IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received March 9, 2020, accepted March 22, 2020, date of publication March 27, 2020, date of current version April 15, 2020.

Digital Object Identifier 10.1109/ACCESS.2020.2983872

Lane Determination of Vehicles Based on a Novel Clustering Algorithm for Intelligent Traffic Monitoring

LIN CAO[1,2], TAO WANG[1,2], DONGFENG WANG[1,3], KANGNING DU[1,2], YUNXIAO LIU[1,2], AND CHONG FU[4]

[1]School of Information and Communication Engineering, Beijing Information Science and Technology University, Beijing 100101, China
[2]Key Laboratory of the Ministry of Education for Optoelectronic Measurement Technology and Instrument, Beijing Information Science and Technology University, Beijing 100101, China
[3]Beijing TransMicrowave Technology Company Ltd., Beijing 100080, China
[4]School of Computer Science and Engineering, Northeastern University, Shenyang 110004, China

Corresponding author: Tao Wang (wt860122@163.com)

This work was supported in part by the National Science Foundation of China under Grant 61671069, in part by the Qin Xin Talents Cultivation Program under Grant QXTCP A201902, in part by the Cross Training of High Level Talents Real Training Plan of the Beijing Municipal Commission of Education, and in part by the General Foundation of the Beijing Municipal Commission of Education under Grant KM202011232021.

ABSTRACT In intelligent traffic monitoring, speed measuring millimeter waves (MMW) radar is one of the most commonly used tools for traffic enforcement. In traffic enforcement field, the radar must provide the evidence of each vehicle belongs to which lane. In this paper, we propose a novel kernel line segment adaptive possibilistic c-means clustering algorithm (KLSAPCM) for lane determination of vehicles. Firstly, the raw measurement data is preprocessed using the extracting method of data adjacent lane centerlines. Secondly, according to the improved minimum radius data search method, outliers are removed and the proposed KLSAPCM algorithm is initialized. Finally, the accuracy of lane determination has been improved by the proposed KLSAPCM clustering algorithm based on adaptive kernel line segment that conforms to the shape features of the measurement data in the actual scene. The experiment results for multiple scenes were: the KLSAPCM algorithm is compared with the DBSCAN, the k-means, the FCM, the PCM, the AMPCM, and the APCM algorithms on real measurement datasets, and the results highlight the classification rate of the proposed algorithm. Meanwhile, the proposed algorithm gets a good real-time performance and strong robustness for some sparse moving vehicle scene applications.

INDEX TERMS MMW radar, radar measurements, lane determination, clustering algorithms.

I. INTRODUCTION

In intelligent transportation systems, lane detection is a research hotspot, including lane routes, road boundaries, and vehicles passing areas [1], [2]. In actual traffic situations, some drivers do not strictly control the speed of the vehicle as required, which may result in overspeed. The problem of speeding in traffic is a major problem that threatens the safety of life. At present, the speed limit signs are used to enforce speed limits on different roads. At the intersection of roads or highways, the radar speedometer is installed to monitor the speed of passing vehicles. However, due to technical limitations, some false detections and missed detections may occur during the speed measurement process, and the intelligent lane separation algorithms can greatly improve this effect [3]–[5]. However, in the application of actual scenes, it is often found that these lane separation algorithms are not accurate in judging the lane of illegal vehicles, which triggers the camera of the wrong lane, resulting in the illegal vehicle capture failure and escape from legal sanctions. Therefore, improving the accuracy of vehicle lane judgment has important practical significance for traffic enforcement.

In the past few years, vision-based detection techniques have been widely used due to the low cost of acquiring large amounts of image and video data [6]. In 2008, Felzenszwalb [7] *et al.* proposed the DPM (Deformable Parts Model) algorithm. It uses root filters and component filters

The associate editor coordinating the review of this manuscript and approving it for publication was Chao Tong.

to extract features of the image and uses LSVM (Latent Support Vector Machine) to train the gradient model which is used as a template to match the target. But the vision-based detection techniques may be affected by environmental conditions such as glare and inclement weather. With the development of civilian radar technology [8]–[10], in recent years, radar has gradually been used for lane detection task. However, there are still very little open literature related to this topic. Xu [11] divides the radar data points into several regions and calculates the random density of the region to detect roadsides. Han *et al.* [12] uses the threshold segmentation method and proposes an IPDAF (Integrated Probabilistic Data Association Filters) algorithm to detect and track road edges. The lane division method based on clustering algorithm is also widely used. Stauffer and Grimson [13] proposed the GMM (Gaussian Mixture Model), and the method of clustering using this model is called soft clustering. The method gives the probability that the sample belongs to each class. Hulle *et al.* [14] proposed the SOM (Self-organizing feature Map), which can give the centers of different classes. In 2016, Xenaki *et al.* proposed a novel adaptive possibilistic c-means clustering algorithm (APCM) for removing redundant clusters problem [15]. In APCM, the parameter η, after their initialization, are properly adapted as the iterations of the algorithm. Compared with other related PCM framework algorithms, the adaptive algorithm for parameter η makes the APCM algorithm more flexible in discovering the underlying clustering structure, especially in unusual datasets such as big difference in their variances or even with those consisting of closely located to each other clusters.

FIGURE 1. The multi-target traffic MMW radar.

The data used in this paper is obtained from a multi-target traffic MMW radar of the actual scene, which is a product of Beijing Trans-Microwave Science and Technology Company, Ltd, as shown in Fig.1. The multi-target traffic MMW radar is designed to capture the speed illegal vehicle vehicles and trigger a camera to take a picture. According to the shape features of the measurement data for the lane determine in the actual scene, we propose an enhanced kernel line segment APCM clustering algorithm (KLSAPCM) based on minimum radius data search method to improve the accuracy of lane determination. After receiving the raw measurement data, we first extract the effective monitoring area data based on the lane centerlines. Then, the outliers is removed and the KLSAPCM algorithm is initialized according to the minimum radius data search method. Finally, the proposed KLSAPCM algorithm is used to classify the data and calculate the similarity between the kernel line segment and the lane to determine the degree of membership between the vehicle and the lane. In the KLSAPCM algorithm, the initial kernel segment direction is the lane centerline direction, the segment center is the initial clustering center, and the segment length is adjusted adaptively according to the length of the maximum distance between two points in the cluster. Therefore, the KLSAPCM algorithm can correctly determine which lane each vehicle belongs to without manually measuring the installation position and installation angle of radar. For radar measurement data, the processing flow is shown in Fig.2.

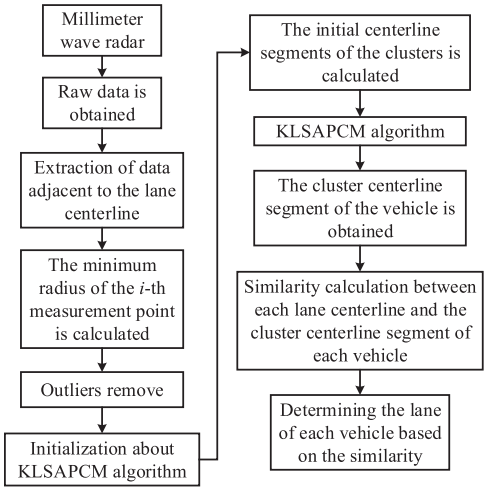

FIGURE 2. Flow chart of measurement data processing.

In summary, the main contributions of the paper are as follows:

1. An extracting monitoring area method (adjacent to the lane centerline) and a minimum radius data search method for the MMW radar measurement data are introduced in traffic detection. These two methods effectively reduce the impact of noise on the clustering algorithm. Meanwhile, the minimum radius data search method can also effectively initialize the KLSAPCM clustering algorithm.

2. A KLSAPCM clustering algorithm is proposed in this paper. The algorithm can correctly determine which lane each vehicle belongs to without manually measuring the installation position and installation angle of radar.

The structure of the paper is as follows. In Section II, we introduce the data acquisition method of multi-target traffic microwave radar. Meanwhile, the k-means, the FCM, the PCM, and APCM clustering algorithms are reviewed. The Section III first preprocesses the measurement data, that is, the extraction of data adjacent lane centerlines. Then, based

on the improved minimum radius data search method, outliers is removed and the proposed KLSAPCM algorithm is initialized. Finally, the principle of KLSAPCM algorithm is described in detail. The Section IV describes the results of several scene experiments, the performance of several algorithms is compared, and the applicability of the algorithm is discussed. Section V summarizes this paper.

II. VEHICLE DATA ACQUIRED BY MMW RADAR AND RELATED ALGORITHM REVIEW

A. THE MMW RADAR IN TRAFFIC MONITORING SCENE

The MMW radar, like the microwave radar, emit electromagnetic waves that are a cone-shaped beam, unlike the laser that are a line. Because the antenna of this band mainly uses electromagnetic radiation as the main method, the large reflective surface makes the millimeter wave radar more reliable, but its resolution is greatly affected. The advantages and disadvantages of millimeter wave radar compared with other kinds of radar are as follows:

· Compared with the centimeter wave radar, the millimeter wave radar has the characteristics of small size, light weight, and high spatial resolution;

· Compared with optical radars such as infrared and laser, the millimeter wave radar has strong ability to penetrate fog, smoke, and dust. Meanwhile, it has the characteristics of long transmission distance and adapting to some extreme weather conditions;

· Stable performance. Not affected by the shape and color of the target.

Therefore, the millimeter wave radar can make up for the shortcomings of sensors such as infrared, laser, ultrasonic, and camera in traffic monitoring applications.

FIGURE 3. The actual scenes. (a) The corresponding camera of the lane. (b) Multi-target millimeter wave radar mounted on the forward direction of three lanes. (c) Multi-target millimeter wave radar mounted on the right side of three lanes. (d) Multi-target millimeter wave radar mounted on the forward direction of three lanes at crossroad.

In this paper, the Beijing Trans-Microwave Science and Technology Company, Ltd. has provided us with a large number of experimental data of actual application scenes. The actual scenes is shown in Fig.3, where figure (a) is the actual scene of the lane cameras, figure (b) is multi-target millimeter wave radar forward installation, figure (c) is multi-target millimeter wave radar side installation, and figure (d) Multi-target millimeter wave radar mounted on the forward direction of three lanes at crossroad.

Besides, the radar measurement period in this paper is 50*ms*. The proposed algorithm usually requires 5 to 15 frames to accumulate data (The frame number can be adjusted based on vehicle speed), and the vehicle driving distance in the monitoring area is about 10 to 30 meters (the higher the speed, the longer the length). The method of accumulating data is: after the radar obtains the measurement data per frame, the system determines whether there are the illegal speeds. If there is illegal speed, the system starts to accumulate data. After the data is accumulated, the proposed algorithm begins to process the data.

B. DEFINE LANE CENTERLINE

After installing the radar, we had to calibrate the number of lanes and the centerline of each lane. The number of lanes can be inputting directly, but the centerline of each lane must be calculated and analyzed by a specific calibrated vehicle equipped with a corner reflector. The acquisition method of the lane centerlines: Firstly, a calibrated vehicle equipped with a radar reflector move slowly along the centerline of each lane at a certain speed. Because the monitoring area is usually not more than 100 meters, the calibration vehicle usually travels from a position 100m away from the radar to the bottom of the radar. Secondly, after the measurement data of the radar calibration vehicle is obtained, the measurement data corresponding to the vehicle speed is extracted and filtered. Finally, the centerline of each lane is calculated by the proposed algorithm in this paper and processed in parallel (the center lines of lanes are parallel to each other). The obtained lane centerlines will be used as the basis for judging the lanes of vehicles. After comparing with the actual manual measurement, the error of the calculated lane centerline based on this method is not more than 10cm. Because the extracted measurement data based on fixed speed is very accurate, the calculated centerlines are very accurate.

Because the accuracy of the lane centerline has a large impact on the performance of the algorithm in this paper, we must obtain an accurate lane centerline. If the calculation of the lane centerline is completed automatically based on the measurement data of vehicles randomly on the lane, the error of the centerline based on this method is much larger than that based on above method. The measurement data of these random vehicles may have the following problems:

(a) Most vehicles may not follow the centerline of the lane;

(b) The road may have an intersection, and most vehicles entering the intersection will deviate to one side of the lane;

(c) Many drivers may have common driving habits that can cause the vehicle's driving trajectory to deviate to one side of the lane.

Therefore, there are many uncertain factors in the measurement data of randomly driven vehicles, which cause

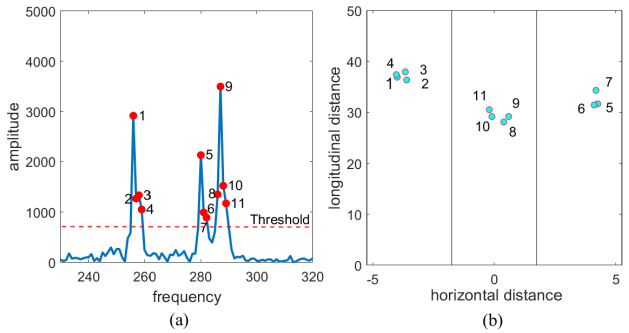

FIGURE 4. Vehicle data transformed from frequency spectrum to scattering points: (a) Frequency spectrum. (b) Vehicles' position.

a large error in the calculated lane centerline. We performed 10 related experiments. The calculated lane centerline error based on this method is not less than 25cm, and the error of the centerline is very unstable.

If the method of training the lane centerline from a large amount of measurement data is used, it takes a long time to accumulate the training data first. Meanwhile, the training data may also have the (b) and (c) problems described above, and the error of the obtained lane centerline is large. In addition, if this method is used in scenarios with low vehicle flow, it will take longer time (hours or days) to accumulate training data, which is much more time consuming than the method in this paper to obtain the centerline.

C. VEHICLE DATA ACQUIRED BY MMW RADAR

The FSK (Frequency-Shift Keying, FSK) radar is used in the paper. The FSK radar alternately transmits two continuous waves of different frequencies at periodic intervals. The Doppler principle is used for velocity measurement, and the phase difference of different carrier frequencies is used to measure distance. According to the principle of the FSK system, it cannot detect the velocity of stationary target. In actual traffic applications, we can use this to shield the information of stationary targets, but also to obtain distance and angle information for stationary targets such as roads, manhole covers, guardrails, and trees.

Raw data containing vehicle information is acquired by the radar in the time domain. In the following, we will introduce how to convert raw data into the vehicle data used in this paper: The spectral amplitude of the vehicle is obtained by using an FFT of the time domain signal. Use the following formula (1) to obtain the distance R between the vehicle and the radar.

$$R = \frac{c \cdot \Delta\phi}{4\pi(f_1 - f_2)} \tag{1}$$

There is a Doppler shift for the radar echo signal of the moving vehicle. The echo signal and the transmitted signal are mixed by a mixer to output the Doppler signal. The frequency (the abscissa of Fig.4(a)) of the Doppler signal depends on the moving velocity, and the amplitude (the ordinate of Fig.4(a)) depends on the distance, target material, and radar cross-section (RCS) of the target. The expression of relative velocity v between the radar antenna and the vehicle is:

$$v = \frac{c \cdot f_d}{2 \cdot f_{Tx} \cdot \cos\alpha} \tag{2}$$

Use the following formula (3) to obtain the angle θ_{angle} between connecting line between target and antenna center and the normal direction of the radar antenna.

$$\theta_{angle} = arcsin(\frac{\lambda \cdot \Delta\phi}{2\pi d}) \tag{3}$$

The radar is designed with ''one transmitting, two receiving'' scheme, and carrier frequencies transmitted alternately are f_1 and f_2, respectively; $\Delta\phi$ is the phase difference between the mixing output of two different echo signals for the same target; f_d is Doppler frequency; f_{Tx} is the center frequency of the radar; α is the angle between the radar beam and the moving direction of the vehicle; λ is the wavelength of the emitted electromagnetic wave; d is the distance between two antennas.

Converting data from a polar coordinate system (R, θ_{angle}) to a Cartesian coordinate system (x, y), the final vehicle data contains three-dimensional information, i.e. velocity v, horizontal distance x, and vertical distance y. The vehicle data for Fig.3(b) is shown in Fig.4, where the figure (a) is the spectrum of the Doppler signal corresponding to the three vehicles for the scene of Fig.3(b), and the 11 red points are considered as the frequency points of the detection data of the vehicles (Firstly, the peak points are searched, such as red dots 1, 5, and 9. Secondly, the points whose amplitude are higher than the threshold are selected at intervals based on the peak points. Finally, the 11 red frequency points are obtained). Based on the 11 frequency points, 11 two-dimensional coordinate measuring points of three vehicles in Fig.3(b) can be calculated, as shown in Fig.4(b). In summary, vehicle data contains many scattering points and each point contains three-dimensional information (x, y, v). In particular, one measurement point does not represent one vehicle (representing a certain point on the vehicle's body), and one vehicle has multiple measurement points. It is difficult to represent the exact position of one vehicle with one measurement point. Therefore, the clustering method in this paper is used to identify all the measurement points of one vehicle and find the center of these points.

D. RELATED ALGORITHM REVIEW

1) THE k-MEANS CLUSTERING ALGORITHM

In this section, the shortcomings of the classical k-means clustering algorithm are first analyzed. Then a solution was proposed: to overcome these shortcomings by using kernel functions.

The k-means algorithm [16] is an efficient and widely-used clustering algorithm that uses iterative ideas to minimize the distortion function. The distortion function J is defined

as follows:

$$J(c, \boldsymbol{\mu}) = \sum_{i=1}^{m} \left\| x^{(i)} - \boldsymbol{\mu}_{c^{(i)}} \right\|^2 \tag{4}$$

where $c^{(i)}$ represents the cluster to which point i-th belongs, and $\boldsymbol{\mu}_j$ is the mean vector of class j corresponding to the center of the cluster.

$$c^{(i)} = \underset{j}{argmin} \left\| x^{(i)} - \boldsymbol{\mu}_j \right\|^2 \tag{5}$$

$$\boldsymbol{\mu}_j = \frac{\sum_{i=1}^{m} \chi(c^{(i)} = j) \cdot x^{(i)}}{\sum_{i=1}^{m} \chi(c^{(i)} = j)} \tag{6}$$

The number of lanes that need to be monitored must be predetermined based on the hardware of the radar. Therefore the parameter k of the k-means algorithm is usually constant and there is no need to bother with it.

However, the k-means algorithm is not suitable for data with a non-spherical distribution because it assumes that the data is subject to a Gaussian distribution.

2) FCM CLUSTERING ALGORITHM

Fuzzy c-means algorithm allocates dataset $\boldsymbol{X} = \{\boldsymbol{x}_1, \cdots, \boldsymbol{x}_i, \cdots, \boldsymbol{x}_n\}$ $(1 \le i \le n)$ into c clusters according to membership degree matrix (fuzzy partition matrix) $U = (\mu_{ij})_{c \times n}$ when the objective function J_{FCM} reaches minimum. The FCM objective function J_{FCM} can be formulated as follows:

$$J_{FCM}(U, \boldsymbol{\theta}) = \sum_{j=1}^{c} \sum_{i=1}^{N} \left(\mu_{ij}^{(l)}\right)^{\alpha} \times d^2\left(\boldsymbol{x}_i, \boldsymbol{\theta}_j^{(l)}\right) \tag{7}$$

where $\alpha \in [1, \infty)$ is the fuzzy weighting exponent for the membership. l is number of iterations. $\boldsymbol{\theta}_j$ is center of the cluster j. μ_{ij} represents the membership degree measures how much the sample $\boldsymbol{x}_i$ belongs to the cluster center $\boldsymbol{\theta}_j$. Here, the c clusters are marked by cluster centers $\boldsymbol{\theta} = \left\{\boldsymbol{\theta}_1, \cdots, \boldsymbol{\theta}_j, \cdots, \boldsymbol{\theta}_c\right\}$ $(1 \le j \le c)$, $\boldsymbol{\theta}$ was randomly selected in the first iteration. Then the membership degree μ_{ij} is calculated as follows:

$$\mu_{ij}^{(l)} \in [0, 1], \quad \sum_{j=1}^{c} \mu_{ij}^{(l)} = 1, \ and \ 0 < \sum_{i=1}^{N} \mu_{ij}^{(l)} < N \tag{8}$$

$$\mu_{ij}^{(l+1)} = \begin{cases} \left(\sum_{h=1}^{c} \left(\dfrac{d^2(\boldsymbol{x}_i, \boldsymbol{\theta}_j^{(l)})}{d^2(\boldsymbol{x}_i, \boldsymbol{\theta}_h^{(l)})} \right)^{(1/(\alpha-1))} \right)^{-1}, \\ \qquad if \ d^2(\boldsymbol{x}_i, \boldsymbol{\theta}_j^{(l)}) > 0 \\ 1, \qquad if \ d^2(\boldsymbol{x}_i, \boldsymbol{\theta}_j^{(l)}) = 0 \end{cases} \tag{9}$$

$$\boldsymbol{\theta}_j^{(l+1)} = \frac{\sum_{i=1}^{n} \left(\mu_{ij}^{(l)}\right)^{\alpha} \boldsymbol{x}_i}{\sum_{i=1}^{n} \left(\mu_{ij}^{(l)}\right)^{\alpha}} \tag{10}$$

the objective function is minimized by continuously updating the membership functions and centers of clusters until $\left\| U^{(l+1)} - U^{(l)} \right\| < \delta$ is satisfied. Similarly, the FCM algorithm is not suitable for data with a non-spherical distribution.

3) PCM CLUSTERING ALGORITHM

The FCM algorithm calculates the membership degree of each measurement point for every cluster, which gives us a calculation method that refers to the reliability of the measurement point classification results. If the membership degree of a measurement point for a certain cluster has an absolute advantage in all membership degrees, it is a very reliable method to assign the measurement point to the cluster. If there are some measurement point with relative average membership for every cluster, we need other methods to process. Besides, the FCM algorithm must give the number of clusters, and cannot adaptively identify the number of clusters in the dataset. Meanwhile, when there is noise in the measurement dataset and the density difference of measurement point between the clusters is large, the classification results may be inaccurate. In order to deal with these disadvantages, Krishnapuram proposed the PCM algorithm based on the FCM algorithm [17], and objective function J_{PCM} added a constraint term to J_{FCM}:

$$J_{PCM}(U, \boldsymbol{\theta}) = \sum_{j=1}^{c} \sum_{i=1}^{n} \left(\mu_{ij}^{(l)}\right)^{\alpha} d^2\left(\boldsymbol{x}_i, \boldsymbol{\theta}_j^{(l)}\right) + \sum_{j=1}^{c} \eta_j \sum_{i=1}^{n} \left(1 - \mu_{ij}^{(l)}\right)^{\alpha} \tag{11}$$

where η_j is a scale parameter, and each one associated with a cluster. More specifically, each η remains unchanged during algorithm execution. The greater the η value, the greater the influence of cluster around $\boldsymbol{\theta}$; on the contrary, the smaller the η value, the smaller the influence. The value of η_j is calculated after running the FCM algorithm:

$$\eta_j = B \frac{\sum_{i=1}^{n} \mu_{ij}^{FCM} d^2\left(\boldsymbol{x}_i, \boldsymbol{\theta}_j^{FCM}\right)}{\sum_{i=1}^{n} \mu_{ij}^{FCM}}, \quad j = 1, \cdots, c \tag{12}$$

where the constant B is usually equal to 1. The update of the membership degree μ_{ij} and cluster center $\boldsymbol{\theta}_j$ is as follows:

$$\mu_{ij}^{(l+1)} = \frac{1}{1 + \left(d^2\left(\boldsymbol{x}_i, \boldsymbol{\theta}_j^{(l)}\right)/\eta_j\right)^{1/\alpha - 1}} \tag{13}$$

$$\boldsymbol{\theta}_j^{(l+1)} = \frac{\sum_{i=1}^{n} \left(\mu_{ij}^{(l)}\right)^{\alpha} \boldsymbol{x}_i}{\sum_{i=1}^{n} \left(\mu_{ij}^{(l)}\right)^{\alpha}} \tag{14}$$

According to the iteration between the two equations, the PCM algorithm gives the update estimates of μ_{ij} and $\boldsymbol{\theta}_j$ at each iteration until the set termination condition is met. From the above description, the PCM algorithm can be summarized as follows:

In [17], Krishnapuram and Keller proposed the second PCM clustering algorithm (PCM_2) in 1996, and objective function J_{PCM_2} is as follows:

$$J_{PCM_2}(U, \boldsymbol{\theta}) = \sum_{j=1}^{c} \sum_{i=1}^{n} \mu_{ij}^{(l)} d^2\left(\boldsymbol{x}_i, \boldsymbol{\theta}_j^{(l)}\right) + \sum_{j=1}^{c} \eta_j \sum_{i=1}^{n} \left(\mu_{ij}^{(l)} \ln \mu_{ij}^{(l)} - \mu_{ij}^{(l)}\right) \tag{15}$$

Algorithm 1 PCM

Require: $\boldsymbol{x}_i$, c, and α.

Ensure: fuzzy partition matrix U, clustering center $\boldsymbol{\theta}$, and scale parameter $\Gamma = \{\eta_1, \cdots, \eta_c\}$.

1: **initialization**: $\boldsymbol{\theta}_j$ from the FCM algorithm, and using (12) initialize Γ;
2: **repeat**
3: using (13) update the membership degree matrix $U^{(l)}$;
4: using (14) update cluster center $\boldsymbol{\theta}^{(l)}$;
5: $l = l + 1$;
6: **until** the difference between $\boldsymbol{\theta}^{(l)}$ and $\boldsymbol{\theta}^{(l+1)}$ is sufficiently small;
7: **return**: results U, $\boldsymbol{\theta}$, and Γ;

Using 15, the small values of the memberships are be penalized based on the last term. Setting to zero the derivatives of $J_{PCM_2}(U, \boldsymbol{\theta})$ with respect to the memberships $\mu_{ij}^{(l)}$:

$$\frac{\partial J_{PCM_2}(U, \boldsymbol{\theta})}{\partial \mu_{ij}^{(l)}} = d^2\left(\boldsymbol{x}_i, \boldsymbol{\theta}_j^{(l)}\right) + \eta_j \ln \mu_{ij}^{(l)} = 0 \quad (16)$$

Then, $\mu_{ij}^{(l)}$ is obtained:

$$\mu_{ij}^{(l)} = \exp\left(-\frac{d^2\left(\boldsymbol{x}_i, \boldsymbol{\theta}_j^{(l)}\right)}{\eta_j}\right) \quad (17)$$

where, the calculation methods of η and $\boldsymbol{\theta}$ are unchanged.

After many years of development, some new classification algorithms based on the PCM algorithm have been proposed. The PCM algorithm described above has no cluster elimination capability, that is, if the number of clusters is overestimated during initialization, they cannot eliminate any clusters in the iteration process. A adaptive possibilistic c-means (APCM) algorithm is proposed for this disadvantage [15]. More specifically, in APCM, the parameter η will be adjusted with the evolution of the algorithm after initialization. Compared with other PCM algorithms, the adaptability of η makes the algorithm more flexible to reveal the underlying clustering structure, especially in dense datasets such as the clusters with large differences in variance or contains clusters that are close to each other. The parameter η of the APCM algorithm is as follows:

$$\eta_j = \frac{\hat{\gamma}}{a}\gamma_j \quad (18)$$

where the parameter γ_j is a measure of the average absolute deviation of cluster c_j, a is a custom positive parameter, and $\hat{\gamma}$ is a constant defined as the minimum of all initial γ_j, the expression is as follows:

$$\hat{\gamma} = \min_j \frac{\sum_{i=1}^{n} \mu_{ij}^{FCM} d^2\left(\boldsymbol{x}_i, \boldsymbol{\theta}_j^{FCM}\right)}{\sum_{i=1}^{n} \mu_{ij}^{FCM}}, \quad j = 1, \cdots, c_{ini} \quad (19)$$

$$\gamma_j^{(l+1)} = \frac{\sum_{\boldsymbol{x}_i : \mu_{ij}^{(l)} = \max_{r=1,\cdots,c^{(l+1)}} \mu_{ir}^{(l)}} d^2(\boldsymbol{x}_i, \boldsymbol{\theta}_j^{(l)})}{n_j^{(l)}}$$
$$j = 1, \cdots, c^{(l+1)} \quad (20)$$

where c_{ini} is the number of clusters at the initial iteration. The objective function J_{APCM} is as follows:

$$J_{APCM}(U, \boldsymbol{\theta}) = \sum_{j=1}^{c}\sum_{i=1}^{n} \mu_{ij}^{(l)} d^2\left(\boldsymbol{x}_i, \boldsymbol{\theta}_j^{(l)}\right) + \frac{\hat{\gamma}}{a}\sum_{j=1}^{c}\gamma_j \sum_{i=1}^{n}\left(\mu_{ij}^{(l)} \ln \mu_{ij}^{(l)} - \mu_{ij}^{(l)}\right) \quad (21)$$

Then, $\mu_{ij}^{(l)}$ is obtained:

$$\mu_{ij}^{(l)} = \exp\left(-\frac{a}{\hat{\gamma}}\frac{d^2\left(\boldsymbol{x}_i, \boldsymbol{\theta}_j^{(l)}\right)}{\gamma_j^{(l)}}\right), \quad j = 1, \cdots, c^{(l)} \quad (22)$$

From the above description, the APCM algorithm can be summarized as Algorithm 2.

Algorithm 2 APCM

Require: $\boldsymbol{x}_i$, c, and a.

Ensure: fuzzy partition matrix U, clustering center $\boldsymbol{\theta}$, and *label*.

1: **initialization**: $\boldsymbol{\theta}_j$ from the FCM algorithm, using (12) initialize γ_j, and set: $\hat{\gamma} = \min_{j=1,\cdots,c_{ini}} \gamma_j^{(l)}$;
2: $c^{(l)} = c_{ini}$;
3: **repeat**
4: using (22) update the membership degree matrix $U^{(l)}$;
5: using (14) update cluster center $\boldsymbol{\theta}^{(l)}$;
6: **for** $i \leftarrow 1$ *to* n *do*
7: $\mu_{ir}^{(l)} = \max_{j=1,\cdots,c^{(l)}} \mu_{ij}^{(l)}$
8: $label(i) = r$
9: **end for**
10: $p = 0$ *//number of removed clusters*;
11: **for** $j \leftarrow 1$ *to* c *do*
12: **if** $j \notin label$ then
13: **Remove** c_j
14: $p = p + 1$
15: **end if**
16: **end for**
17: $c^{(l+1)} = c^{(l)} - p$
18: using (20) update $\gamma_j^{(l+1)}$;
19: $l = l + 1$;
20: **until** the difference between $\boldsymbol{\theta}^{(l)}$ and $\boldsymbol{\theta}^{(l+1)}$ is sufficiently small;
21: **return**: results U, $\boldsymbol{\theta}$, and *label*;

III. KERNEL LINE SEGMENT ADAPTIVE POSSIBILISTIC C-MEANS CLUSTERING ALGORITHM

This section first preprocesses the measurement data, that is, the extraction of data adjacent lane centerlines. Secondly, based on the improved minimum radius data search method, outliers are removed and the proposed KLSAPCM algorithm

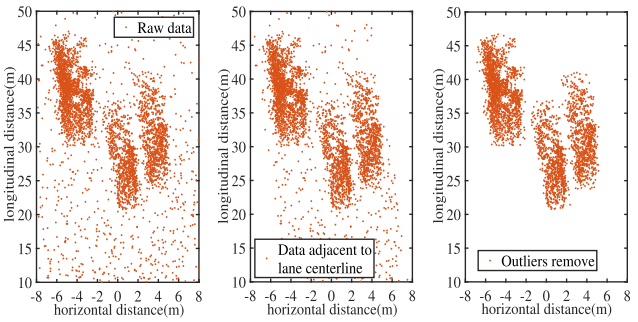

FIGURE 5. The measurement data processing in actual scene of the Fig.3(b). (a) Raw measurement data. (b) Data Adjacent to the Lane Centerline. (c) Outliers Remove.

is initialized. Finally, the principle of KLSAPCM algorithm is described in detail.

A. EXTRACTION OF DATA ADJACENT TO THE LANE CENTERLINE

The obtaining method of the lane centerline for each scene and the raw measurement data from the MMW radar are described above. For example, the raw measurement data in actual scene of the Fig.3(b) is shown in Fig.5(a). Based on the lane centerline, we can effectively extract the interesting measurement data of the adjacent lane. When the raw data from the traffic radar measurement is obtained, we first need to extract the measurement data from the centerline of the adjacent lane. Firstly, we calculate the vertical distance from each data point to the centerline of each lane, extract the minimum distance, and then remove the measurement data where the minimum distance is greater than the set threshold d'. Because the lane width is generally about $3.5m$, and the error of the radar measurement is less than $0.2m$, we generally set it to $1.95m$. Finally, the measurement data of the traffic surveillance area was obtained.

B. OUTLIERS REMOVE AND INITIALIZATION BASED ON SAMPLE DENSITY FEATURE

After the extraction of data adjacent lane centerlines, the vehicle data can be visualized as follows. As shown in Fig.5(b), there is still some noise in the data. In the following, the coordinate information of the vehicle data will be used to eliminate the influence of outliers.

Let us represent dataset $\boldsymbol{X} = \{\boldsymbol{x}_i = (x_i, y_i), i = 1, \cdots, M\}$ containing all the coordinate information of the processed data. Where $\boldsymbol{x}_i$ is a two-dimensional vector and M is the total number of scattering points in this dataset. Since dense points are more likely to be produced by the vehicle, the density characteristics are calculated. And dense points are surrounded by outliers with low local density.

The minimum radius τ_i [18] of the i-th measurement point is defined as follows:

$$\tau_i = \min_{\rho_i > C}(d_{ij}) \tag{23}$$

where d_{ij} is the Euclidean distance between $\boldsymbol{x}_i$ and $\boldsymbol{x}_j$. The local density ρ_i can be interpreted as the number of points closer to the point i in the neighborhood. When ρ_i is greater than constant C, minimum neighborhood radius is treated as the minimum radius τ_i of the point i. Fig.5(c) shows the result of simulated dataset (the radius has been normalized to give a better intuition). In practice, any measurement point greater than the $\beta \times \bar{\tau}$ is considered an outlier, where β is a proportional coefficient, and $\bar{\tau}$ the average of the minimum radii of all points.

Next, we extract the previous data points with the minimum radius value less than the threshold d_r, and merge the neighboring extraction points into one data point (that is, if the mutual distance between the extraction points is less than the threshold d_{r1}, we will merge these points. The position of the merged point is the average of these extracted points in the neighborhood). Then, the remaining extraction and merged points are considered to be the initial clustering center $\boldsymbol{\theta}$.

Finally, the initialization of $\boldsymbol{\theta}_j$ is carried out using the final cluster representatives obtained from the above method. Taking into account that the above method is very likely to drive the representatives to dense in data regions (since $c_{ini} > c$), the probability of at least one of the initial $\boldsymbol{\theta}_j$ to be placed in each dense region (cluster) of the dataset increases with c_{ini}. Besides, we combine the longest distance between the two elements in the cluster with the shape of the measuring vehicle to determine whether there are missing clusters. This provides better initialization results for APCM. After the initialization of $\boldsymbol{\theta}_j$, we can calculate the corresponding μ_{ij} value based on (22). Then, the initialization of γ_j is as follows:

$$\gamma_j = \frac{\sum_{i=1}^{n} \mu_{ij}^{ini} d\left(\boldsymbol{x}_i, \boldsymbol{\theta}_j^{ini}\right)}{\sum_{i=1}^{n} \mu_{ij}^{ini}}, \quad j = 1, \cdots, c_{ini} \tag{24}$$

where $\boldsymbol{\theta}_j^{ini}$ s and μ_{ij}^{ini} in (24) are calculated.

C. KLSAPCM ALGORITHM

As well known, all clusters obtained from moving vehicles for several frames will have a large number of clusters that are elongated in shape. However, the APCM algorithm has a good classification effect for clustering with circular shape, and has poor classification performance for slender clustering. If the length of clustering shape is too large, the APCM algorithm may be divided into two or more clusters. In this part, we design a KLSAPCM algorithm based on adaptive line segment kernel according to the characteristics of the measurement dataset. After the measurement data is processed by the above method, the initial clustering result is obtained. The adaptive line segment along lane centerline for each clustering data are used to update the kernel function. In the proposed KLSAPCM, the Kernel function K is improved to accommodate the classification of traffic radar measurement data. The Kernel function K is as follows:

$$K\left(\boldsymbol{x}_i, L_j^{(l+1)}\right) = \min \hat{d}^2\left(\boldsymbol{x}_i, L_j^{(l+1)}\right) \tag{25}$$

where

$$L_j^{(l+1)} = \begin{bmatrix} \hat{x}_i^{(l+1)} \\ \hat{y}_i^{(l+1)} = a_1 \hat{x}_i^{(l+1)} + a_2 \end{bmatrix},$$
$$y_i^{(l+1)} - \alpha\delta_i^{(l+1)} \le \hat{y}_i^{(l+1)} \le y_i^{(l+1)} + \alpha\delta_i^{(l+1)} \quad (26)$$

and

$$x_j^{(l+1)} = \frac{\sum_{i=1}^{n} \left(\mu_{ij}^{(l)}\right)^{\alpha} x_i}{\sum_{i=1}^{n} \left(\mu_{ij}^{(l)}\right)^{\alpha}}, \quad y_j^{(l+1)} = \frac{\sum_{i=1}^{n} \left(\mu_{ij}^{(l)}\right)^{\alpha} y_i}{\sum_{i=1}^{n} \left(\mu_{ij}^{(l)}\right)^{\alpha}} \quad (27)$$

where $L_j^{(l+1)}$ is the center line segment of the cluster j, $\hat{d}$ is the distance from the measurement point to the point on the center line segment. $\hat{x}_i^{(l+1)}$ and $\hat{y}_i^{(l+1)}$ are the abscissa and ordinate value of the center line segment, respectively. $x_j^{(l+1)}$ and $y_j^{(l+1)}$ are the abscissa and ordinate value of the center point of the center line segment for the cluster j, respectively. a_1 and a_2 are the coefficients of center line segment expression, whose values can be obtained by the lane centerline. $\delta_j^{(l+1)}$ is the length of the cluster j along Y axis, α is the coefficient that determines the length of the cluster center line segment. $\hat{\gamma}$ is a constant defined as the minimum of all initial γ_j, the expression is as follows:

$$\hat{\gamma} = \min_j \frac{\sum_{i=1}^{n} \mu_{ij}^{ini} d^2\left(\boldsymbol{x}_i, \boldsymbol{\theta}_j^{ini}\right)}{\sum_{i=1}^{n} \mu_{ij}^{ini}}, \quad j = 1, \cdots, c_{ini} \quad (28)$$

$$\gamma_j^{(l+1)} = \frac{\sum_{\boldsymbol{x}_i : \mu_{ij}^{(l)} = \max_{r=1,\cdots,c^{(l+1)}} \mu_{ir}^{(l)}} K(\boldsymbol{x}_i, L_j^{(l)})}{n_j^{(l)}}$$
$$j = 1, \cdots, c^{(l+1)} \quad (29)$$

The objective function J_{APCM} is as follows:

$$J_{APCM}(U, L) = \sum_{j=1}^{c} \sum_{i=1}^{n} \mu_{ij}^{(l)} K\left(\boldsymbol{x}_i, L_j^{(l)}\right) + \frac{\hat{\gamma}}{a} \sum_{j=1}^{c} \gamma_j \sum_{i=1}^{n} \left(\mu_{ij}^{(l)} \ln \mu_{ij}^{(l)} - \mu_{ij}^{(l)}\right) \quad (30)$$

Then, $\mu_{ij}^{(l)}$ is obtained:

$$\mu_{ij}^{(l)} = \exp\left(-\frac{a}{\hat{\gamma}} \frac{K\left(\boldsymbol{x}_i, L_j^{(l)}\right)}{\gamma_j^{(l)}}\right), \quad j = 1, \cdots, c^{(l)} \quad (31)$$

From the above description, the proposed KLSAPCM algorithm can be summarized as Algorithm 3.

When the iteration is ended, the best individual in the current generation is the global optimum solution. Partitioning dataset by the best cluster line segments, we can get the classification results finally.

Algorithm 3 KLSAPCM

Require: γ_j, c_{ini}, α, and a.
Ensure: fuzzy partition matrix U, clustering center $\boldsymbol{\theta}$, and *label*.

```
initialization: θ_j from the initialization method, using (24) initialize γ_j, and set: γ̂ = min_{j=1,···,c_ini} γ_j^(l);
c^(l) = c_ini;
repeat
calculating the δ_j^(l+1) of each cluster;
using (27) update the cluster centers θ^(l);
using (26) update the center line segment L^(l) of each cluster;
using (25) update the Kernel function K^(l);
using (31) update the membership degree matrix U^(l);
  for i ← 1 to n do
    μ_ir^(l) = max_{j=1,···,c^(l)} μ_ij^(l)
    label(i) = r
  end for
p = 0 //number of removed clusters;
  for j ← 1 to c do
    if j ∉ label then
      Remove c_j
      p = p + 1
    end if
  end for
  c^(l+1) = c^(l) − p
using (29) update γ_j^(l+1);
l = l + 1;
until the difference between L^(l) and L^(l+1) is sufficiently small;
return: results U, θ, and label;
```

IV. EXPERIMENT RESULTS

In this section, the proposed clustering method is compared with the DBSCAN, the k-means [19]–[21], the FCM [22], the PCM [5], the AMPCM [20], [23], and the APCM in a typical scene. Then, the experimental results for the scenes in the Fig.3(b) show that the proposed clustering method is superior than the other several clustering algorithms. Finally, the proposed clustering method had better robustness against some special scenes.

A. COMPARISON WITH EXPERIMENTAL RESULTS OF CLUSTERING ALGORITHMS

This section gives experimental comparisons of the DBSCAN, the k-means, the FCM, the PCM, the AMPCM, the APCM, and the KLSAPCM algorithms in the scene of Fig.3(b). In the scene of Fig.3(b), highway speed limit: the speed range of miniature vehicle in the driving lane is $60 \sim 100km/h(16.67 \sim 27.78m/s)$, and the speed range of big vehicle in the driving lane is $60 \sim 80km/h(16.67 \sim 22.22m/s)$.

The scene 1 description: Two trucks are running on three lanes. The length and width of the truck in the middle lane are about 8 m and 2.4 m respectively (From the type of truck in video surveillance). The velocity is about $25m/s$. The length and width of the truck in the left lane are about $11m$ and 2.4 m respectively. The velocity is about 24 m/s. The two trucks were moving at closing speed in the adjacent lane.

(a) (b)

FIGURE 6. The actual scenes. (a) and (b) The scene 1 and 2 in the Fig.3(b).

The scene 2 description: One microbus and one car are running on three lanes. The length and width of the microbus in the middle lane are about 4.5m and 1.8m respectively. The velocity is about 34 m/s. The length and width of the car in the right lane are about 4.2m and 1.8m respectively. The velocity is about 33 m/s. The two vehicles were moving at closing speed and very close together in the adjacent lane. Meanwhile, the microbus was suspected of dangerous driving.

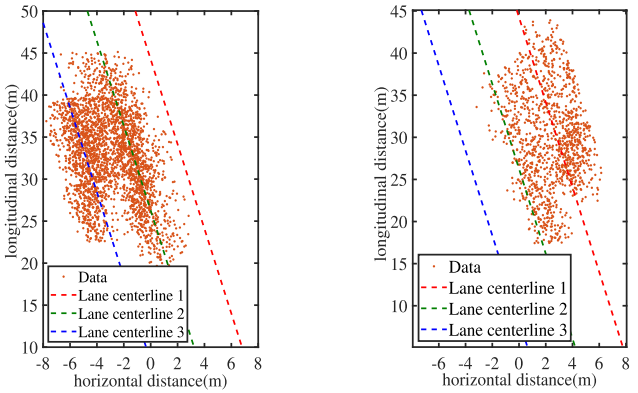

FIGURE 7. The measurement data after removing the outliers. (a) and (b) The scene 1 and 2 in the Fig.6.

Firstly, we extracted measurements in the adjacent lane centerline and filtered out the outliers from the raw data to obtain the effective measurement data of the vehicles, as shown in Fig.7. By observing the measurement data after removing the outliers, it is known that the measurement of the two vehicles is difficult to distinguish about the 2 scenes. Here, we find that when the outliers removed from raw data, the different cluster centers are very clear in the Fig. 5, but the vehicle data are mixed together and hard to decide its cluster center in this section. The main reasons are: in the real scene Fig. 3(b) of the Fig. 5, the distance between the three vehicles is large, and there is no mirror interference between each other during radar measurement. However, in this Section, we deliberately selected two scenarios where the vehicles are very close to each other to show the performance of the algorithm. We observe the two scenes of the Fig. 6, because the distance between the vehicles is close, there will be a lot of mirror interference between the vehicles, making the measurement data of the two vehicles mixed together and difficult to distinguish. Aiming at this problem, the traditional clustering methods cannot distinguish the two measurement datasets correctly. Therefore, this paper proposed the KLSAPCM algorithm to deal with the problem.

Next, the noise-filtered data is processed by the DBSCAN, the k-means, the FCM, the PCM, the AMPCM, and the APCM classification algorithms, and the results are shown in Fig. 8. It can be seen that the DBSCAN, the k-means, the FCM, the PCM, the AMPCM, and the APCM algorithms are not suitable for data with non-spherical distribution. When the DBSCAN algorithm is executed, the algorithm will also filter out low-density measurement data, as shown by Outliers in Fig.8 (a) and (g). When the PCM algorithm is executed, because the PCM algorithm can greatly eliminate the influence of low-density data (outliers) on the clustering result, we first remove the low-density data (outliers) before executing the PCM clustering algorithm, as shown by Outliers in Fig.8 (d) and (j). The clustering results of the PCM algorithm obtained by this method are very close to those obtained by directly executing the PCM algorithm, and the computational efficiency of the algorithm is improved. The Fig.9 shows the clustering results of the proposed algorithm, and the results are obviously better than other algorithms.

TABLE 1. CR and running time of clustering algorithms for the scene 1 and 2.

Algorithms	Scene 1		Scene 2	
	CR (%)	Time (s)	CR (%)	Time (s)
DBSCAN	52.36	0.098	51.33	0.082
k-means	51.94	0.071	50.17	0.070
FCM	50.89	0.092	54.63	0.085
PCM	50.29	0.131	59.91	0.126
AMPCM	52.36	10.399	51.33	8.015
APCM	56.69	0.162	62.53	0.145
KLSAPCM	96.52	0.119	95.91	0.104

In the following, the proposed classification algorithm is compared with the DBSCAN, the k-means, the FCM, the PCM, the AMPCM, and the APCM algorithms on real measurement datasets from the scenes of Fig.6 with respect to the classification rate (CR) index and the time consumption. Table 1 and Fig. 10 shows the CR values obtained by the DBSCAN, the k-means, the FCM, the PCM, the AMPCM, the APCM, and the KLSAPCM algorithms in the two scenes. As it can be deduced from the Table 1 and the Fig. 10, the KLSAPCM algorithm has optimal CR index. It is worth noting that the time consumption of the KLSAPCM algorithm is more than that of the DBSCAN, the k-means, and the FCM algorithms, but less than the PCM, the AMPCM, and the APCM algorithms. Because the KLSAPCM algorithm is first initialized through the improved minimum radius data search method, and then the KLSAPCM algorithm is run. In terms of time complexity, it is necessary to calculate the minimum radius of each data point and find the data point with the smallest minimum radius in the neighborhood.

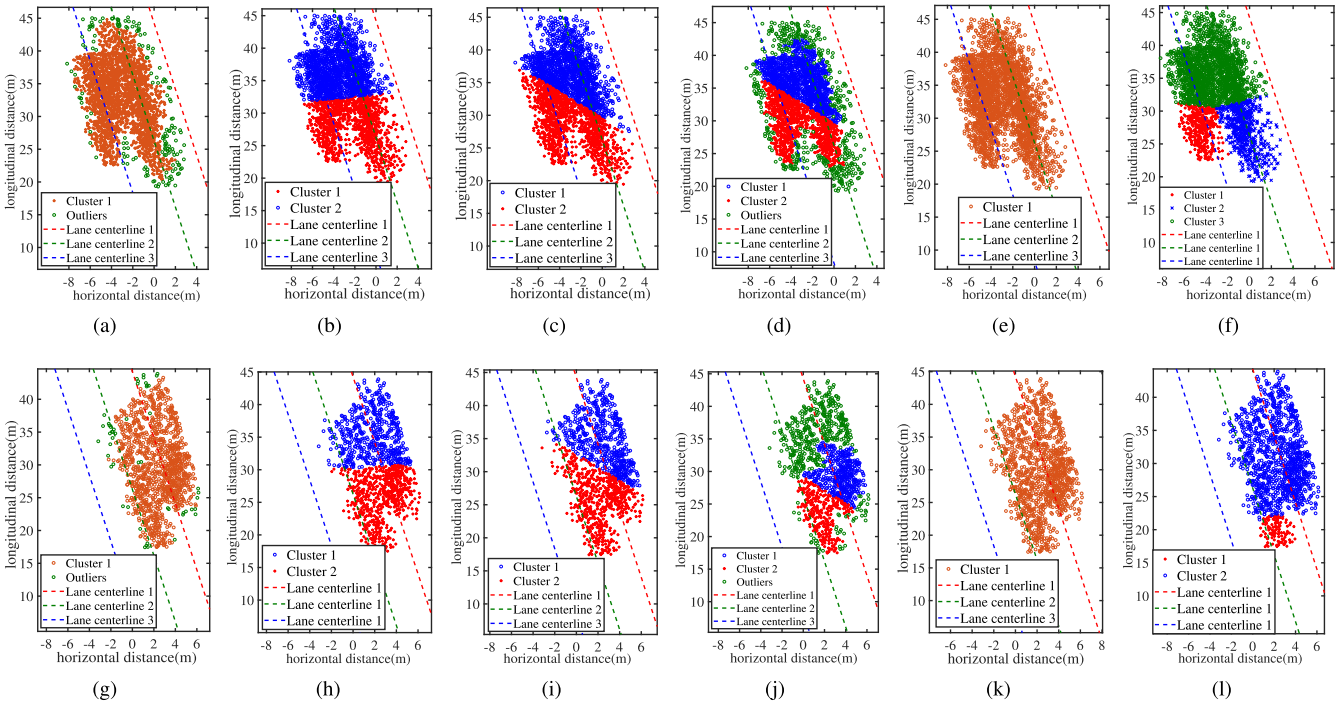

FIGURE 8. The clustering results. (a)∼(f) The results of the DBSCAN, the *k*-means, the FCM, the PCM, the AMPCM, and the APCM classification algorithms for the scene 1. (g)∼(l) The results of the DBSCAN, the *k*-means, the FCM, the PCM, the AMPCM, and the APCM classification algorithms for the scene 2.

But proposed algorithm can remove many outliers by the minimum radius data search method, and reduce the number of processing points of the subsequent algorithm. In addition, the lane centerline is known. When the KLSAPCM clustering algorithm is run, the initial cluster centerline segments calculated based on the lane centerline are already close to the true cluster center, thereby greatly reducing the number of iterations of the algorithm. Therefore, the proposed algorithm improves the clustering accuracy, the time consumption is within the acceptable range, and is less than some new clustering algorithms proposed in recent years. The time consumption comparison is shown in Fig. 11.

In summary, while ensuring real-time application, the classification performance of the proposed algorithm is better than that of the DBSCAN, the *k*-means, the FCM, the PCM, the AMPCM, and the APCM classification algorithms in the lane determination of the vehicle.

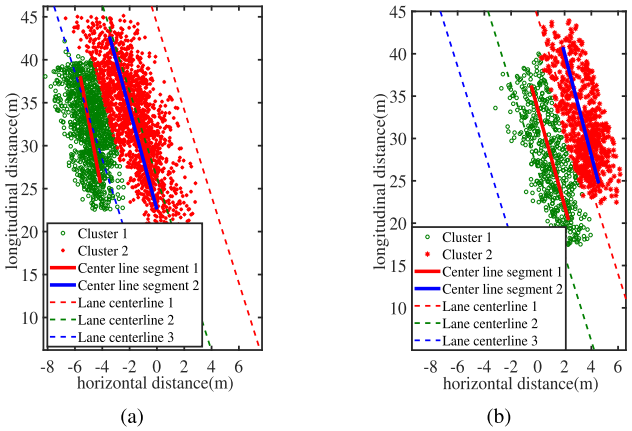

FIGURE 9. The clustering results of the proposed algorithm. (a) and (b) The scene 1 and 2.

B. COMPARISON OF EXPERIMENTAL RESULTS FOR NEAREST LANE CENTERLINE METHOD

As we all know, the higher the accuracy of capturing illegal vehicles on the expressway (the lower the rate of missing and wrong capturing, that is, the more accurate the law enforcement), the more it will help reduce the incidence of traffic accidents. Besides, the Section II-C of this paper introduces the data acquisition of the MMW radar. Based on the Section, we can know that one measurement point does not represent one vehicle (representing a certain point on the vehicle's body, one vehicle has multiple measurement points) and vehicle measurement data of different densities can be obtained by adjusting relevant thresholds. If we use the lane centerline and the vehicle position to determine the lane of a vehicle for the scenes of the Fig.12, law enforcement errors may occur. Next, the distance between the position of the vehicle and the centerline of the lane (nearest lane centerline method) is used to determine the lane of the vehicle, the measurement data processing steps and possible problems are as follows:

a) When the measurement data is sparse (there are only a few or one measurement points per vehicle, and even some vehicles do not have measurement points). Although

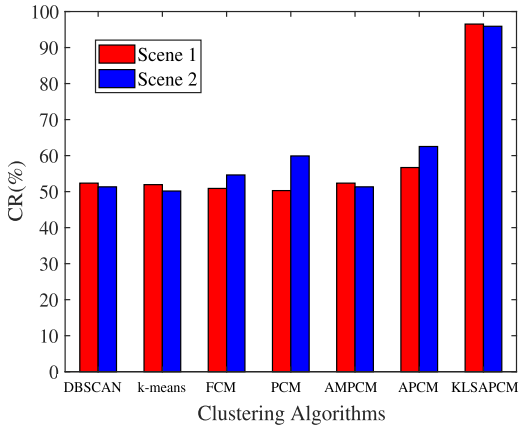

FIGURE 10. The CR of the different clustering algorithms for the scene 1 and 2.

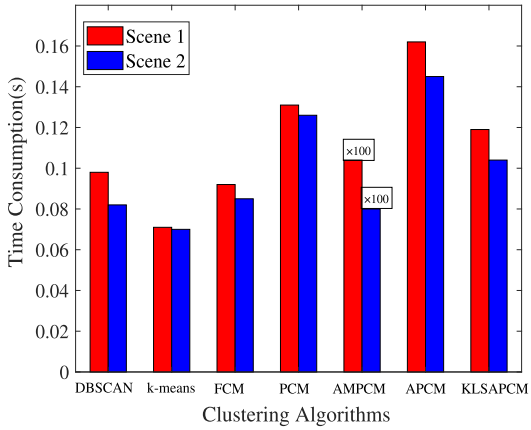

FIGURE 11. The time consumption of the different clustering algorithms for the scene 1 and 2.

FIGURE 12. The actual scenes. (a) and (b) The scene 3 and 4 in the Fig.3(b).

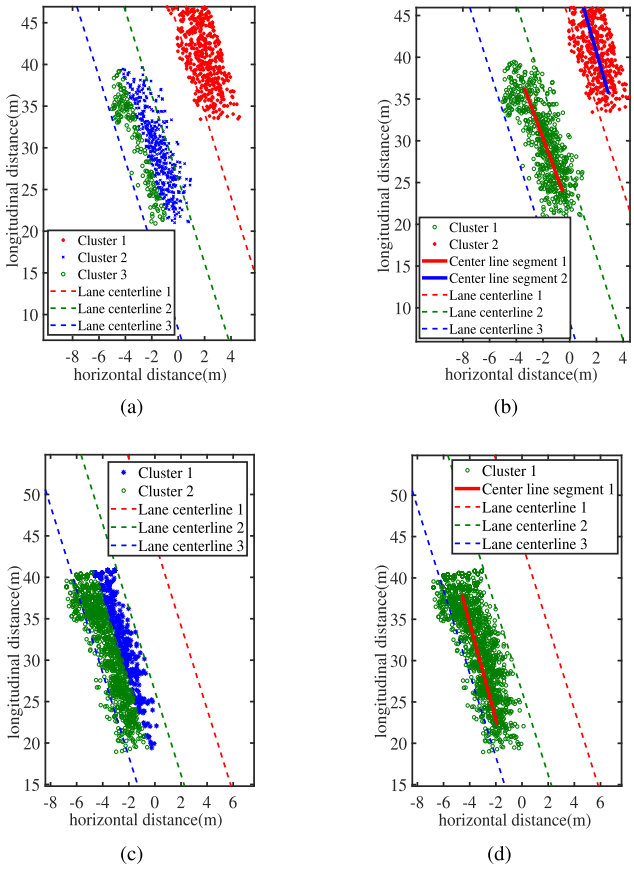

FIGURE 13. The experimental results. (a) The nearest lane centerline method for the scene 3. (b) the proposed method for the scene 3. (c) The nearest lane centerline method for the scene 4. (d) the proposed method for the scene 4.

we can easily obtain the number of vehicles in all lanes, these measured points are from the position where the radar wave reflects strongly on the vehicle, that is, it is likely to be on one side of the vehicle, such as the vehicles in Fig.12 (a) and (b). The measurement points of the vehicle for several consecutive cycles may be in the wrong lane, and the position of the vehicle may be inaccurate. Therefore, the position of the vehicle may be in the wrong lane, and finally causing the vehicle's lane to be determined incorrectly. In addition, some vehicles will be missed detection in this mode, that is, there is no measurement in several cycles and some illegal vehicles may be missed.

b) When the measurement data is dense (every vehicle has many measurement points). 1) We must first calculate the number of vehicles included in the measurement data, and then calculate the position of each vehicle. The clustering algorithm can well calculate the number of vehicles, then the related algorithm is used for accurate positioning, and finally the distance between the position of the vehicle and the lane centerline is calculated to define the lane of the vehicle. Its computational cost is greater than the computational cost of the proposed method in this paper, and it also performs the clustering algorithm. 2) It is assumed that the clustering algorithm and the positioning algorithm are not used in lane determination. Without knowing the number of vehicles, the lane of each measurement point is directly calculated based on the Euclidean distance between the point and the centerline of each lane. For example, in the scenes of Fig.12 (a) and (b), one vehicle is directly driving between lane 2 and lane 3, respectively. The main body (number plate) of the vehicle in Fig.12 (a) and (b) is in the middle lane and the lane 3, respectively. In this way, the measurement data of the vehicle is divided into two parts (one in lane 3 and the other in lane 2), and the system will mistakenly consider them as two vehicles, as shown in Fig. 13 (a) and (c).

In summary, the dense measurement data mode in this paper is used (although this mode has a large amount of data, but the vehicle measurement is stable and there will be no missed detection vehicle). Then, the proposed clustering algorithm can be used to achieve the accurate determination of vehicle lanes, as shown in Fig. 13 (b) and (d). The accuracy of lane determination is higher than the method of judging vehicle lanes by using the distance between the position of the vehicle and the centerline of the lane.

C. CLUSTER PERFORMANCE EVALUATION FOR MULTIPLE APPLICATION SCENES

More to illustrate the performance of the algorithm, we extract data from 4 different scenes in Fig.3 (c) and (d). The 4 scenes are as follows:

The scene 5 description: In Fig.3(c), highway speed limit: light rain, visibility less than $200m$, less than $60km/h(16.67m/s)$. In the scene, two cars (the car's length and width are about $4m$ and $1.8m$ respectively) are driven on three lanes at about $21m/s$ and $17m/s$ respectively, and the weather is light rain.

The scene 6 description: In Fig.3(d), highway speed limit: the speed of entering the intersection is less than $40km/h(11.11m/s)$. In the scene, three vehicles travel at speeds of about $23m/s$, $18m/s$(entering the intersection), and $21m/s$, respectively. On the right side of the lane, a car runs normally along the lane. In the middle lane close to the radar, the vehicle is the medium-sized buses with a width of about $2.5m$ and a length of about $10m$, which runs normally along the middle lane. In the middle lane away from the radar, a car drives from the middle lane into the leftmost lane at the fork.

The scene 7 description: In Fig.3(c), highway speed limit: less than $120km/h(33.33m/s)$. In the scene, two high-speed vehicles travel at speeds of about $28m/s$ and $39m/s$, respectively. The vehicle in the rightmost lane is speeding vehicle.

The scene 8 description: In Fig.3(d), highway speed limit: the speed in the driving lane (the rightmost lane) should not be less than $60km/h(16.67m/s)$. In the scene, a vehicle in the right-most lane slowed down, changed lanes, veered into the left-most lane, and accelerated.

We use the proposed algorithm to analyze the measured data of the four scenes, and the experimental results are shown in Fig.14. The experimental results of the four scenes are all correct. For the experimental result of the first scene, as shown in Fig.14(a), our system operates normally in an appropriate light rain environment. The experimental result of the second scene show that our algorithm can adapt to the simultaneous detection of multiple vehicles, as shown in Fig.14(b). The experimental result of the third scene show that the proposed algorithm can accurately judge the lane of the overspeed vehicle without affecting the real-time performance, as shown in Fig.14(c). The experimental result of the last scene, as shown in Fig14(d), show that the low-speed

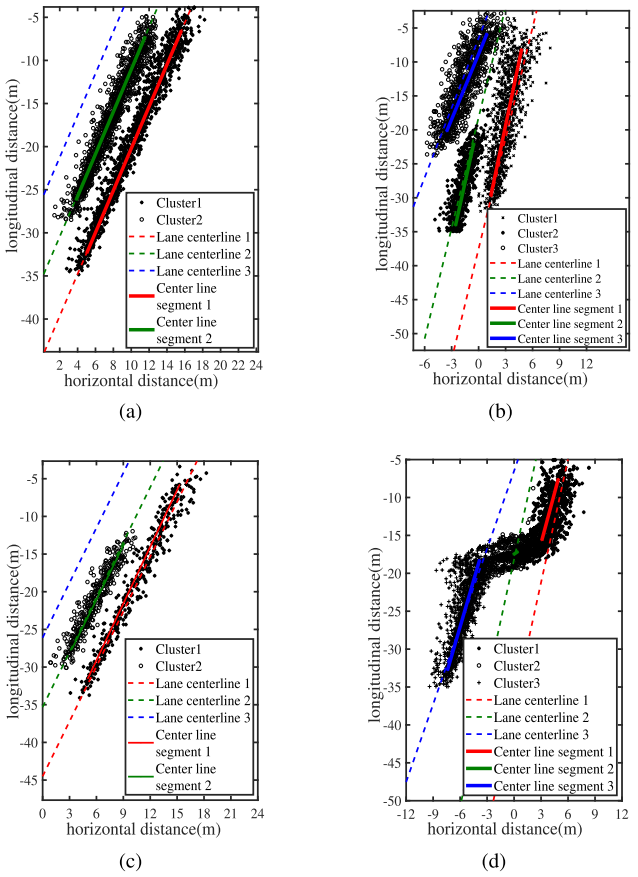

FIGURE 14. The experimental results of the different scenes. (a)∼(d) The experimental result of the scene 5∼8.

lane-changing vehicle appears in three lanes simultaneously. The system will capture illegal vehicles in three lanes at the same time.

In addition, we extracted experimental data of 42 minutes, 35 minutes, and 68 minutes in the three different scenes of the Fig.3 (b), (c), and (d), respectively, and 372 vehicles, 416 vehicles and 397 vehicles were monitored respectively. The analysis results of the proposed algorithm were compared with the monitoring video, and three vehicles were missed. The KLSAPCM algorithm can achieve the accuracy of 99.58% in the three different scenes of the Fig.3 (b), (c), and (d).

D. DISCUSS

The proposed algorithm in this paper is used to capture illegal vehicles on speed-limited highways. The speed limit of highway is generally between $40km/h(11.11m/s)$ and $120km/h(33.33m/s)$, the safe driving distance is more than $30m$, and the lane length of the monitoring area from accumulate measurement data of 5 to 15 frames in this paper is generally 4 to 30 meters. When the car is running at high speed, it is difficult for vehicles to complete the whole lane-changing process based on about $20m$. Therefore, we believe

that vehicle in the monitoring area was driving in the fixed lane. When a speeding vehicle appears, the lane of the speeding vehicle is identified and captured. When there are single or sparse illegal ultra-low-speed vehicles, the frame number of measurement data for low-speed vehicles in the monitoring area is more. If the vehicle changes lanes, the vehicle is simultaneously captured by cameras in multiple lanes before and after lane changes. When the vehicle is parked in the monitoring area, it is difficult to distinguish the measurement points of the vehicle from the road surface measurement points. The algorithm needs other types of algorithms to judge, such as data association algorithm. The innovation of the algorithm in this paper is on data classification, so there is no specific description of how to judge the driving vehicle parking in the monitoring area. In the event of traffic congestion and ultra-low speed driving, the radar can determine whether it is a traffic congestion mode by the running speed of multiple vehicles and the number of vehicles in the monitoring area. In this case, the algorithm stops running and the camera stops capturing. If there are weather conditions such as dust storms, hail, torrential rain, heavy snow, fog, ice and so on, the traffic management department closes the system by remote control. However, if there are some special scene such as some cars with the same speed, the big car shielding the small car, and the continuous driving of many low-speed illegal vehicles, the algorithm will fail, and the system needs other corresponding means to deal with them.

V. CONCLUSION

In this paper, a novel KLSAPCM clustering algorithm was developed in intelligent transportation monitoring system. The main works of the paper can be concluded as follows: 1) The extracting monitoring area method (adjacent to the lane centerline) and the minimum radius data search method for the MMW radar measurement data are introduced in traffic detection. These two methods effectively reduce the impact of noise on the clustering algorithm. Meanwhile, the minimum radius data search method can also effectively initialize the KLSAPCM clustering algorithm; 2) A KLSAPCM clustering algorithm is proposed in this paper. The algorithm can correctly determine which lane each vehicle belongs to without manually measuring the installation position and installation angle of radar; And 3) the detailed experiments of multiple scenes were presented and the results were: the KLSAPCM algorithm is compared with the DBSCAN, the k-means, the FCM, the PCM, the AMPCM, and the APCM algorithms on real measurement datasets, and the results highlight the classification rate of the proposed algorithm. Meanwhile, the proposed algorithm has good real-time performance and strong robustness for some sparse moving vehicle scene applications. As described in the discussion section, our future research content will be carried out to improve classification accuracy for more complex scenes, and increasing the scope of application of the system.

REFERENCES

[1] H.-T. Wu and G.-J. Horng, "Establishing an intelligent transportation system with a network security mechanism in an Internet of vehicle environment," *IEEE Access*, vol. 5, pp. 19239–19247, 2017.
[2] S. Bao, Y. Cao, A. Lei, P. Asuquo, H. Cruickshank, Z. Sun, and M. Huth, "Pseudonym management through blockchain: Cost-efficient privacy preservation on intelligent transportation systems," *IEEE Access*, vol. 7, pp. 80390–80403, 2019.
[3] M. Beyeler, F. Mirus, and A. Verl, "Vision-based robust road lane detection in urban environments," in *Proc. IEEE Int. Conf. Robot. Autom. (ICRA)*, May 2014, pp. 4920–4925.
[4] X. Du, K. Kiong Tan, and K. Ko Ko Htet, "Vision-based lane line detection for autonomous vehicle navigation and guidance," in *Proc. 10th Asian Control Conf. (ASCC)*, Zurich, Malaysia, May 2015, pp. 1–5.
[5] X. He, T. Wang, W. Liu, and T. Luo, "Measurement data fusion based on optimized weighted least-squares algorithm for multi-target tracking," *IEEE Access*, vol. 7, pp. 13901–13916, 2019.
[6] Z. Yang, H. Li, S. Ali, Y. Ao, and S. Guo, "Lane detection by combining trajectory clustering and curve complexity computing in urban environments," in *Proc. 13th Int. Conf. Semantics, Knowl. Grids (SKG)*, Aug. 2017, pp. 240–246.
[7] P. F. Felzenszwalb, R. B. Girshick, D. McAllester, and D. Ramanan, "Object detection with discriminatively trained part-based models," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 32, no. 9, pp. 1627–1645, Sep. 2010.
[8] D. Felguera-Martin, J.-T. Gonzalez-Partida, P. Almorox-Gonzalez, and M. Burgos-García, "Vehicular traffic surveillance and road lane detection using radar interferometry," *IEEE Trans. Veh. Technol.*, vol. 61, no. 3, pp. 959–970, Mar. 2012.
[9] M. Thuy and F. León, "Lane detection and tracking based on Lidar data," *Metrol. Meas. Syst.*, vol. 17, no. 3, pp. 311–321, Jan. 2010.
[10] F. Janda, S. Pangerl, and A. Schindler, "A road edge detection approach for marked and unmarked lanes based on video and radar," in *Proc. 16th Int. Conf. Inf. Fusion*, Istanbul, Turkey, Jul. 2013, pp. 871–876.
[11] Z. Xu, "Laser rangefinder based road following," in *Proc. IEEE Int. Conf. Mechatronics Autom.*, Jul./Aug. 2005, pp. 713–717.
[12] J. Han, D. Kim, M. Lee, and M. Sunwoo, "Enhanced road boundary and obstacle detection using a downward-looking LIDAR sensor," *IEEE Trans. Veh. Technol.*, vol. 61, no. 3, pp. 971–985, Mar. 2012.
[13] C. Stauffer and W. E. L. Grimson, "Adaptive background mixture models for real-time tracking," in *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit.*, vol. 2, Jun. 1999, pp. 246–252.
[14] M. M. V. Hulle, *Self-Organizing Maps*. Berlin, Germany: Springer, 2012.
[15] S. D. Xenaki, K. D. Koutroumbas, and A. A. Rontogiannis, "A novel adaptive possibilistic clustering algorithm," *IEEE Trans. Fuzzy Syst.*, vol. 24, no. 4, pp. 791–810, Aug. 2016.
[16] A. Rodriguez and A. Laio, "Machine learning. Clustering by fast search and find of density peaks," *Science*, vol. 344, no. 6191, pp. 1492–1496, 2014.
[17] R. Krishnapuram and J. M. Keller, "The possibilistic C-means algorithm: Insights and recommendations," *IEEE Trans. Fuzzy Syst.*, vol. 4, no. 3, pp. 385–393, Aug. 1996.
[18] W. S. Wijesoma, K. R. S. Kodagoda, and A. P. Balasuriya, "Road-boundary detection and tracking using ladar sensing," *IEEE Trans. Robot. Autom.*, vol. 20, no. 3, pp. 456–464, Jun. 2004.
[19] C. L. Chowdhary and D. P. Acharjya, "A hybrid scheme for breast cancer detection using intuitionistic fuzzy rough set technique," *Int. J. Healthcare Inf. Syst. Informat.*, vol. 11, no. 2, pp. 38–61, Apr. 2016.
[20] N. R. Pal, K. Pal, J. M. Keller, and J. C. Bezdek, "A possibilistic fuzzy c-means clustering algorithm," *IEEE Trans. Fuzzy Syst.*, vol. 13, no. 4, pp. 517–530, Aug. 2005.
[21] T. Chaira, "A novel intuitionistic fuzzy c means clustering algorithm and its application to medical images," *Appl. Soft Comput.*, vol. 11, no. 2, pp. 1711–1717, Mar. 2011.
[22] T. Wang, X. Wang, W. Shi, Z. Zhao, Z. He, and T. Xia, "Target localization and tracking based on improved Bayesian enhanced least-squares algorithm in wireless sensor networks," *Comput. Netw.*, vol. 167, Feb. 2020, Art. no. 106968.
[23] M.-S. Yang and C.-Y. Lai, "A robust automatic merging possibilistic clustering method," *IEEE Trans. Fuzzy Syst.*, vol. 19, no. 1, pp. 26–41, Feb. 2011.

LIN CAO received the B.Eng. degree in telecommunication engineering from Northeastern University, China, in 1999, and the Ph.D. degree in signal and information processing from the Institute of Electronics, Chinese Academy of Sciences, in 2005. He is currently a Professor with the Department of Electronic Engineering, Beijing Information Science and Technology University (BISTU). He teaches courses on digital signal processing, digital image processing, and soft design fundamentals. He is also the Dean of the School of Information and Communication Engineering and the Deputy Director of the Key Laboratory of the Ministry of Education for Optoelectronic Measurement Technology and Instrument, BISTU. He has published over forty articles on image processing and pattern recognition. His research interests include radar signal processing and image understanding and recognition. He is a member of the Institute of Electronics, Information and Communication Engineers (IEICE) and the China Education Society of Electronics.

TAO WANG received the Ph.D. degree from the School of Electronic and Information Engineering, Beihang University, in 2019. He is currently a Teacher of electronic engineering with Beijing Information Science and Technology University. His research interests include radar signal processing, data fusion, and target localization and tracking in intelligent transportation systems or vehicle intelligent assistance systems.

DONGFENG WANG received the Ph.D. degree in signal and information processing from the Institute of Electronics, Chinese Academy of Sciences, in 2003. He worked with the Institute of Electronics, Chinese Academy of Sciences, successively as a Research Assistant Fellow and an Associate Research Fellow, from 2003 to 2008. Since 2008, he has been the General Manager with Beijing TransMicrowave Technology Company Ltd. He is also a part-time Professor with Beijing Information Science and Technology University. His research interest is in radar signal processing.

KANGNING DU received the B.Sc. degree in telecommunication engineering from Beijing Information Science and Technology University, in 2011, and the Ph.D. degree in communication and information system from the Institute of Electronics, Chinese Academy of Sciences, in 2016. He is currently a Teacher of electronic engineering with Beijing Information Science and Technology University. His research interests include radar signal processing, and image understanding and recognition.

YUNXIAO LIU received the B.Sc. degree in electronic and information engineering from the University of Science and Technology Beijing, in 2017. He is currently pursuing the Master of Science degree in electronic and communication engineering with Beijing Information Science and Technology University. His research interest is in radar signal processing.

CHONG FU received the M.S. degree in telecommunication and information systems and the Ph.D. degree in computer software and theory from Northeastern University, Shenyang, China, in 2001 and 2006, respectively. He joined Northeastern University, in 2001, where he is currently a Professor with the Department of Communication and Electronics Engineering, School of Computer Science and Engineering. In 2010, he spent three months as a Visiting Researcher with the Department of Electronics Information Engineering, The Hong Kong Polytechnic University. His research interests include multimedia security and computer vision.