

A Sensor-Fusion Drivable-Region and Lane-Detection System for Autonomous Vehicle Navigation in Challenging Road Scenarios

Qingquan Li, Long Chen, Ming Li, Shih-Lung Shaw and Andreas Nüchter

Abstract—Autonomous vehicle navigation is challenging since various types of road scenarios in real urban environments have to be considered, especially when only perception sensors are used, without position information. This paper presents a novel real-time optimal-drivable-region and lane detection system for autonomous driving based on the fusion of Light Detection and Ranging (LIDAR) and vision data. Our system uses a multisensory scheme to cover the most drivable areas in front of the vehicle. We propose a feature-level fusion method for LIDAR and vision data and an optimal selection strategy for detection of the best drivable region. Then a conditional lane detection algorithm is selectively executed depending on an automatic classification of the optimal drivable region. Our system successfully handles both structured and unstructured roads. The results of several experiments are provided to demonstrate the reliability, effectiveness, and robustness of the system.

Index Terms—Autonomous vehicles, drivable-region detection, lane detection, multilevel feature fusion, LIDAR, vision.

I. INTRODUCTION

ROAD/LANE detection is a challenging task and is a critical issue for autonomous vehicle navigation. Particularly in situations where no position information is available, the navigation system must be aware of different kinds of terrain and road situations without the need for user input. This paper presents a real-time-capable road and lane detection system that deals with various kinds of challenging situations in real-world urban scenarios.

The complexity of urban environments is mainly due to the following factors:

- 1) Structured and unstructured roads occur alternately. Fig. 1(b) and (c) show structured roads, and the other subfigures present examples of roads without lane markings.

Copyright (c) 2013 IEEE. Personal use of this material is permitted. However, permission to use this material for any other purposes must be obtained from the IEEE by sending a request to pubs-permissions@ieee.org.

Qingquan Li is with the Shenzhen Key Laboratory of Spatial Smart Sensing and Services, Shenzhen University, 3688 Nanhai Avenue, Nanshan District, Shenzhen, Guangdong, China. (e-mail: liqq@szu.edu.cn).

Long Chen is with the School of Mobile Information Engineering, Sun Yat-sen University, Zhuhai, Guangdong, P.R.China (corresponding author to provide phone: (+86756)3668567; e-mail: chenl46@mail.sysu.edu.cn).

Ming Li is with the Computer School, Wuhan University, Hubei, China (e-mail: liming751218@gmail.com).

Shih-Lung Shaw is with the Department of Geography, University of Tennessee, Knoxville, TN, USA (e-mail: sshaw@utk.edu).

Andreas Nüchter is with the Robotics and Telematics group at University of Würzburg, Am Hubland, D-97074 Wrzburg, Germany (e-mail: andreas@nuechti.de).

- 2) Pavement uniformity cannot always be taken as given: there is interference from many causes, such as heavy shadow, cf. Fig. 1(d) and (h), pavement distress, dirt, and puddles. Fig. 1(e) shows a road where oval stones and concrete are present, and Fig. 1(b) shows a road that has different colors and some dirt on it.
- 3) The appearance of a road may change frequently because of weather conditions, for example rain, snow as in Fig. 1(g), and it also changes between daytime, dusk as in Fig. 1(c), and night.
- 4) The curvature of a road is not always as low as it is in highway scenarios. Here, we use “highway” in the sense of a paved, main, direct road, in contrast to a minor road. Figure 1(a) shows a sharp turn where a camera at the front does not cover all of the turn.

Focusing on these challenging situations, we propose a multi-cue fusion-based system. By efficiently using several laser scanners and cameras, our perception system figures the optimal drivable region out and detects lane markings if necessary. Our real-time road and lane detection system is distinguished from related approaches in the following ways:

- Our system deals reliably with challenging urban environments including both structured and unstructured roads in real time, we estimate whether or not to need to do lane detection based on proposed fused method, without manual switching or using information from a Global Positioning System (GPS) and a Geographic information system (GIS). In the case of structured roads, the lanes and the road edges are located. For unstructured roads, the system detects the drivable region and the boundaries of the road.
- A fusion-based method is proposed. Feature-level fusion is used for drivable-region detection. The lane detection method is restricted to the optimal drivable region and is only applied when the road is estimated to be wide enough.
- The proposed strategy extracts optimal drivable regions in front of the vehicle instead of recognizing every pixel of the road surface, which we believe is to be too time-consuming and unnecessary for autonomous vehicle driving.

The methods developed here have potential application in cost-effective driver assistance systems. We used a simple setup and fused algorithms, since we did not want to rely on expensive hardware such as the high-end Velodyne HDL-

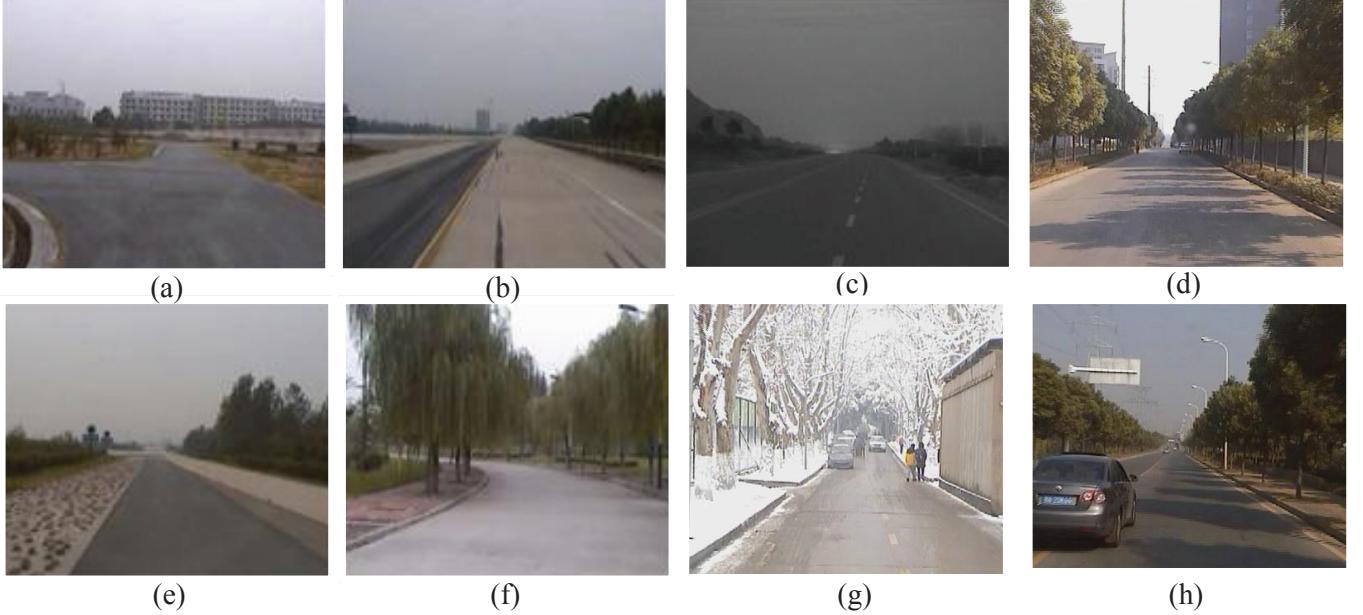


Fig. 1. Challenging urban road scenarios: there may be both structured and unstructured roads; interference from dirt; faded (FIXME) lane markings; different road materials, including asphalt and concrete; different illumination conditions depending on the time of day and season; and interference from shadows and vehicles ahead.

64ES2 sensor. However, we use this sensor to benchmark our approach.

Our system was successfully tested in the China Future Challenge 2010, an intelligent-vehicle competition. The car running the software described here was the only system that completed the course in the allotted time.

II. STATE OF THE ART

A. Lane detection

Several effective lane detection systems were proposed over a decade ago, such as AWSTM, AutoVue, RALPH ([1], [2], [3]), AURORA [4], SCAR [5], GOLD ([6], [7]), and LOIS [8].

More recently, some researchers have focused on lane detection on highways and other structured roads ([9], [10], [11], [12]). Wang *et al.* [13] constructed a Catmull–Rom spline-based lane model to describe the effect of perspective on parallel lines. Jung and Kelber [14] proposed a linear–parabolic model for lane following, where a linear model was designed for the near-vision field and a quadratic model was used for the far field. This method provides a good description of roads with a curb; its main limitation is related to the significant occlusion of lane markings due to vehicles in front of the camera. To solve this problem, Cheng *et al.* [11] applied a color-information-based method that utilizes size, shape, and motion information to distinguish occlusions from real lane marks. In addition, Zhou *et al.* [15] introduced a deformable-template model of the projection of the lane boundaries, and then the lane detection problem was formulated as a problem of maximization of an a posteriori estimate. The parameters of the model completely determine the position of the vehicle inside the lane, its heading direction, and the local structure of the lane. In [12], steerable filters were used for robust and accurate lane detection. These filters provide an efficient method

for detecting circular-reflector markings, solid-line markings, and segmented-line markings under varying lighting and road conditions. For more challenging structured roads, Kim [16] introduced a robust real-time lane-detection algorithm based on Random Sample Consensus (RANSAC), and this was combined with a particle-filtering-based tracking algorithm using a probabilistic grouping framework. Gopalan *et al.* [17] also proposed a learning-based approach to address several challenging issues such as appearance variations in lane markings caused by factors such as occlusion, shadows and changing lighting conditions of the scene. Recently, some researchers have also focused on lane-keeping systems, e.g., [18]. Laser sensors have been used for lane detection using reflection intensity information, e.g., [19]. But purely reflection-based lane detection is not sufficient, and vision information is needed, such as [20]. There have been laser scanner-based systems with high levels of accuracy and reliability for lane detection [21], [22], [23]. Laser scanners are more suitable for detecting road surfaces using distance information [24].

B. Road detection

Although these algorithms are effective enough for structured roads, they are not satisfactory in complicated real-world urban road environments where lane markings are not always present. More recently, the research priority has shifted to the detection of unstructured roads, i.e., secondary roads and off-road conditions [25], [26], [27], [28], [29], [30]. In the absence of lane markings, color and texture features are mainly used. He *et al.* [25] proposed two steps: first, boundaries are estimated, and then road areas are detected based on a full-color image. Tarel and Bigorgne [31] also proposed a region-growing segmentation with a dedicated pre-smoothing of the

image, and this method was improved by the use of radiometric calibration. But this is an offline algorithm that cannot be used for autonomous driving. Álvarez and Lopez [28] used a shadow-invariant feature space in combination with a model-based classifier. The model was built online to improve the adaptability of the algorithm to the current lighting conditions and the presence of other vehicles in the scene. Stereo vision has also been applied frequently, e.g., [26], [27], [32]. In [26], the authors presented a complex trinocular vision system for obstacle and path detection, in which stereo system is used for pitch estimation and obstacle detection and a monocular vision system is used to detect the drivable path. Wedel *et al.* [27] modeled the road surface as a parametric B-spline curve and used stereo measurements to estimate the surface parameters. Fang *et al.* [29] also used multi-views camera, but the aim is to build an accuracy map instead of the real-time road detection. Danescu *et al.* [32] uses the information from stereovision and grayscale image through a particle filtering framework to detect and track lanes in challenging roads. Because of the ability of stereo to obtain depth information of the scenarios, which is beneficial for detecting obstacles on the road in sensing data. Most pure vision-based road detection approaches to obtaining accurate and complete road areas are complex and time-consuming because the methods must overcome the luminance problem, which is a difficult problem in machine vision. Although stereo vision offers some assistance, the short effective range restricts its applicability.

In addition to the vision sensors, active sensors have increasingly been used, such as millimeter-wave radar sensors [33] and laser scanners ([33], [34], [35], [36], [37], [38]). Cramer *et al.* [35] applied a single-scan system to the detection of the road border, then a tracking algorithm is responsible for border estimation. Based on [35], Fardi *et al.* [36] fused reflectivity information from the laser to detect the road border. [37] presents another single-scan system, where a new laser feature based on equal angular separation by three consecutive laser data is introduced. This is also used as one laser feature in our method. [38] proposed a laser-scanner based method for perceiving a large dynamic urban environment, it emphasises on solving SLAM problems instead of road detection. Because of their characteristics, most active-sensor-based methods focus only on road extraction and boundary detection for country roads or off-road conditions; for lane markings, it is hard for active sensors to do better than vision even with the help of reflection feature evaluation. Besides, one frame of data from one single-line laser scanner is typically not enough to represent most of the area in front of a vehicle. Thus, some researchers have used the GPS in conjunction with a GIS to provide supplementary information. However, the GPS has limited resolution. In addition, GIS are often absent, especially in developing countries, or just not updated frequently enough. Since all one-class sensor approaches have limitations, Multi-sensor fusion-based road detection has also been proposed [39], [40]. In [39], LIDARs, cameras, and the GPS were used. Thrun *et al.* [40] added GIS information, but the difficulty remains as the GPS is unavailable in many places.

In the work described in the present paper, we have addressed the problem of obtaining accurate road and lane

information for autonomous driving of vehicles in an unknown challenging urban scenario. Our methods focus on detecting optimal drivable regions and lanes within the optimal drivable region in front of an intelligent vehicle, without using any position sensors. We did not aim at extracting every pixel belonging to the road. Instead, inspired by the work described above, we built our system using laser scanners, vision, and a clever fusion-based method. Our method is somewhat similar to those of [41], [42] and our previous work [43]. Broggi *et al.* [41] adopted a data-level fusion method, where each sensor module generates a bird's-eye view bitmap, encoding in each pixel a grayscale color proportional to the chance that an obstacle is present in that area. A static weighting is then applied to the bitmaps based on the reliability of the sensor in the region of interest. This is a good method, especially for obstacle detection in off-road scenarios. However, for structured roads and for roads without boundaries or obstacles, these methods cannot identify explicit drivable regions and lane markings. Manz *et al.* [42] used a decision-level fusion method, in which the scans were rated independently in a way that is appropriate to each sensor and the information gathered was subsequently fused during the process of path selection. Both, cameras and laser scanners have been used in [43]. But the video data and the point cloud data were processed independently and not fused. Different from these three methods, we propose a method based on a feature-level fusion combined with a conditional lane detection method.

According to the information presented above, the road and lane detection system dealing with both structured roads, i.e., roads with lane markings, and unstructured roads, i.e., roads without lane markings is necessary for autonomous driving. Moreover such road situations are frequently found around us. Actually, there have been several solutions to solve this problem, such as [44], [45]. Rauskolb *et al.* [44] proposed two great methods to detect lane and road area respectively. However they did not give us a clear method about how to classify these two situations. In addition, an Inertial Measurement Unit (IMU) and GPS were also been used for data acquisition, which were not used in our system. Another effective method in [45] is similar to our method in the process flow. There is a key problem that the average process time of the method for unstructured frame is over 1.3s and thus is unable to meet the real-time requirement. Consequently, this paper aims at proposing a system that implements detection of both structured and unstructured roads in real-time without manual selection or GPS/GIS data.

C. Proposed system

According to the above summary, a perception system for autonomous driving needs two specific abilities if the aim is that unmanned vehicles are driven on real urban roads: (1) the system must be able to handle both structured and unstructured roads; and (2) it must be able to overcome the problems presented by various kinds of interference on the road.

With reference to Fig. 2(a), if an unmanned vehicle drives from position 1 to 3 through 2, it has to determine when it

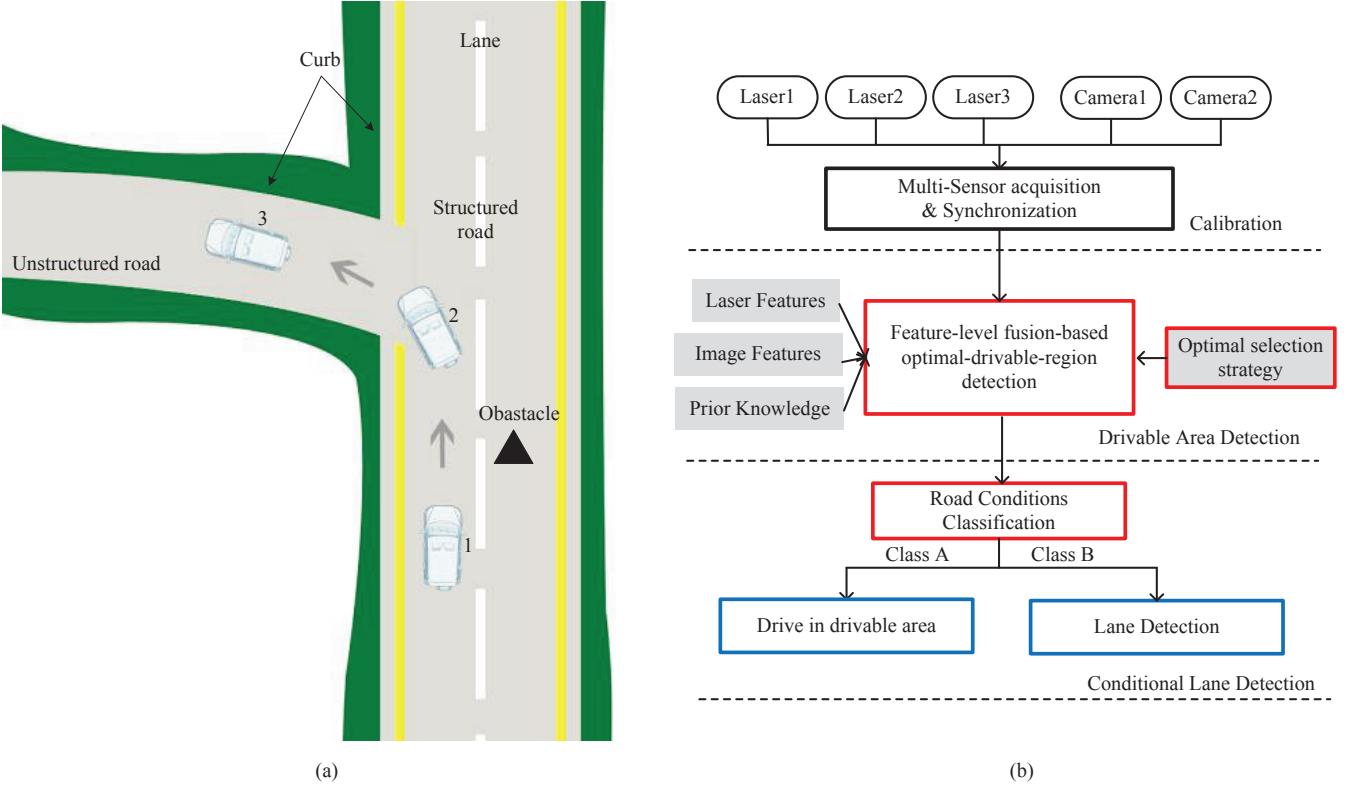


Fig. 2. (a) Scene description. (b) Flowchart of algorithm.

should follow a lane and when it should not. Instead of basing the analysis on the complete road conditions, we classify the current drivable region obtained by our proposed fusion algorithm into two classes: narrow regions, where there is no need to detect lanes, and wide regions, where there is a need to detect lanes. This means that even when there are lanes on the road, our system may not detect them at all if the drivable region is not wide enough.

In addition to this, our system has its unique way of dealing with interference on the road. Instead of striving to detect the road based on a global model, we have chosen a method that proceeds from local optimization to a global optimization. First we detect curb points, and then we filter and fuse segments divided by the curb points into drivable lines. We have also designed an optimal selection strategy to deal with cases where there is more than one drivable line. The optimal drivable region is formed by three optimal drivable lines. According to this idea, our system generates an autofit region instead of using the whole set of road conditions.

The algorithm of our system has two main parts: first, feature-level fusion-based optimal-drivable-region detection, and second, conditional lane detection. The first step ensures that the vehicle is driven in a safe area, while the second step limits the lane detection to the drivable region.

Fig. 2(b) shows the overall design of our system. We first determined the sensor layout and performed calibration. After data acquisition, the system fuses laser features, image features, and additional prior knowledge for the purpose of optimal-drivable-area detection. Then, the system classifies the present road based on previous results. Based on this

classification, the algorithm decides whether or not it should detect lane markings. If lane detection is done, the method is applied in the drivable area only.

III. SYSTEM DESCRIPTION

The system described in this paper was developed to be used on an automatic vehicle of Wuhan University called SmartV-II (Fig. 3(a)). SmartV-II is equipped with eight laser scanners and seven cameras, three lasers and two cameras of them were used for road detection.

Fig. 3 shows the layout of the laser scanners and cameras. The laser scanners were SICK-LMS291 single-line scanners, which generates a vector of 181 range values, spaced 0.5° apart over a field of view of 90° , every 26 ms. Fig. 3(a) shows the three lasers, labeled L1, L2, and L3. L1 was mounted on the roof, and L2 and L3 were mounted in the head of vehicle. All of the laser scanners were tilted downward to scan the road ahead. We adjusted the pitch angles ρ_{l_1} , ρ_{l_2} , and ρ_{l_3} to ensure that the lasers scanned the area in front of the vehicle at different distances. The advantage of using different scan distances is that the three single-line scanners form a plane in front of the vehicle and thereby increased the accuracy of the system. The laser with the longest detection distance plays also a warning role in case there is no drivable area ahead.

For vision, we used two analog cameras (C1 and C2) with a frame rate of 25 fps and an image resolution of 640×480 pixels. Both cameras were mounted on the car roof, but with different pitch angles ρ_{c_1} , ρ_{c_2} and different roll angles θ_{c_1} , θ_{c_2} . These angles are adjusted to cover the optimum range.

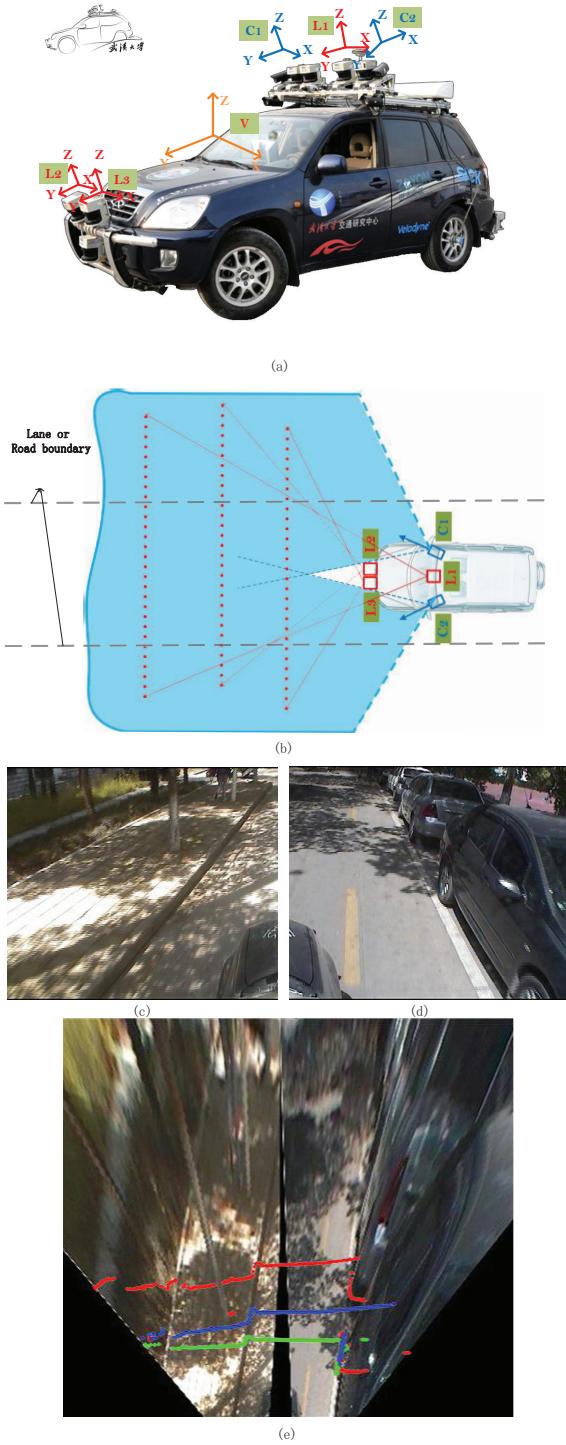


Fig. 3. SmartV-II sensor layout and data fusion. (a) Photograph of the intelligent vehicle SmartV developed by Wuhan University. L1, L2, and L3 are laser scanners; their coordinate systems are shown by X, Y, and Z in red. C1 and C2 are cameras; their coordinate systems are shown by X, Y, and Z in blue. V indicates the vehicle; its coordinate system is shown by X, Y, and Z in orange. (b) Coverage of the system in a bird's-eye view. The area in blue is the field of view of the two cameras; the three red dashed lines show the point cloud obtained from the laser scanners. (c), (d) Real images captured by the left and right cameras. (e) All of the data in one bird's-eye view image after calibration; the red, blue, and green lines show the points from the three laser scanners.

In our system, we ensured that the field of view could cover 6 m to the right and left of the vehicle and more than 10 m ahead of the vehicle. The camera coordinate systems for C1, C2 and the laser scanner coordinate systems for L1, L2 and L3 are given in Fig. 3(a). Calibration was performed using OPENCV functions [46] and the Camera Calibration Toolbox for MATLAB; the algorithm applied was mainly the one from [47].

All of the data were transformed into the vehicle coordinate frame (denoted by V in Fig. 3(a)), in which y is the direction in which the vehicle is moving, and z is the upward-pointing vertical axis. After calibration, all data are transformed to a bird's-eye view (Fig. 3(c) and Fig. 3(d)). Fig. 3(e) presents registered camera images and point clouds from the laser scanner. After these transformations, our multicue fusion method for drivable-area detection is started.

IV. SENSOR FUSION ALGORITHM

The algorithm for our system has two main parts: first, feature-level fusion-based optimal-drivable-region detection, and second, conditional lane detection.

A. Feature-level fusion-based optimal-drivable-region detection method

The so-called drivable region is where “the vehicle moves on safely in a physical sense, but without taking symbolic information into account (e.g., one-way-streets and traffic signs)” [48]. In most cases, the drivable region is a part of a road area divided by obstacles and curbs where a vehicle can pass through safely. In addition to finding all the drivable regions in an area, our system identifies an optimal region within these regions. In this paper, a method based on feature-level fusion of laser-scanner and video data is proposed for detection of the optimal drivable region. The flow of the algorithm is shown in Fig. 4(a). Besides the feature-level fusion method, a multilayer perception strategy is used. In layer 1, point cloud features and image features are fused for detection of curb points. “Curb points” refer here to the edges of obstacles and to road curbs, which divide the road into several segments. The segments are then filtered in layer 2, based on point cloud features and prior knowledge. After that, color features located in a square around the segments are used for segment clustering to obtain the drivable area in layer 3.

1) *Curb point detection*: In addition to different elevation characteristics, color variability will also occur in most road curbs owing to shadows from the curb and differences in color between the road surface and the roadside. To detect curb points with high reliability, we use feature-level fusion of laser features and color features in combination with a classifier-based method. Each point from the region where the laser and vision data overlap is represented by a feature vector made up of geometric and color features extracted from the point clouds and the images. This vector forms the input to the later stage of classification.

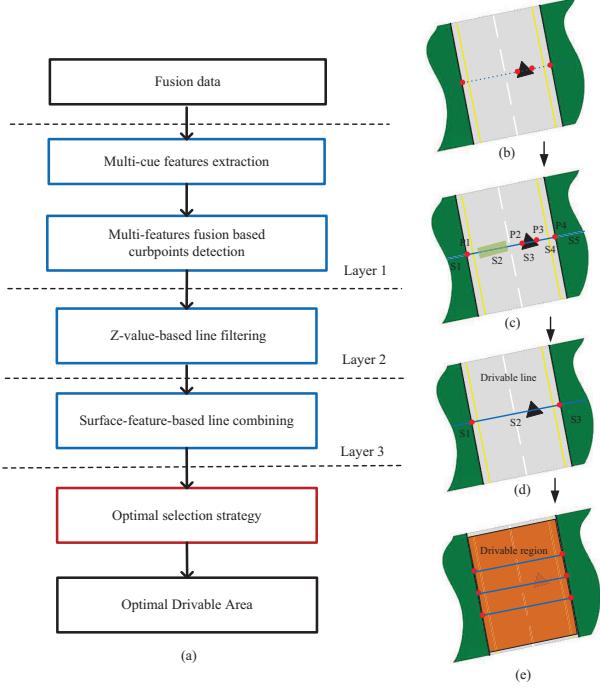


Fig. 4. (a) The flowchart of the feature-level, fusion-based, and drivable-region detection algorithm. (b)–(e) Step-by-step schematic illustration of the optimal-drivable-region detection method.

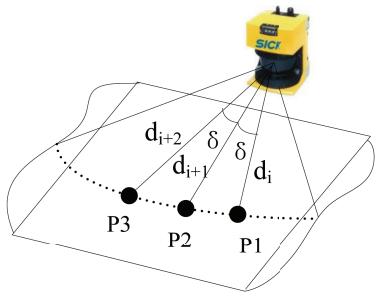


Fig. 5. Three consecutive laser data points on a flat road surface [37].

a) *Laser features*: An “angle bisector feature” [37] was chosen as one of the laser features to be used in the analysis (Fig. 5). Consider the three points P_1 , P_2 , and P_3 , on the line segment denoted by the dotted line, at ranges from the LIDAR measuring system of d_i , d_{i+1} , and d_{i+2} , respectively. The range measurements are obtained at equal angular separations of δ degrees. We see that

$$d_{i+2} = \frac{d_i d_{i+1}}{2d_i \cos \delta - d_{i+1}}. \quad (1)$$

The difference between the actual value of the range and the desired value,

$$A = d'_{i+2} - d_{i+2}, \quad (2)$$

where d'_{i+2} denotes the actual value of the range, is defined as the angle bisector feature. The Z -value, Y -value, Z -variance, and Y -variance of the line segment around the points are also

chosen as features, i.e.,

$$\bar{Z} = \frac{1}{9} \sum_{k=i-4}^{i+4} Z_k \quad (3)$$

$$\bar{Y} = \frac{1}{9} \sum_{k=i-4}^{i+4} Y_k \quad (4)$$

$$\text{Var}(Z) = \frac{1}{9} \sum_{k=i-4}^{i+4} (Z_k - \bar{Z})^2 \quad (5)$$

$$\text{Var}(Y) = \frac{1}{9} \sum_{k=i-4}^{i+4} (Y_k - \bar{Y})^2 \quad (6)$$

where $\text{Var}(Z)$ and $\text{Var}(Y)$ are the Z -variance and Y -variance of nine consecutive points, \bar{Z} and \bar{Y} are the mean value of the nine points.

b) *Color features*: In general, there will be a color change around a curb point, due to such things as dark shadows and green grass. Because of small inaccuracies in data registration, we have taken a 5×5 block of image information in the present version instead of only one-pixel features. Based on [49], a unique combination of color channel inputs was chosen for each block. The mean values of red and green channels in the RGB color space and the H channel in the HSV color space were chosen. Empirically, the red/green combination gives a strong, yet differentiable response to both yellow and white (road markings) in addition to capturing the properties of road-to-roadside transitions in urban and rural environments. To represent color changes, the gradients of each channel were also chosen as features, i.e.,

$$R_{kj} = \frac{1}{25} \sum_{k=i-2}^{i+2} \sum_{l=j-2}^{j+2} r_{kj} \quad (7)$$

$$G_{kj} = \frac{1}{25} \sum_{k=i-2}^{i+2} \sum_{l=j-2}^{j+2} g_{kj} \quad (8)$$

$$H_{kj} = \frac{1}{25} \sum_{k=i-2}^{i+2} \sum_{l=j-2}^{j+2} h_{kj} \quad (9)$$

and

$$\nabla R_{kj} = \left(\frac{\partial R}{\partial k}, \frac{\partial R}{\partial j} \right)^T \approx (DR_k, DR_j)^T \quad (10)$$

$$|\nabla R_{kj}| \approx |DR_k| + |DR_j| \quad (11)$$

$$\nabla G_{kj} = \left(\frac{\partial G}{\partial k}, \frac{\partial G}{\partial j} \right)^T \approx (DG_k, DG_j)^T \quad (12)$$

$$|\nabla G_{kj}| \approx |DG_k| + |DG_j| \quad (13)$$

$$\nabla H_{kj} = \left(\frac{\partial H}{\partial k}, \frac{\partial H}{\partial j} \right)^T \approx (DH_k, DH_j)^T \quad (14)$$

$$|\nabla H_{kj}| \approx |DH_k| + |DH_j|, \quad (15)$$

where R_{kj} , G_{kj} and H_{kj} are the mean value of red, green and hue channel of the block, ∇R_{kj} , ∇G_{kj} and ∇H_{kj} refer the gradient of red, green and hue channel respectively. DR_k refers the gradient of red channel in the direction of k . And so on, for DG_k , DH_k , DR_j , DG_j , DH_j .

c) *Fused feature vector and classifiers:* The concatenated feature vector V of one point is a 11-D vector and is denoted by

$$V = \{A, Z, \text{Var}(Z), Y, \text{Var}(Y), R, G, H, |\nabla R|, |\nabla G|, |\nabla H|\} \quad (16)$$

Finally, the problem of curb-point detection using these feature descriptor vectors is formulated as a supervised machine-learning two-class classification problem with labeled classes. We utilize several classification approaches to illustrate the potential performance of an optimized noise-tolerant classification approach, using an artificial neural network (ANN) [50], a support vector machine (SVM) [51], and adaptive boosting (AdaBoost) [52].

Our ANN was a multilayer network of interconnected perceptrons trained to discriminate between specified output classes using a training methodology derived from gradient descent, i.e., backpropagation [50]. The use of ANNs is a well-established method; see [50] for further details.

SVM is a generation learning system based on recent advances in statistical learning theory. SVMs are based on the structural risk minimization principle, closely related to regularization theory. An SVM model is a representation of the examples as points in space, mapped such that the examples in separate categories are divided by a wide region around a hyperplane. New examples are then mapped into that same space and predicted to belong to a category based on which side of the plane they fall [53].

AdaBoost is adaptive in the sense that subsequent classifiers that are built are tweaked in favor of those instances misclassified by previous classifiers. AdaBoost is sensitive to noisy data and outliers. However, in some problems it is less susceptible to the overfitting problem than most learning algorithms.

2) *Z-value-based line filtering:* Curb points divide one scan line into several segments. In this step, we judge whether the vehicle can drive through each segment, based on the

following equation:

$$\begin{cases} \text{if } Z_{S_i \max} < Z_{th}, \text{ segment } i \text{ is drivable.} \\ \text{else, segment } i \text{ is not drivable.} \end{cases} \quad (17)$$

Here $Z_{S_{max}}$ refers to the maximum Z -coordinate in a particular line segment. Z_{th} refers to the threshold for line filtering. In Fig. 4(c), segment S3 satisfies the above condition in Eq. (17). Thus, it is considered as part of the drivable line, and will be connected to other parts of the drivable line. Because of the black obstacle in the road, the road is divided into several parts. Thus, we combine the segments, for example S2, S3, and S4.

3) *Surface-feature-based line combining:* This step combines adjacent segments based on surface features. The combined segments are considered as either drivable lines or undrivable lines. In the system described in this paper, segments are combined if the points on them satisfy the following equation:

$$\begin{cases} |\bar{Z}_{S_i} - \bar{Z}_{S_{i+1}}| < Z'_{th} \\ |\bar{H}_{S_i} - \bar{H}_{S_{i+1}}| < H'_{th} \end{cases} \quad (18)$$

Only when both, the height values and the hue values are similar enough, we consider them as a combined segment. \bar{Z}_{S_i} and \bar{H}_{S_i} refer to the mean height and the mean hue of segment s_i segment s_i respectively. The Z'_{th} and H'_{th} refer to the corresponding threshold and is assigned according to the prior knowledge. Because there is sometimes grass beside the curb at the same height as the road, combination based purely on the height value is not accurate enough. In this step, S2, S3, and S4 (see Fig. 4(d)) are combined.

4) *Optimal selection strategy:* Many times, more than one drivable line is detected for each laser scanner. For autonomous driving, one does not need to know all the drivable lines; detecting the optimal one is sufficient. To find the optimal line out of all the detected drivable lines, we propose an optimal selection strategy based on three properties of each line: the mean of the Z -values, \bar{Z}_{L_j} ; the distance between the midpoint of the line and the front of the vehicle, dis_{L_j} ; and the width of the line, $width_{L_j}$. Here L_j means the j th line. We consider a drivable line as the optimal drivable line when its height is as close as possible to the ground, the distance between the line and the vehicle is as small as possible, and the width of the line is as large as possible, taking account of the continuity of a road. The reason of choosing the nearer line is that the laser with the shortest detection distance was least affected by interference and could obtain the most accurate results. To evaluate each drivable line with these three rules, we designed an evaluation function, shown in the following equation:

$$\begin{aligned} \text{Index} = & m_1 * e^{-|(Z_{L_j} - Z_l)/Z_l|} \\ & + m_2 * e^{-|(dis_{L_j} - dis_{min})/dis_{min}|} \\ & + m_3 * e^{-|(width_{L_j} - width_{max})/width_{max}|}, j \in \text{drivable lines} \end{aligned} \quad (19)$$

In this equation, we assign weights (m_1, m_2, m_3) to each rule to adjust the influence of the rules. Here, Z_l is the Z -value of the surface in the last frame and the initial value is assigned

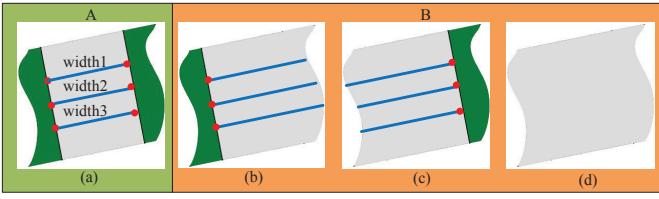


Fig. 6. Road classification based on minimum width of optimal drivable region. (a) belongs to A, where there is no need for lane detection; (b), (c), and (d) belong to B, and need lane detection.

based on priori information, dis_{\min} is the minimum distance between the midpoint of each drivable line and the front of the vehicle, and width_{\max} is the maximum width of all of the drivable lines. The line that maximizes the value of Index is the optimal drivable line. In our system, we set $m_1 = 0.85$, $m_2 = 0.1$, and $m_3 = 0.05$ on the basis of experience.

For each data frame, the optimal drivable region can be constructed only when all the three optimal drivable lines are obtained. By connecting the left curb points of each optimal drivable line one by one, we can construct the left edge of the drivable region, and the right edge similarly. Then the region between the two edges is the optimal drivable region, which is what we are interested in.

Using this optimal drivable region, we ensure safe driving of the autonomous vehicle on either structured or unstructured roads. To follow lanes on structured roads, our vehicle first determines whether lane detection is necessary in the current road scenario. Our system uses a simple classification method based on the detected optimal drivable lines. Based on the minimum width of the optimal drivable region, we classify the current road into two types (A and B in Fig. 6). If $\min(\text{width}_1, \text{width}_2, \text{width}_3) > \text{width}_{th}$, we detect lanes in the region (Figs. 6(b), (c), (d)); otherwise, we assume that there are no lane markings on the road (Fig. 6(a)). Usually, we chose width_{th} equal to 1.5 times the standard width of a lane based on experience.

B. Conditional lane detection

Instead of considering a complete image, we focus on the optimal drivable region computed as described above. To get more lane information from the image, we extend the top and bottom of the optimal drivable region to the top and bottom of the image. Fig. 7 presents the flow of the lane detection algorithm. First, a top-hat transform and a Prewitt vertical-gradient operator are used for preprocessing, and then an adaptable-threshold method is used to obtain a binary image, followed by a progressive probabilistic Hough transform (PPHT), which, in turn, is used for line detection.

1) *Top-hat transform for preprocessing:* We use a top-hat transform [54] as preprocessing to enhancing contrast and reduce the interference of some non-lane markings. In the field of mathematical morphology, the chosen top-hat transform is an operation that extracts brighter small elements and details from given images. The formula for this operation is

$$h = f - (f \circ b) \quad (20)$$

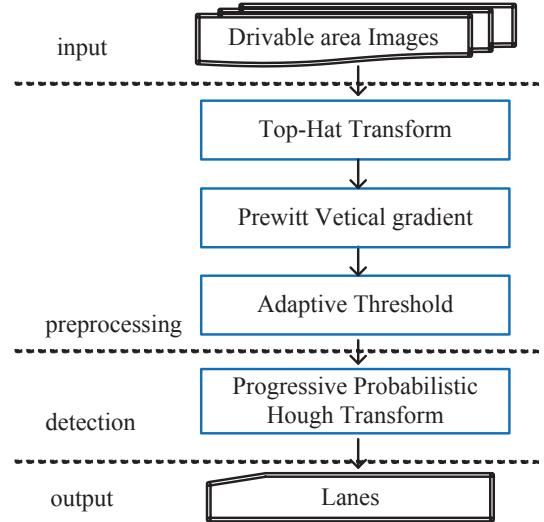


Fig. 7. Flowchart of lane detection algorithm.

where f denotes the source image, h denotes the result image obtained after applying the top-hat transform, and \circ denotes an opening operation that is performed by the top-hat transform and is controlled by the choice of the structuring element b . The larger the size of the structuring element b , the larger the elements extracted are. The top-hat transformation is aimed at suppressing slow trends, and therefore it enhances the contrast of some features in the images, according to a size or shape criterion. The top-hat transformation extracts the objects that have not been eliminated by the opening operation. That is, it removes objects larger than the structuring element. In Fig. 8(c), one sees that the top-hat transform enhances the contrast and removes some nonlane markings effectively. In most cases, the vehicle is parallel to a lane, and so vertical-gradient detection is used.

2) *Local OTSU thresholding:* Applying a threshold is a key preprocessing operation in lane detection. To deal with the illumination problem, adaptation of the threshold value is needed. In the bird's-eye view, most lanes are located within a narrow rectangle, so we cut the image into two rectangles (see Fig. 8(e), where the image is divided into left and right regions), and the OTSU algorithm [55] was used in each rectangle for threshold segmentation. Fig. 8(e) shows the result obtained from our threshold method.

3) *PPHT:* The PPHT [56] is a variation of the standard Hough transform (HT) [57]. PPHT minimises the amount of computation needed to detect lines by exploiting the difference in the fraction of votes needed to reliably detect lines with different numbers of supporting points. The fraction of points used for voting need not be specified ad hoc or using a priori knowledge, as in the Probabilistic Hough Transform. The algorithm is ideally suited for real-time applications with a fixed amount of available processing time, since voting and line detection are interleaved. The most salient features are likely to be detected first. Fig. 8(f) presents a result obtained by applying the PPHT. Finally, we obtain lanes based on the location and scope of the lines (see Fig. 8(g)).

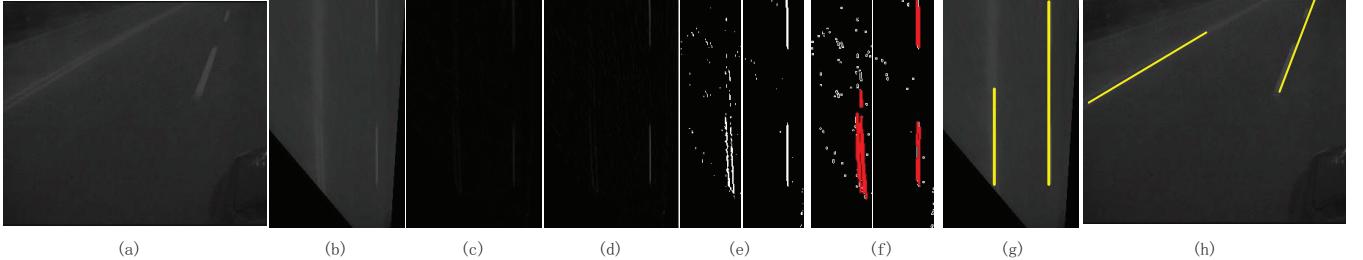


Fig. 8. Vision-based lane detection results. (a) Original image; (b) bird's-eye view after homography transform; (c) result of top-hat transform; (d) result of vertical-gradient detection by Prewitt operator; (e) result after local OTSU thresholding, after the image was divided into two parts; (f) result after PPHT; (g) result after line combining; (h) final result.

V. EXPERIMENTS AND RESULTS

To verify the effectiveness of the proposed method, we performed experiments where we applied our algorithm to registered video and point cloud data offline. We present these experiments first. We then present the performance of our vehicle in the China Future Challenge using the proposed system.

A. Experiments with offline video and point cloud data

We now present experiments in which our algorithm was applied to three offline registered video and point cloud data sequences. In the three data sequences chosen, there were structured roads, unstructured roads, and transitions between these two types of roads. All three data sequences were acquired by the SmartV vehicle (see Fig. 3(a)).

Data sequence 1 was taken at Chang'an University, Xi'an, China. Its total length was 3 min 50 s. It contains campus roads and highways, and transitions between them. Data sequence 2 was taken at the campus of Wuhan University, Wuhan, China. Its total length was 2 min 45 s. It contains mainly roads without markings. Data sequence 3 was taken on Shuiguo Lake Road, Wuhan. Its total length was 4 min 20 s. This sequence contains many multilane roads, mostly urban roads. In all of these sequences, the frame rate was 10 frames/s, and thus the three sequences contained 2300, 1650, and 2600 frames, respectively.

1) *Curb point detection results:* The set of manually labeled examples used for training the three classification approaches was made up of 1000 points (300 positive points and 700 negative points) collected from above data sets. The experiment presented here aimed at choosing the optimum classifier for our fused features. We have tested our fused features with an ANN, an SVM, and Adaboost. We use receiver operator characteristic (ROC) curves, which show how the number of correctly classified positive examples varies with the number of incorrectly classified negative examples, to represent the classification results. The curb point detection results from the three learning machines are shown in Figs. 9 and 10. For the ANN, we have employed a classical two-layer network topology with H hidden nodes, one input node per feature vector entry, i.e., 10 inputs, and one output node per class. The ANN was trained using I iterations of the backpropagation algorithm [58]. Considering the general ranges of the parameters $H = 5, 10, 20, 30$ and $I = 1000, 2000$, we present

only a subset of the ANN results over these parameter ranges in Table I, based on testing the above three video sequences. The best classification result was achieved when $H = 10$ and $I = 1000$, and the worst result occurred when $H = 5$ and $I = 1000$. When the result was close to the peak result, increasing the values of H and I did not improve the result.

Similar experiments have been carried out using AdaBoost. The classification performance was 82% which was a little better than that of the ANN 80%, because AdaBoost needs a large number of training examples, and so only 1000 examples might be a little unsatisfactory. For SVM, the radial basis function (RBF) is applied with the form $K(u, v) = \exp(-\gamma * |u - v|^2)$, where the kernel parameter γ and soft margin C are selected from $[2^{-2}, 2^{-1}, 1, 2, 2^2]$ and $[10^{-1}, 1, 10]$, respectively, by using 3-folds cross validation on the training set. Based on this SVM, we can get classification performance at 86%. The use of an SVM after genetic-algorithm-based parameter optimization gave the best results among the three types of methods. Genetic algorithms (GAs) are a general adaptive optimization search methodology based on a direct analogy to Darwinian natural selection and genetics in biological systems, and are a promising alternative to conventional heuristic methods. The accuracy of the GA-based SVM method reached 88%, which is a reasonable accuracy for complicated road conditions because some false points are discarded in the next step. Please refer to [59] for more details on GA-based parameter optimization.

2) *Optimal-drivable-area detection results:* We tested 6550 frames of fused data, including 2584 frames of structured road and 2966 frames of unstructured road. In the former situation, the accuracy reached 94.2%, and in the latter situation the accuracy was 91.7%. We see that the accuracy of curb point detection was not very high, but after line filtering and combining, some false positives were discarded. See Table I for details.

Our system handles very challenging unstructured road scenarios. Some results for drivable-area detection using fused data are listed in Fig. 11. Here, the three laser point cloud lines are rendered as green, blue, and red; the black points denote curb points detected by our system, and the curb points form the drivable areas shown in dark red. Fig. 11(a), (b), (c), and (d) show campus roads with heavy shadows: (a) shows a road with an irregular boundary, and (d) shows a road with a static vehicle in front. In all four situations, road detection is

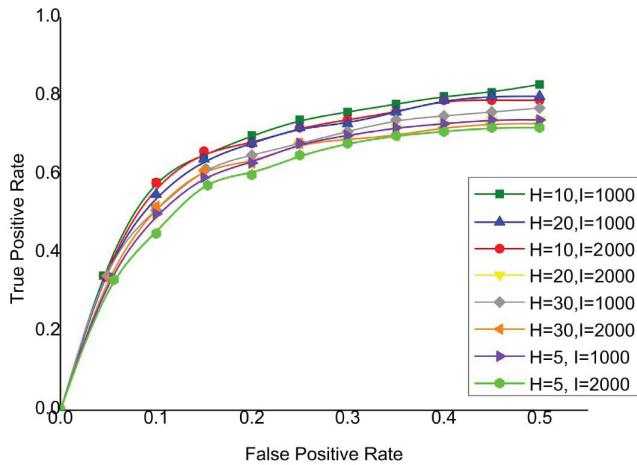


Fig. 9. ROC of curb point detection by ANN with different numbers of hidden nodes and iterations. The labels of the curves are sorted by the true positive rate at a false positive rate of 0.5.

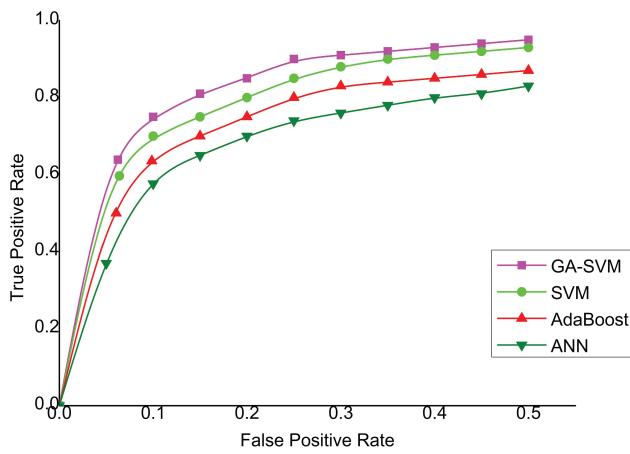


Fig. 10. ROC of curb point detection by GA-SVM, SVM, AdaBoost, and ANN. The labels of the curves are sorted by the true positive rate at a false positive rate of 0.5.

nearly impossible using only vision-based methods. But with our system, the correct drivable area was detected accurately. The accuracy benefits from the proposed fusion method and the optimal selection strategy.

3) *Lane detection results:* In addition to working with unstructured roads, our system does well in lane detection. Because of the previous step, lane detection becomes much simpler. From Table II, we see that the accuracy of lane detection for the three data sets reaches 90.6%, 93%, and

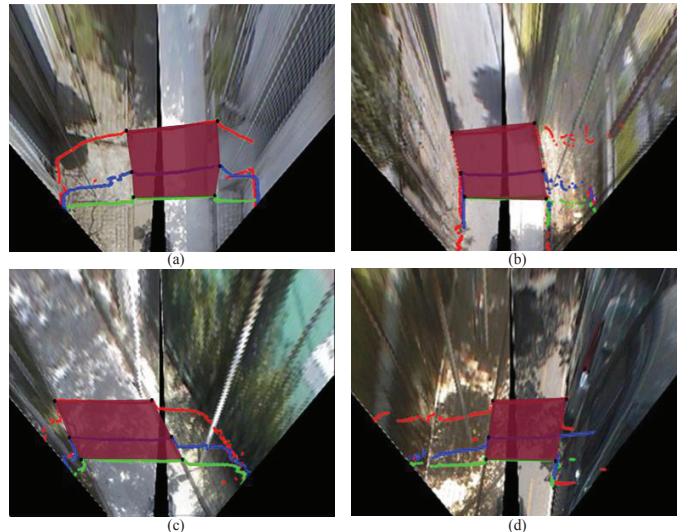


Fig. 11. Results of optimal-drivable-region detection in several challenging road scenarios. All four scenarios contain heavy shadow. (a) A road with an irregular shape, (b) a road with a slight turn, (c) a road after a turn, (d) a road with a static vehicle ahead. The red, blue, and green lines denote laser points, the black points denote the detected curb points, and the dark red area denotes the optimal drivable region.

93.7%; the result based on the totals is 92.9%. In highly structured urban environments, the accuracy is better than that of any lane detection system based only on vision. Detecting lanes is made easier by using the laser scanners. In addition, we obtain better results in dusk conditions with our proposed lane detection algorithm. To show the superiority of our methods, we shall now compare some results obtained from our algorithm and from the CHEVP algorithm [60].

The main steps of CHEVP are:

- 1) Canny-based edge pixel extraction [61].
- 2) Straight-line detection by Hough transform.
- 3) Horizon and vanishing point detection.
- 4) Estimating the midline of the road and the parameter k from the detected road lines.
- 5) Calculating the control points of the lane model to approach the detected midline.

In Fig. 12, we present three rows of images: the first row was recorded in the daytime and the other two in the evening. One sees that our algorithm finds the correct lanes in all three cases. The first step of CHEVP is difficult. Without self-adaptation of the thresholds for the Canny operator, we cannot get satisfactory results for all three images. If we use the threshold that is best for daytime roads, the algorithm does not work anymore in the evening, and vice versa. The accuracy of lane detection for three data sets using CHEVP algorithm has also been listed in Table II, in which “P” refers to the proposed algorithm and “C” refers to CHEVP algorithm. The accuracy with CHEVP are 84.7%, 75.1%, and 82.4% according to the three data sets; the result based on the totals is 82.7%. From this comparison, we see that our lane detection algorithm has better adaptability to the time of the day. This is mainly due to our local self-adaptation threshold.

4) *Road and lane detection results:* The most important contribution of our system is its ability to handle both struc-

TABLE I
DRIVABLE-AREA DETECTION RESULTS

	True area	Positive area	Accuracy
Data sequence 1	2300	2098	91.2%
Data sequence 2	1650	1537	93.2%
Data sequence 3	2600	2461	94.7%
Structured road	3584	3376	94.2%
Unstructured road	2966	2720	91.7%
Total	6550	6096	93.1%

TABLE II
LANE DETECTION RESULTS

	True lanes	Positive lanes(C)	Positive lanes(P)	Accuracy(C)	Accuracy(P)
Data sequence 1	3328	2819	3016	84.7%	90.6%
Data sequence 2	498	374	463	75.1%	93%
Data sequence 3	9456	7791	8862	82.4%	93.7%
Total	13282	10984	12341	82.7%	92.9%

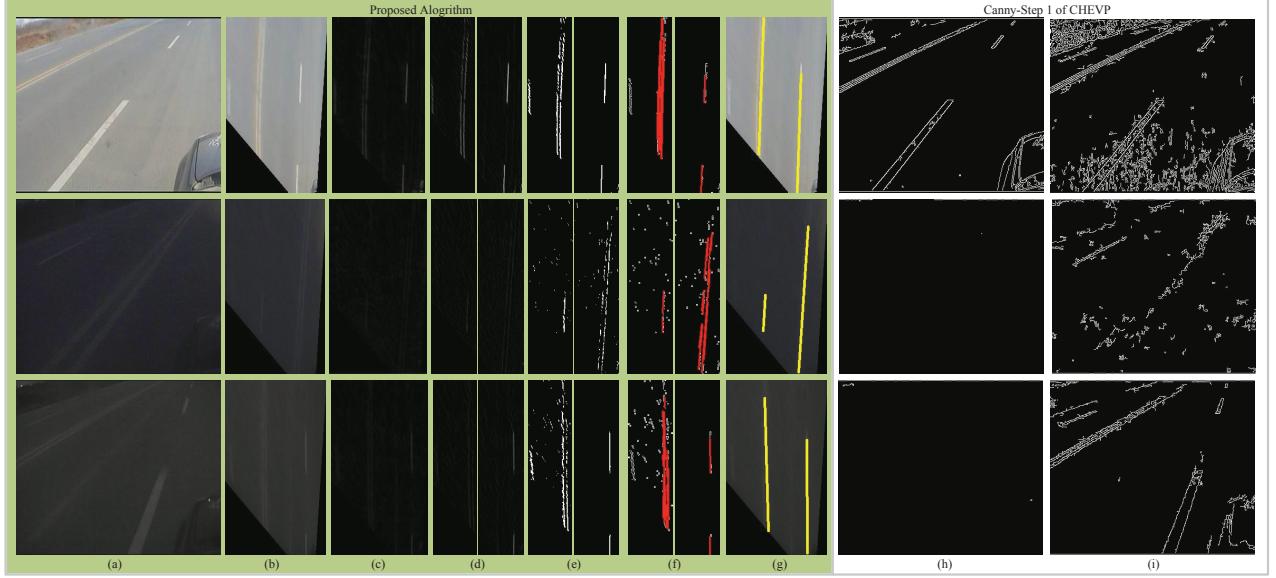


Fig. 12. Lane detection results from proposed algorithm and CHEVP algorithm. The images in the dark green rectangular box are the original images and the results from our proposed algorithm, and the images in the white rectangular box are the results from the Canny edge detection algorithm, which is the first step of the CHEVP algorithm. (a) Original image; (b) bird's-eye view after homography transform; (c) result of top-hat transform; (d) result of vertical-gradient detection by Prewitt operator; (e) result after local OTSU thresholding, where one image is divided into two parts; (f) result after PPHT; (g) result after line combining; (h) result after Canny edge detection with appropriate parameters for a daytime image (first row), (i) result after Canny edge detection with appropriate parameters for dusk image (third row).

tured and unstructured roads without the help of artificial switching or position information. We now present some experiments on these two types of road situations. Fig. 13 lists some results for drivable-area detection and lane detection. The blue points are the detected curb points, the dark red area is the drivable area, and the yellow lines relate to the detected lanes. Fig. 13(a) and (b) show curved campus roads with a large and a small turn radius. Fig. 13(c) and (d) show campus roads with heavy shadow and static vehicles. The algorithm avoided the mistake of taking the road in Fig. 13(c) to be a crossing.

In such road situations, methods based only on vision hardly work well, but our method easily obtained correct results. Because these situations were classified as type A (see Fig. 6), lane detection was not necessary. Fig. 13(e) and (f) show campus roads with snow present. Fig. 13(g), (h), (i), and (j) show structured roads with multiple lanes. Fig. 13(g) shows a road with interference from other markings; Fig. 13(h) and (i) show multilane road situations captured in daytime and at nightfall, respectively. Fig. 13(j) shows another part of the same road, made of two different materials: the back left part is asphalt and the remainder is concrete, and there is also some interference from dirt on the road. From these results, we see

that our system deals well with two completely different road situations even when various kinds of interference are present.

Fig. 14 show the performance of the drivable area detection algorithm on the roads with some obstacles. When the obstacles are not near the vehicle, the algorithm will not be affected (see Fig. 14(a)(b)). Short detection distance is a limitation of the algorithm, especially when the speed of the vehicle is fast. Fig. 14(c) is a successful detection on the situation that there are other vehicle and pedestrian in front of the vehicle. When the detection area does not meet drivable conditions, the vehicle will stop until there is a drivable area in front. The black area in Fig. 14(d) means the area is not drivable because the most distant line is not drivable.

Although we does not use expensive sensors, our system obtains results that are nearly as good as those that can be obtained using expensive sensors. To demonstrate this, we tested our algorithm, using one forward cameras fused with three SICK laser scanners, in comparison with a simple road boundary detection method based on a high-end sensor, namely the Velodyne HDL-64ES2 sensor, a dense 64-beam scanning LIDAR that provides 360° coverage at 10 Hz, generating just over 1 million 3D points with associated infrared reflectance values per second.

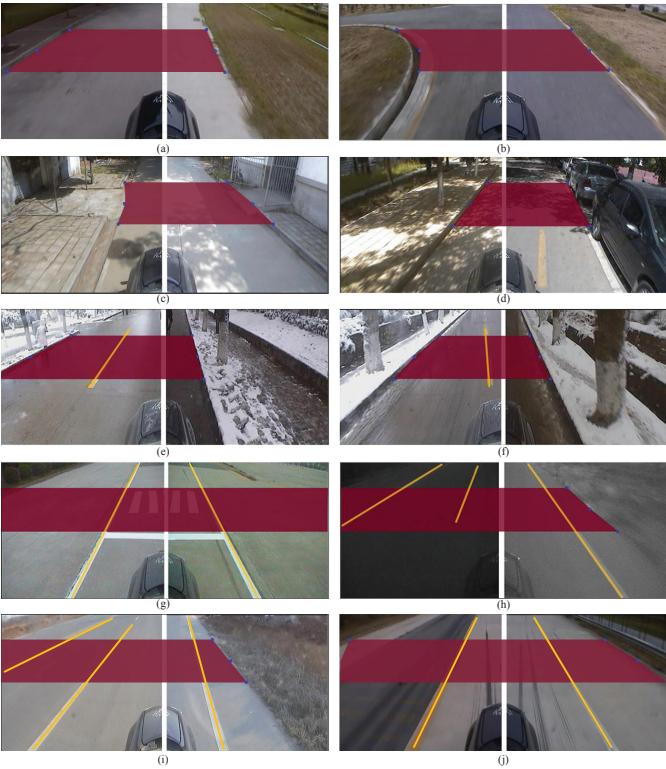


Fig. 13. Results of optimal-drivable-region and lane detection in several challenging road scenarios. The blue points denote the detected curb points, the dark red areas shows the optimal drivable regions, and the yellow lines show the detected lanes.



Fig. 14. Results of optimal-drivable-region detection in roads with obstacles.

First, we calculate the local Z -variance for every point ahead of the vehicle, and then traversed these points from right to left. These points were identified as boundary points when the local Z -variance of two consecutive points was more than a threshold value. The RANSAC algorithm [62] was then applied to fit a line to the boundary points. Fig. 15 shows the results from the two methods. Undeniably, the Velodyne sensor has a longer detection distance and a wider coverage area, but our system obtains accurate results for autonomous driving in a more cost-effective manner. In all six scenarios, our algorithm finds the correct optimal drivable regions, but the boundary detection result based on the Velodyne sensor in the sixth scenario is erroneous because of a vehicle in front. Our algorithm, using fusion of low-level sensors, works as well as the high-level sensor. This demonstrates the effectiveness of our fusion algorithm.

B. Performance of proposed system in the China Future Challenge

In addition to the above experiments on offline data sequences, the reliability of our system has also been verified in a real road challenge, the China Future Challenge (CFC) 2010. The CFC 2010 was a challenge for unmanned autonomous vehicles held by the National Natural Science Foundation in Xi'an in 2010. A total of 10 teams from China's universities and research institutions took part in this challenge. The CFC was similar to the Defense Advanced Research Projects Agency (DARPA) Grand Challenges, which were held in US in 2004, 2005, and 2007. The distance covered in the CFC (less than 4 km) was much shorter than that in the DARPA Grand Challenges (more than 100 km), but the two main differences from the DARPA Grand Challenges were the following:

- The CFC included different road conditions in one test: the first half was on campus roads at Chang'an University and the second half used several different road situations (including both structured and unstructured roads) at a proving ground at Chang'an University. The DARPA Grand Challenges of 2004 and 2005 used off-road (desert) scenarios, and that of 2007 was carried out on urban roads with lane markings.
- The use of the GPS was forbidden in the CFC but it was permissible in the DARPA Grand Challenges. So the requirements for road perception in the CFC were different from those in the DARPA tests.

These two characteristics of the CFC implied a high requirement for road perception. We now present the performance of our vehicle in this challenge.

The different kinds of roads are shown in Fig. 16. The top two images show campus roads; the bottom two show roads in the proving ground. We see that there are roads and lanes of different colors, that the roads are made from two different materials, and that there is interference of various kinds on the roads. In the challenge, our vehicle could deal with all these roads using the proposed fusion system without manual switching. In some special scenarios such as the junction between a road segment and an island shown in Fig. 17, our vehicle could find the drivable region successfully.

Finally, Fig. 18 shows the route of the China Future Challenge 2010. The challenge included a part made up of campus roads (from A to G, nearly 1.2 km) and a part made up of different kinds of roads (from G to H, nearly 2.3 km). On the campus roads, there were no obvious lane markings but there were several sharp turns (D, E, and F), one U-turn (B), and one island (C). In this challenge, our vehicle's speed reached 10 km/h on the campus roads, and on the highway the speed reached 30 km/h. Our vehicle, running the algorithm presented here, was the only one that crossed the finishing line within the required time. The following links illustrate the performance of our vehicle in this challenge:
http://v.youku.com/v_show/id_XNTc4MzU2MTU2.html (or
<http://youtu.be/cbf3OViCHas>).

VI. CONCLUSION

In this paper, a multi-level fusion-based road detection system for unmanned vehicle navigation in challenging road

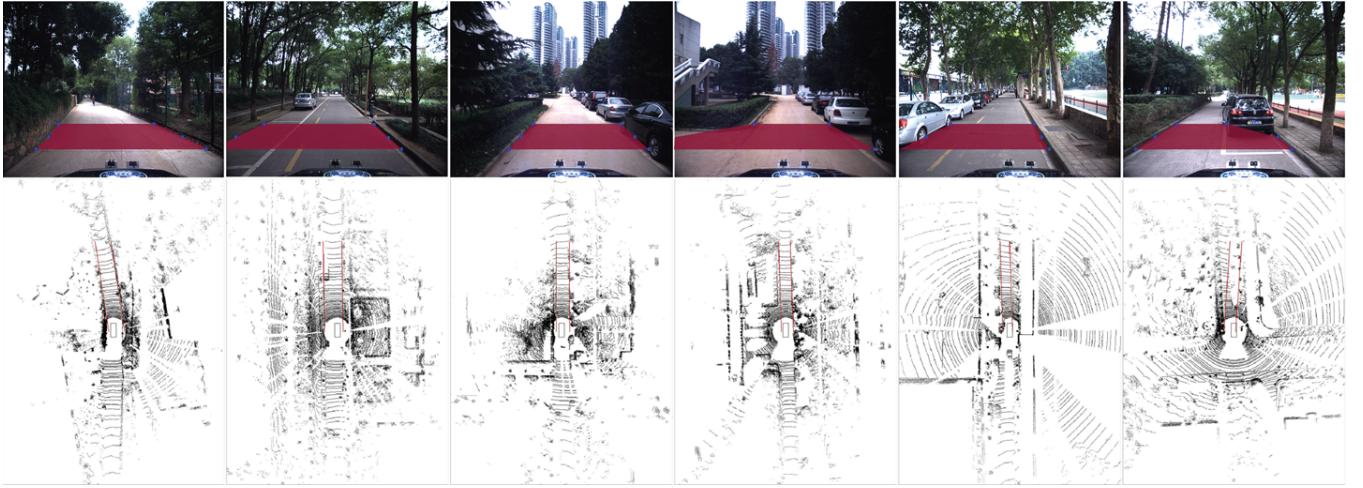


Fig. 15. Results from our algorithm with one forward camera and three SICK laser scanners compared with a boundary detection method based on Velodyne data.



Fig. 16. Different kinds of roads in the CFC 2010.



Fig. 17. Performance of our vehicle on a junction between a road segment and island.

conditions, including structured roads, unstructured roads, and transitions between them, has been presented. This system makes use of active sensing by lasers and color information obtained by cameras. A feature-level fusion-based method is used for optimal-drivable-region detection, and then conditional lane detection method is applied only when optimal drivable region is estimated to be wide enough. Our system ensures that the vehicle is driven safely in any kind of road conditions without knowing the terrain beforehand. This research effort generated many innovations:

- Our system deals reliably with challenging urban environments, including structured and unstructured roads, without manual switching.
- A fusion-based method was proposed. Feature-level fusion is used for drivable-region detection, and following

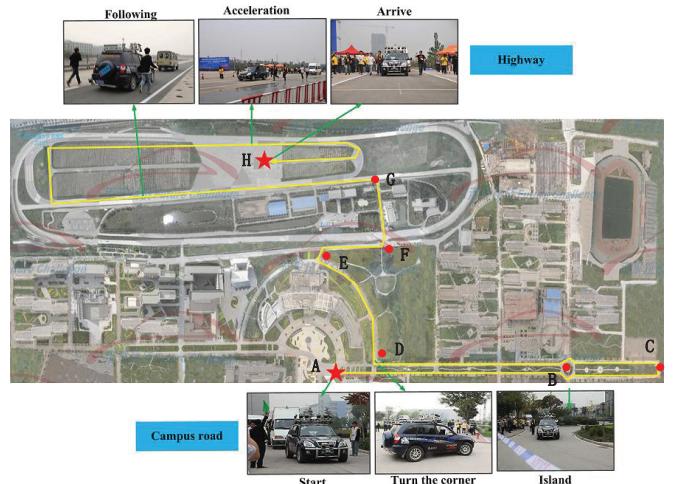


Fig. 18. Performance of SmartV in the China Future Challenge 2010. A map of the route is shown in the middle, where the five-pointed stars indicate the start and end points. The small images show the performance of the vehicle at various points, marked by red dots on the map. The total length of the challenge was more than 3 km. Several different road scenarios, including structured and unstructured roads, occurred in the competition.

with a conditional lane detection method.

- The proposed strategy extracts optimal drivable regions in front of the vehicle instead of recognizing every pixel of road surface.

Many offline experimental results and the performance of our vehicle in the China Future Challenge 2010 show that our system is reliable in challenging urban environments, especially under complicated lighting conditions such as those created by hard shadows, snow, and evening conditions. Even so, there is much left to be done. On roads with densely packed vehicles, our system will encounter problems because points on the vehicles will confuse our fused features so that we cannot pick out the correct curb points. To address this problem, we plan to develop a dynamic obstacle recognition system based on vision, LIDAR, or a fusion of the two. Furthermore, we are also considering how to enlarge the

scope of perception for the purpose of high-speed autonomous driving; this might possibly be achieved by fusion of data from a Velodyne sensor and a high-resolution camera.

ACKNOWLEDGMENT

The work described in this paper was supported by the National Natural Science Foundation of China (Grant No. 91120002) and Shenzhen Dedicated Funding of Strategic Emerging Industry Development Program (No. J-CYJ20121019111128765). Furthermore, we thank the German Academic Exchange Service (PPP VR China 54368573) and the Sino-German Center for Research Promotion GZ 692.

REFERENCES

- [1] P. Batavia, D. Pomerleau, and C. Thorpe, "Predicting lane position for roadway departure prevention," in *in Proc. IEEE Intell. Veh. Symp.*, 1998.
- [2] P. Batavia, "Driver adaptive warning systems," Technical Report CMU-RI-TR-98-07, Carnegie Mellon University, Tech. Rep., 1998.
- [3] M. Bertozzi, A. Broggi, M. Cellario, A. Fasoli, P. Lombardi, and M. Porta, "Artificial vision in road vehicles," in *Proc. IEEE*, vol. 90, no. 7, pp. 1258–1271, 2002.
- [4] M. Chen, T. Jochem, and D. Pomerleau, "AURORA: A vision-based roadway departure warning system," in *in Proc. IROS*, vol. 1, 1995, pp. 243–248.
- [5] D. Pomerleau, C. Thorpe, and L. Emery, "Performance specification development for roadway departure collision avoidance systems," in *in Proc. 4th World Congress Intel. Transp. Syst.* Citeseer, 1997.
- [6] M. Bertozzi and A. Broggi, "Vision-based vehicle guidance," *Computer*, vol. 30, no. 7, pp. 49–55, 1997.
- [7] ———, "GOLD: A parallel real-time stereo vision system for generic obstacle and lane detection," *IEEE Trans. Image Process.*, vol. 7, no. 1, pp. 62–81, 1998.
- [8] K. Kluge and S. Lakshmanan, "A deformable-template approach to lane detection," in *in Proc. IEEE Intell. Veh. Symp.*, 1995, pp. 54–59.
- [9] Y. Yim and S. Oh, "Three-feature based automatic lane detection algorithm (TFALDA) for autonomous driving," *IEEE Trans. Intell. Transp. Syst.*, vol. 4, no. 4, pp. 219–225, 2003.
- [10] Q. Li, N. Zheng, and H. Cheng, "Springrobot: A prototype autonomous vehicle and its algorithms for lane detection," *IEEE Trans. Intell. Transp. Syst.*, vol. 5, no. 4, pp. 300–308, 2004.
- [11] H. Cheng, B. Jeng, P. Tseng, and K. Fan, "Lane detection with moving vehicles in the traffic scenes," *IEEE Trans. Intell. Transp. Syst.*, vol. 7, no. 4, pp. 571–582, 2006.
- [12] J. McCall and M. Trivedi, "Video-based lane estimation and tracking for driver assistance: Survey, system, and evaluation," *IEEE Trans. Intell. Transp. Syst.*, vol. 7, no. 1, pp. 20–37, 2006.
- [13] Y. Wang, D. Shen, and E. Teoh, "Lane detection using spline model," *Pattern Recognition Letters*, vol. 21, no. 8, pp. 677–689, 2000.
- [14] C. Jung and C. Kelber, "Lane following and lane departure using a linear-parabolic model," *Image Vision Comput.*, vol. 23, no. 13, pp. 1192–1202, 2005.
- [15] Y. Zhou, R. Xu, X. Hu, and Q. Ye, "A robust lane detection and tracking method based on computer vision," *Meas. Sci. Technol.*, vol. 17, p. 736, 2006.
- [16] Z. Kim, "Robust lane detection and tracking in challenging scenarios," *IEEE Trans. Intell. Transp. Syst.*, vol. 9, no. 1, pp. 16–26, 2008.
- [17] R. Gopalan, T. Hong, M. Shneier, and R. Chellappa, "A learning approach towards detection and tracking of lane markings," *IEEE Trans. Intell. Transp. Syst.*, vol. 13, no. 3, pp. 1088–1098, 2012.
- [18] A. Amditis, M. Bimpas, G. Thomaidis, M. Tsogas, M. Netto, S. Mammar, A. Beutner, N. Möhler, T. Wirthgen, S. Zipser *et al.*, "A situation-adaptive lane-keeping support system: Overview of the SAFELANE approach," *IEEE Trans. Intell. Transp. Syst.*, vol. 11, no. 3, pp. 617–629, 2010.
- [19] T. Ogawa and K. Takagi, "Lane recognition using on-vehicle lidar," in *in Proc. IEEE Intell. Veh. Symp.*, 2006, pp. 540–545.
- [20] F. Homm, N. Kaempchen, and D. Burschka, "Fusion of laserscanner and video based lanemarking detection for robust lateral vehicle control and lane change maneuvers," in *in Proc. IEEE Intell. Veh. Symp.*, 2011, pp. 969–974.
- [21] K. Dietmayer, N. Kaempchen, K. Fuerstenberg, J. Kibbel, W. Justus, and R. Schulz, "Roadway detection and lane detection using multilayer laserscanner," *Advanced Microsystems for Automotive Applications 2005*, pp. 197–213, 2005.
- [22] S. Kammler, J. Ziegler, B. Pitzer, M. Werling, T. Gindel, D. Jagzent, J. Schröder, M. Thuy, M. Goebel, F. Hundelshausen *et al.*, "Team annieway's autonomous system for the 2007 darpa urban challenge," *J. Field Robot.*, vol. 25, no. 9, pp. 615–639, 2008.
- [23] P. Lindner, E. Richter, G. Wanielik, K. Takagi, and A. Isogai, "Multi-channel lidar processing for lane detection and estimation," in *in Proc. IEEE Conf. Intell. Transp.*, 2009, pp. 1–6.
- [24] A. von Reyher, A. Joos, and H. Winner, "A lidar-based approach for near range lane detection," in *in Proc. IEEE Intell. Veh. Symp.*, 2005, pp. 147–152.
- [25] Y. He, H. Wang, and B. Zhang, "Color-based road detection in urban traffic scenes," *IEEE Trans. Intell. Transp. Syst.*, vol. 5, no. 4, pp. 309–318, 2004.
- [26] C. Caraffi, S. Cattani, and P. Grisleri, "Off-road path and obstacle detection using decision networks and stereo vision," *IEEE Trans. Intell. Transp. Syst.*, vol. 8, no. 4, pp. 607–618, 2007.
- [27] A. Wedel, H. Badino, C. Rabe, H. Loose, U. Franke, and D. Cremers, "B-spline modeling of road surfaces with an application to free-space estimation," *IEEE Trans. Intell. Transp. Syst.*, vol. 10, no. 4, pp. 572–583, 2009.
- [28] J. Álvarez and A. Lopez, "Road detection based on illuminant invariance," *IEEE Trans. Intell. Transp. Syst.*, no. 99, pp. 1–10, 2010.
- [29] H. Fang, C. Wang, M. Yang, and R. Yang, "Ground-texture-based localization for intelligent vehicles," *IEEE Trans. Intell. Transp. Syst.*, vol. 10, no. 3, pp. 463–468, 2009.
- [30] C. Guo, S. Mita, and D. McAllester, "Robust road detection and tracking in challenging scenarios based on markov random fields with unsupervised learning," *IEEE Trans. Intell. Transp. Syst.*, vol. 13, no. 3, pp. 1338–1354, 2012.
- [31] J. Tarel and E. Bigorgne, "Long-range road detection for off-line scene analysis," in *in Proc. IEEE Intell. Veh. Symp.*, 2009, pp. 15–20.
- [32] R. Danescu and S. Nedevschi, "Probabilistic lane tracking in difficult road scenarios using stereovision," *IEEE Trans. Intell. Transp. Syst.*, vol. 10, no. 2, pp. 272–282, 2009.
- [33] M. Nikolova and A. Hero III, "Segmentation of a road from a vehicle-mounted radar and accuracy of the estimation," in *in Proc. IEEE Intell. Veh. Symp.*, 2000, pp. 284–289.
- [34] J. Sparbert, K. Dietmayer, and D. Streller, "Lane detection and street type classification using laser range images," in *in Proc. IEEE Intell. Veh. Symp.*, 2001, pp. 454–459.
- [35] H. Cramer and G. Wanielik, "Road border detection and tracking in non cooperative areas with a laser radar system," in *in Proc. German Radar Symp.* Bonn, Germany, 2002, pp. 24–29.
- [36] B. Fardi, U. Scheunert, H. Cramer, and G. Wanielik, "Multi-modal detection and parameter-based tracking of road borders with a laser scanner," in *in Proc. IEEE Intell. Veh. Symp.*, 2003, pp. 95–99.
- [37] W. Wijesoma, K. Kodagoda, and A. Balasuriya, "Road-boundary detection and tracking using ladar sensing," *IEEE Trans. Robot. Autom.*, vol. 20, no. 3, pp. 456–464, 2004.
- [38] H. Zhao, M. Chiba, R. Shibasaki, X. Shao, J. Cui, and H. Zha, "A laser-scanner-based approach toward driving safety and traffic data collection," *IEEE Trans. Intell. Transp. Syst.*, vol. 10, no. 3, pp. 534–546, 2009.
- [39] J. Trinkle and Y. Matsuo, "Robotics: Science and systems," *AI Magazine*, vol. 31, no. 1, pp. 99–100, 2010.
- [40] S. Thrun, M. Montemerlo, H. Dahlkamp, D. Stavens, A. Aron, J. Diebel, P. Fong, J. Gale, M. Halpenny, G. Hoffmann *et al.*, "Stanley: The robot that won the DARPA Grand Challenge," *J. Field Robot.*, vol. 23, no. 9, pp. 661–692, June 2006.
- [41] A. Broggi, S. Cattani, P. Porta, and P. Zani, "A laserscanner-vision fusion system implemented on the Terramax autonomous vehicle," in *in Proc. IROS*, 2006, pp. 111–116.
- [42] M. Manz, M. Himmelsbach, T. Luettel, and H. Wuensche, "Fusing LIDAR and vision for autonomous dirt road following," *Autonome Mobile Systeme 2009*, pp. 17–24, 2009.
- [43] L. Chen, Q. Li, M. Li, L. Zhang, and Q. Mao, "Design of a multi-sensor cooperation travel environment perception system for autonomous vehicle," *Sensors*, vol. 12, no. 9, pp. 12 386–12 404, 2012.
- [44] F. Rauskolb, K. Berger, C. Lipski, M. Magnor, K. Cornelisen, J. Effertz, T. Form, F. Graefe, S. Ohl, W. Schumacher *et al.*, "Caroline: An autonomously driving vehicle for urban environments," *J. Field Robot.*, vol. 25, no. 9, pp. 674–724, 2008.

- [45] H. Cheng, C. Yu, C. Tseng, K. Fan, J. Hwang, and B. Jeng, "Environment classification and hierarchical lane detection for structured and unstructured roads," *IET Comput Vis*, vol. 4, no. 1, pp. 37–49, 2010.
- [46] G. Bradski and A. Kaehler, *Learning OpenCV: Computer vision with the OpenCV library*. O'Reilly Media, 2008.
- [47] Z. Zhang, "Flexible camera calibration by viewing a plane from unknown orientations," in *in Proc. ICCV*, vol. 1, 1999, pp. 666–673.
- [48] T. Michalke, R. Kastner, M. Herbert, J. Fritsch, and C. Goerick, "Adaptive multi-cue fusion for robust detection of unmarked inner-city streets," in *in Proc. IEEE Intell. Veh. Symp.*, 2009, pp. 1–8.
- [49] I. Tang and T. Breckon, "Automatic road environment classification," *IEEE Trans. Intell. Transp. Syst.*, no. 99, pp. 1–9, 2011.
- [50] T. Mitchell, *Machine Learning*. McGraw-Hill, 1997.
- [51] C. Burges, "A tutorial on support vector machines for pattern recognition," *Data Min. Knowl. Discov.*, vol. 2, no. 2, pp. 121–167, 1998.
- [52] Y. Freund and R. Schapire, "A desicion-theoretic generalization of online learning and an application to boosting," *J. Comput. Syst. Sci.*, vol. 55, no. 1, p. 119C139, 1997.
- [53] N. Cristianini and J. Shawe-Taylor, "An introduction to support vector machines," *History*, vol. 47, no. 2, pp. 1–15, 2000.
- [54] J. Serra, *Image Analysis and Mathematical Morphology*. London: Academic Press, 1982.
- [55] N. Otsu, "A threshold selection method from gray-level histograms," *Automatica*, vol. 11, pp. 285–296, 1975.
- [56] J. Matas, C. Galambos, and J. Kittler, "Robust detection of lines using the progressive probabilistic hough transform," *Comput. Vis. Image Und.*, vol. 78, no. 1, pp. 119–137, 2000.
- [57] R. Duda and P. Hart, "Use of the Hough transformation to detect lines and curves in pictures," *Comm. ACM*, vol. 15, no. 1, pp. 11–15, 1972.
- [58] P. Werbos, "Beyond regression: New tools for prediction and analysis in the behavioral sciences," 1974.
- [59] C. Huang and C. Wang, "A GA-based feature selection and parameters optimization for support vector machines," *Expert Syst. Appl.*, vol. 31, no. 2, pp. 231–240, 2006.
- [60] Y. Wang, E. Teoh, and D. Shen, "Lane detection and tracking using B-Snake," *Image Vision Comput.*, vol. 22, no. 4, pp. 269–280, 2004.
- [61] J. Canny, "A computational approach to edge detection," *IEEE Trans. Pattern Anal. Mach. Intell.*, no. 6, pp. 679–698, 1986.
- [62] M. Fischler and R. Bolles, "Random sample consensus: A paradigm for model fitting with applications to image analysis and automated cartography," *Comm. ACM*, vol. 24, no. 6, pp. 381–395, 1981.



Qingquan Li received the M.S. degree in engineering and the Ph.D degree in photogrammetry and remote sensing from Wuhan University, Wuhan, China, in 1988 and 1998, respectively.

From 1988 to 1996, he was an Assistant Professor with Wuhan University, where he became an Associate professor from 1996 to 1998 and has been a Professor with the State Key Laboratory of Information Engineering in Surveying, Mapping, and Remote Sensing since 1998. He is currently the President of Shenzhen University and the Director of

Shenzhen Key Laboratory of Spatial Information Smart Sensing and Services, Shenzhen University. He is an expert in Modern Traffic with the National 863 Plan and an Editorial Board Member of the *Surveying and Mapping Journal* and the *Wuhan University Journal-Information Science Edition*. His research interests include photogrammetry, remote sensing, and intelligent transportation systems.

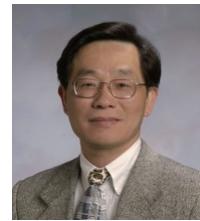


Long Chen received the B.Sc. degree in communication engineering in 2007 and Ph.D degree in signal and information Processing in 2013 from Wuhan University, Wuhan, China. He was a visiting student at National University of Singapore from Oct.2011 to Nov.2012.

He is currently an assistant professor with School of Mobile Information Engineering, Sun Yat-sun University. His areas of interest include computer vision, point cloud processing, and unmanned autonomous vehicles. From 2008 to 2013, He was in charge of environmental perception system for autonomous vehicle SmartV-II with the Intelligent Vehicle Group, Wuhan University.



Ming Li received the Ph.D degree in Photogrammetry and Remote Sensing from the Wuhan University. He is currently an associate professor in the Computer School of Wuhan University. He is also an associate professor in National Engineering Research Center for Multimedia Software. His research interests are multi-sensor fusion, point cloud processing and driving environment sensing.



Shih-Lung Shaw is Professor of Geography at the University of Tennessee, Knoxville. His research interests cover transportation geography, transportation analysis and modeling, geographic information systems for transportation (GIS-T), and space-time geographic information analysis. Dr. Shaw is a Fellow of the American Association for Advancement of Science (AAAS) and received the Edward L. Ullman Award for Outstanding Contributions to Transportation Geography from the Association of American Geographers (AAG). Dr. Shaw holds Alvin and Sally Beaman Professorship and Arts and Sciences Excellence Professorship at the University of Tennessee.



Andreas Nüchter is professor of computer science (telematics) at University of Wrzburg. Before summer 2013 he headed as assistant professor the Automation group at Jacobs University Bremen. Prior he was a research associate at University of Osnabrück. Further past affiliations were with the Fraunhofer Institute for Autonomous Intelligent Systems (AIS, Sankt Augustin), the University of Bonn, from which he received the diploma degree in computer science in 2002 (best paper award by the German society of informatics (GI) for his thesis) and the Washington

State University. He holds a doctorate degree (Dr. rer. nat) from University of Bonn. His thesis was shortlisted for the EURON PhD award. Andreas works on robotics and automation, cognitive systems and artificial intelligence. His main research interests include reliable robot control, 3D environment mapping, 3D vision, and laser scanning technologies, resulting in fast 3D scan matching algorithms that enable robots to perceive and map their environment in 3D representing the pose with 6 degrees of freedom. The capabilities of these robotic SLAM approaches were demonstrated at RoboCup Rescue competitions, ELROB and several other events. He is a member of the GI and the IEEE.