th} landmark $[\mathcal{L}_i]$. This link (thanks to (23), (24), (26), and (27)) imposes constraints on $[\mathcal{X}_k]$ and $[\mathcal{L}_i]$, which can be used to update both the vehicle pose and landmark position.

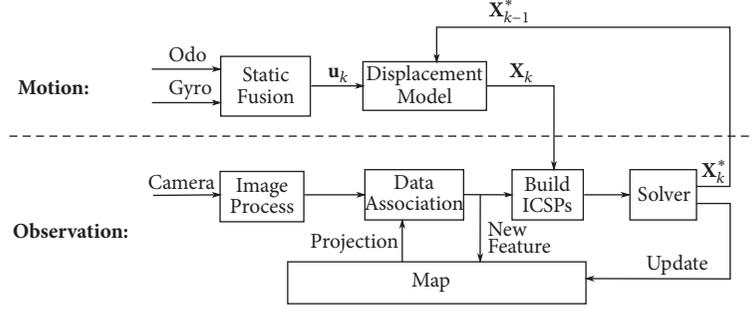


FIGURE 4: Mapping framework.

Based on the above derivation, an ICSP composed of 2 top level constraints and a set of 9 variables can be set up. For each matched landmark i , the $\text{ICSP}_k(V_k, D_k, C_k)$ is generated and defined as follows:

(i) A set of 9 variables:

$$V_k = \{x_k, y_k, \theta_k, x_o, y_o, z_o, \alpha_i, \varphi_i, d_i\} \quad (28)$$

(ii) A set of 9 domains:

$$D_k = \{[x_k], [y_k], [\theta_k], [x_o], [y_o], [z_o], [\alpha_i], [\varphi_i], [d_i]\} \quad (29)$$

(iii) A set of 2 top level constraints:

$$C_k = (C_k^1, C_k^2)^T \quad (30)$$

$$C_k^1 \iff [u_{pre}^i] = [u_{obs}^i] \quad (31)$$

$$C_k^2 \iff [v_{pre}^i] = [v_{obs}^i]$$

If the data association process found N matched landmarks, then we get a multiple ICSP to solve:

$$v\text{ICSP}_k = (\text{ICSP}_k^1, \text{ICSP}_k^2, \dots, \text{ICSP}_k^N) \quad (32)$$

Each ICSP contains constraints between the vehicle pose and a unique landmark position. The good point is that they are all linked together via the same variables x_k, y_k , and θ_k (the vehicle pose at current time k). This correlation makes ICP algorithms very powerful to jointly contract the domains of all variables.

3.7. System Overview. Figure 4 depicts the framework of the mapping stage. It is an interval SLAM process which is cast into two parts: motion and observation. In the motion part, the preliminary pose of the vehicle is predicted by fusing the measurements of the odometers and gyro with bounded error model; in the observation part, features are firstly extracted from the image, data association task aims at finding out the correspondences between the features and the map, new features will be initialized as new landmarks and added to the map, and old ones will be processed to generate new ICSPs; then these ICSPs are jointly solved by an ICP algorithm, which correct the vehicle pose and update the map.

```

(1) while Localization is required do
(2)   Input:  $[\mathcal{X}_{k-1}], [\mathcal{U}_k], [\mathcal{M}], I_k$ 
(3)   Output:  $[\mathcal{X}_k]$ 
(4)    $[\mathcal{X}_k] \leftarrow f_{[]}([\mathcal{X}_{k-1}], [\mathcal{U}_k])$ 
(5)    $S_k \leftarrow \text{ExtractFeature}(I_k)$ 
(6)    $[(u, v)_{obs}, \mathcal{L}]^N \leftarrow \text{DataAssociate}([\mathcal{X}_k], [\mathcal{M}], S_k)$ 
(7)   for  $j \leftarrow 1$  to  $N$  do
(8)      $\text{ICSP}_j \leftarrow \text{Build}([\mathcal{X}_k], [(u, v)_{obs}, \mathcal{L}]^j)$ 
(9)   end for
(10)   $v\text{ICSP}_k = \bigcup_{j=1}^N \text{ICSP}_j$ 
(11)   $[\mathcal{X}_k^*] \leftarrow \text{Solver}(v\text{ICSP}_k)$ 
(12) end while

```

ALGORITHM 1: Localization.

The localization stage is quite similar to the mapping stage. There are two main differences: high cost sensors are no longer used and no new feature is added to the map. Algorithm 1 gives the pseudocode of our localization method:

- (i) Line (4): motion estimation. The current vehicle pose is calculated based on the motion function $f_{[]}$ and input vector $[\mathcal{U}_k]$.
- (ii) Line (5): image processing. A set of SURF S_k are extracted from the new image I_k .
- (iii) Line (6): data association. The correspondence between the features and the map is found.
- (iv) Line (8): ICSP generating. New ICSP are generated based on the matched landmarks.
- (v) Line (10)-(11): Solving. $v\text{ICSP}_k$ is composed by the ICSPs and solved by an ICP algorithm.

4. Simulation Results

A simulation experiment is set up to evaluate the performance of our proposed ICSP based mapping and localization method. The vehicle starts from the origin (see Figure 5) and moves straight forward in the positive direction of x-axis with constant linear and angular velocity ($v = 0.1m \cdot s^{-1}$, $\omega = 0$). 5 landmarks are positioned in the surrounding environment without any prior knowledge of their positions. At each time

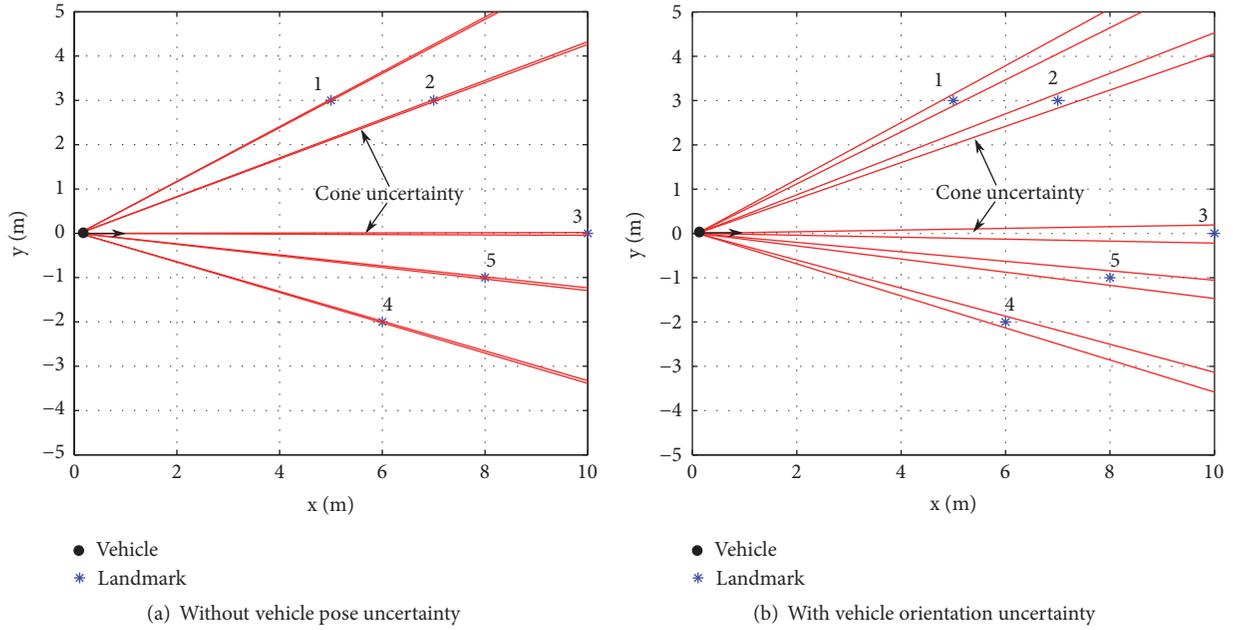


FIGURE 5: Interval based landmark initialization.

step, the vehicle detects the 5 landmarks with a simulated camera model $[f, k_u, k_v, c_u, c_v] = [40, 8, 8, 320, 240]$ and focuses on estimating their position. The experiment is first carried out without odometric sensor noise (the vehicle pose is perfectly known) to verify the feasibility of our method. This condition (no odometric noise) is then released in the experiment presented in Section 4.3.

4.1. Landmark Initialization. At time step $k = 1$, the 5 landmarks are first detected and the initialization procedure is executed. Figure 5(a) shows the top view of the initialization result of landmarks (see Section 3.3). Each landmark's uncertainty is initialized as an infinite cone because of the lacking information about the landmark depth. As there is no uncertainty on vehicle pose, the landmarks are well initialized and the opening angles which indicate the observation uncertainty are small (due to the bounded observation noise which is 1 pixel). If the vehicle has an uncertainty on its orientation (the heading angle is not perfectly defined: $[\theta_k] = [-\pi/180, \pi/180]$), then the cones of uncertainty are much larger than the previous ones (see Figure 5(b)). This is reasonable since the cone uncertainty is the fusion of observation uncertainty and vehicle pose uncertainty. Because our bounded error hypothesis and the use of interval analysis, the initialization result is guaranteed to contain the true landmark position.

4.2. Landmark Uncertainty Evolution. As the vehicle moves, the landmarks are repeatedly observed with different parallaxes. By building and solving the associated ICSP (see Section 3.6), each of the landmark's parameters can be estimated. Figure 7 shows, since initialization, the evolution of the poses uncertainties of landmarks 1, 2, and 5 in top view. Each row plots the estimation result at time steps 1, 5, 10, and

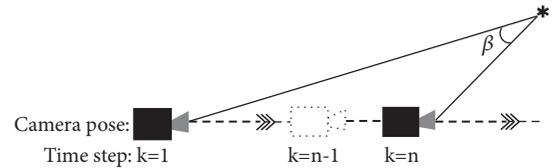


FIGURE 6: Parallax observations.

20. The parallax β (defined as the angle difference between the initial observation and current observation, see Figure 6) is displayed as the title of each subfigure. With the estimation proceeds, the landmark uncertainty estimate improves. As we can see, the bigger the parallax is, the better the contraction of the landmark uncertainty will be. When enough parallax is eventually available, the infinite interval uncertainty is significantly reduced. The uncertainty of landmark 3 keeps infinite because it locates on axis x and no parallax has been detected during the horizontal translation.

To visualize the evolution of a landmark state, it is necessary to convert the 6-dimensional representation to a global 3D point using (21): $(x_o, y_o, z_o, \alpha_i, \phi_i, d_i)^T \rightarrow (x_i, y_i, z_i)^T$. The interval uncertainty (IU) can be evaluated by computing the box volume [31]: $U(i) = w([x_i]) \times w([y_i]) \times w([z_i])$. As the depth of each landmark is initialized as $[0, +\infty]$, the x , y , and z coordinates have infinite value, and the uncertainty turns out to be infinite at the first hand. After multiple observations (at different times) from different parallaxes, the landmark's state is updated. Figure 8 depicts the evolution of the uncertainty associated with the landmarks 1, 2, 4, and 5. The 4 landmarks encounter different observation parallaxes, so the contraction of their uncertainties is not simultaneous. However, they exhibit the same performance:

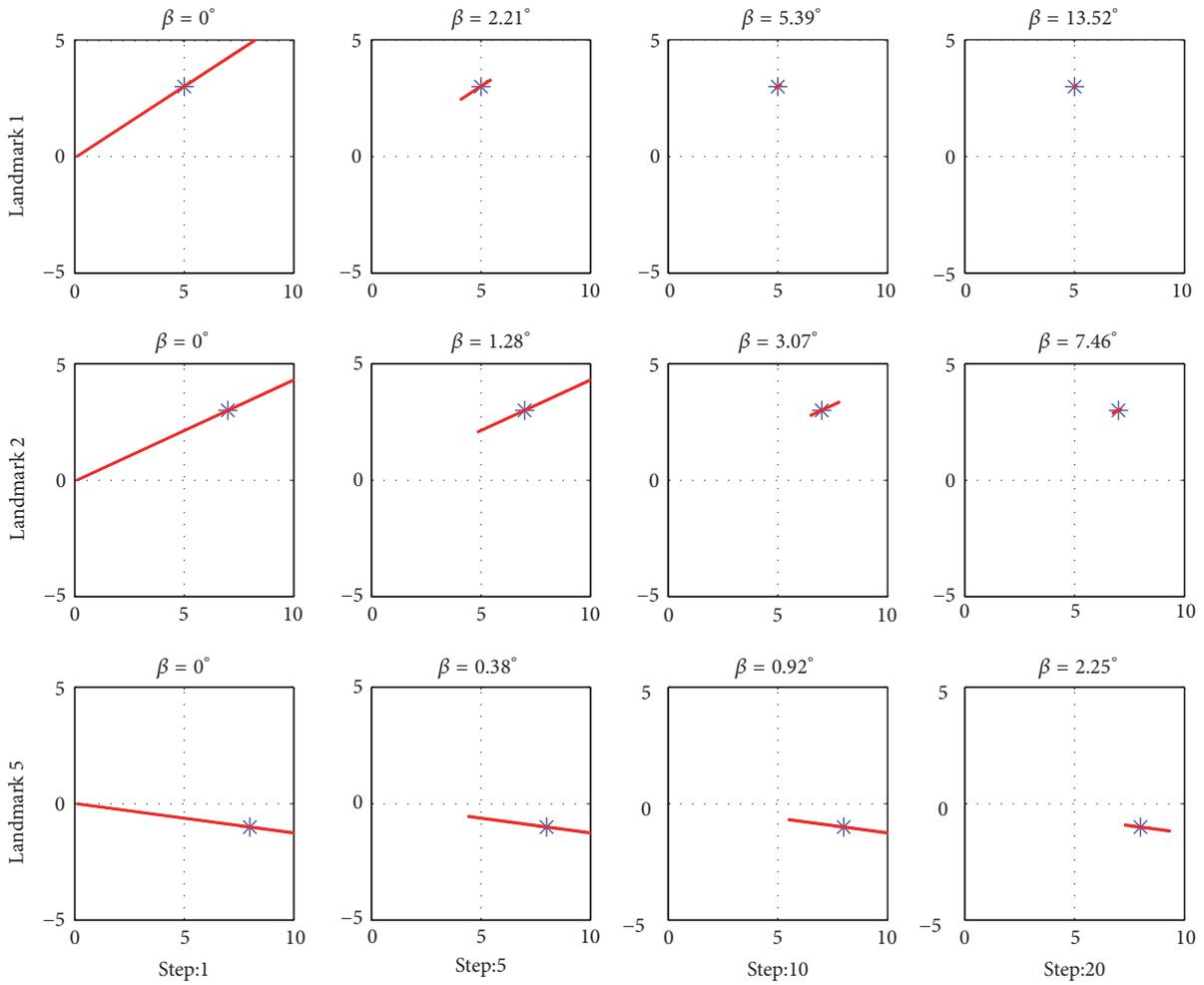


FIGURE 7: Interval uncertainty evolution.

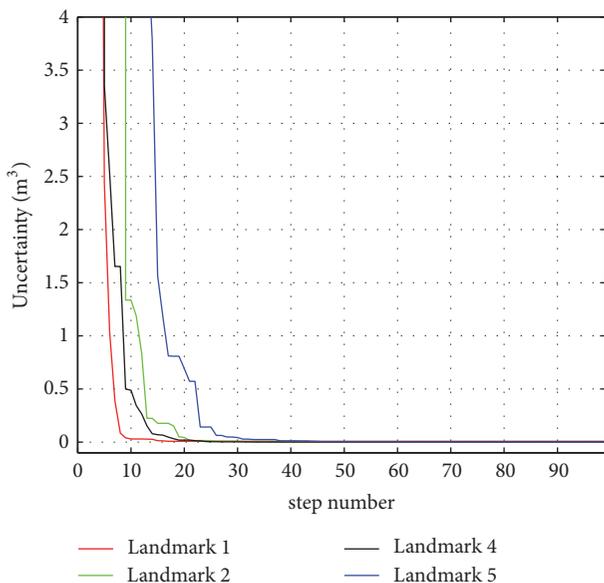


FIGURE 8: Landmark position uncertainty evolution.

the interval uncertainty of each landmark is well propagated and decreases exponentially with the vehicle navigating in the environment.

4.3. Estimation Accuracy. In real world, the vehicle pose is always accompanied with uncertainties due to the cumulative odometric noise. The vehicle pose uncertainty will affect the estimation accuracy of the landmarks (a smaller vehicle pose uncertainty means higher landmarks estimation accuracy). In our simulation, the error bound of the odometric measurement is supposed to be proportional to elementary displacement, such that $w([\mathcal{U}_k]) \leq \gamma \cdot \mathcal{U}_k + \Gamma$, $\Gamma = [0.0001m, 0.01^\circ]^T$, with \mathcal{U}_k being the real odometric measurement, γ denoting the ratio value (bigger γ value indicates a relative bigger uncertainty of odometric measurements), and Γ accounting for the steady system error. The average position uncertainty of the 4 landmarks has been computed at each time step. Figure 9 displays the results with different γ value ($\gamma = 0.01, 0.1$ and 0.2). As it is shown, the estimation accuracy decreases when γ value becomes bigger. This result supports the fact that the landmark estimation accuracy is strongly correlated with the vehicle pose uncertainty. Bigger

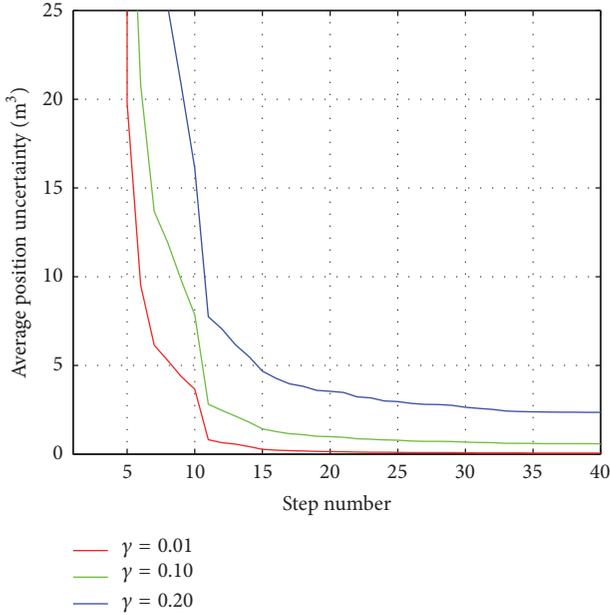


FIGURE 9: Average landmark uncertainty under different γ values.

odometric uncertainty will result in a worse estimation of vehicle pose; as a result, the landmark estimation accuracy decreases.

5. Real Data-Set Experiment

5.1. Experiment Set-Up. To evaluate the performance of our method in realistic circumstance, we used the experimental data set from Institute Pascal [38]. It contains multisensory timestamped data which can be used in a large variety of applications. It includes data from low-cost sensors, ground truth, and multiple sequences as well as data for calibration. Data were collected on the VIPALAB platform (see Figure 10(a)), driving on the PAVIN track (see Figure 11(a)), an experimental site located on the French campus of Blaise Pascal University. The environment is composed of scaled street, road markings, functional traffic lights, painted walls, buildings vegetation, and so on (see Figure 11(b)). The platform was equipped with multiple sensors (see Figure 10(b)). In our experiment, we need to use 4 sets of timestamped sensor data:

- (i) Odometers: each rear wheel is equipped with an odometer in the form of a ring gear providing 64 “top” per wheel revolution.
- (ii) Gyro (Melexis MLX90609-N2): it measures the rotational speed in an inertial reference system. Vibrating silicon structures, that use the Coriolis force, output the yaw rate from which we can deduce the elementary rotation.
- (iii) Camera (FOculus FO432B camera + Pentax C418DX lens): it provides information about the perceived environment. We use the camera which located in the left-front of the vehicle and looking ahead.

TABLE 1: Average bounding boxes size.

| Parameter | x | y | z |
|----------------|-------|-------|-------|
| Mean width (m) | 0.207 | 0.221 | 0.013 |

- (iv) RTK-GPS: the data acquired with an embedded ProFlex500 RTK-GPS receiver from NavtechGPS coupled with a Sagitta Magellan GPS base station are considered as the ground truth. This system provides an accurate absolute localization measurement to within 2 cm (± 1 cm).

5.2. Mapping Stage. A mapping stage is first conducted by our method in order to build a map of the environment. The vehicle follows 6 simple consecutive loops around the PAVIN environment. Figure 12(a) shows the reference trajectory that the vehicle passed by and Figure 12(b) details the 6 loops in order. As we have discussed in Section 4.3, the estimation of landmark is strongly correlated with the vehicle pose uncertainty. So it is necessary to maintain the pose uncertainty in a reasonable size during the mapping stage in order to pursue a precise map.

Figure 13 shows the vehicle position uncertainty (calculated by $(w([x])/2) \cdot (w([y])/2)$) with and without RTK-GPS correction during the mapping process. As we can see, without correction, the position uncertainty is cumulative and increases very quickly until loop closure is realized. This is because when navigating, the odometric noise is cumulative since there is no absolute measurement. Interval analysis is a pessimistic method maintaining all possible solutions, the cumulative errors will result in interval expansion, and the pose uncertainty thus becomes larger and larger. This is a hinder to pursue a good map. To overcome this problem, our method makes use of the RTK-GPS measurement, and the GPS error bounds are characterized as the maximum imprecision ($\leq 2cm$) of the GPS receiver. The cumulative errors can be eliminated and a precise estimation of the vehicle position is obtained (see Figure 13(b)). The result is assumed to be consistent since only ICP technique was used.

At the end of the loop, 3450 images are processed and a map of 1140 landmarks is created. Figure 14 shows part of the map projected on $x - y$ plane. The black bounding boxes denote the estimated landmark’s position. The average size of these bounding boxes is shown in Table 1, which can be regarded as a factor to evaluate the quality of the map. The z coordinate gets much higher precision than x and y ; this is because the estimation of z is only correlated with the observation elevation angle which is invariant w.r.t the vehicle’s heading angle, while x and y are affected by the heading uncertainty. A good way to improve the map is to get a better estimation of the vehicle’s orientation. Note that the map building process could be done offline such that computation intensive ICP algorithms and manually guidance could be used to obtain a better map.

5.3. Localization Result. The localization stage benefits from the map data which offer position constraints. With the built map, the vehicle seeks consistent localization using the

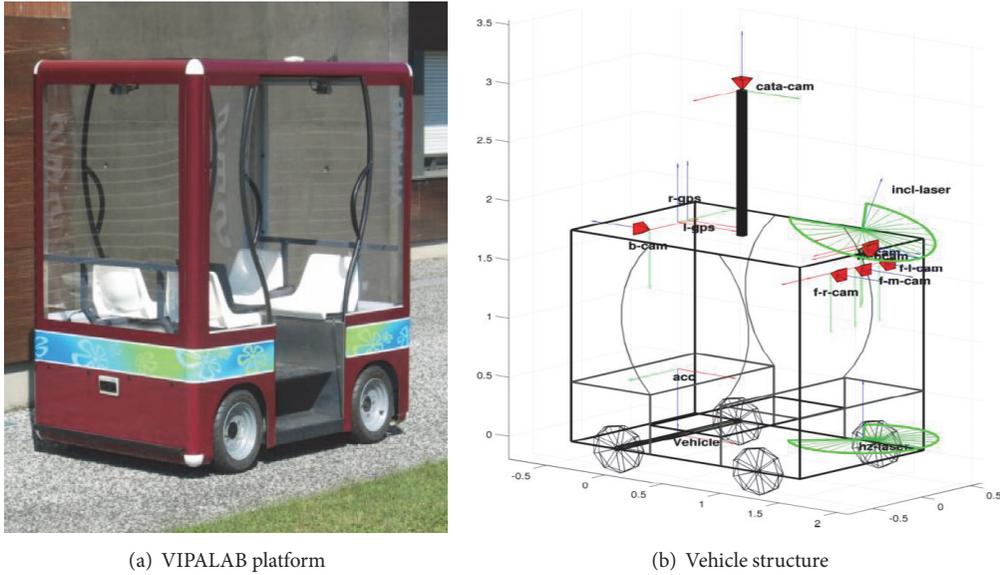


FIGURE 10: VIPALAB Experimental platform.

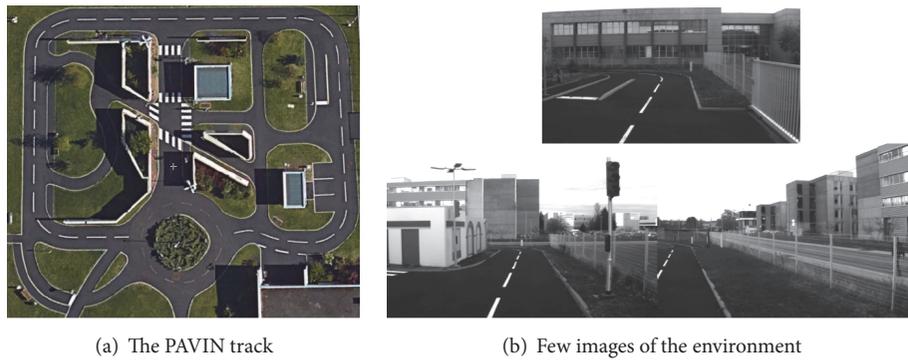


FIGURE 11: Institute Pascal Experimental Environment.

current image captured by the monocular camera. Three sequences of data from the whole data set are used to test our localization method. The reference path (RTK-GPS) of the vehicle when these data are collected is illustrated in Figures 15, 18(a), and 18(b).

Our proposed localization method is firstly tested on Sequence 1, and the vehicle follows a nearly 200m trajectory; see Figure 15. The blue path is the reference trajectory from RTK-GPS, the red path is the dead reckoning result, and the black squares are the localization boxes output by our proposed interval method. The dead reckoning result is obtained via a geometric evolution model using the odometer and steering angle data. It can be seen from the figure that the localization boxes followed the reference trajectory in a consistent way (intersect with the blue line). Table 2 (DR: dead reckoning; IM: interval method) presents the localization result of both methods at $t = 200$ and $t = 400$ and at the end of the track. By fusing the map, our interval

method gives a consistent estimation of the vehicle pose: the localization boxes well enclose the ground truth.

To compare the vehicle position estimated by our proposed method and the dead reckoning, the root of sum square error (RSSE) is computed at each time step. It is expressed by the following formula:

$$RSSE_k = \sqrt{E_k^2(x) + E_k^2(y)} \quad (33)$$

where $E_k(x)$ and $E_k(y)$ are the estimation error of $\text{mid}([x])$ position and $\text{mid}([y])$ position at time step k . RTK-GPS data is regarded as the ground truth.

Figure 16 shows the RSSE results of both methods. As it could be expected, the cumulative odometric error of dead reckoning is significant and caused an increasing discrepancy (at the end, the discrepancy reaches about 3 meters). Our method takes advantage of the map data and corrects the odometric error, providing a very small discrepancy.

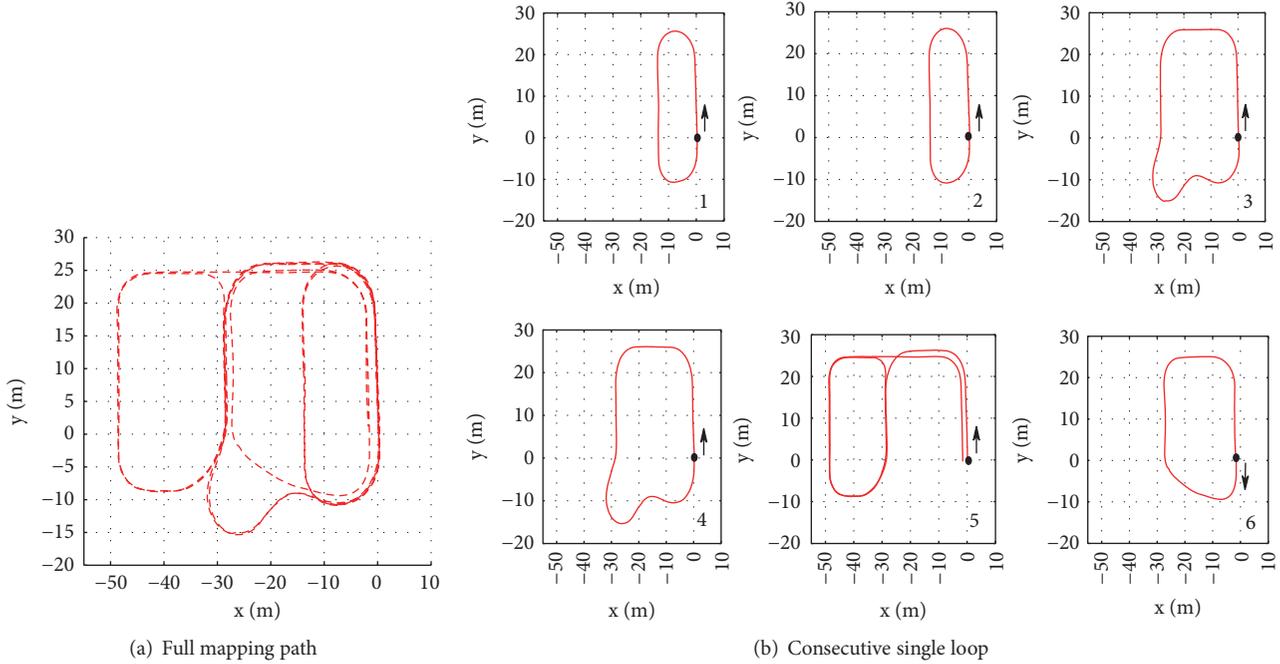


FIGURE 12: The reference path (RTK-GPS) of mapping sequence.

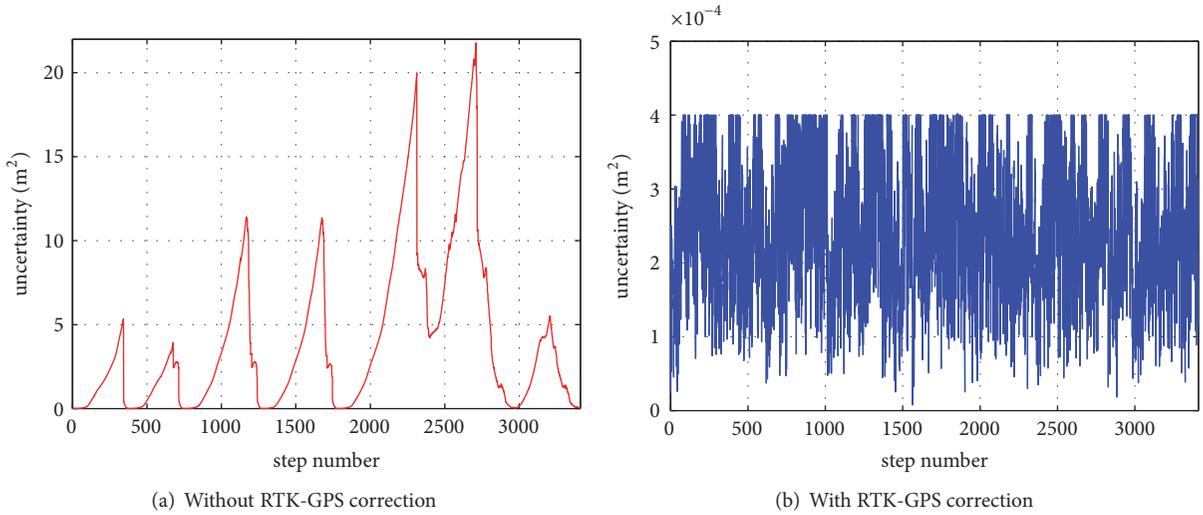


FIGURE 13: Vehicle position uncertainty.

In order to verify the consistency of the localization results, interval error is defined by the upper and lower bounds of the estimated state minus the corresponding reference (real value), mathematically,

$$[E_k(\alpha)] = [\underline{\alpha}_k - \alpha_k^*, \overline{\alpha}_k - \alpha_k^*] \quad (34)$$

where $[\underline{\alpha}_k, \overline{\alpha}_k]$ is the estimated state of variable α at time step k and α_k^* is the corresponding reference. The estimated state is said to be consistent if and only if $0 \in [E_k(\alpha)]$ is a true statement. Moreover, an estimation method is consistent and

precise if $\forall k \in \{1, 2, \dots, \infty\}, 0 \in [E_k(\alpha)]$ is always satisfied and the interval error width $w([E_k(\alpha)])$ is thin.

We compute the real-time interval error of our localization result. At each time instant k when the new captured image had been processed, $E_k(x)$ and $E_k(y)$ are calculated, respectively. Figure 17 depicts the interval corridors consisting of the upper and lower bounds of E_k . The zero line is well included by those corridors, proving that the localization results are consistent all along the track. This is a significant result in vehicle localization field where safety is a crucial issue.

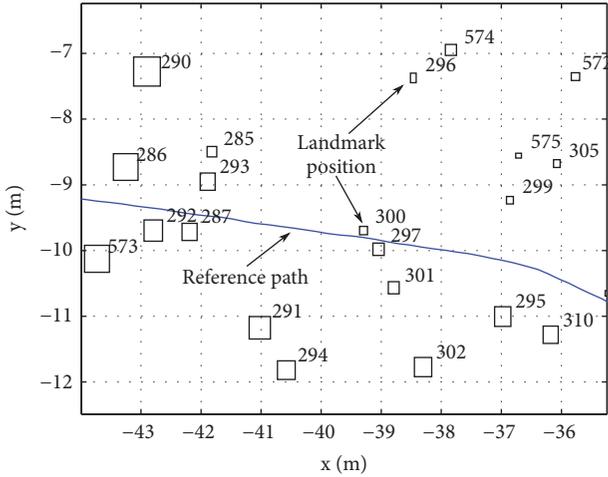


FIGURE 14: Map building result.

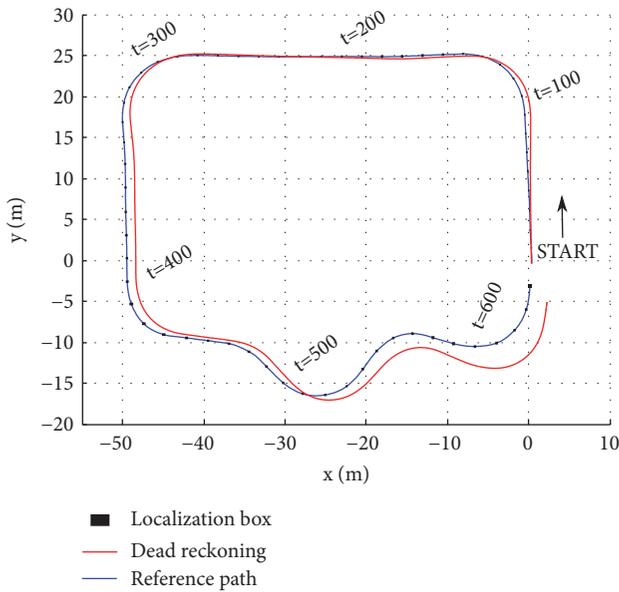


FIGURE 15: Localization trajectory of Sequence 1.

TABLE 2: Localization result.

| Time | Position | RTK | DR | IM |
|-------|----------|--------|--------|------------------|
| t=200 | x | -22.32 | -22.72 | [-22.35, -22.26] |
| | y | 25.17 | 25.23 | [25.14, 25.22] |
| t=400 | x | -49.23 | -48.13 | [-49.29, -49.18] |
| | y | -2.51 | -2.31 | [-2.63, -2.45] |
| End | x | 0.14 | 1.89 | [0.10, 0.18] |
| | y | -3.26 | -5.23 | [-3.32, -3.21] |

Similar results are obtained when performing the localization process on Sequences 2 and 3 with the same map; the output localization boxes by our method and the corresponding reference trajectory are displayed in Figure 18. It shows the robustness of our method providing consistent

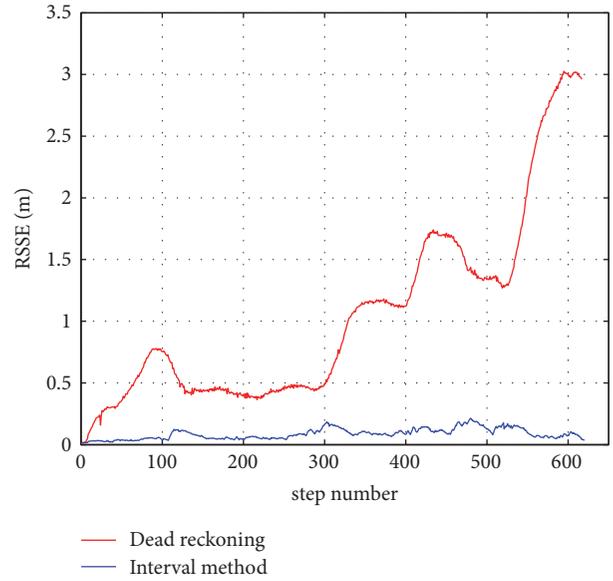


FIGURE 16: The error of localization results by DR and IA.

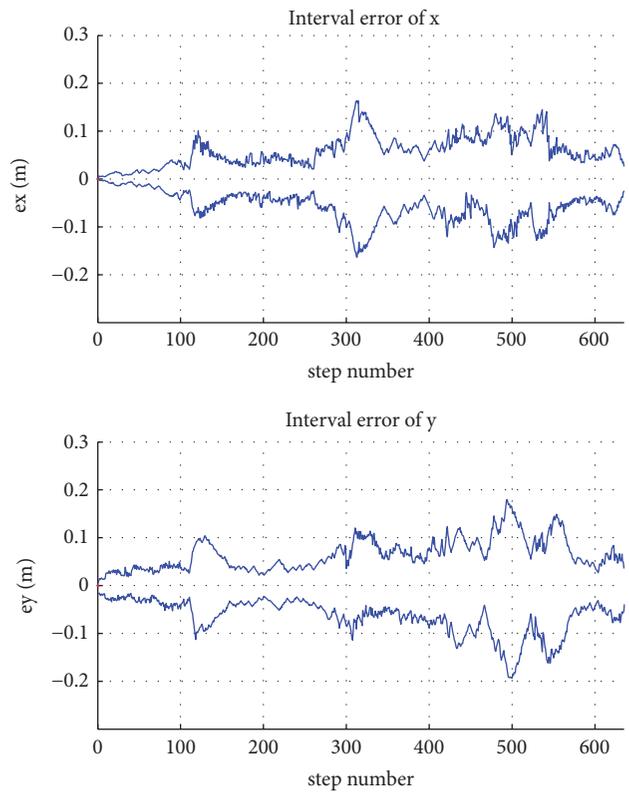


FIGURE 17: Interval error of real-time localization result.

localization result. The width of the interval error corridors (see Figure 17) with respect to the vehicle's x position and y position is displayed in Figure 19, reporting a maximum value of 0.6 m. The average interval error width and mean RSSE are computed to evaluate the localization accuracy. Table 3 details the results of the 3 experiments. Computing an average value for the 3 experiments gives 12.7 cm accuracy for x and

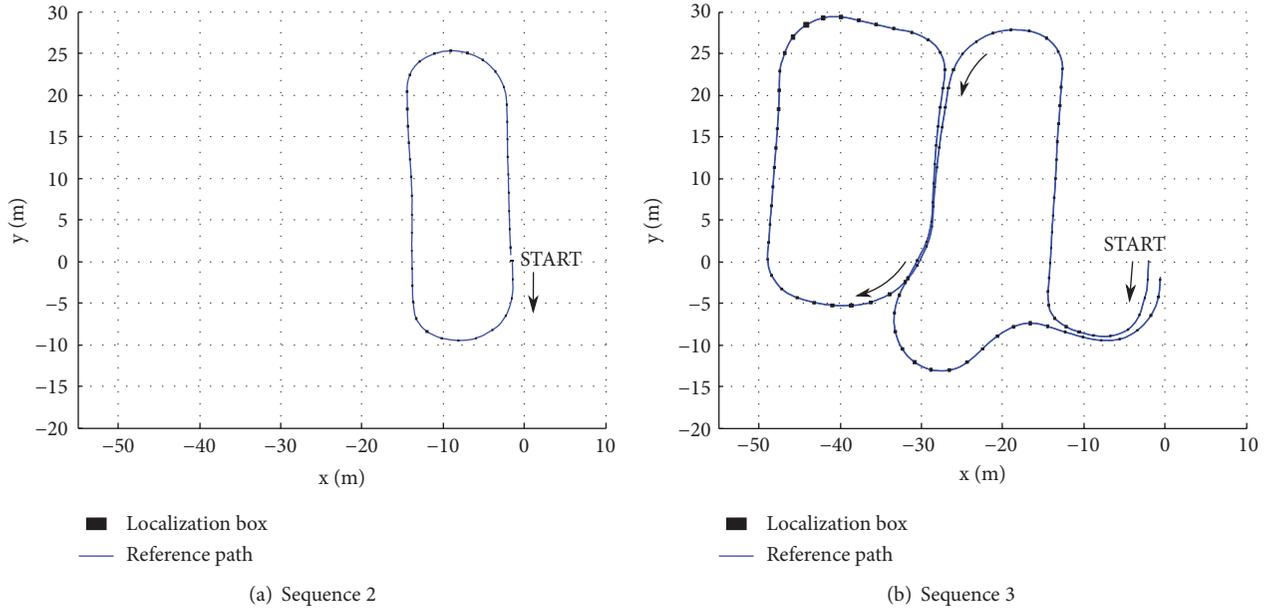


FIGURE 18: Localization boxes compare with reference path.

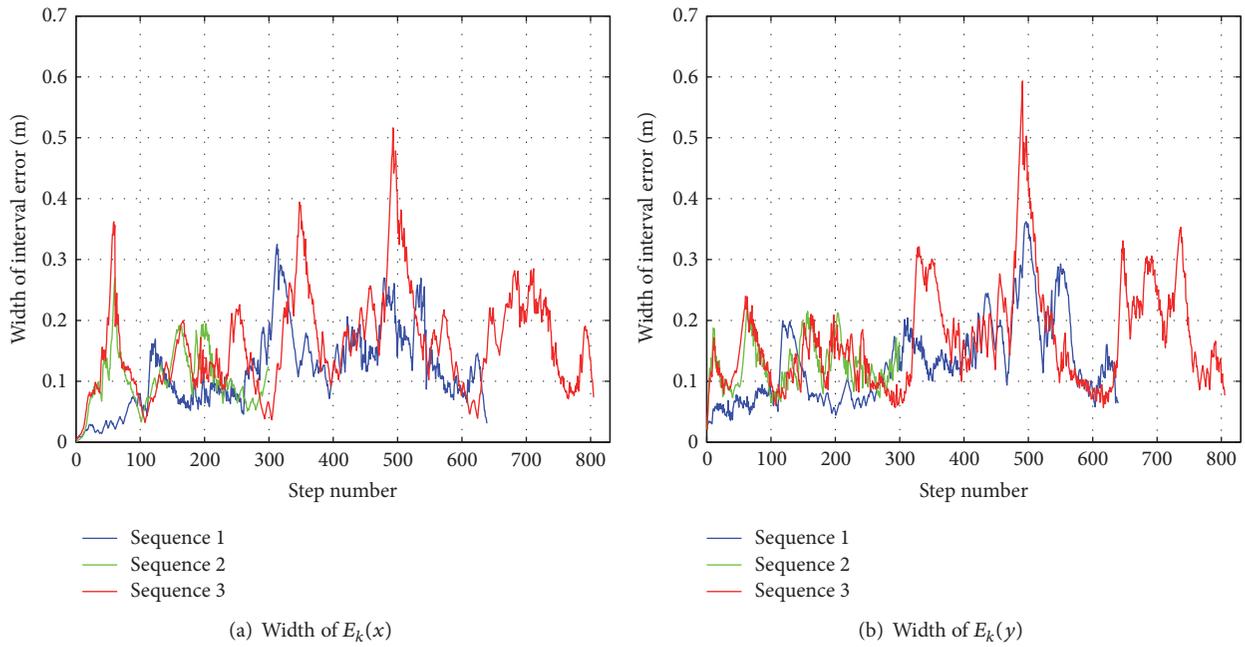


FIGURE 19: Interval error width of localization result.

14.4 cm for y . The average RSSE value is about 9.66 cm (less than 10 cm). Note that these localization results are consistent, and the localization boxes are guaranteed to enclose the real positions of the vehicle.

5.4. Discussion. In this paper, we present a complete methodology for localizing a vehicle within a prebuilt map. We bring forward a two-stage framework (mapping and localization) to solve the problem. Closer approaches have already been proposed by different researchers [39–41]. They all involve a

TABLE 3: Mean localization interval error and RSSE.

| Experiment | x (cm) | y (cm) | RSSE (cm) |
|------------|--------|--------|-----------|
| Sequence 1 | 11.7 | 13.4 | 8.90 |
| Sequence 2 | 10.3 | 12.7 | 8.29 |
| Sequence 3 | 16.1 | 17.2 | 11.8 |

visual learning step to reconstruct a map of the environment, and they use this map for localization and navigation task.

Eric Royer et al. [39] uses bundle adjustment in mapping step and the localization results are obtained via Newton iteration method. Hyon Lim et al. [41] uses a structure from motion (SFM) algorithm to reconstruct the 3D environment; then in the localization step, two discrete Kalman filters are employed to estimate the camera trajectory. Courbon's method [40] does not involve the 3D reconstruction procedure during the map building step; instead, it records some key views and the performed path as references and uses them for future navigation missions. Our method uses totally different theory and techniques from these approaches, we cast the two-stage problem into an ICSP, and deterministic techniques (ICP) are used in order to find the solution.

The localization accuracy of these works is close to ours. The paper [39] has an average localization error of 15 cm over 6 localization experiments on 3 outdoor sequences. The reported mean indoor localization errors of [41] are ranged between 5.6 cm and 10.8 cm among 6 experiments. The outdoor mean position error is less than 15 cm. The paper [40] presents an average lateral error of 25 cm on an urban vehicle navigating along a 750 m trajectory. Our method has been tested on 3 outdoor sequences; the position error (RSSE) ranges between 8.29 cm and 11.8cm. The strength of our method is that the localization boxes are guaranteed to contain the real position of the vehicle. Such a guarantee is due to the modeling of the noises as intervals and the computing of constraints using interval functions.

The formulas of our algorithm work whatever the camera manufacturer is. As for different cameras, the observation model always holds; only the camera intrinsic parameters (f, k_u, k_v, c_u, c_v) change. Furthermore, the feature points extracted from images of various cameras would be the same since the SURF algorithm is scale invariant, rotation invariant, and affine transformation invariant [33]. This property makes it possible to find correct correspondence between two images from different cameras. This feasibility has been shown by [39] in two sequences of images taken by a pair of cameras: one sequence was used in the map learning step, and the other one was used for vehicle localization; the result turns out to be robust.

6. Conclusion

In order to overcome the inconsistency problem of probabilistic methods, this paper proposed a consistent vehicle localization method based on interval analysis. We propose a 2-stage process: mapping and localization. The map is built by a leader vehicle equipped with a sensor or a set of sensors providing an accurate positioning. The built map can be beneficial for a fleet of vehicles intending to achieve a consistent localization in the environment. By fusing low-cost dead reckoning sensors and map data, the proposed localization method is able to locate the vehicle consistently in the environment using the current image taken by the camera. The cumulative odometric error is eliminated and the localization consistency is maintained. Our method is validated with an outdoor car like vehicle equipped with odometers, gyro, and monocular camera. The experiments highly illustrate the consistency of the localization results

provided by our method. Future work will deal with a formal proof of the obtained consistency.

Our methodology can be used to solve various problems, even problems with higher DOF. Indeed, the used ICP algorithms express a polynomial time with respect to the number of variables in the ICSP [42], so our ICP based method is capable of scaling up to higher DOF cases without high CPU load. In order to adapt to other applications, we have only to define a new motion model.

Data Availability

Data used in this manuscript are openly available at "http://ipds.univ-bpclermont.fr/"; see [38] for more details.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

References

- [1] M. Adams, S. Zhang, and L. Xie, "Particle filter based outdoor robot localization using natural features extracted from laser scanners," in *Proceedings of the Proceedings- 2004 IEEE International Conference on Robotics and Automation*, pp. 1493–1498, May 2004.
- [2] A. Gning and P. Bonnifait, "Guaranteed dynamic localization using constraints propagation techniques on real intervals," in *Proceedings of the Proceedings- 2004 IEEE International Conference on Robotics and Automation*, pp. 1499–1504, May 2004.
- [3] D. Gruyer, R. Belaroussi, and M. Revilloud, "Map-aided localization with lateral perception," in *Proceedings of the 25th IEEE Intelligent Vehicles Symposium, IV 2014*, pp. 674–680, June 2014.
- [4] A. Howard, M. Matarik, and G. Sukhatme, "Localization for mobile robot teams using maximum likelihood estimation," in *Proceedings of the IROS 2002: IEEE/RSJ International Conference on Intelligent Robots and Systems*, vol. 1, pp. 434–439, Lausanne, Switzerland, 2002.
- [5] D. Gruyer, A. Lambert, M. Perrollaz, and D. Gingras, "Experimental comparison of bayesian positioning methods based on multi-sensor data fusion," *International Journal of Vehicle Autonomous Systems*, vol. 12, no. 1, pp. 24–43, 2014.
- [6] S. Thrun, W. Burgard, and D. Fox, "A Probabilistic Approach to Concurrent Mapping and Localization for Mobile Robots," *Autonomous Robots*, vol. 5, no. 3-4, pp. 253–271, 1998.
- [7] W. Burgard, D. Fox, D. Hennig, and T. Schmidt, "Estimating the absolute position of a mobile robot using position probability grids," in *Proceedings of the 1996 13th National Conference on Artificial Intelligence*, pp. 896–901, August 1996.
- [8] F. Dellaert, D. Fox, W. Burgard, and S. Thrun, "Robust Monte Carlo localization for mobile robots," *Artificial Intelligence*, vol. 128, no. 1-2, pp. 99–141, 2001.
- [9] S. Huang and G. Dissanayake, "A critique of current developments in simultaneous localization and mapping," *International Journal of Advanced Robotic Systems*, vol. 13, no. 5, pp. 1–13, 2016.
- [10] M. Di Marco, A. Garulli, A. Giannitrapani, and A. Vicino, "A set theoretic approach to dynamic robot localization and mapping," *Autonomous Robots*, vol. 16, no. 1, pp. 23–47, 2004.
- [11] R. E. Moore, *Interval Analysis*, Prentice-Hall, Englewood Cliffs, NJ, USA, 1966.

- [12] A. Gning and P. Bonnifait, "Constraints propagation techniques on real intervals for guaranteed fusion of redundant data. application to the localization of car-like vehicles," *Automatica*, vol. 42, no. 7, pp. 1167–1175, 2006.
- [13] L. Jaulin, *Applied Interval Analysis: with Examples in Parameter And State Estimation, Robust Control And Robotics*, Springer, 2001.
- [14] T. Rassi, N. Ramdani, and Y. Candau, "Set membership state and parameter estimation for systems described by nonlinear differential equations," *Automatica*, vol. 40, no. 10, pp. 1771–1777 (2005), 2004.
- [15] M. Kieffer, L. Jaulin, É. Walter, and D. Meizel, "Robust autonomous robot localization using interval analysis," *Reliable Computing. An International Journal Devoted to Reliable Mathematical Computations Based on Finite Representations and Guaranteed Accuracy*, vol. 6, no. 3, pp. 337–362, 2000.
- [16] L. Jaulin, M. Kieffer, E. Walter, and D. Meizel, "Guaranteed robust nonlinear estimation with application to robot localization," *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)*, vol. 32, no. 4, pp. 374–381, 2002.
- [17] M. Kieffer, E. Seignez, A. Lambert, E. Walter, and T. Maurin, "Guaranteed robust nonlinear state estimator with application to global vehicle tracking," in *Proceedings of the 44th IEEE Conference on Decision and Control, and the European Control Conference, CDC-ECC '05*, pp. 6424–6429, esp, December 2005.
- [18] E. Seignez, M. Kieffer, A. Lambert, E. Walter, and T. Maurin, "Real-time bounded-error state estimation for vehicle tracking," *International Journal of Robotics Research*, vol. 28, no. 1, pp. 34–48, 2009.
- [19] A. Lambert, D. Gruyer, B. Vincke, and E. Seignez, "Consistent outdoor vehicle localization by bounded-error state estimation," in *Proceedings of the 2009 IEEE/RSJ International Conference on Intelligent Robots and Systems, IROS 2009*, pp. 1211–1216, October 2009.
- [20] I. K. Kueviakoe, A. Lambert, and P. Tarroux, "A real-time interval constraint propagation method for vehicle localization," in *Proceedings of the 2013 16th International IEEE Conference on Intelligent Transportation Systems: Intelligent Transportation Systems for All Modes, ITSC 2013*, pp. 1707–1712, nld, October 2013.
- [21] V. Drevelle and P. Bonnifait, "Robust positioning using relaxed constraint-propagation," in *Proceedings of the 23rd IEEE/RSJ 2010 International Conference on Intelligent Robots and Systems, IROS 2010*, pp. 4843–4848, twm, October 2010.
- [22] K. Lassoued, P. Bonnifait, and I. Fantoni, "Cooperative Localization with Reliable Confidence Domains Between Vehicles Sharing GNSS Pseudorange Errors with No Base Station," *IEEE Intelligent Transportation Systems Magazine*, vol. 9, no. 1, pp. 22–34, 2017.
- [23] B. Desrochers and L. Jaulin, "A minimal contractor for the polar equation: Application to robot localization," *Engineering Applications of Artificial Intelligence*, vol. 55, pp. 83–92, 2016.
- [24] L. Jaulin, A. Stancu, and B. Desrochers, "Inner and outer approximations of probabilistic sets," in *Proceedings of the American Society of Civil Engineers (ASCE) Second International Conference on Vulnerability and Risk Analysis and Management and Sixth International Symposium on Uncertainty Modelling and Analysis*, pp. 13–16.
- [25] L. Jaulin, "Range-only SLAM with occupancy maps: A set-membership approach," *IEEE Transactions on Robotics*, vol. 27, no. 5, pp. 1004–1010, 2011.
- [26] P. T. I. K. Kueviakoe and A. Lambert, "Vehicle localization based on interval constraint propagation algorithms," in *Proceedings of the IEEE International Conference on Cyber Technology in Automation, Control and Intelligent Systems*, pp. 179–184, May 2013.
- [27] R. E. Moore, R. B. Kearfott, and M. J. Cloud, *Introduction to Interval Analysis*, Society for Industrial Mathematics, 2009.
- [28] F. Le Bars, E. Antonio, J. Cervantes, C. De La Cruz, and L. Jaulin, "Estimating the trajectory of low-cost autonomous robots using interval analysis: Application to the eurathlon competition," in *Marine Robotics and Applications*, pp. 51–68, Springer, 2018.
- [29] E. Hyvönen, "Constraint reasoning based on interval arithmetic: the tolerance propagation approach," *Artificial Intelligence*, vol. 58, no. 1-3, pp. 71–112, 1992.
- [30] A. Lambert and N. Le Fort-Piat, "Safe task planning integrating uncertainties and local maps federations," *International Journal of Robotics Research*, vol. 19, no. 6, pp. 597–611, 2000.
- [31] B. Vincke, A. Lambert, D. Gruyer, A. Elouardi, and E. Seignez, "Static and dynamic fusion for outdoor vehicle localization," in *Proceedings of the 11th International Conference on Control, Automation, Robotics and Vision, ICARCV 2010*, pp. 437–442, December 2010.
- [32] J. M. M. Montiel, J. Civera, and A. J. Davison, "Unified inverse depth parametrization for monocular SLAM," in *Proceedings of the 2nd International Conference on Robotics Science and Systems (RSS '06)*, pp. 81–88, August 2006.
- [33] 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.
- [34] A. Schmidt, M. Kraft, and A. Kasiński, "An Evaluation of Image Feature Detectors and Descriptors for Robot Navigation," in *Computer Vision and Graphics*, pp. 251–259, 2010.
- [35] R. Hartley and A. Zisserman, *Multiple View Geometry in Computer Vision*, Cambridge University Press, 2nd edition, 2003.
- [36] Z. Dan and H. Dong, "A robust object tracking algorithm based on surf," in *Proceedings of the IEEE International Conference on Wireless Communications & Signal Processing (WCSP)*, pp. 1–5, 2013.
- [37] J. P. Lewis, "Fast normalized cross-correlation," *Vision Interface*, vol. 10, pp. 120–123, 1995.
- [38] H. Korrapati, J. Courbon, S. Alizon, and F. Marmoiton, "The institut pascal data sets: un jeu de données en extérieur, multicapteurs et datées avec réalité terrain, données d'étalement et outils logiciels," in *Orasis, Congrès des jeunes chercheurs en vision par ordinateur*, 2013.
- [39] E. Royer, M. Lhuillier, M. Dhome, and T. Chateau, "Localization in urban environments: Monocular vision compared to a differential GPS sensor," in *Proceedings of the 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, CVPR 2005*, vol. 2, pp. 114–121, June 2005.
- [40] J. Courbon, Y. Mezouar, and P. Martinet, "Autonomous navigation of vehicles from a visual memory using a generic camera model," *IEEE Transactions on Intelligent Transportation Systems*, vol. 10, no. 3, pp. 392–402, 2009.
- [41] H. Lim, S. N. Sinha, M. F. Cohen, M. Uyttendaele, and H. J. Kim, "Real-time monocular image-based 6-DoF localization," *International Journal of Robotics Research*, vol. 34, no. 4-5, pp. 476–492, 2015.
- [42] G. Chabert and L. Jaulin, "Contractor programming," *Artificial Intelligence*, vol. 173, no. 11, pp. 1079–1100, 2009.