IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received July 27, 2018, accepted September 1, 2018, date of publication September 6, 2018, date of current version September 28, 2018.

Digital Object Identifier 10.1109/ACCESS.2018.2868976

Robust Lane Detection for Complicated Road Environment Based on Normal Map

CHANG YUAN[1], HUI CHEN[1], JU LIU[1], (Senior Member, IEEE), DI ZHU[1], AND YANYAN XU[2,3]

[1]School of Information Science and Engineering, Shandong University, Jinan 250100, China
[2]Department of Civil and Environmental Engineering, Massachusetts Institute of Technology, Cambridge, MA 02139, USA
[3]Department of City and Regional Planning, University of California at Berkeley, Berkeley, CA 94720, USA

Corresponding authors: Hui Chen (huichen@sdu.edu.cn) and Yanyan Xu (yanyanxu@mit.edu)

This work was supported in part by the Key Project of National Natural Science Foundation of China under Grant 61332015, in part by the Natural Science Foundation of Shandong Province of China under Grant ZR2013FM302 and Grant ZR2017MF057, and in part by the Science and Technology Innovation Action Plan Project of Shanghai Science and Technology Commission under Grant 18511104202.

ABSTRACT Detection of road or lane is indispensable for the environmental perception of advanced driver assistance systems. It has been an active field of research with a wide application prospect. However, due to the complex illumination and interferences, such as vehicles and shadows in the real driving environment, lane detection is still a challenging task today. To address these issues, a robust method for road segmentation and lane detection based on a normal map is proposed. The first step of this approach is to generate the normal map by using the depth information and then extract a segmented road pavement without vehicles and buildings based on the normal map. Second, we improve an adaptive threshold segmentation method and denoising operations to enhance the lane markings. Third, the combination of Hough transform and vanishing point makes it more accurate to determine the starting points of host lanes, and then the lanes in the following image sequence can be detected in the adaptive region of interest. Compared with the state-of-the-art methods, the experimental results on the data sets in two countries demonstrate that our approach produces more credible performance under various light conditions or dense traffic.

INDEX TERMS Advanced driver assistance systems, complicated road environment, lane detection, road segmentation, normal map.

I. INTRODUCTION

With the rapid development of transportation, the traffic safety has aroused people's great attention. Hence, intelligent transportation is an inevitable trend to avoid or reduce the occurrence of traffic accidents. So advanced driver assistance systems (ADAS) [1], [2] which are designed to help drivers in driving process emerge as the times require.

Lane-mark detection is one of the most important parts of ADAS and autonomous driving cars, and it's also a precondition for lane departure warning (LDW) [3]–[5]. The objective of lane detection is to locate and track the lane boundaries in road images so that the vehicle can be maintained to run along the host lane. We focus on the geometric computer vision-based approaches as they are cheaper in computational cost for lane detection. In the past two decades, researchers have made considerable progress in the vision-based approaches [2], [3], while facing several major challenges, especially in attaining robustness under complex lighting conditions and dense traffic. There are many studies of lane detection, we concentrate on reviewing related literature in the recent years, especially on methods of preprocessing.

Generally, visual-based lane detection approaches mostly follow this major pipeline: image preprocessing, feature extraction, lane model fitting, and lane tracking. The lane itself can be expressed in a simple model or analytical formula, like linear [4], [6], [7], parabolic [8], [9], hyperbolic [10], spline-based [7], [11] and so on. And for lane tracking, Kalman filter [6], [11]–[15] and particle filter [9], [16], [17] are the two tools widely adopted. Among the four steps in the pipeline, image preprocessing is the most basic and crucial step, which directly determines the quality of the follow-up lane feature extraction. The image preprocessing can be further subdivided into the following three steps:

1) Generation of the region of interest (ROI). Extracting the ROI is a simple but effective method to reduce redundant image data quantity, and some works select the bottom side of the image as a rough ROI [4], [11], [12], [18], while some others generate the ROI based on vanishing point detection techniques [5], [10], [14]. After the ROI is confirmed, some of these methods apply warp perspective mapping (WPM) [16] or inverse perspective mapping (IPM) [10], [13], [17], [19], [20] to yield the bird's eye view of the road image based on the assumptions of parallel lane boundaries and flat roads. In this way, most of the noises can be eliminated and lanes can be extracted conveniently. However, the ROI generated by above methods may still contain noises. For instance, once the host lane is covered by a large area of interferences, like vehicles or road signs, the methods are likely to detect the lane incorrectly [10], [18].

2) Enhancement of lane information. It is a necessary yet challenging step to preserve and enhance the lane information. According to different feature extraction methods, compatible preprocessing of images should be performed accordingly. There are variant features, such as color, edge, and geometric shape can be used to represent and separate the lanes from the background pixel by pixel. For the color-based methods, in order to extract the bright white or yellow lanes, some researchers perform a variety of color-space transformations, e.g., YCbCr [5], [21], HSI [17], [22], LAB [23], and others. However, most color models are sensitive to the varying illumination, while edges-based methods seem to be more robust against challenging light conditions. For the edge-based methods, some classical edge detectors, like Sobel detector [9], [24], [25] and Canny detector [12], [18], [21], [26], [27], are commonly adopted to enhance the lane edge. For the hybrid-based methods, color, edge, and width, as well as the vehicle speed, are taken into consideration to extract the lane marking information [18]. After edge detection, threshold segmentation methods can be used to reduce computational time and improve the result of the lane detection algorithm, among them, the Otsu's method [6], [18] and local adaptive threshold segmentation [12] are widely used. However, they can only accommodate traffic scenes with good or single lighting conditions.

3) Elimination of non-lane information. The aforementioned generation of ROI has removed some of the non-lane information, but the noises, like vehicles, shadows, and stains, still need to be further eliminated. To this end, some filters and other denoising methods have been introduced [15], [7], [27]. Nan *et al.* employed the crossing point filter and the structure triangle filter [7], which introduce spatial structure constraints and temporal location constraints into lane detection, to filter out the noisy line segments. Wu *et al.* proposed a new lane-mark extraction algorithm to divide the boundary image into sub-images to calculate the local edge-orientation of each block [12], while the edges with abnormal orientations are removed to eliminate noise edges efficiently by defining a scope of the direction of lane edge. All of the above methods are disturbed by complex illumination and noises in the same direction as the lane edges.

By surveying most of the literature related to vision-based lane detection methods and testing the commonly used preprocessing approaches, we observed that the accuracy of lane detection mainly depends on whether the lane can be completely preserved during the image preprocessing stage. Therefore, this paper targets at a robust image preprocessing method. Our major contributions are in the following three aspects: (1) The road is segmented as the ROI by merging the normal map with the raw image; (2) A modified local adaptive threshold segmentation method is presented, which is able to binarize the image under various light conditions. In addition, the proposed denoising operation plays an important role in removing mostly non-lane information; (3) The position of the initial lane is determined accurately by using Hough transform in combination with the vanishing point.

The remaining sections of this paper are organized as follows. Section 2 presents the details of our lane detection method, including road segmentation, lane feature enhancement, and identification of the lane. In Section 3, the optimal configuration of parameters, experimental results, and the limitations of our method are presented. Lastly, conclusions are given in Section 4.

II. PROPOSED METHOD

For vision-based lane detection algorithm, images are initially captured with road-facing camera and then the preprocessing is carried out to extract valuable features. Images recorded with binocular cameras preserve not only lane lines, but also the depth information. Therefore, many middle and high-end cars are equipped with binocular cameras to determine the distance between the obstacles and the vehicles using the inferred depth information. In the lane detection problem, to effectively remove the interferences of vehicles and other non-lane information, we take full advantage of this depth information to segment the road pavement. The lane detection algorithm of this paper can be divided into three functional modules: road segmentation, lane feature enhancement and identification of the lane. The flowchart is shown in Fig. 1.

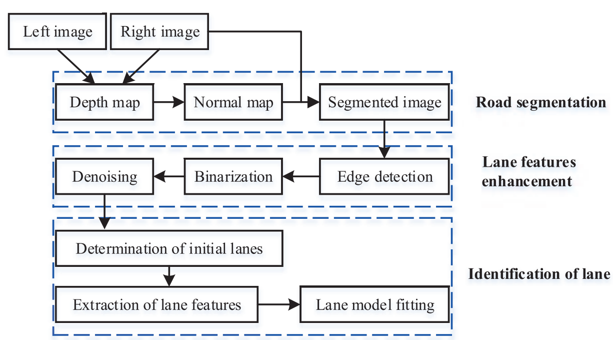


FIGURE 1. Flowchart of proposed lane detection method.

A. ROAD SEGMENTATION

Images recorded from road often contain lots of useless information considering lane detection, such as the sky, trees, vehicles, buildings, etc, namely non-lane information. To reduce the influence of such non-lane information, normal map inferred from the stereo image pair is utilized to segment the road based on the knowledge that the pixels of pavement have the same normal vector as they are in the same plane. Specific steps of road segmentation are presented as follows.

1) ACQUISITION OF DEPTH MAP

Stereo vision can obtain the disparity map from two or more images with different viewpoints. An accurate disparity map is fundamental to better reconstructing the normal map. In this paper, the depth image dataset is provided from KITTI dataset. For our dataset, the stereo matching algorithm [28] is used to acquire the disparity map, as it is reliable and easy implemented. Then, the depth map can be calculated according to the relationship between disparity and depth, as shown in the following equation,

$$Z = \frac{B * f}{d} \tag{1}$$

where Z denotes the actual distance from the object in the scene to the baseline of the camera, namely depth. B denotes the distance of optical centers of two cameras, namely baseline. f is the focal length of the camera, and d refers to the disparity of the object. Two cameras are assumed to be with the same internal parameters.

2) ESTIMATION OF NORMAL MAP

After 3D information is extracted, the plane normal map consisting of normal vectors of each pixel in the original image can be generated with the following steps.

- The intrinsic parameter matrix $\mathbf{K}$ of camera is obtained by camera calibration,

$$\mathbf{K} = \begin{bmatrix} f_x & 0 & u_0 \\ 0 & f_y & v_0 \\ 0 & 0 & 1 \end{bmatrix} \tag{2}$$

  where, $f_x = f/d_x$ and $f_y = f/d_y$, represent the scale factor of d_x and d_y directions, respectively. u_0 and v_0 are the centers of the image planes.
- With the transformation relation between coordinate systems in the camera model, the points in depth map can be converted from image pixel coordinates (u, v) to camera coordinates (X_C, Y_C),

$$\begin{bmatrix} X_C \\ Y_C \\ 1 \end{bmatrix} = \begin{bmatrix} \frac{1}{f_x} & 0 & -\frac{u_0}{f_x} \\ 0 & \frac{1}{f_y} & -\frac{v_0}{f_y} \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} \tag{3}$$

- We traverse each pixel in the depth map and search its $N \times N$ neighboring pixels to form vectors (such as $\vec{a}$ and $\vec{b}$ in Fig. 2). Taking the 3×3 neighborhood as an example, the normal vector $\vec{c}$ of this pixel can be obtained by a cross product, as follows,

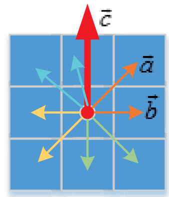


FIGURE 2. Illustration of normal vector estimation.

$$\vec{c} = \vec{a} \times \vec{b} \tag{4}$$

$$\begin{cases} \vec{a} = [X_C - X_{Ci}, Y_C - Y_{Ci}, Z_C - Z_{Ci}] \\ \vec{b} = [X_C - X_{Cj}, Y_C - Y_{Cj}, Z_C - Z_{Cj}] \\ i, j \in (1, 8), i \neq j \end{cases} \tag{5}$$

To avoid the interference of noises, we traverse the pixels in the depth map. For each pixel, we find two clockwise adjacent pixels in its $N \times N$ neighborhood and form two vectors, as the $\vec{a}$ and $\vec{b}$ shown in Fig. 2. Then we use Equation (4) to calculate the normal vector, and finally average the normal vector of this pixel. The schematic diagram of the normal vector estimation is illustrated in Fig. 2, and the experimental results of the normal map are shown in Fig. 3.

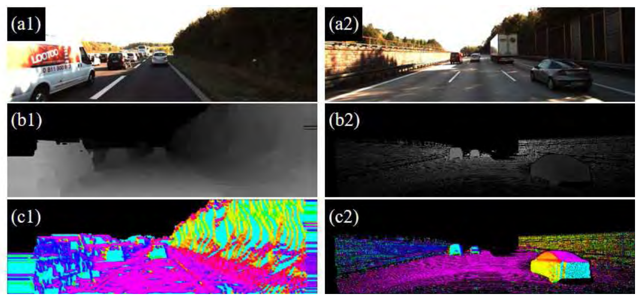


FIGURE 3. The results of disparity map and normal map. (a1) and (a2) are the original road images from the KITTI dataset; (b1) and (b2) are the disparity maps generated by the stereo matching algorithm [28] and provided directly by KITTI dataset, respectively; (c1) and (c2) are the corresponding normal maps.

3) NORMAL MAP OPTIMIZATION AND SEGMENTATION

The normal vector (x, y, z) of each point can correspond to (r, g, b) in the color map one by one. Hence, to optimize the normal map, morphological operations are first utilized to merge pixels with similar colors into regions. Next, we choose the front of the vehicle as the seed region, and the region with the same normal vector as the seed region is regarded as the road pavement. When the size of this road region is less than a certain value, we can determine that there is no safe area ahead and an alarm arises.

Teichmann et al. proposed a network architecture named "MultiNet" [29], which is able to perform road segmentation, car detection, and street classification at the same time. For road segmentation, their algorithm can segment

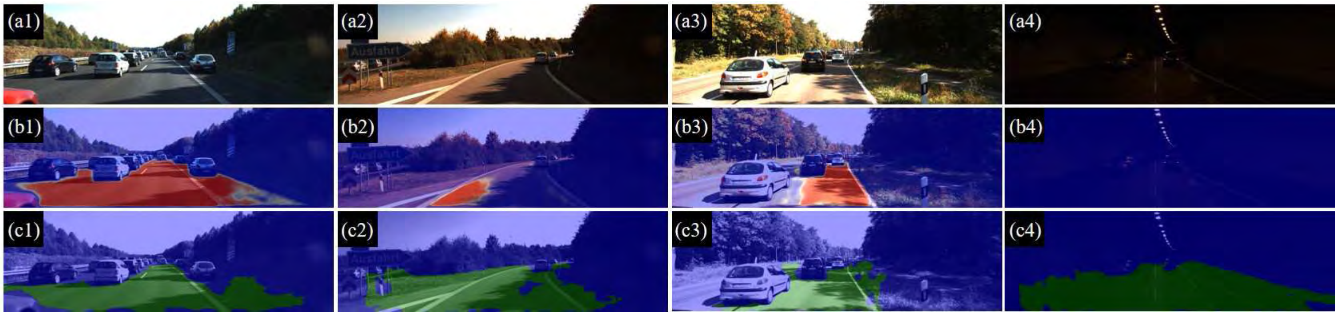


FIGURE 4. Sample results of road segmentation. (a1)~(a4) are the original images; (b1)~(b4) are the results of "MultiNet" [29]; (c1)~(c4) are the results of the proposed method.

the road more accurately without the roadside structures, but requires a large number of different training data to ensure the accuracy of the algorithm and fails under certain lighting conditions, as illustrated in Fig. 4. Our work only needs to segment the pavement area to remove the interference caused by the vehicles and buildings. The normal map is not sensitive to color and illumination, and thus it can properly extract the dominant plane in the scene. Therefore, it is a simple and effective method to separate the road pavement and can be coupled with other lane detection algorithms based on vision.

B. LANE FEATURE ENHANCEMENT

According to the natural characteristics of the road scene, the lane in the image usually has a certain slope angle, so the Sobel operator [24] is employed to enhance the lane edges, which is proved to be of simplicity and credible performance. After performing the edge detection process, the binarization is applied to preserve valid lane features and reduce data processing.

1) ADAPTIVE THRESHOLD SEGMENTATION

Reasonable threshold is not only the key to binarization, but also the basis for accurate lane detection. There are two widely used methods. One is Otsu's method [30] which is a kind of global threshold method. Otsu's method reaches relative good performance if the histogram of the image can be assumed to have bimodal distribution with a deep and sharp valley between the two peaks. Otherwise, in the cases of low visibility of objects or a large amount of noise in the image, the possibly incorrect threshold determined by Otsu's method results in the segmentation error. Another method is the local adaptive threshold segmentation [31], which relies on the distribution of the pixel values in the adjacent blocks. However, the performance of local thresholding method is limited by uneven illumination and noises. To deal with these issues, we propose a modified local adaptive threshold segmentation method. For image $f(x, y)$, we assume that the average of gray value of the pixels in a certain $w \times w$ block is m and the standard deviation is σ. Then the constant difference to correct the threshold can be defined in Equation (6), which depends on the statistical characteristic of the pixel values in the adjacent blocks.

$$C = \max(k1 * \sigma, k2 * m) \tag{6}$$

where, $k1$ and $k2$ are two constants. Therefore, the binarization threshold for a certain $w \times w$ block centered on the pixel (i, j) can be expressed as,

$$T_{i,j} = \frac{\sum_{w \times w} f(i, j)}{w \times w} - C \tag{7}$$

In the proposed threshold segmentation method, there are three key parameters, $\{w, k1, k2\}$, affecting the performance of the lane detection algorithm, which will be discussed in detail in the experimental description of Section 3A.

2) ELIMINATION OF EDGE NOISE

The proposed adaptive threshold segmentation approach can be well adapted to complex illumination, but the segmented results still contain some noises, especially in the case of uneven illumination. Therefore, proper denoising operations must be employed.

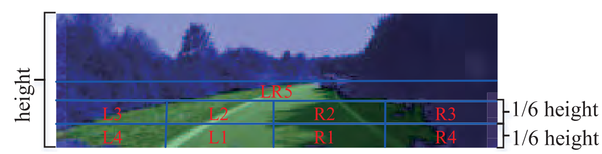


FIGURE 5. Elimination of edge noise.

Generally, the lanes in the images captured by the road-facing cameras often have relatively fixed shapes and orientations, so elimination of noise can be carried out in two aspects: (1) Area (S) constraint, namely defining an area size threshold of the connected region; (2) Orientation (θ) constraint, namely limiting the angle of the lane. In this research, we regard the lanes as a set of discrete line segments, and extract nine sub-images ($L1, L2, L3, L4, R1, R2, R3, R4, LR5$) in segmented image, as illustrated in Fig. 5. If the lines meet some predefined conditions, as shown in Equation (8), it will be considered as the line in a lane, and we preserve the line and all the points on its extension which exist in the binary map. Fig. 6 shows the results of image preprocessing

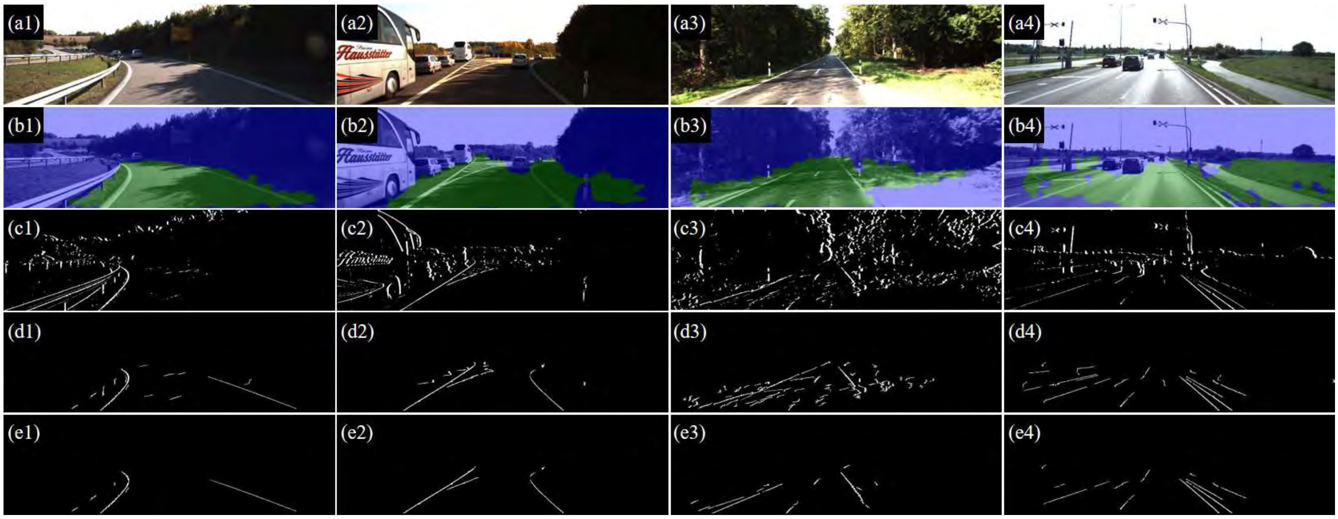


FIGURE 6. **Sample results of image preprocessing. (a1)~(a4) are the original images; (b1)~(b4) are the results of road segmentation; (c1)~(c4) are the results of binarization with Otsu's method [30]; (d1)~(d4) are the results of proposed threshold segmentation method; (e1)~(e4) are the results after removing noise.**

and comparison with Otsu's method. As seen in the figure, the Otsu's method works well only when lane-mark edges are clearly visible and the illumination distribution is even. However, in the cases of complex environments, the lane-mark edges will be affected by various conditions, such as sunlight, shadows, or dirt. In contrast, the proposed threshold segmentation method and denoising operation can eliminate vehicles and other obstacles on the road efficiently.

$$line = \begin{cases} True, & S > 50;\ \theta \in [15^\circ, 165^\circ] \\ False, & otherwise \end{cases} \tag{8}$$

C. IDENTIFICATION OF LANE

With the candidate lane edges obtained through the previous image preprocessing, the host lanes are identified from the candidates into three steps, (1) finding the initial position of the lane; (2) making linear prediction of the location of the adaptive regions of interest (AROI) of lanes, and then searching the inner edge points of the lane in the AROI; (3) fitting the lanes with the least square method.

1) DETERMINATION OF INITIAL LANES

The determination of the initial lanes is the key factor to the lane detection. For the first frame, a comprehensive approach combining Hough transform and vanishing point is used to extract a accurate initial position of lanes. First, Hough transform is applied to detect the lines, and then the intersections of these lines are calculated. However, it is hard to get a unique intersection point when more than two lines exist. Thus, optimization procedure should be employed, and the total squared distance from all the lines as a cost function is defined in Equation (9),

$$I = \frac{1}{2}\sum \left(n_i^T\ (V_P - p_i)\right)^2 \tag{9}$$

where p_i denotes the point on line i detected by Hough transform and n_i is the unit normal to line i. Then the vanishing point (V_P) is selected as the point whose cost function is minimal. To find the minimum, the cost function is differentiated with respect to the V_P, so that the expression of vanishing point can be obtained with the following equation.

$$V_P = \left(\sum n_i n_i^T\right)^{-1}\left(\sum n_i n_i^T p_i\right) \tag{10}$$

In the case of complex illumination (intense light, uneven illumination, etc.) or serious lack of the lane lines, the Hough transform may not be able to detect all the lines of lanes. At this time, estimation of the vanishing point tends to fail, then the center of image is taken as the vanishing point.

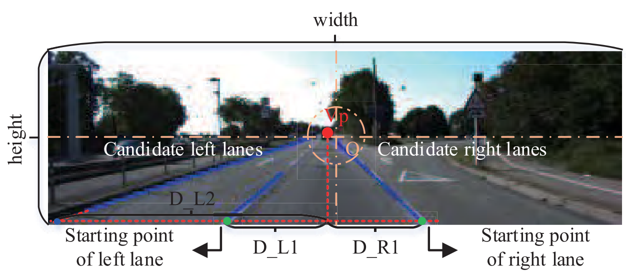


FIGURE 7. **Determination of starting point of lanes.**

Once the candidate lines and vanishing point are found, we extend all these candidate lines to find their intersection points with the horizontal line at the bottom of the image. Then the distance between each intersection point and the vertical line traversing the vanishing point can be calculated to assist in finding the initial position of the lanes, like D_L1, D_L2, D_R1 in Fig. 7. Generally, when the vehicle camera is fixed, the width between the left and right lanes is constant. Based on this knowledge, we define $\Delta d = D\_Li + D\_Ri$.

For each candidate line, if the aforementioned constraints are reached, this line is labeled as a candidate lane. And then the starting point of host lane is regarded as the intersection point who has the shortest distance to the vertical line traversing the vanishing point, as shown in Fig. 7, D_L1 and D_R1 are the starting point of host lane. It is efficient to predict the starting position of lane marking in the next frame by using the information of the lane position obtained from the previous frame, and then all the lane features can be detected from the bottom to upper in the AROIs. Considering that lane change in the driving, the starting point of the lanes is re-identified around 20 frames.

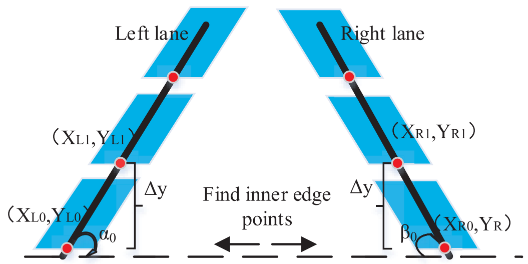


FIGURE 8. **Illustration of AROI.**

2) EXTRACTION OF LANE FEATURES

According to the information of the starting point of lane obtained above, the first ROI close to the host lane can be determined, and then the points of inner lane edge are searched in the ROI. By this way, the noise disturbance can be eliminated further. Then, we apply the least square method to fit these points into a short line segment. The angle of this line segment, the position of the previous ROI, and the interval Δy can be used to determine the next ROI, as shown in Equation (11), which takes the left lane as an example. This allows a linear prediction of the ROI along the direction of the lane, as illustrated in Fig. 8.

$$\begin{cases} X_{L1} = X_{L0} + \dfrac{\Delta y}{\tan \alpha_0} \\ Y_{L1} = Y_{L0} - \Delta y \end{cases} \tag{11}$$

In order to achieve the flexible tracking of the straight or curve lane, the discrete AROIs are introduced, whose size and position can be adjusted dynamically according to the curvature change of lane and the speed of the vehicle. Experimental results imply that the curvature change of the straight lane is relatively small, so we gradually decrease the width of ROI and properly increase the interval between ROIs. While the curvature of the curve lane varies greatly, we increase the width of the ROI and reduce the interval between ROIs to avoid missing detection.

3) LANE MODEL FITTING

After the lanes are extracted, a compact high-level representation of the path can be obtained, which is used for decision making. The angles of lane lines obtained from each ROI are utilized to identify the lane as straight or curve. When the change of angle is greater than a certain value (10° in our experiments), the least square curve model is adopted to fit the lane. Otherwise, the least squares linear model is adopted.

III. EXPERIMENTS AND DISCUSSIONS

The experiments are performed on a PC equipped with 3.20 GHz Intel Core i5 CPU and 8 GB RAM. In order to quantify the robustness of the proposed method, the performance of the proposed algorithm and other algorithms are estimated on the KITTI dataset and our dataset, separately. It's noteworthy that the images we used are all based on the criteria that the lane markings should be visible to the naked eye. The performance of lane detection varies with a set of parameters. Hence, we analyze the effects of different parameters on the performance of the lane detection and confirm the optimal configuration of parameters finally.

A. OPTIMAL CONFIGURATION OF PARAMETERS

For edge-based methods, the degeneration of a gray level image to a binary image is the most critical but challenging step, and directly determines whether the lane information can be retained and separated from the background. The performance of our modified threshold segmentation method primarily affected by three parameters, $\{w, k1, k2\}$. Inappropriate selection of the parameters would lead to failure of lane preservation or excessive noises in the binarized images. To this end, 500 road images are randomly selected from the two datasets, then different combinations of the $\{w, k1, k2\}$ are tested. We use the accuracy rate, R, which is defined as the fraction between the number of correctly detected frames and the total number of frames, to measure the performance of the detection algorithm. For each configuration, the corresponding accuracy rate is calculated, as displayed in Fig. 9. It is apparent that when the window size is 9, the overall performance is the best among the candidate sets. In the local window of this size, the values of $\{k1, k2\}$, are varied between the ranges illustrated in Fig. 9(d). It can also be observed that the highest R value, 93.60%, is achieved with the configuration $\{9, 0.2, 0.2\}$. Of course, we may find a better configuration, but this one is acceptable.

B. COMPARISON WITH THE STATE-OF-THE-ARTS

We measure the performance of the proposed lane detection algorithm using the "optimal" configuration of parameters and other four baseline methods, i.e., IPM + hyperbolic model [10], superparticle method [16], lane detection with two-stage feature extraction (LDTFE) method [26], and a state-of-the-art fully convolutional networks (FCN) method [32]. These four different methods are relatively new and practical in current vision-based lane detection algorithms. We evaluate the performance using the metrics of accuracy rate, detailed statistics on the KITTI dataset and our dataset, which are depicted below and shown in Table 1 and Table 2, respectively.

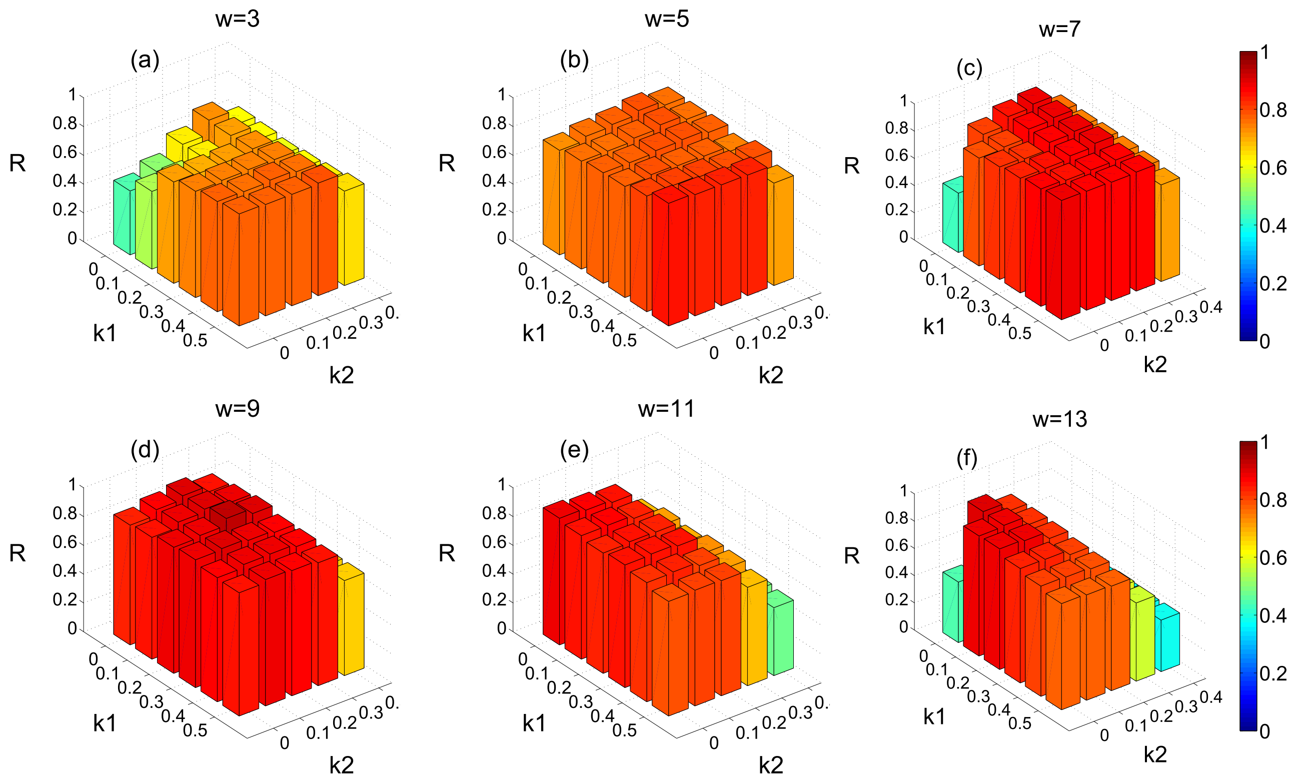


FIGURE 9. The accuracy rate of different configurations of parameters for different window sizes. (a) $w = 3$; (b) $w = 5$; (c) $w = 7$; (d) $w = 9$; (e) $w = 11$; (f) $w = 13$.

TABLE 1. The performances of different lane detection algorithms on KITTI dataset.

Scenario	Total lane markings	Accuracy rate (%)				
		IPM + hyperbolic model [10]	Superparticle [16]	LDTFE [26]	FCN [32]	Proposed method
Daytime	1234	93.92	94.73	94.00	90.44	95.22
Dusk/dense traffic	837	70.97	84.71	85.07	91.04	96.77
Shadow	957	89.86	84.12	88.51	87.04	97.49
Strong light reflection	390	57.44	90.51	86.67	65.38	91.28

TABLE 2. The performances of different lane detection algorithms on our dataset.

Scenario	Numbers	Accuracy rate (%)				
		IPM + hyperbolic model [10]	Superparticle [16]	LDTFE [26]	FCN [32]	Proposed method
Daytime	584	76.03	93.15	90.07	94.18	95.55
Night	613	42.90	60.20	87.11	70.31	92.82
Dense traffic	563	44.23	92.72	87.21	92.18	96.80
Stains/scratches	607	83.19	89.62	89.95	89.95	90.12

1) KITTI DATASET

KITTI is a challenging real-world computer vision dataset [33]. In our experiment, four different types of environments are tested to quantify the performance of different lane detection algorithms. They are, (i) day-time scenarios, which are well illuminated with vehicles and stains; (ii) dusk scenarios, where the brightness of the lanes is inconsistent because of the existence of the shadow, and dense traffic makes the detection of lanes more difficult; (iii) shadows scenarios, where intense light and tree shadows create the illumination inconsistencies and deteriorate the visibility; (iv) light reflection scenarios, where the optical reflections on the road reduce the contrast between lane markings and pavement. The accuracy rates of different lane detection algorithms on KITTI dataset are shown in Table 1. Sample detection results of four scenarios on KITTI dataset are illustrated in Fig. 10.

2) OUR DATASET

In order to further test the robustness of the proposed algorithm in the different urban environment. We create our dataset of road images captured by a binocular camera,

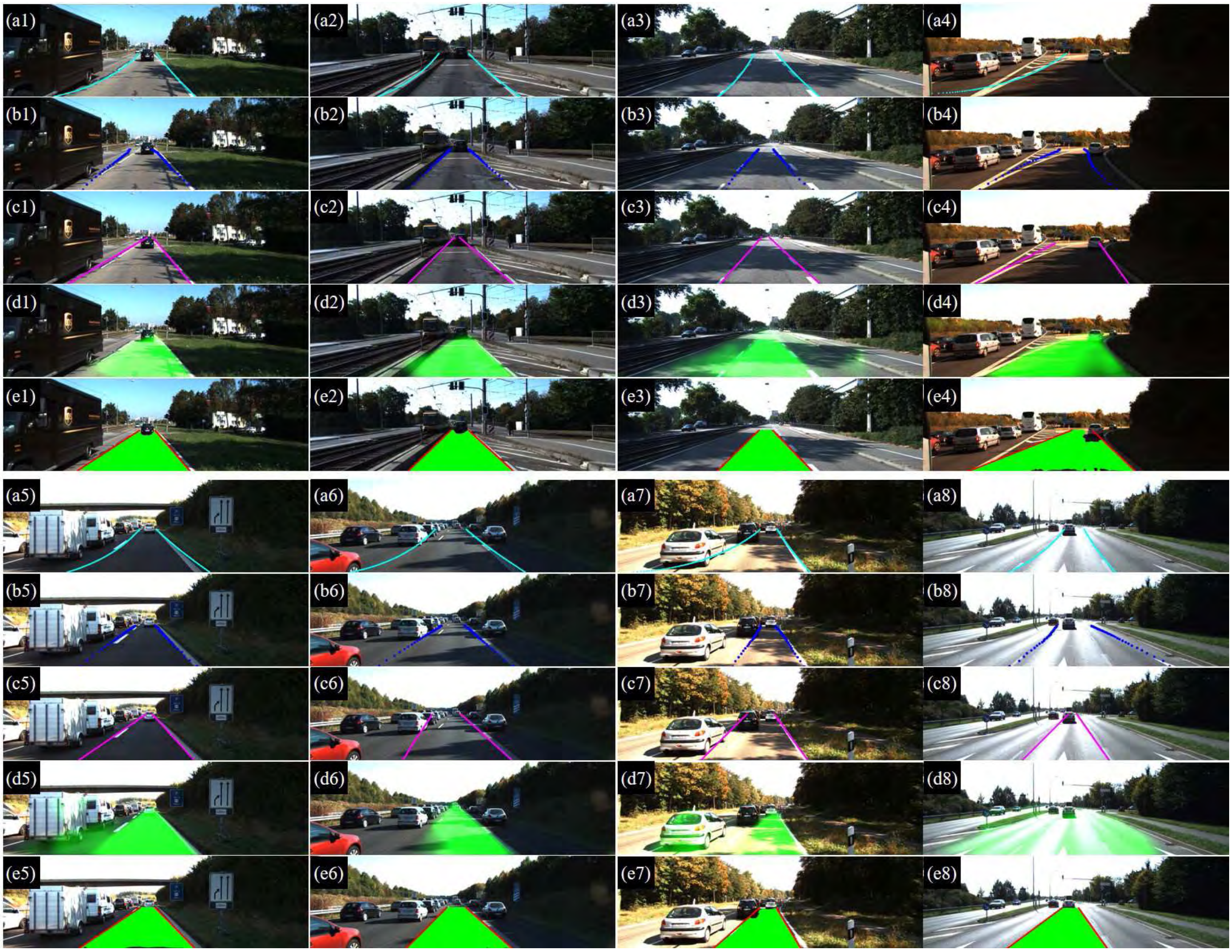


FIGURE 10. Sample detection results on KITTI dataset.(a1)∼(a8) are the results of IPM + hyperbolic model [10], cyan lines refer to the detected lanes; (b1)∼(b8) are the results of superparticle [16], blue points refer to the detected lanes; (c1)∼(c8) are the results of LDTFE [26], magenta lines refer to the detected lanes; (d1)∼(d8) are the results of FCN [32], green area refers to the host lane; (e1)∼(e8) are the results of the proposed method, red lines refer to the detected lanes, and green area refers to the host lane.

including four different scenarios: daytime, night, dense traffic, and stains/scratches. The daytime scenarios are well illuminated, while the brightness and visibility of the lanes in the night scenarios are relatively low; For scenarios with dense traffic or stains and scratches, the interference factors around lanes increase the difficulty of lane detection. The performances of different lane detection algorithms on our dataset are shown in Table 2. Sample detection results of four scenarios on our dataset are shown in Fig. 11.

As can be observed from the Fig. 10 and Fig. 11, despite there exist heavy shadows or uneven illumination, the proposed method can successfully identify the lanes. While the reference methods are easy to fail faced with the complex environment. Besides, we inspect the reasons for the failure of the reference methods in typical scenarios. In the IPM + hyperbolic model, Xu *et al.* chose 95% of the maximum pixel values as the global threshold [10], which is not feasible in diverse scenes. As can be observed in Table 1, Table 2, Fig. 10(a4), (a6), (a7), and Fig. 11(a2), (a4), (a5), (a7), (a8), (a10), this method collapses when facing complex light condition and dense traffic, e.g., the strong light reflection scenario in KITTI dataset, and the night and dense traffic scenarios in our dataset. In the superparticle method [16], the oriented distance transform (ODT) applied on the binary edge map is sensitive to occurrences of noisy pixels, especially when the lanes are discontinuous or missing. In our experiments, the superparticle method collapse in the night scenario in our dataset with only 60.20% accuracy rate. Some failures can be observed in Fig. 10(b4), (b6), (b7), (b8), and Fig. 11(b4), (b6), (b8), (b10). In the LDTFE method [26], despite Niu *et al.* attempted to introduce the Canny detector to reduce the impact of uneven light conditions, it still can not work well with dense traffic, as demonstrated in Fig. 10(c6), (c7), (c8), and Fig. 11(c2), (c4), (c7), (c8), (c10). In the FCN method [32], a fully convolutional network is used for semantic segmentation of the host lane. As a deep learning method, FCN requires a large amount of data to train

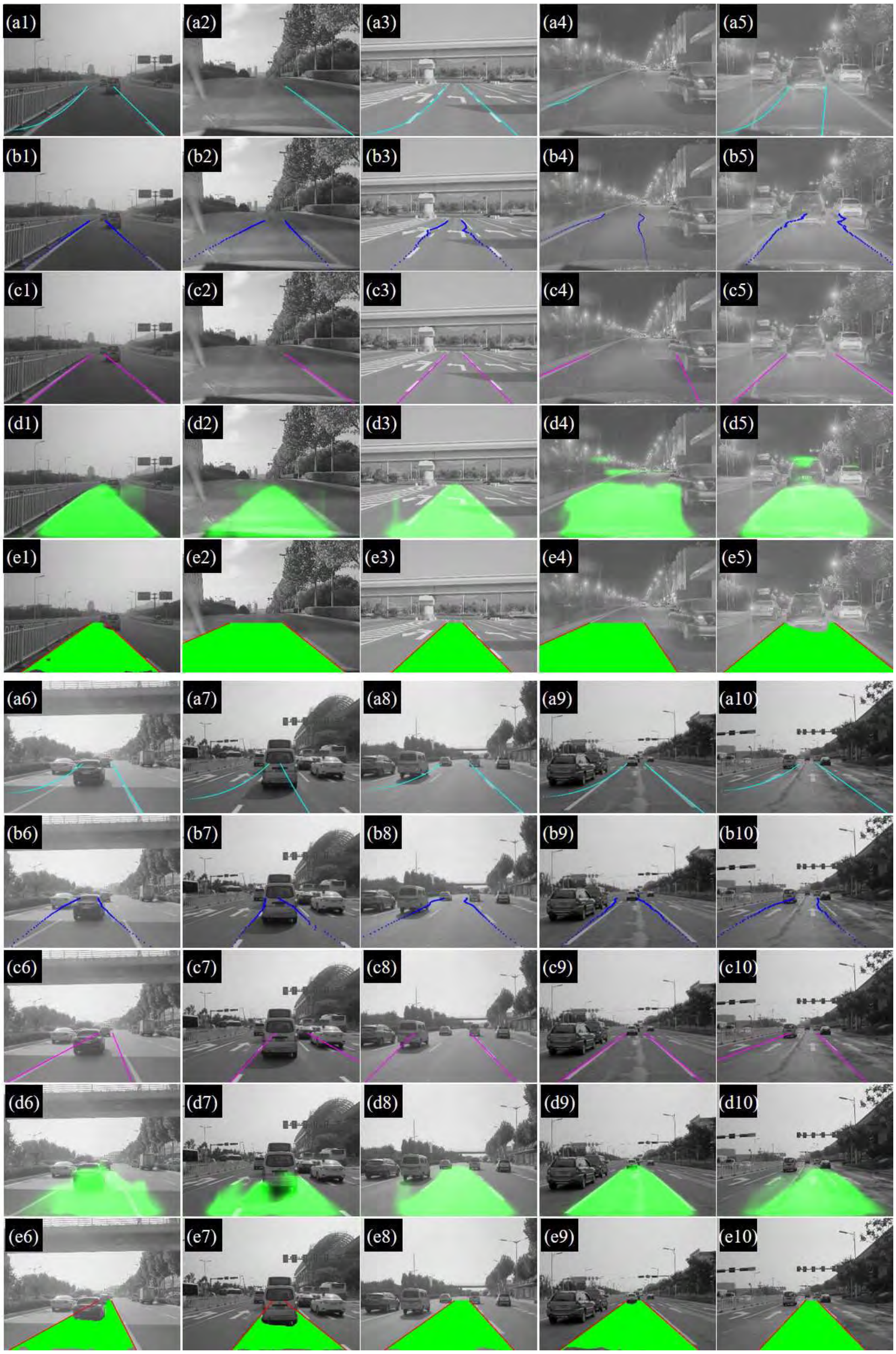


FIGURE 11. Sample detection results on our dataset. (a1)~(a10) are the results of IPM + hyperbolic model [10], cyan lines refer to the detected lanes; (b1)~(b10) are the results of superparticle [16], blue points refer to the detected lanes; (c1)~(c10) are the results of LDTFE [26], magenta lines refer to the detected lanes; (d1)~(d10) are the results of FCN [32], green area refers to the host lane; (e1)~(e10) are the results of the proposed method, red lines refer to the detected lanes, and green area refers to the host lane.

a powerful model. However, it trends to fail when tackling the scene which is significantly different from the training datasets, as can be seen in Fig. 10(d3), (d5), (d7), (d8), and Fig. 11(d4), (d7), (d10).

C. LIMITATIONS OF PROPOSED APPROACH

Despite the proposed model reaches reliable performances in typical scenes, there are still some limitations. Firstly, the proposed algorithm relies on the depth map to accurately

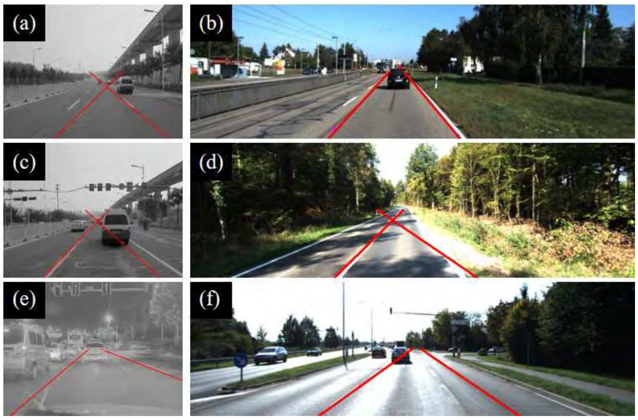


FIGURE 12. Failed detection cases. (a), (c) and (e) are the failed detection results on our dataset; (b), (d) and (f) are the failed detection results on KITTI dataset.

segment the road pavement, but the development of stereo matching algorithm can facilitate this requirement. Secondly, the generation of normal map consumes most of the time in the proposed lane detection algorithm and needs to be optimized and accelerated to save the computing resources. Thirdly, the proposed method is only applicable to detect lanes in structured roads with lane markings and clear road boundaries. Fourthly, lane-like scratch (as shown in Fig. 12(a)∼(c)) and low visibility of lanes caused by light (as shown in Fig. 12(d)∼(f)) are challenges for the proposed algorithm, which are also the further work to improve and overcome.

IV. CONCLUSIONS

In this paper, a road segmentation and lane detection algorithm based on the normal map has been described. Unlike the deep learning methods need to be trained on large data sets, our model is based on geometric computer vision and ''plug and play''. Especially, we focus on dealing with the complex illumination and dense traffic which are common in the real driving environment. To that end, first of all, a method based on the normal map to segment the road is proposed. This method can effectively remove the interference of buildings and vehicles on the basis of a higher-precision depth map. And it also can be used to avoid collision on unstructured rural roads. Secondly, the traditional local adaptive threshold segmentation method is improved to adapt to a variety of complex illumination conditions. The ''optimal'' configuration of key parameters is found through a large number of experiments and works in a variety of lighting conditions. Thirdly, in order to accurately infer the starting point of the lane, Hough transform and vanishing point are combined to jointly identify lane boundaries. To demonstrate the strength of proposed approach, three state-of-the-art lane detection methods are compared on two databases. Experimental results indicate that the proposed algorithm works robustly and accurately under various challenging situations, in spite of some limitations stated above, such as lane-like scratch and blur lanes. In the near future, we can optimize the time cost for the mobile devices and improve the detection rate in more complex scenes, and our research will also be followed up with lane tracking.

REFERENCES

[1] K. Bengler, K. Dietmayer, B. Farber, M. Maurer, C. Stiller, and H. Winner, ''Three decades of driver assistance systems: Review and future perspectives,'' *IEEE Intell. Transp. Syst. Mag.*, vol. 6, no. 4, pp. 6–22, Oct. 2014.

[2] A. B. Hillel, R. Lerner, D. Levi, and G. Raz, ''Recent progress in road and lane detection: A survey,'' *Mach. Vis. Appl.*, vol. 25, no. 3, pp. 727–745, 2014.

[3] S. P. Narote, P. N. Bhujbal, A. S. Narote, and D. M. Dhane, ''A review of recent advances in lane detection and departure warning system,'' *Pattern Recognit.*, vol. 73, pp. 216–234, Jan. 2018.

[4] V. Gaikwad and S. Lokhande, ''Lane departure identification for advanced driver assistance,'' *IEEE Trans. Intell. Transp. Syst.*, vol. 16, no. 2, pp. 910–918, Apr. 2015.

[5] J. Son, H. Yoo, S. Kim, and K. Sohn, ''Real-time illumination invariant lane detection for lane departure warning system,'' *Expert Syst. Appl.*, vol. 42, pp. 1816–1824, Oct. 2014.

[6] M. Liang, Z. Zhou, and Q. Song, ''Improved lane departure response distortion warning method based on Hough transformation and Kalman filter,'' *Informatica*, vol. 41, no. 3, pp. 283–288, 2017.

[7] Z. Nan, P. Wei, L. Xu, and N. Zheng, ''Efficient lane boundary detection with spatial-temporal knowledge filtering,'' *Sensors*, vol. 16, no. 8, p. 1276, 2016.

[8] S. Kwon, D. Ding, J. Yoo, J. Jung, and S. Jin, ''Multi-lane dection and tracking using dual parabolic model,'' *Bull. Netw., Comput., Syst., Softw.*, vol. 4, no. 1, pp. 65–68, 2015.

[9] Y. Wang, N. Dahnoun, and A. Achim, ''A novel system for robust lane detection and tracking,'' *Signal Process.*, vol. 92, no. 2, pp. 319–334, 2012.

[10] S. Xu, P. Ye, S. Han, H. Sun, and Q. Jia, ''Road lane modeling based on RANSAC algorithm and hyperbolic model,'' in *Proc. 3rd Int. Conf. Syst. Inform. (ICSAI)*, Nov. 2016, pp. 97–101.

[11] J. Deng and Y. Han, ''A real-time system of lane detection and tracking based on optimized RANSAC B-spline fitting,'' in *Proc. Res. Adapt. Convergent Syst.*, 2013, pp. 157–164.

[12] P.-C. Wu, C.-Y. Chang, and C. H. Lin, ''Lane-mark extraction for automobiles under complex conditions,'' *Pattern Recognit.*, vol. 47, no. 8, pp. 2756–2767, Aug. 2014.

[13] A. Borkar, M. Hayes, and M. T. Smith, ''A novel lane detection system with efficient ground truth generation,'' *IEEE Trans. Intell. Transp. Syst.*, vol. 13, no. 1, pp. 365–374, Mar. 2012.

[14] C. Lee and J.-H. Moon, ''Robust lane detection and tracking for real-time applications,'' *IEEE Trans. Intell. Transp. Syst.*, to be published.

[15] X. An, E. Shang, J. Song, J. Li, and H. He, ''Real-time lane departure warning system based on a single FPGA,'' *EURASIP J. Image Video Process.*, vol. 2013, no. 1, p. 38, 2013.

[16] B.-S. Shin, J. Tao, and R. Klette, ''A superparticle filter for lane detection,'' *Pattern Recognit.*, vol. 48, no. 11, pp. 3333–3345, 2015.

[17] X. Du and K. K. Tan, ''Comprehensive and practical vision system for self-driving vehicle lane-level localization,'' *IEEE Trans. Image Process.*, vol. 25, no. 5, pp. 2075–2088, May 2016.

[18] D. Cáceres Hernández, L. Kurnianggoro, A. Filonenko, and K. H. Jo, ''Real-time lane region detection using a combination of geometrical and image features,'' *Sensors*, vol. 16, no. 11, p. 1935, 2016.

[19] S. Jung, J. Youn, and S. Sull, ''Efficient lane detection based on spatiotemporal images,'' *IEEE Trans. Intell. Transp. Syst.*, vol. 17, no. 1, pp. 289–295, Jan. 2016.

[20] M. Aly, ''Real time detection of lane markers in urban streets,'' in *Proc. IEEE Intell. Vehicles Symp.*, Jun. 2008, pp. 7–12.

[21] H. Yoo, U. Yang, and K. Sohn, ''Gradient-enhancing conversion for illumination-robust lane detection,'' *IEEE Trans. Intell. Transp. Syst.*, vol. 14, no. 3, pp. 1083–1094, Sep. 2013.

[22] T.-T. Tran, C.-S. Bae, Y.-N. Kim, H.-M. Cho, and S.-B. Cho, ''An adaptive method for lane marking detection based on hsi color model,'' in *Proc. Int. Conf. Intell. Comput.* Springer, 2010, pp. 304–311.

[23] C. Ma and M. Xie, ''A method for lane detection based on color clustering,'' in *Proc. 3rd Int. Conf. Knowl. Discovery Data Mining*, Jan. 2010, pp. 200–203.

[24] D.-H. Shen, L.-C. Zhang, and E. Xu, "An improved edge detection algorithm based on Sobel operator," *Inf. Technol.*, vol. 4, p. 6, Apr. 2015.

[25] U. Ozgunalp, R. Fan, X. Ai, and N. Dahnoun, "Multiple lane detection algorithm based on novel dense vanishing point estimation," *IEEE Trans. Intell. Transp. Syst.*, vol. 18, no. 3, pp. 621–632, Mar. 2017.

[26] J. Niu, J. Lu, M. Xu, P. Lv, and X. Zhao, "Robust lane detection using two-stage feature extraction with curve fitting," *Pattern Recognit.*, vol. 59, pp. 225–233, Nov. 2016.

[27] S. Srivastava, M. Lumb, and R. Singal, "Improved lane detection using hybrid median filter and modified hough transform," *Int. J. Adv. Res. Comput. Sci. Softw. Eng.*, vol. 4, no. 1, pp. 30–37, 2014.

[28] C. Çiğla, "Recursive edge-aware filters for stereo matching," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. Workshops (CVPRW)*, Jun. 2015, pp. 27–34.

[29] M. Teichmann, M. Weber, M. Zoellner, R. Cipolla, and R. Urtasun. (2016). "Multinet: Real-time joint semantic reasoning for autonomous driving." [Online]. Available: https://arxiv.org/abs/1612.07695

[30] N. Otsu, "A threshold selection method from gray-level histograms," *IEEE Trans. Syst., Man, Cybern.*, vol. SMC-9, no. 1, pp. 62–66, Jan. 1979.

[31] T. R. Singh, S. Roy, O. I. Singh, T. Sinam, and K. M. Singh. (2012). "A new local adaptive thresholding technique in binarization." [Online]. Available: https://arxiv.org/abs/1201.5227

[32] J. Long, E. Shelhamer, and T. Darrell, "Fully convolutional networks for semantic segmentation," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2015, pp. 3431–3440.

[33] A. Geiger, P. Lenz, C. Stiller, and R. Urtasun, "Vision meets robotics: The KITTI dataset," *Int. J. Robot. Res.*, vol. 32, no. 11, pp. 1231–1237, 2013.

CHANG YUAN received the B.S. degree in optoelectronic materials and devices from Harbin Engineering University, Harbin, China, in 2016. She is currently pursuing the M.S. degree in information and communication engineering with the School of Information Science and Engineering, Shandong University, Shandong, China.

Her research interests include computer vision and image processing.

HUI CHEN received the B.S. and M.S. degrees in electronic engineering from Shandong University, Jinan, China, in 1984 and 1987, respectively, and the Ph.D. degree in computer science from The University of Hong Kong, China, in 2002. She is currently a Professor with the School of Information Science and Engineering, Shandong University.

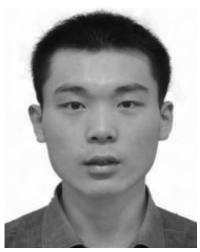
Her research interests include computer vision, 3D morphing, and virtual reality.

JU LIU received the B.S. and M.S. degrees in electronic engineering from Shandong University (SDU), Jinan, China, in 1986 and 1989, respectively, and the Ph.D. degree in Signal Processing from Southeast University (SEU), Nanjing, China, in 2000. Since 1989, he has been with the Department of Electronic Engineering, SDU. From 2002 to 2003, he was a Visiting Professor with the Department of Signal Theory and Communication, Polytechnic University of Catalonia and Telecommunications Technological Centre of Catalonia, Barcelona, Spain. From 2005 to 2006, funded by DAAD, he was a Visiting Researcher with the Department of Communication Engineering, University of Bremen, and the Department of Communication Systems, University of Duisburg-Essen, Germany. In 2009 and 2012, he was a Senior Research Fellow with the Department of Electrical Engineering, University of Washington, Seattle, WA, USA, and also with the Department of Electrical and Computer Engineering, University of California, San Diego, CA, USA, respectively. He has authored or co-authored over 200 journal and conference papers. He is the holder of the Program for New Century Excellent Talents in University of China, and the holder of three conference Best Paper Awards, and five National and Local Academic Awards in communication and signal processing. He is an Associate Editor of the *International Journal of Digital Crime and Forensics*. He is a member of editorial committees of the *Journal of Swarm Intelligence Research* and the *Journal of Electronics & Information Technology* in China. He has been an active Reviewer of the IEEE Transactions on Signal Processing, the IEEE Signal Processing Letters, the IEEE Transactions on Wireless Communications, and the IEEE Communications Letters.

His current research interests include space-time processing in wireless communication, blind signal separation, and multimedia communications.

DI ZHU received the B.S. degree in information science and engineering from Shandong University, Shandong, China, in 2018, where he is currently pursuing the M.S. degree in information and communication engineering with the School of Information Science and Engineering, Shandong University.

His research interests include computer vision, image processing, and ADAS.



YANYAN XU received the B.S. and M.S. degrees from the School of Information Science and Engineering, Shandong University, in 2007 and 2010, respectively, and the Ph.D. degree in pattern recognition and intelligent systems from the Department of Automation, Shanghai Jiao Tong University, China, in 2015. He is currently a Post-Doctoral Associate at the Human Mobility and Networks Lab, Department of Civil and Environmental Engineering, Massachusetts Institute of Technology, a Visiting Scholar at the Department of City and Regional Planning, UC Berkeley, and a Guest Post-Doctoral Fellow at the Energy Analysis and Environmental Impacts Division, Lawrence Berkeley National Laboratory, Berkeley, CA, USA.

His research interests are human mobility and urban computing, with particular emphasis placed on the use of massive trajectory data in intelligent transportation systems, urban planning, environment, and energy from the interdisciplinary perspective. His works have been published in nature energy, the IEEE Transactions on Intelligent Transportation Systems, and the *Journal of the Royal Society Interface*.