Detection of road traffic participants using cost-effective arrayed ultrasonic sensors in low-speed traffic situations

Downloaded from: https://research.chalmers.se, 2026-11-10 06:33 UTC

Citation for the original published paper (version of record):
Li, G., Li, S., Zou, R. et al (2019). Detection of road traffic participants using cost-effective arrayed ultrasonic sensors in
low-speed traffic situations. Mechanical Systems and Signal Processing, 132: 535-545.
http://dx.doi.org/10.1016/j.ymssp.2019.07.009

N.B. When citing this work, cite the original published paper.

research.chalmers.se offers the possibility of retrieving research publications produced at Chalmers University of Technology. It covers all kind of research output: articles, dissertations, conference papers, reports etc. since 2004. research.chalmers.se is administrated and maintained by Chalmers Library

(article starts on next page)

Mechanical Systems and Signal Processing 132 (2019) 535–545

Contents lists available at ScienceDirect

Mechanical Systems and Signal Processing

journal homepage: www.elsevier.com/locate/ymssp

Detection of road traffic participants using cost-effective arrayed ultrasonic sensors in low-speed traffic situations

Guofa Li [a,b], Shengbo Eben Li [b,*], Ruobing Zou [b], Yuan Liao [c], Bo Cheng [b]

[a] Institute of Human Factors and Ergonomics, College of Mechatronics and Control Engineering, Shenzhen University, Shenzhen 518060, China
[b] State Key Lab of Automotive Safety and Energy, School of Vehicle and Mobility, Tsinghua University, Beijing 100084, China
[c] Division of Physical Resource Theory, Department of Space, Earth and Environment, Chalmers University of Technology, Gothenburg 41296, Sweden

ARTICLE INFO

Article history:
Received 22 March 2019
Received in revised form 23 June 2019
Accepted 6 July 2019

Keywords:
Driver assistance systems
Driving safety
Ultrasonic sensor
Object detection

ABSTRACT

Effective detection of traffic participants is crucial for driver assistance systems. Traffic safety data reveal that the majority of preventable pedestrian fatalities occurred at night. The lack of light at night may cause dysfunction of sensors like cameras. This paper proposes an alternative approach to detect traffic participants using cost-effective arrayed ultrasonic sensors. Candidate features were extracted from the collected episodes of pedestrians, cyclists, and vehicles. A conditional likelihood maximization method based on mutual information was employed to select an optimized subset of features from the candidates. The belonging probability to each group along with time was determined based on the accumulated object type attributes outputted from a support vector machine classifier at each time step. Results showed an overall detection accuracy of 86%, with correct detection rate of pedestrians, cyclists and vehicles around 85.7%, 76.7% and 93.1%, respectively. The time needed for detection was about 0.8 s which could be further shortened when the distance between objects and sensors was shorter. The effectiveness of arrayed ultrasonic sensors on objects detection would provide all-around-the-clock assistance in low-speed situations for driving safety.

© 2019 Published by Elsevier Ltd.

1. Introduction

Walking and bicycling are fundamental modes of transportation. However, more than 2000 cyclists and 7000 pedestrians were killed annually in the European Union, accounted for 27% of all fatalities [1]. The situation was even worse in the emerging and developing countries like India and China [2]. Their rapid economic growths increase travel demand, but overwhelm the limited transportation infrastructure. The traffic in these developing countries is characterized by a wide mix of vehicle types and pedestrians that differ substantially in their dimensions, performance capabilities and driver behavior. These situations greatly challenge the availability and feasibility of Advanced Driver Assistance Systems (ADAS). To avoid crashes with cyclists and pedestrians in mixed traffic, it is critical to detect the moving objects to help car drivers notice them when they are around the vehicle [3]. It is noted that current studies always treated pedestrians or cyclists separately without distinguishing the differences between them in a single study. However, their mobility flexibilities and patterns do differentiate. With a clear understanding of the object types around the vehicle, ADAS could carry out different strategies to keep drivers being aware of the imminent conflicts or to help drivers prevent crashes when they are not aware of the dangers [4,5].

* Corresponding author.
E-mail addresses: guofali@szu.edu.cn (G. Li), lishbo@tsinghua.edu.cn (S.E. Li), yuan.liao@chalmers.se (Y. Liao), chengbo@tsinghua.edu.cn (B. Cheng).

https://doi.org/10.1016/j.ymssp.2019.07.009
0888-3270/© 2019 Published by Elsevier Ltd.

To understand the surrounding environment, most of the current techniques developed to detect vehicles, pedestrians and cyclists are based on visible-light cameras [5–7]. Histograms of Oriented Gradients (HOG) features combining with Support Vector Machine (SVM) classification algorithms are commonly used to establish the object detection methods based on camera systems [8]. Baek et al. proposed a developed SVM algorithm to detect pedestrians using HOG features [9]. Results indicated promising applications in detection accuracy and computation time. A HOG + SVM method was proposed to detect vehicles on road in [10] and results showed competitive performance in field applications. Similarly, cyclists were successfully extracted from naturalistic driving videos using HOG pattern recognition methods in [11].

Recent adoption of deep learning methodologies, convolutional neural networks (CNN) in particular, further improves the object detection technologies based on camera systems [8,12]. By training networks like ResNet [13], VGG [14] and GoogleNet [15] for classification, models learn to extract powerful features from raw pixels for object detection. Ouyang et al. proposed a joint deep network architecture for pedestrian detection [16]. The architecture jointly learns feature extraction, deformation handling, occlusion handling, and classification to maximize the strength of each component. Results showed an average miss rate of 8.57% on the examined dataset [17] proposed a model that included multiple build-in subnetworks to detect pedestrians with scales from disjoint ranges. The final detection decision were made from an adaptive combination of the outputs from the subnetworks. Competitive results were reported to prove the state-of-the-art performance.

However, night time situation may cause dysfunction of these camera systems. Because of the lack of light, dynamically changing light distribution and increasing image noise at night time, the extracted features from images will be severely degraded [18]. On rural roads without lighting facilities, it is hardly for camera systems to detect objects, even for human eyes. Traffic safety statistics show that the majority (77%) of preventable pedestrian fatalities occurred at night in 2015 in the U.S. [19]. This study also compared the fatalities elimination effectiveness of different sensor technologies. Results showed that using cameras only cannot give a satisfactory performance on pedestrian fatalities elimination at night. In summary, although camera systems are cost-effective and show high detection accuracy in daytime, nighttime situation is a major disadvantage which strongly limits their practical applications. Alternatively, ultrasonic, radar or lidar systems may contribute, independent of the lighting condition [19].

Based on a radar system, Hyun et al. proposed a coherent phase difference method to detect pedestrians in complex environments [20]. Results showed that the method could effectively distinguish pedestrians from the background environment. Based on a 3D lidar system, Asvadi et al. proposed an environment perception system by employing a ground surface estimation and an object detection model to recognize cyclists and pedestrians [21]. Results demonstrated the effectiveness and efficiency of the proposed method based on the KITTI database. To solve the sparse point cloud problem of departing pedestrians in lidar detection, [22] proposed a density enhancement method to rehabilitate the shapes of pedestrians. Naturalistic driving data were used to examine the effectiveness of the enhancement method. However, although both radar and lidar systems can be effective in object detection, these approaches are much more expensive than camera or ultrasonic systems [19].

Patole et al. reviewed various aspects of sensor technologies in automotive applications (e.g., pedestrian and cyclist detection) [7]. Considering the influence of lighting and cost in practical applications, ultrasonic sensor may be a promising candidate for object detection in low-speed traffic situations. However, none of the previous studies using ultrasonic sensors have tried to distinguish the object types which is commonly analyzed in camera systems. Liang and Juang proposed a SVM approach to classify moving objects (pedestrians, vehicles, bicycles, and motorcycles) using a combination of static appearance features and spatial and temporal entropy values from camera systems [23]. Zangenehpour et al. presented a SVM classification method to detect pedestrians, cyclists and vehicles using extracted HOG features from video images collected in low-speed traffic situations [24].

Accordingly, this study aims to develop an automated road user detection methodology using arrayed ultrasonic sensors. Moving objects are categorized into three groups: pedestrians, cyclists, and motor vehicles. A prior analysis on the performance of the used ultrasonic sensors was presented in [25]. That paper focuses on an introduction of the dynamical tracking capability of the employed sensors. This paper further extends the previous work to recognize the types of the detected objects. The contributions of this study include: (1) We utilize a scheme to effectively detect pedestrians, cyclists and vehicles based on arrayed ultrasonic sensors. Although this methodology may not be entirely novel in the field of computer science, this work is the first to combine and use this method in ultrasonic sensor application in automotive engineering. (2) The effective features for object detection using ultrasonic sensors are learned from the original features by approximating iterative maximizers of the conditional likelihood. Therefore, far more candidate features can be proposed to cover more information of the target objects from the original signals without worrying the effectiveness of the potential features. (3) A decision making method based on Dempster-Shafer theory (DST) is developed to determine the object type along with time. Outputs of the SVM classifier may not be consistent at each time step. Addressing how to correctly recognize the object type along with time would help enhance the design of driver assistance systems based on ultrasonic sensors.

2. Scheme of the detection algorithm

The scheme for moving objects detection is shown in Fig. 1. Following the original data module as the foundation of the proposed algorithm, three phases are organized sequentially:

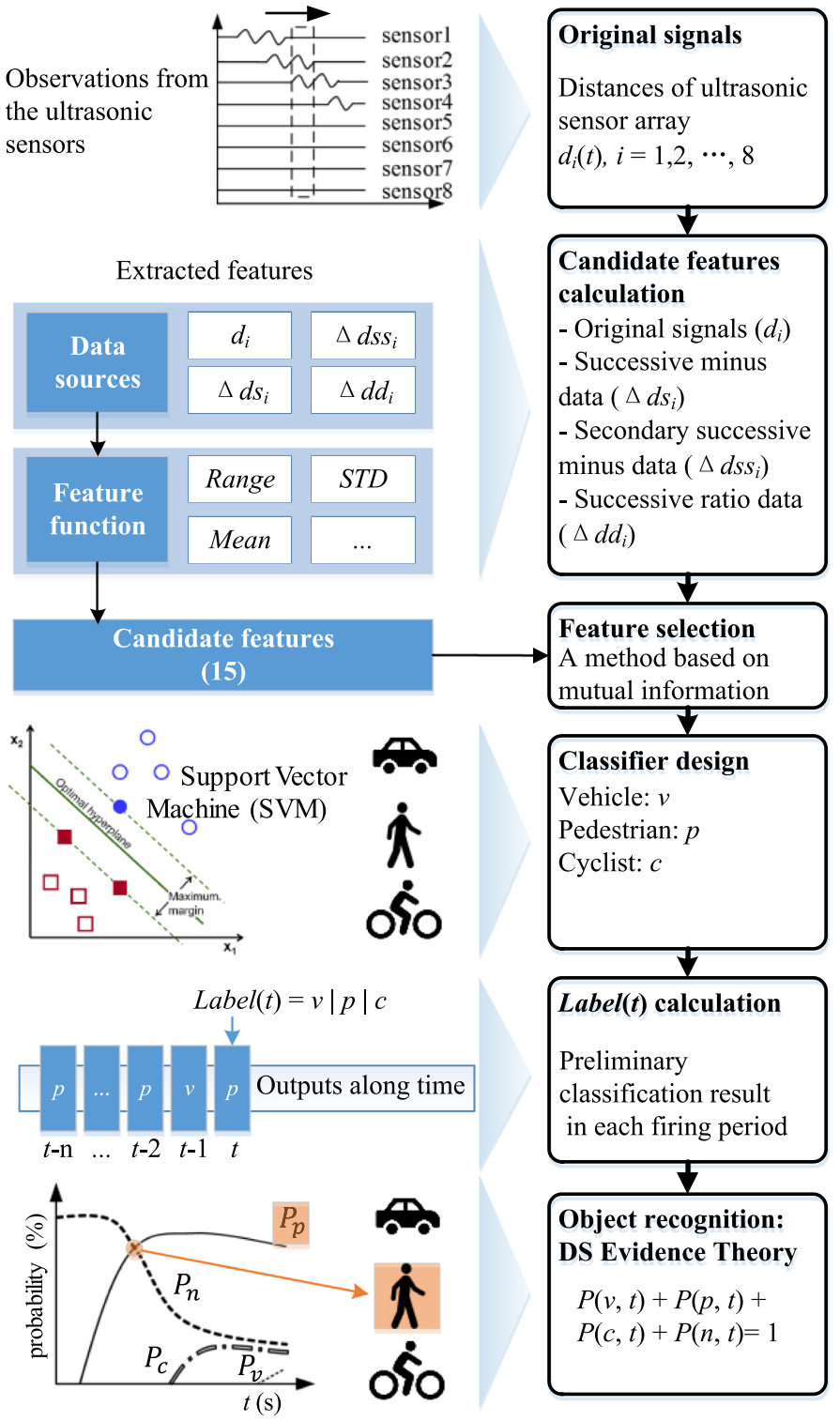

Fig. 1. Scheme of the detection algorithm.

(1) Selecting the optimized feature subset: The original signals are the measured distances d_i between the ultrasonic sensors and objects. Candidate features are derived from the measured distances. A conditional likelihood maximization method based on mutual information is adopted to reduce the feature dimensionality for an optimized subset.

(2) Classifying the digital appearance at each time step: Support Vector Machine (SVM) was employed for this task. At each time step, the SVM classifier generates a classification output of the detected object using the selected features as inputs to describe the 'digital appearance' of the detected object.
(3) Computing the belonging probability to each group along with time: The classification results are fed into an object type recognition algorithm. The algorithm computes the belonging probability to each group along with time based on the Dempster-Shafer theory of evidence.

3. Method

3.1. System configuration and data collection

The detectable distance range of the ultrasonic sensors used in this study is about 5 m, and the detectable angle range is about 120 degrees. Because of the detection scope limitation of a single ultrasonic sensor, multiple ones are needed for applications in vehicles to detect surrounding objects. Therefore, eight sensors with an interval of 50 cm between each were arrayed linearly to detect objects. Two typical firing sequences, serial firing and mutual firing, were applicable, while the serial firing sequence was employed in this study because of its relatively better performance as reported in [25]. The firing period T was 50 ms. An additional camera was used for data synchronization and verification.

Real traffic data were collected on the main roads in Tsinghua University. See Fig. 2 for the examples of the collected images. To prepare the training and testing data for the SVM classifier, a manual coding method was used to label the object type (pedestrians, cyclists or vehicles) in each episode by visual identification. The beginning of each episode was defined as the time when the first sensor received the first returning echo from the detected object, and the ending as the time when the last sensor received the last returning echo. The duration of each episode varied from 3 to 7 s with one object appeared in each. Although detection of a single road user using ultrasonic sensors is quite straight-forward, it would still contribute to the automotive industry as this cost-effective method has never been reported before in this research community. The detection and localization of multiple objects in one episode using ultrasonic sensors will be conducted in our further extended work based on this study.

3.2. Candidate features

Each sensor measured and returned the distance between an object and the sensor itself. The arrayed system would return a vector including eight measured distances at each time step. Four function groups were designed to describe the characteristics of detected objects, including successive minus data, secondary successive minus data, ratio of successive minus data and the original data. See (1)–(3) for the calculation formulas of the employed functions. In each firing period of the arrayed system, the original data, successive minus data, secondary successive minus data, and successive ratio data would be collected or computed. Six features was derived from the original data. See Table 1. The mean, standard deviation, and range of the other three datasets were computed to collect nine features. All these 15 features were combined together to form a feature vector for the following phase. Although the arrayed sensors cannot directly measure sizes and velocities of the detected objects, the differences of these attributes among the object types would result in the differences of the proposed candidate features.

The successive minus data Δds is given as:

$$\Delta ds_i = d_i - d_{i+1} \tag{1}$$

where d_i is the measured distance to the detected object from sensor i.

The secondary successive minus data Δdss is given as:

$$dss_i = \Delta ds_i - \Delta ds_{i+1} \tag{2}$$

(a) Vehicle (b) Pedestrian (c) Cyclist

Fig. 2. Examples of the collected episodes.

Table 1
Candidate features.

Data sources	#	Descriptions	Features	Unit
Original data (d)	1	Number of valid data	*Num*	–
	2	Standard deviation	*Std*	cm
	3	Range, i.e. max–min	*Range*	cm
	4	(Fist measure – last measure)/Range	*FLPR*	–
	5	(Last measure – min)/Range	*LMPR*	–
	6	(First measure – min)/Range	*FMPR*	–
The successive minus data (Δds)	7	Mean	*MeanS*	cm
	8	Standard deviation	*StdS*	cm
	9	Range	*RangeS*	cm
The secondary successive minus data (Δdss)	10	Mean	*MeanSS*	cm
	11	Standard deviation	*StdSS*	cm
	12	Range	*RangeSS*	cm
The successive ratio data (Δdd)	13	Mean	*MeanD*	–
	14	Standard deviation	*StdD*	–
	15	Range	*RangeD*	–

The successive ratio data Δdd is given as:

$$\Delta dd_i = d_i / d_{i+1} \tag{3}$$

3.3. Feature selection

To select an optimized subset from the 15 candidate features, a conditional likelihood maximization method based on mutual information is employed [26]. This method incorporates the feature relevancy and redundancy concepts together to select the optimized subset by approximate iterative maximizers of the conditional likelihood. The optimized subset is a best balance between the relevancy and redundancy of the candidate features. The criterion of feature selection using the joint mutual information (JMI), which has been proved to be the best tradeoff in terms of accuracy, stability, and flexibility with small data samples [26], is:

$$J_{JMI}(X_k) = \sum_{X_j \in \mathbb{s}} I(X_k X_j; C) = I(X_k; C) - \frac{1}{|\mathbb{s}|} \sum_{X_j \in \mathbb{s}} \left[I(X_k; X_j) - I(X_k; X_j | C) \right] \tag{4}$$

where $X_k, X_j \in$ the candidate features, $C \in$ {pedestrian, cyclist, vehicle}, $\mathbb{s}$ is the set of features already selected, $|\mathbb{s}|$ is the cardinality of the $\mathbb{s}$ set. The algorithm selects the kth feature X_k that maximizes the left-hand-side of the formulas. All entries of the right-hand-side of the formulas above can be divided into three categories according to their function: $I(X_k; C)$measures the correlation between terms $X_k$and object type label C, $I(X_k; X_j)$measures the redundancy between terms $X_k$and X_j, and$I(X_k; X_j | C)$ measures the complementariness between terms $X_k$and X_j.

3.4. Object type recognition based on classifier outputs

To date, SVM has been commonly used in supervised learning tasks. "SVMs are based on statistical learning techniques and can be used for pattern classification and inference of nonlinear relationships between variables" [27]. As SVM is originally a binary classifier, a one-versus-rest strategy was utilized in the model training for each group (vehicle, pedestrian or cyclist). The same number of episodes were randomly selected from the other two groups to generate a balanced negative dataset in each training process. The performance of a SVM classifier with Gaussian kernel relies on the selection of regularization parameter R and kernel spread gamma (γ). The best combination of R and γ was selected using a grid search method. Typically, each combination of the parameters was verified using a five-fold cross validation method, and the combination with the highest cross-validation accuracy was selected. The introduction to LIBSVM MATLAB toolbox used in this study can be found in [28].

Using the chosen descriptors from feature selection as inputs of the classifier, a classified object type would be outputted at each time step. However, the output classification result may not be consistent along with time, which has not been analysed well in the existed literature. Thus, how to correctly recognize the type of the moving objects along with time is the key element in our proposed method. The following algorithm has been proposed to solve this problem.

The temporal likelihood of the three defined categories is calculated according to (5). If only one returned echo measurement has been received in total from the beginning of the episode, the temporal likelihood will be denoted as 'not specified'. Otherwise, it will be calculated according to the majority outputs of the historical data in the episode.

$$P(C|t) = \begin{cases} not\ specified, & if\ only\ 1\ echo\ is\ received \\ n_{Ct}/N_t & otherwise \end{cases} \tag{5}$$

where $C \in \{pedestrian, cyclist, vehicle\}$, $P(C|t)$ is the temporal likelihood of the examined episode at time t, n_{Ct} is the number of accumulated data points categorized into type C by the classifier at time t in the episode, N_t is the total number of data points at time t in the episode. Note that $P(notspecified|t) + P(pedestrian|t) + P(cyclist|t) + P(vehicle|t) = 1$.

The category label at time t was determined based on the Dempster-Shafer theory (DST) of evidence, a promising improvement on 'traditional' approaches to decision making [29]. The DST, based on the use of probabilities with upper and lower bounds, has subsequently been popularized in artificial intelligence and expert systems with particular emphasis on combining evidences from different sources [30].

Set the uncertainty set $\Theta = \{pedestrian, cyclist, vehicle\}$ and define the basic probability assignment (bpa) function as:

$$m_t(C) = \begin{cases} 0 & t = 0 \\ P(C|t) & otherwise \end{cases} \tag{6}$$

where $C \epsilon \Theta$. The bpa function is also known as a mass function, similar to the posterior probability in the Bayesian theory. Then the belief function accepting C and the plausibility function not rejecting C can be defined as:

$$Bel_t(C) = \sum\nolimits_{A \subset C} m_t(A) \tag{7}$$

$$Pl_t(C) = 1 - Bel_t(\overline{C}) \tag{8}$$

Thus, the confidence interval of C is $[Bel_t(C), Pl_t(C)]$. Regarding the sequential outputs from classifier as evidence sources, the combination of these mass function is defined according to Dempster rule:

$$(m_1 \oplus m_2 \oplus \cdots \oplus m_t)(\mathrm{C}) = \frac{1}{K} \sum_{C_1 \cap C_2 \cap \cdots \cap C_t = C} m_1(C_1) \cdot m_2(C_2) \cdots m_t(C_t) \tag{9}$$

$$K = \sum_{C_1 \cap C_2 \cap \cdots \cap C_t \neq \varnothing} m_1(C_1) \cdot m_2(C_2) \cdots m_t(C_t) \tag{10}$$

The coefficient K is known as the normalization factor, while $1 - K$ indicates the conflict among the evidences.

The final decision on the category label at time t is determined by the proposed method as shown in (11).

$$Label(t) = \begin{cases} not - spicified, & if \ \max\limits_{C} Bel_t(C) < Bel_t(\Theta) \\ arg \ \max\limits_{C} Bel_t(C), & otherwise \end{cases} \tag{11}$$

where Label(t) is the detection result of the examined episode at time t, $Bel_t(C)$ is the lower bound of the confidence interval of labelling the object as C.

4. Results and discussion

In total, 210 episodes of pedestrians, 236 of cyclists and 320 of vehicles were collected. Multi-features were derived from the originally collected signals. A conditional likelihood maximization method based on mutual information was employed to select an optimized subset of features from the candidates. The belonging probability to each group along with time was determined based on the outputs from classifiers using the proposed recognition method. The distance effect on the detection performance was examined by a simulation experiment. Practical implications and limitations of this study were then proposed.

4.1. Selected features and detection results

Eight features were selected to achieve the highest classification accuracy. The eight features were: *Num*, *Range*, *StdD*, *RangeS*, *MeanSS*, *StdS*, *Std*, *RangeD*. Among the eight features, three of them were derived from the original data, two from the successive minus data, one from the secondary successive minus data and two from the successive ratio data. See Fig. 3 for the differences of the selected features between different object types. Statistical significances were found among the object type groups in all the selected features. Take Fig. 3(c) for an example, the standard deviation of the successive ratio data describes the shape and size of the detected objects. If a detected object is covered with a flat surface (e.g., a vehicle), the successive ratio data will not fluctuate much, thus the value of its standard deviation will be smaller as illustrated. Larger size of vehicles also contributes to lower level of *Range, RangeD,* and *Std.* A SVM classifier was then employed to recognize the detected objects based on the eight selected features. A five-fold method was applied for cross validation. The correct classification rate could then be measured.

By applying the recognition algorithm based on DST, the probabilities to be recognized as pedestrians, cyclists and vehicles along with time are shown in Fig. 4. The DST takes the outputs from (11) at each time step as evidence and calculates the lower bounds of belief functions for the three target object types. The probability to accept the hypothesis that the detected object belongs to the specific category will increase along with time. That is because more evidences have been collected for

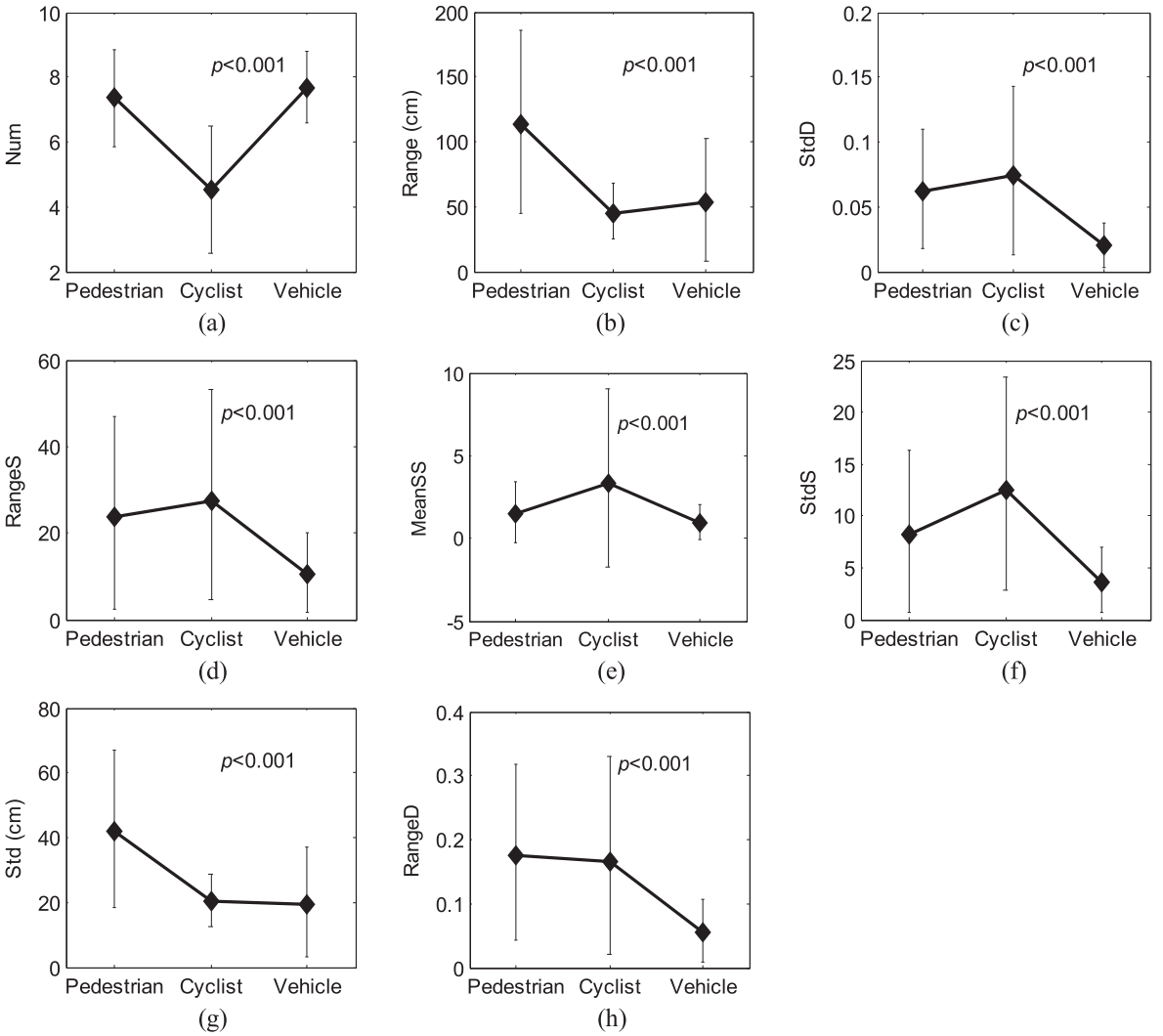

Fig. 3. Differences of the selected features between object types.

decision-making on the object type. Compared with the original outputs regarding the proportional results from the SVM classifier, the method proposed in this paper boasts the ability of more stable and robust detection since the uncertainty from sensors and the fluctuation from the SVM classifier is tolerable.

If the probability to be recognized as the correct object type is higher than the other two types and never falls behind after the first achievement, the episode is considered to be detected correctively which will count for the detection accuracy. In summary, the overall detection accuracy was 86.0%. Some 85.7% (180 out of 210) of the pedestrians were detected correctly. The numbers were 76.7% for cyclists, and 93.1% for vehicles. Taking pedestrians and cyclists as a vulnerable road user group, it was easy to distinguish the vulnerable group from the vehicles. The detection accuracy for the vulnerable group and the vehicles group were 98.7% and 93.1%, respectively. Only 6 vulnerable group episodes were detected as vehicles, and 22 vehicle episodes were detected as vulnerable road users. However, pedestrians and cyclists were in a way similar to each other so that the misdetection rate between these two types was high. About 13.3% of the pedestrians were detected as cyclists, and 21.6% of the cyclists were detected as pedestrians.

As for the time needed for detection, it is defined to be the first time onset when the probability to be recognized as the correct object type is higher than the other two types. As shown in Fig. 4, the time needed for detection were all about 0.8 s for either pedestrians, cyclists, or vehicles. As indicated in previous studies, drivers' brake response time to imminent dangers was about 1.0 s [31]. The detection time of the proposed method in this study is shorter than drivers' brake response time, which may help a lot in the design of ADASs. As timely detection of moving objects would significantly improve ADASs performance especially when the objects are in blind zones, a higher data collection frequency may further shorten the time needed for detection. In the following simulation experiment in Section 4.3, a higher data collection frequency was adopted to verify the frequency effect on the detection performance.

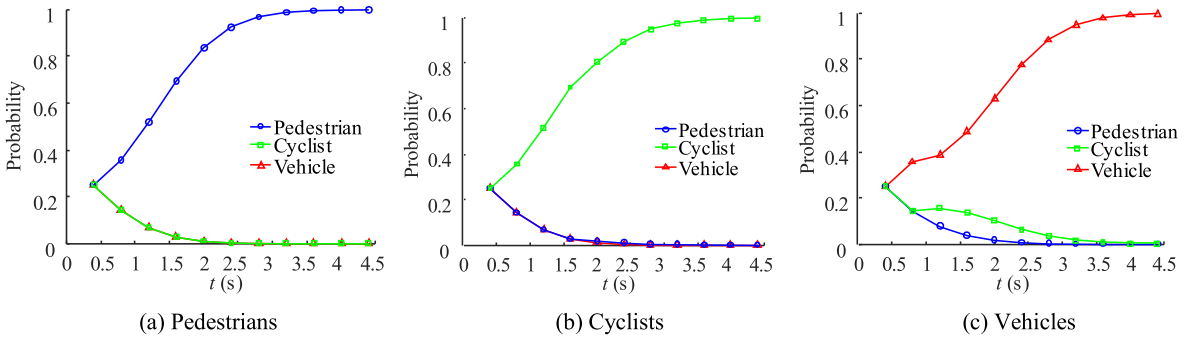

Fig. 4. The probability to be recognized as pedestrians, cyclists and vehicles along with time.

4.2. Causation on the detection performance of pedestrians and cyclists

The developed methodology achieved an overall detection accuracy of 86%. However, the detection accuracy varies across object types and is the highest for vehicles and lower for pedestrians and cyclists. Similar results were found in [24], in which HOG was used to extract features and SVM was used to recognize moving objects based on camera systems. Due to the similarity in appearance between cyclists and pedestrians (a cyclist consists of a bicycle and a human who rides the bicycle), [24] reported that cyclists were the most difficult for classification.

Due to the various shapes and surface materials of the detected objects, the number of sensors with valid measured values vary, thus the detection performance would be affected. Pedestrians can be considered as moving cylinders with cloth surface, vehicles as flat metal surface, and cyclists as a mixture of shapes and surface materials [25]. According to the sensor models built in [25], shape and surface material of objects would affect the chance of detection, thus the sensor measurements would be affected. As shown in Fig. 5(a), about $2 \sim 4$ sensors measurements were with valid values in each firing period when pedestrians passed by the arrayed system. The regular fluctuation of each sensor's measurements was similar to each other. For cyclists, the number of sensors with valid measurements mainly ranged from 2 to 6. See Fig. 5(b). The fluctuation of each sensor's measurements was similar to that of pedestrians, which may lead to the relatively high misclassification rates between pedestrians and cyclists. However, the fluctuation of cyclist's episodes was much more irregular which would contribute to its lower detection accuracy. For vehicles, the valid sensor number mainly ranged from 5 to 8. See Fig. 5(c). The length of a vehicle is about 5 m so that most of the sensors could be covered when the vehicle was passing by the arrayed system. The more valid sensor measurements and regular fluctuations would contribute to the better detection performance.

4.3. Influence of distance on the detection performance

Besides the attributes of moving objects (e.g., shape and surface material), distance between sensors and objects would also affect the detection performance. However, the distance can never be accurately controlled in naturalistic situations. To further examine the influence of distance on the detection performance of the arrayed ultrasonic sensors, a simulation platform was employed [25]. The detection results were updated every 160 ms. A cloth covered cylinder model with an 18 cm

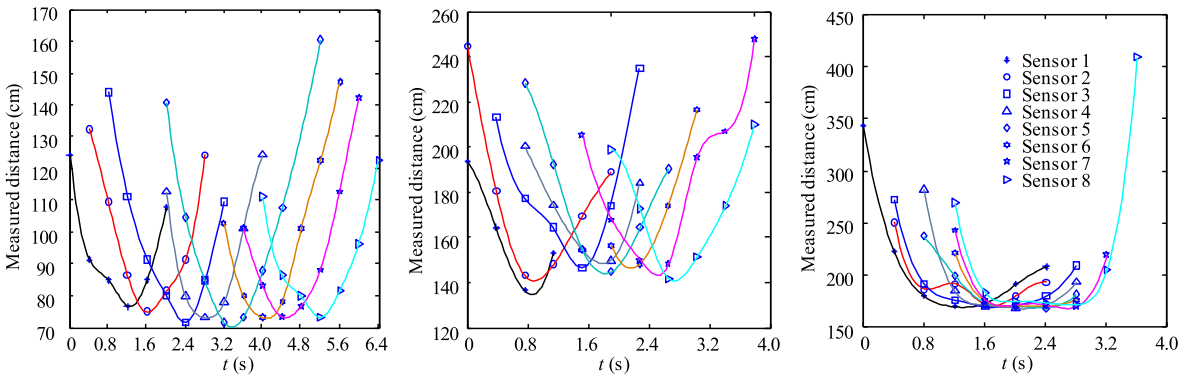

Fig. 5. Examples of measurements for pedestrians, cyclists and vehicles.

radius was used as pedestrians in the simulation. Two distances, 180 cm and 320 cm, were examined. See Fig. 6 for the detection performance.

When the distance between sensors and the moving cylinder was 180 cm, it was easily for the arrayed system to detect and correctly recognize the object type. The corresponding detection time was about 0.32 s. However, when the distance was 320 cm, the detection time delays about 0.5 s. As shown in Fig. 6(b), the moving object cannot be correctly recognized in the first 0.8 s. As for the influence of data collection frequency on the detection performance, results illustrated in Fig. 6(a) show that the detection time was significantly shortened when compared with the results illustrated in Fig. 4(a). Therefore, it can be concluded that both distance and collection frequency affect the detection performance, which needs to be considered in further applications.

4.4. Practical implications of this study

Although the detection range of ultrasonic sensor is limited, it can be an alternative or assistant approach for camera systems in low-speed traffic situations with mixed traffic participants around vehicles. See the left bottom pictures in Fig. 7 for a general understanding on the mixed traffic situations which will happen every day all around Beijing, China. Vehicles are very close to each other, both longitudinal and lateral, with pedestrians and cyclists around vehicles with short distances where ultrasonic sensors would be helpful for safety enhancement.

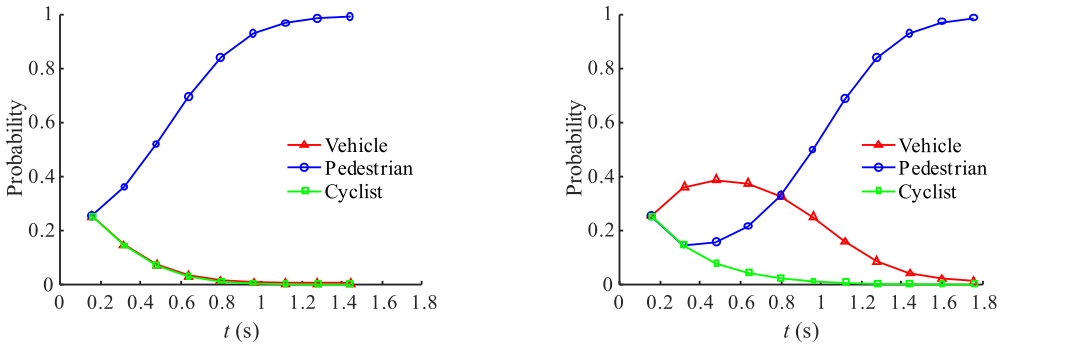

Fig. 6. Detection performance of pedestrians in different distance situations.

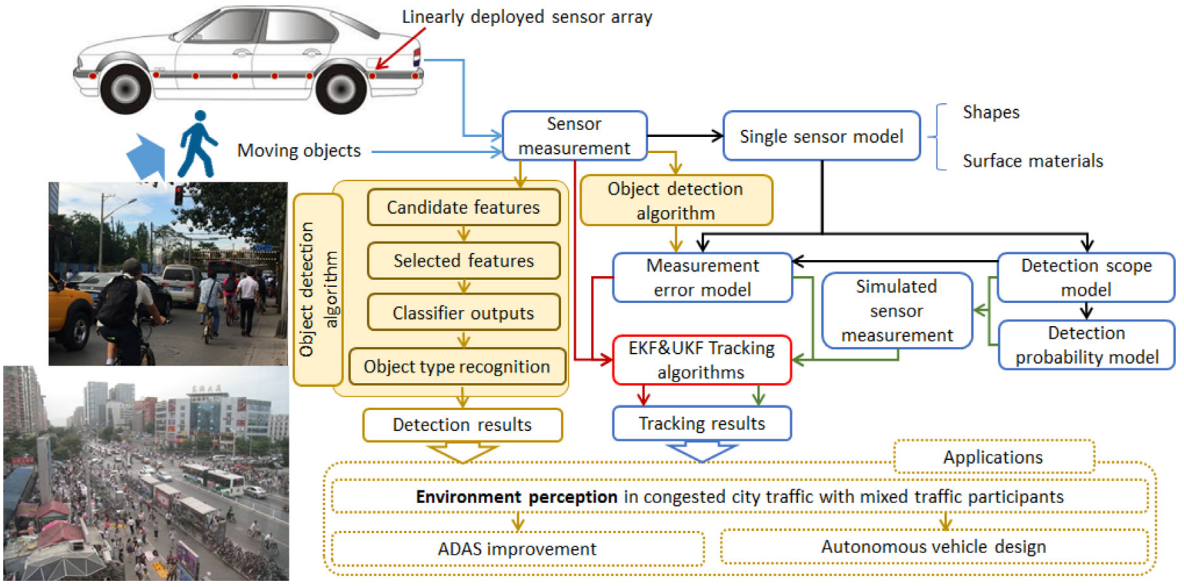

Fig. 7. A possible sensor configuration on vehicles and the extended applications (EKF for Extended Kalman Filter [25], UKF for Unscented Kalman Filter [2532]).

A possible sensor configuration on vehicles and the extended applications are proposed in Fig. 7. The linearly arrayed sensors provide a wider view span for assistance on environment perception around vehicles. The elements connected with solid black lines, green lines and red lines have already been presented in [25]. The work presented in this paper further extends the previous work on the key issue of object detection. Marked as dash lines in Fig. 7, object detection method and tracking method based on arrayed ultrasonic sensors can be applied in the environment perception in congested city traffic with low speed and mixed traffic participants. Current ADAS systems intensively focus on vehicle safety countermeasures in middle- or high-speed traffic situations, such as forward collision warning systems, lane departure warning systems and so on. However, lateral safety countermeasures in low-speed and night driving situations have not drawn much attention. The work presented in this paper would provide a novel perspective to improve ADAS development and to strengthen the environment perception module in autonomous vehicle design. Based on the better environment perception capability, automatic steering systems could be further developed to help avoid crashes [33].

For applications in traffic management, this work can be used to count pedestrians, cyclists and vehicles to optimize the design of road infrastructures to provide better transportation services and solutions. Other possible applications include road access control, parking lot management, automatic toll collection, etc. The advantages of cost-effective and independence of light distribution will encourage ultrasonic systems to be deployed in the above-mentioned systems.

4.5. Contributions, limitations and future work

There are many promising algorithms in the artificial intelligence community and they are quickly updated. Researchers will not always use the up-to-date-best algorithm in their studies. SVM has been widely used in machine learning tasks with sound performance. That's why we use SVM here in this study, but which classifier we use is not the focus of this work. Further efforts can be inspired to adopt other advanced algorithms (e.g., CNN, RNN, etc.) for better performances following the proposed methodology in this study. Our main contributions focus on the successful employment of arrayed ultrasonic sensors for moving objects detection in automotive applications, which may not be novel in computer science or robotics but would improve the development of automotive industry. The extracted features to effectively describe the moving objects and the time-series decision making method based on SVM outputs are the other two contributions which have not been used for objects detection based on arrayed ultrasonic sensors.

The main limitation of this study is the miss of multiple objects detection in a single episode. As this is our first attempt to detect moving objects based on arrayed ultrasonic sensors, we cannot solve all the challenging tasks in a single study. Correct detection of multiple objects is important for traffic safety enhancement. In the future efforts, we will focus on the solution to multiple objects detection based on arrayed ultrasonic sensors.

As reported in previous studies, using knowledge from multiple kinds of sensor sources is an effective way to enhance the detection accuracy. Suhr and Jung proposed a parking slot detection system that fused camera video and ultrasonic sensors [34]. It is reported that the parking slot marking detection could achieve 96% and the occupancy classification rate could achieve 98%. Prioletti et al. reviewed the literature on pedestrian detection technologies to find out that although computer vision technologies were still the mainstream in pedestrian detection, approaches based on fusion of multiple sensor sources were becoming more attractive to both researchers and engineers, due to the higher detection accuracy and better robustness [35]. If information from cameras and ultrasonic sensors could be fused in future work, it would greatly improve the detection performance in a cost-effective way.

5. Conclusion

This paper proposed a method to detect pedestrians, cyclists, and vehicles using arrayed ultrasonic sensors. Results showed an overall detection accuracy of 86.0%. Vehicles were relatively easier to be detected, whereas vulnerable road users (e.g., pedestrians and cyclists) were difficult to be distinguished because of their similar appearance. The time needed for detection was about 0.8 s which is not very fast but should work for a better understanding on the traffic environment in slow traffic situations (e.g., mixed congestion with cyclists and pedestrians around vehicles on branch roads). Shorter distance and higher data collection frequency would contribute to a faster detection. The effectiveness of ultrasonic sensors on objects detection would improve the performance of environment perception module in ADAS development and autonomous vehicle design by jointly working with camera systems to provide all-around-the-clock assistance in future applications.

Acknowledgement

This study is supported by National Key R&D Program in China with 2017YFB0102603, NSF China with 51575293, 51622504 and 51805332, Natural Science Foundation of Guangdong Province with 2018A030310532, the Young Elite Scientists Sponsorship Program by China Society of Automotive Engineers, and the State Key Laboratory of Automotive Safety and Energy under Project No. KF1801. The authors also would like to thank Long Xin from Tsinghua University for his help on the improvement of this work.

References

[1] J. Strandroth, S. Sternlund, A. Lie, C. Tingvall, M. Rizzi, A. Kullgren, M. Ohlin, R. Fredriksson, Correlation between Euro NCAP pedestrian test results and injury severity in injury crashes with pedestrians and bicyclists in Sweden, Stapp Car Crash J. 58 (2014) 213–231.
[2] G. Asaithambi, V. Kanagaraj, T. Toledo, Driving behaviors: models and challenges for non-lane based mixed traffic, Transport. Dev. Econ. 2 (2016) 2–16.
[3] R.O. Chavez-Garcia, O. Aycard, Multiple sensor fusion and classification for moving object detection and tracking, IEEE Trans. Intell. Transp. Syst. 17 (2016) 525–534.
[4] G. Li, S. Eben Li, B. Cheng, Field operational test of advanced driver assistance systems in typical Chinese road conditions: The influence of driver gender, age and aggression, Int. J. Automotive Technol. 16 (2015) 739–750.
[5] C. Bila, F. Sivrikaya, M.A. Khan, S. Albayrak, Vehicles of the future: A survey of research on safety issues, IEEE Trans. Intell. Transp. Syst. 18 (2017) 1046–1065.
[6] Y. Li, F.-Y. Wang, Vehicle detection based on and–or graph and hybrid image templates for complex urban traffic conditions, Transport. Res. Part C: Emerg. Technol. 51 (2015) 19–28.
[7] S.M. Patole, M. Torlak, D. Wang, M. Ali, Automotive radars: a review of signal processing techniques, IEEE Signal Process Mag. 34 (2017) 22–35.
[8] A. Brunetti, D. Buongiorno, G.F. Trotta, V. Bevilacqua, Computer vision and deep learning techniques for pedestrian detection and tracking: A survey, Neurocomputing 300 (2018) 17–33.
[9] J. Baek, J. Kim, E. Kim, Fast and efficient pedestrian detection via the cascade implementation of an additive Kernel support vector machine, IEEE Trans. Intell. Transp. Syst. 18 (2017) 902–916.
[10] Y. Xu, G. Yu, Y. Wang, X. Wu, Y. Ma, A hybrid vehicle detection method based on Viola-Jones and HOG + SVM from UAV images, Sensors 16 (2016) 1325.
[11] C. Liu, R. Fujishiro, L. Christopher, J. Zheng, Vehicle–bicyclist dynamic position extracted from naturalistic driving videos, IEEE Trans. Intell. Transp. Syst. 18 (2017) 734–742.
[12] M. Braun, S. Krebs, F. Flohr, D. Gavrila, EuroCity persons: A novel benchmark for person detection in traffic scenes, IEEE Trans. Pattern Anal. Machine Intell. (2019) 1–18 (early access).
[13] K. He, X. Zhang, S. Ren, J. Sun, Deep residual learning for image recognition, in: 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), IEEE, Las Vegas, NV, USA, 2016, pp. 770–778.
[14] K. Simonyan, A. Zisserman, Very deep convolutional networks for large-scale image recognition, ArXiv:1409.1556 [Cs]. (2014). (accessed March 6, 2019).
[15] C. Szegedy, Wei Liu, Yangqing Jia, P. Sermanet, S. Reed, D. Anguelov, D. Erhan, V. Vanhoucke, A. Rabinovich, Going deeper with convolutions, in: 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), IEEE, Boston, MA, USA, 2015, pp. 1–9.
[16] W. Ouyang, H. Zhou, H. Li, Q. Li, J. Yan, X. Wang, Jointly learning deep features, deformable parts, occlusion and classification for pedestrian detection, IEEE Trans. Pattern Anal. Mach. Intell. 40 (2018) 1874–1887.
[17] J. Li, X. Liang, S. Shen, T. Xu, J. Feng, S. Yan, Scale-aware fast R-CNN for pedestrian detection, IEEE Trans. Multimedia 20 (2018) 985–996.
[18] J. Shen, G. Li, W. Yan, W. Tao, G. Xu, D. Diao, P. Green, Nighttime driving safety improvement via image enhancement for driver face detection, IEEE Access 6 (2018) 45625–45634.
[19] T.S. Combs, L.S. Sandt, M.P. Clamann, N.C. McDonald, Automated vehicles and pedestrian safety: Exploring the promise and limits of pedestrian detection, Am. J. Prev. Med. 56 (2019) 1–7.
[20] E. Hyun, Y.S. Jin, J.H. Lee, Design and development of automotive blind spot detection radar system based on ROI pre-processing scheme, Int. J. Automot. Technol. 18 (2017) 165–177.
[21] A. Asvadi, C. Premebida, P. Peixoto, U. Nunes, 3D Lidar-based static and moving obstacle detection in driving environments: An approach based on voxels and multi-region ground planes, Rob. Auton. Syst. 83 (2016) 299–311.
[22] K. Li, X. Wang, Y. Xu, J. Wang, Density enhancement-based long-range pedestrian detection using 3-D range data, IEEE Trans. Intell. Transport. Syst. 17 (2016) 1368–1380.
[23] C.-W. Liang, C.-F. Juang, Moving object classification using a combination of static appearance features and spatial and temporal entropy values of optical flows, IEEE Trans. Intell. Transp. Syst. 16 (2015) 3453–3464.
[24] S. Zangenehpour, L.F. Miranda-Moreno, N. Saunier, Automated classification based on video data at intersections with heavy pedestrian and bicycle traffic: Methodology and application, Transport. Res. Part C: Emerg. Technol. 56 (2015) 161–176.
[25] S.E. Li, G. Li, J. Yu, C. Liu, B. Cheng, J. Wang, K. Li, Kalman filter-based tracking of moving objects using linear ultrasonic sensor array for road vehicles, Mech. Syst. Sig. Process. 98 (2018) 173–189.
[26] G. Brown, A. Pocock, M.-J. Zhao, M. Luján, Conditional likelihood maximisation: a unifying framework for information theoretic feature selection, J. Machine Learn. Res. 13 (2012) 27–66.
[27] Y. Liang, M.L. Reyes, J.D. Lee, Real-time detection of driver cognitive distraction using support vector machines, IEEE Trans. Intell. Transport. Syst. 8 (2007) 340–350.
[28] C.-C. Chang, C.-J. Lin, LIBSVM: A library for support vector machines, ACM Trans. Intell. Syst. Technol. 2 (2011) 1–27.
[29] D.H. Krantz, H.C. Kunreuther, Goals and plans in decision making, Judg. Decision Making 2 (2007) 137–168.
[30] M. Beynon, B. Curry, P. Morgan, The Dempster-Shafer theory of evidence: An alternative approach to multicriteria decision modelling, Omega 28 (2000) 37–50.
[31] G. Li, W. Wang, S.E. Li, B. Cheng, P. Green, Effectiveness of flashing brake and hazard systems in avoiding rear-end crashes, Adv. Mech. Eng. 2014 (2014) 1–12.
[32] K. Jiang, F. Yan, H. Zhang, Hydrothermal aging factor estimation for two-cell Diesel-engine SCR systems via a dual time-scale unscented Kalman filter, IEEE Trans. Indust. Electron. (2019), https://doi.org/10.1109/TIE.2019.2896030, Early access.
[33] H. Zhang, J. Wang, Active steering actuator fault detection for an automatically-steered electric ground vehicle, IEEE Trans. Veh. Technol. 66 (2017) 3685–3702.
[34] J.K. Suhr, H.G. Jung, Sensor fusion-based vacant parking slot detection and tracking, IEEE Trans. Intell. Transp. Syst. 15 (2014) 21–36.
[35] A. Prioletti, A. Mogelmose, P. Grisleri, M.M. Trivedi, A. Broggi, T.B. Moeslund, Part-based pedestrian detection and feature-based tracking for driver assistance: Real-time, robust algorithms, and evaluation, IEEE Trans. Intell. Transp. Syst. 14 (2013) 1346–1359.