Color-Based Road Detection in Urban Traffic Scenes

Yinghua He, Hong Wang, and Bo Zhang

***Abstract*—Road detection is a key issue for autonomous driving in urban traffic. In this paper, after a brief overview of existing methods, we present a road-area detection algorithm based on color images. This algorithm is composed of two modules: boundaries are first estimated based on the intensity image and road areas are subsequently detected based on the full color image. In the first module, an edge image of the scene is analyzed to obtain the candidates for the left and right road borders and to delimit the area that will subsequently be used to compute the mean and variance of the Gaussian distribution, assumed to be obeyed by the color components of road surfaces. The second module effectively extracts the road area and reinforces boundaries that most appropriately fit the road-extraction result. The combination of these modules can overcome basic problems due to inaccuracies in edge detection based on the intensity image alone and due to the computational complexity of segmentation algorithms based on color images. Experimental results on real road scenes have substantiated the effectiveness of the proposed method.**

***Index Terms*—Color-based segmentation, intelligent vehicle, road detection.**

I. Introduction

ROAD following is a key requirement for the successful development and employment of intelligent vehicles. In the past decades, several approaches for road following were proposed and successfully demonstrated [1], [2]. Most were aimed at lateral vehicle guidance on highways or other well-structured roads. In these environments, road following is always reduced to the estimation of lane markings painted on the road surface [3]–[11]. The generic obstacle and lane-detection (GOLD) system developed at Parma University, Parma, Italy, reprojects the image ahead of the vehicle onto the ground plane and extracts "left–middle–right" groups of road markings [3], [4]. The rapidly adapting lateral position handler (RALPH) system also involves the transformation between image and ground planes and estimates curvatures of the road [6]. The yet another road follower (YARF) system pays even more attention to the modeling of road geometry and thereby estimates road-model parameters [10], based on the road-marking extraction result. Obviously, these methods are not appropriate in many real-life situations in which the roads have no clear road markings, such as roads in urban and campus environment. At the same time, there also are some algorithms to work on the unstructured road [12]–[24]. The supervised classification applied to road following (SCARF) and unsupervised clustering applied to road following (UNSCARF) systems at Carnegie Mellon University (CMU), Pittsburgh, PA, have been built to deal with roads in parks, with no lane markings and considerable difference in color between the road area and surroundings, which is the basis for road detection [12], [13]. Although these systems have achieved impressive accuracy and robustness on roads in park, they are still too slow to meet the real-time requirement, which is crucial in dealing with urban traffic. Rasmussen utilized the texture cues to navigate on rural roads that lack road marking and a homogeneous road surface [15], [16]. Onoguchi *et al.* face this problem using a planar projection stereopsis (PPS) method for road extraction [17], which exploits the fact that height information is also useful in distinguishing the road from the surroundings. However, in many urban environments, the border between the road area and surroundings is rather fuzzy and consists only of road shoulders, whose height above the road surface may be too small to ensure success with PPS methods. Another category of methods delimited the search range in advance by knowledge to search road boundaries or road markings [18], [19]. Although they can work in urban traffic on some occasions, the clear boundaries in gray image are the basis of this method. Paetzold *et al.* provide some further ideas for road extraction in the urban environment [20]; however, they use only the intensity edge image, which is far from enough to extract the true road boundaries due to the unstructured and often quite dense presence of vehicles and pedestrians on the road.

The features of roads in urban and campus environments are markedly different from those of highways and roads in parks. Here, there may be no road markings at all; the border between road and its surroundings is often characterized by low and irregular road shoulders and, in some places, may be spatially fuzzy and, as stated, the numbers and types of moving vehicles and of pedestrians is higher than elsewhere. To cope with the problem of road detection in urban and campus environments, a better solution is to combine edge detection based on the intensity of images with a road-area extraction process based on the color components of an image. Since the former can give us some initial information on the boundaries of road area and the latter can distinguish the road area from the surroundings based on color homogeneity, the two processes should combine and interact in efficiently yielding good quality and efficient segmentations. Although, in the past few years, some methods have been proposed to detect a road based on combining boundaries and area information, few considered the features in urban traffic scenes [21]–[24]. Gibb *et al.* fused the result of edge detection, road segmentation, and white lane follower to acquire the final result [21]. Because the fusion module regarded three modules

Manuscript received December 1, 2003; revised July 16, 2004 and August 1, 2004. This work was supported by the National High Technology Development Program of China under Project 863. The Associate Editor for this paper was N. Zheng.

The authors are with the State Key Laboratory, Intelligent Technology and System, Department of Computer Science and Technology, Tsinghua University, Beijing 100084, P.R. China (e-mail: heyinghua01@mails.tsinghua.edu.cn; wanghong@mail.tsinghua.edu.cn; dcszb@mail.tsinghua.edu.cn).

Digital Object Identifier 10.1109/TITS.2004.838221

independently, it is not only time consuming, but also offers low reliability for lacking knowledge of road scenes. The method presented by Lutzeler processed edge detection in the low part of the image and image segmentation in the upper part of the image, which cannot solve the problem of blurred road marks in urban traffic scenes [23].

This paper is organized as follows. First, we analyze the characteristics of roads in color images of urban and campus environments. Then, an algorithm is proposed to extract the candidates of road boundaries and subsequently combining the results of boundary detection with the color information in the image captured, we present a method to precisely extract the road areas. Finally, experimental results for real road scenes are presented to substantiate the effectiveness of the proposed method.

II. Model of a Road Scene

Aiming to develop a system that has both high accuracy and robustness, we first analyze factors that may affect road detection in urban and campus environments. The factors we considered are shown in Fig. 1 and as follows.

1) The road area is not very wide and may have no clear road marking. There may be some moving cars and pedestrians, which may be quite near our intelligent vehicle, where the camera acquiring the scene is placed.
2) The surrounding area is always distinguished from the road area by some road shoulder or by a different color.
3) Trees and buildings exist in the scene and cast shadows on both the road area and elsewhere, but their presence can obviously be used as evidence to distinguish the road from its surroundings.
4) The position of the sun will determine the direction of incident light and its spectrum too.
5) The autonomous vehicle can obviously provide the current position and speed information, thereby reducing the search range for locating road boundaries.

Factors 1)–4) pertain to the color of the images captured, so let us first consider the photometric relationship between points lying on real-world surfaces and images in the image plane. Suppose that a point with coordinates (u, v) in the image plane correspondents to the point (x, y, z) in the world. The color image acquired with a charge-coupled device (CCD) camera consists of the outputs of three RGB sensors and the intensity of sensor $i \in \{\mathrm{RGB}\}$ can be represented as [25]

$$E_i(u,v) = c(\varphi,\theta,\gamma,x,y,z)C_i(x,y,z), i \in \{\mathrm{RGB}\} \quad (1)$$

where $c(\varphi,\theta,\gamma,x,y,z)$ depends on the geometry of the reflection and $C_i(x,y,z)$ is related to the properties of the surface. Therefore, the total intensity image can be represented as

$$\begin{aligned} I(u,v) &= \frac{1}{3}\big(R(u,v) + G(u,v) + B(u,v)\big) \\ &= \frac{1}{3}C(\varphi,\theta,\gamma,x,y,z)(C_R + C_G + C_B)(x,y,z). \end{aligned} \quad (2)$$

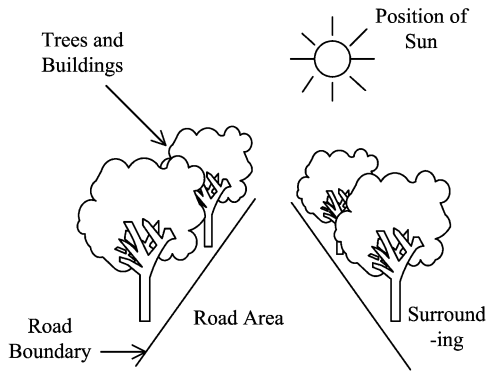


Fig. 1. Road scene.

Several methods have been proposed to solve the road-detection problem based solely on intensity images [3]–[11], [14], [17]–[20], but (2) shows that it is expected to be quite difficult to get good results without clearly visible road markings painted on road surfaces and in conditions of cluttered traffic. This is clearly so since the edges of shadows; some random road markings painted on the middle of road (which often exists in urban environments); and edges of vehicles, pedestrians, etc. can hopelessly clutter the true road boundaries as they will appear as edges in the intensity image. Furthermore, road areas with different colors from the surroundings may appear quite similar in intensity in the graylevel images.

Therefore, in order to cope with some of these problems, it is wise to heavily rely on the full-color image in order to robustly and efficiently detect the roads from their surroundings. We will use the idea of combining intensity images with information in color images and elementary set theory, as proposed by Birchfield [26], in order to develop a framework to solve our problem. The road in the captured image can formally be divided in two sets: the boundary and interior, which are complementary. The boundary and interior can be labeled as road border and road surface, respectively. Hence, we will use a four-tuple to represent road model in the image captured

$$R = (l, r, c, u) \quad (3)$$

where l and r are the left and right boundaries of road, respectively; c is the curvature of road lane; and u is the mean value of the color of road surface. In order to retrieve the parameters of R, the information that we can use may be as follows.

1) *Real-Life Boundaries.* These may be defined by road shoulders or borders between road area and the surroundings. Equation (2) shows us that although the color of the road surface and shoulders may be same, the differences in geometry may enable us to detect them as edges in the intensity image. Therefore, we may and will use the intensity image to locate the boundaries.
2) *Road Surface Itself.* We called it the Interior, which hopefully has a distinctive color. As discussed before, we know that it will be necessary to use the color image to estimate the interior in the images we will acquire. We argue that road-area extraction based on color images can often compensate for the shortcoming of road-boundary detection based on gray levels alone.

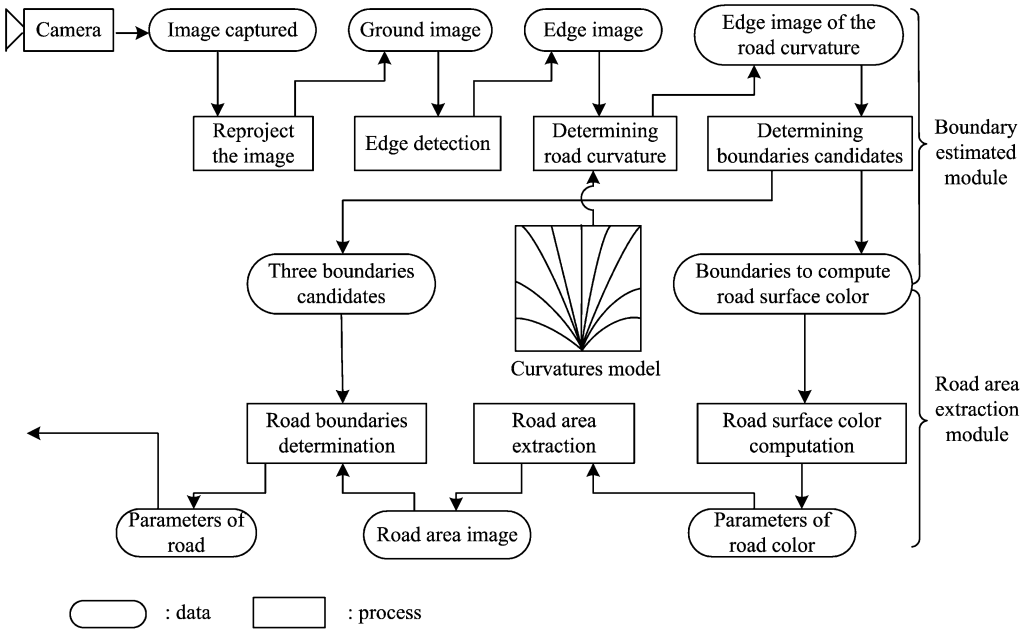


Fig. 2. Block diagram.

This division is the formal basis for the method that we will describe in the following sections.Fig. 2 shows the flowchart of the method presented in this paper.

III. Boundaries Estimated Module

Although we cannot obtain accurate boundaries from graylevel images alone, we can certainly obtain information from edge detectors based on intensity. Therefore, three tasks are defined in this first module of our proposed method, leaving some other for the area extraction module. These tasks will be:

1) determining the curvature of road;
2) approximately delineating an area in the image plane that will subsequently be used to estimate the color of road surface;
3) providing plausible candidates for the road boundaries.

According to (3), we must determine c and approximate the parameters l and r.

Since road boundaries are always pairs of parallel lines in the three-dimensional (3-D) world lying on a planar surface, the road plane (which is assumed to be Z $= 0$ in the 3-D "world") is more convenient to aim to detect the road boundaries in the road plane than in image plane. Under the assumption that the road is indeed planar, the projective transformation between the road plane (x, y) and the image plane (u, v) can be represented as [3], [6], [17]

$$\rho[x \quad y \quad 1]^T = M[u \quad v \quad 1]^T. \tag{4}$$

Here, M is only a function of the camera calibration parameters. We will further assume that in 3-D the curve of the road boundaries are two "parallel" planar circular arcs, as shown in Fig. 3, so the relationship among points (x, y) lying in one of road boundaries in the road plane can be represented as

$$(x - O_x)^2 + (y - O_y)^2 = r^2 \tag{5}$$

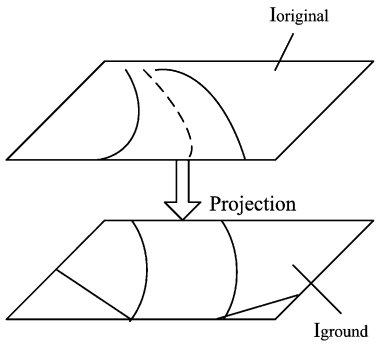


Fig. 3. Projection relationship.

where O_x and O_y is the circle center of the circular arc and r is the radius.

We denote the original and projection images by $I_{\text{original}}(x, y)$ and $I_{\text{ground}}(u, v)$, respectively, as shown in Fig. 3. Basic mathematical morphology operators—erosion and dilation—are then used to extract the edges of the image $I_{\text{ground}}(u, v)$. In order to remove noise and simplify the computation in subsequent steps, we process the edge image with a threshold T_B and get a binary image $I_{\text{binary}}(u, v)$. Next, we do a curvature estimation and determine boundary candidates based on the binary image $I_{\text{binary}}(u, v)$.

A. Road-Curvature Determination

Several methods have been proposed to cope with the problem of road-curvature determination [6], [10], [15], [21]. Klage and Thorpe precisely estimate the curvature of roads, based on reliable location of road markings [10]. Pomerleau and Jochem hypothesize five curvatures for the projection image and determine the straightest image after shifting by an offset determined by the assumed curvatures [6]. All of these methods aim at processing highway images, which have left, middle, and right road markings that are clearly visible on road

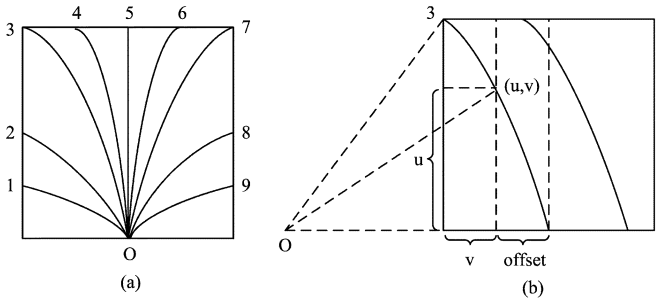


Fig. 4. Road curvatures. (a) nine curvatures and (b) curvature 3.

surfaces. Due to the more complex situations faced by vehicles in an urban environment, our problem is much more difficult. The basic idea presented by Pomerleau [6] can provide a rough estimate of curvature. However, in an urban environment, road surfaces have no road markings to enable us to find borders as clear sharp discontinuities; the road may follow a variety of curves and the presence of cars and pedestrians in the scene will result in highly complex edge images. Therefore, keeping the assumption of having the circular arcs as road boundaries on the ground, we base our computations on the consideration of nine kinds of curvature, shown in Fig. 4(a). Each curve is assumed to be part of a circular arc and the circle's center and radius can be computed offline. For each assumed curvature, we then use an algorithm that differentially shifts the rows of image $I_{\text{binary}}(u, v)$ to effectively undo the curvature and straighten out the assumed boundaries of the road. For instance, the shift distance corresponding to curvature 3 in Fig. 4(a) can be computed by the following equation, as shown in Fig. 4(b):

$$\text{offset} = \sqrt{(u + 3W/4)^2 + v^2} - (u + 3W/4), \qquad (6)$$

where W is the width of projection image. Notice that the width and height of the image $I_{\text{binary}}(u, v)$ is assumed to be the same, which considerably simplifies the computation of shifting offsets.

After shifting the image $I_{\text{binary}}(u, v)$ based on each of the assumed curvatures in Fig. 4(a), we will obtain nine images denoted by $\mathrm{U}_1, \ldots, \mathrm{U}_9$, corresponding to curvature 1–9. In this way, the search for the true road curvature is reduced to the determination of the most "straight-lined image" among the nine images $\mathrm{U}_1, \ldots, \mathrm{U}_9$. Hence, if the curvature hypothesis i is the winning hypothesis in this process, it will then yield a transformed image having the straightest features. The features that are parallel to the road, such as road boundaries, road markings, edges of buildings, etc. will also yield "maximum votes" in some image column.

Therefore, to address the problems of finding the image with most straight lines, we simply follow the following curvature-finding process.

Step 1) Compute the sum of the number of pixels that have been set to 1 in the columns of each image $\mathrm{U_i}(\mathrm{i} = 1, \ldots, 9)$ and store the results in the vector $V_i(\mathrm{i} = 1, \ldots, 9)$ (each V_i is a row matrix of dimension $1 \times W$, where W is the width of projection image).

Step 2) Order the vectors $V_i(\mathrm{i} = 1, \ldots, 9)$ according to the maximal value in each $V_i(\mathrm{i} = 1, \ldots, 9)$ and call the reordered vectors $V_{ki}(\mathrm{i} = 1, \ldots, 9)$. Then set the $k1, k2, k3$ to be the leading candidates for determining subsequent curvature.

Step 3) Determine the "meaningful range" in vectors $V_i(\mathrm{i} = 1, \ldots, 9)$. The intent of this step is to exclude the part of the vectors $\mathrm{V_i}$ the corresponding part in $\mathrm{U_i}$, of which have no detected edge corresponding points in the original image captured, as shown in the left part of Fig. 5. We have to do this only for $\{\mathrm{V_{k1}}, \mathrm{V_{k2}}, \mathrm{V_{k3}}\}$. Denote the left and right boundaries of the range in the vector $\mathrm{V_{ki}}$ as $\mathrm{Rl_i}$ and $\mathrm{Rr_i}$.

Step 4) Set $b_1 = \min(V_{k1}[Rl_1 : Rr_1]), b_2 = \min(V_{k2}[Rl_2 : Rr_2]), b_3 = \min(V_{k3}[Rl_3 : Rr_3])$ and take $b = \min(b_1, b_2, b_3)$.

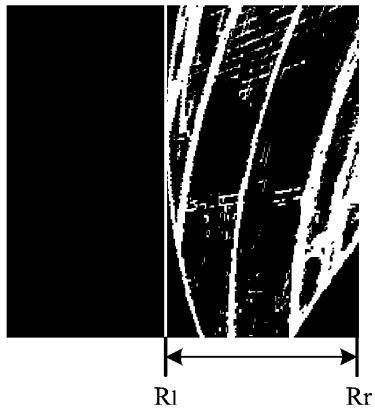


Fig. 5. Meaningful range.

The curvature then is determined as

$$\text{curvature} = \begin{cases} k1 & \text{if } b = b_1 \\ k2 & \text{if } b = b_2 \\ k3 & \text{if } b = b_3. \end{cases} \qquad (7)$$

B. Estimating the Position of Road Boundaries

The task of this section is not to accurately determine boundaries of the road, but to approximately locate left and right boundaries as our input for the road-area extraction module and at the same time obtain rough "left" and "right" boundaries candidates for the road. Let i be the "winning" curvature in the voting process described in Section III-A. We will obtain three candidate positions for the left and right boundaries from the vector V_i, denoted by l_1, l_2, l_3 and $\mathrm{r}_1, r_2, \mathrm{r}_3$, respectively.

Fig. 6(a) and (b) and shows the left and right range of an example used to find the candidates for the boundaries. These values can be acquired by finding the maximum values in the V_i vector. And at the same time, the following relationship between the numbers l_j and l_k must be satisfied as

$$|l_j - l_k| \geq T_W, \qquad |r_j - r_k| \geq T_W, \qquad j, k \in \{1, 2, 3\} \quad (8)$$

where T_W is a threshold that guarantees that the distance between the chosen candidates is not too small.

In order to remove the l_j and r_k, which are unlikely to be candidate positions, we take no account of the candidate whose

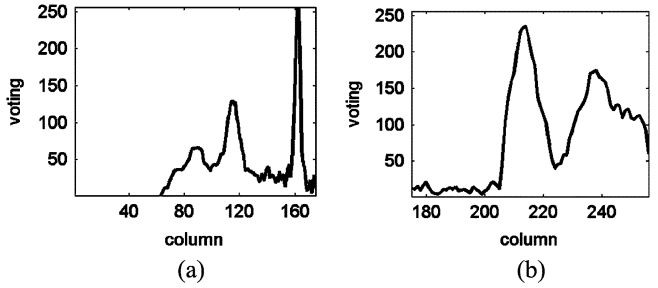


Fig. 6. Left and right range used to find the candidates for the boundaries. (a) Left and (b) right voting.

value of voting is lower than T_m/K, where $T_m = \max(V_i)$ and K is always chosen to be 3 or 4.

Next, we will yield approximations for the left and right boundaries for road-area extraction. It is ideal to select the region that lies in the middle part of the road area. Therefore, for the left boundaries, we choose the rightmost one but, for right boundaries, we choose leftmost one. And at the same time, left and right boundaries must satisfy the relation

$$|\text{left} - \text{right}| \geq T_b \tag{9}$$

where *left* and *right* are the results when using the road-area extraction module and T_b is again the threshold that guarantees that the chosen area is not so small.

IV. Road-Area Extraction Module

Here, we will work under the assumption that the color of road is approximately homogeneous. Since pedestrians and cars above the road surface will appear elongated in the projection image $I_{\text{ground}}(u, v)$, which has been defined in Section III, we will use the original image $I_{\text{original}}(x, y)$ to extract the road area. Though some cars, pedestrians, and road markings on the road may also affect color in the road area, they often cover only a small ratio of the road surface. The borders we choose from the algorithm presented in Section IV-B are often too conservative and narrower than the true boundaries, shown in Fig. 7(a) and (b). First, we use the shadow area described in Fig. 7(a) to compute the color that the road surface may satisfy. Then, we find the area in the original image that has color similar to the area selected and denote this area as road area. Finally, we select the appropriate boundaries that are matching the area extraction result well.

A. Road-Area Extraction

We assume that the color components of the road surface satisfy a multivariate Gaussian distribution. We use μ_n and Σ_n to represent the mean value and variance of this Gaussian distribution at frame n as

$$P(x) = \frac{1}{(2\pi)^{\frac{d}{2}}|\Sigma_n|^{\frac{1}{2}}} e^{-\frac{1}{2}(x-\mu_n)^T\Sigma_n^{-1}(x-\mu_n)}. \tag{10}$$

If a point $I_{\text{original}}(i, j)$ belongs to road area, the color is similar to that of μ_n. The probability of the pixel value $I_{\text{original}}(i, j)$ belonging to road area is $P(I_{\text{original}}(i, j))$. If $P(I_{\text{original}}(i, j)) \leq T$ (where T is a threshold), then the probability that this point is on the road area is very low; otherwise, we regard that $I_{\text{original}}(i, j)$ belongs to the road area. In this way, we acquire a binary image I_{result} that denotes the road-area extraction result.

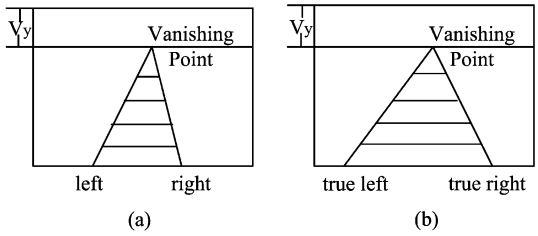


Fig. 7. Relationship between true road area and the area that is used for computing the road color. (a) Area used for computing color model and (b) true road area.

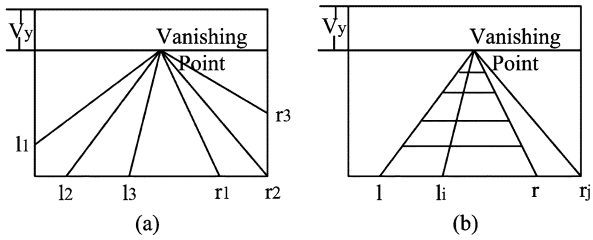


Fig. 8. Boundaries determination. (a) Left and right boundaries candidates and (b) ratio computing.

According to the assumption of the road being planar, the road area is always below the vanishing point in the image, as shown in Fig. 7. Therefore, it is enough for us to only compute the probability values of the pixels bellow the "horizon" of the vanishing points. The vertical position of the vanishing point can be computed offline and is denoted by $\mathrm{V_y}$ in Fig. 7.

In Section III-B, we presented the process for the estimation of the left and right boundaries of the road, which were denoted by *left* and *right*. The area between *left* and *right* will now be used to compute the mean value and variance for our Gaussian distribution model. These values can be obtained by a maximum likelihood estimation procedure, which yields

$$\begin{aligned}\hat{\mu}_n &= \frac{1}{m-1}\sum_{(i,j)\in\Omega} I_{\text{original}}(i, j) \\ \hat{\Sigma}_n &= \frac{1}{m-2}\sum_{(i,j)\in\Omega}\left(I_{\text{original}}(i, j) - \hat{\mu}_n\right) \\ &\quad \times \left(I_{\text{original}}(i, j) - \hat{\mu}_n\right)^T\end{aligned} \tag{11}$$

where Ω is the area delimited by the left and right in original image, as the shadow area in Fig. 7(a), and m is the total pixels number of the area Ω.

B. Finding Boundaries of Road

Linking the results of edge detection, we can obtain a reasonable approximation of road boundaries. In Section III-B, three candidates of the left and right boundaries have been defined, as shown in Fig. 8(a). Next, we will compute which combination of the left and right boundaries provide a good fit to the road-area extraction process described in Section IV-A.

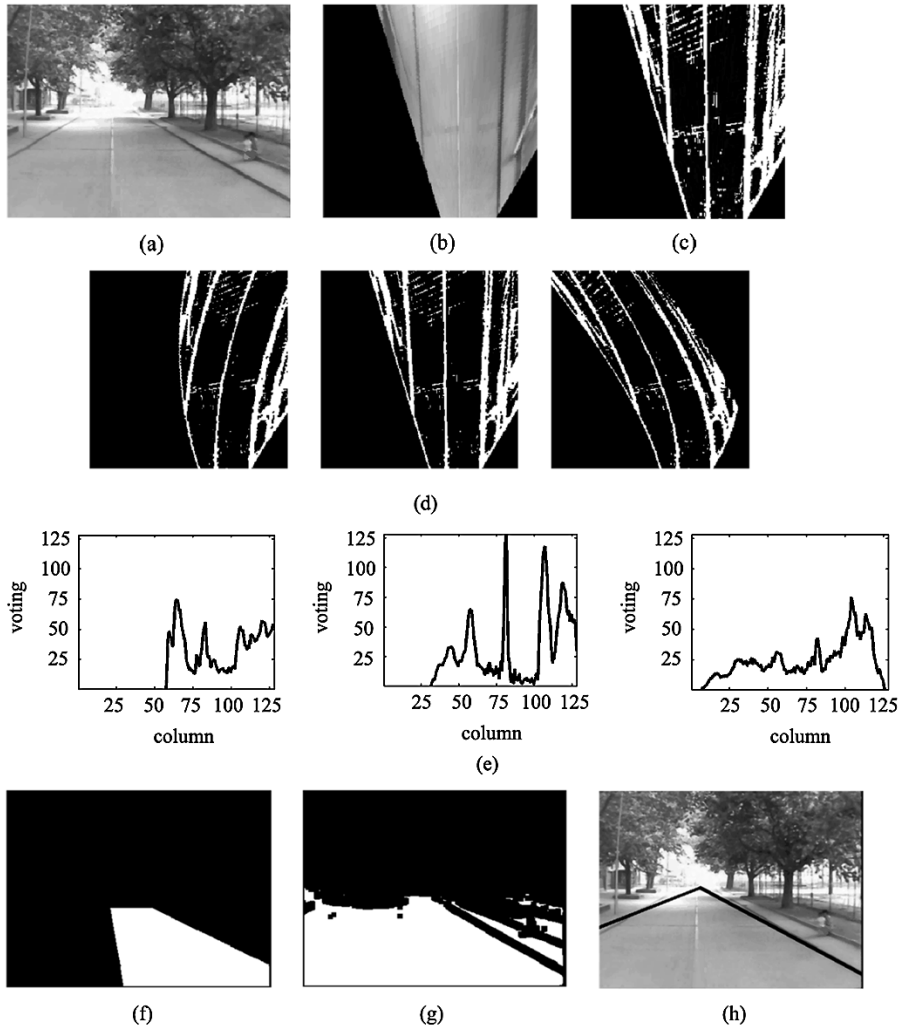


Fig. 9. Result of road detection in a straight road. (a) Original image. (b) Projection image. (c) Edge image. (d) Edge images of the three winners after shifting by nine curvatures shown in Fig. 4(a) (corresponding to curvatures 4, 5, and 6). (e) Voting images of edge images in (d). (f) Area used for computing road color. (g) Result of road extraction. (h) Final result.

Just as shown in Section III.B, we have computed three candidate positions for the left and right boundaries, denoted by l_1, l_2, l_3 and r_1, r_2, r_3, respectively. Suppose that the road area extracted in the last section is Λ. We use random variables L, R, and S to denote the position of left, right boundaries, and the road area. We will use Bayes rule to compute the probability of the candidate boundaries.

$$P(L = l_i, R = r_j \,|\, S = \Lambda) = \frac{P(S = \Lambda \,|\, L = l_i, R = r_j) P(L = l_i, R = r_j)}{P(S = \Lambda)} \quad (12)$$

where $P(S = \Lambda) = \sum_{i,j=1}^{3} P(S = \Lambda \,|\, L = l_i, R = r_j) P(L = l_i, R = r_j)$.

For any pair of l_i, r_j, we assume that $P(L = l_i, R = r_j)$ is a constant. Therefore, the computation of $P(L = l_i, R = r_j \,|\, S = \Lambda)$ is reduced to compute $P(S = \Lambda \,|\, L = l_i, R = r_j)$. In order to compute the item $P(S = \Lambda \,|\, L = l_i, R = r_j)$, we will show how the road-area extraction in Section IV-A uses the area between the "candidate" left and right boundaries. As shown in Fig. 8(b), let us denote by l_i and r_j the candidate position of boundaries that will be evaluated and by l and r is the true position of the road boundaries. We denote the area bounded by l_i and r_j by $S_{c,i,j}$ and the total shadow area that is included in $S_{c,i,j}$ by $S_{o,i,j}$. By counting the number of the pixels that have been set to 1 in image $I_{\text{result}}(x, y)$ delimited by area $S_{c,i,j}$, we can obtain $S_{o,i,j}$. We define the ratio as

$$\text{ratio}(l_i, r_j) = \frac{S_{o,i,j}}{S_{c,i,j}}. \quad (13)$$

We set the another ratio *ratio_non* as

$$\text{ratio\_non}(l_i, r_j) = 1 - \frac{S_{o,i,j}}{N_\Lambda} \quad (14)$$

where N_Λ is the total pixels number of the area Λ.

Because the voting values of the candidate boundaries always also affect the value of $P(S = \Lambda \,|\, L = l_i, R = r_j)$, we use the following equation to compute this item:

$$\text{vote}(l_i, r_j) = \frac{V(l_i) \times V(r_j)}{N_\Lambda} \quad (15)$$

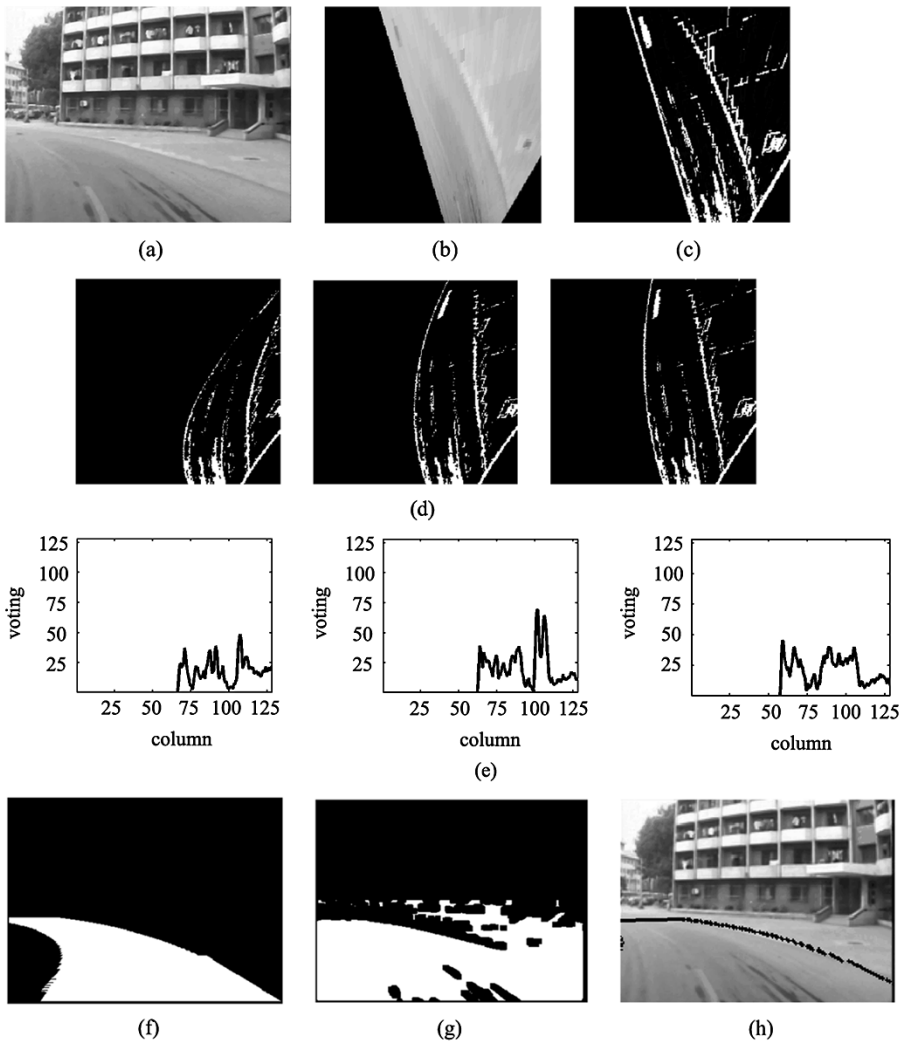


Fig. 10. Result of road detection in a curved road. (a) Original image. (b) Projection image. (c) Edge image. (d) Edge images of the three winners after shifting by nine curvatures shown in Fig. 4(a) (corresponding to curvatures 2, 3, and 4). (e) Voting images of edge images in (d). (f) Area used for computing road color. (g) Result of road extraction. (h) Final result.

where $V(l_i)$ and $V(r_j)$ are the voting value of the candidate boundaries l_i and r_j computed in Section III-B. N_V is a normalized factor that guarantees that $\sum_{i,j=1}^{3} \text{vote}(l_i, r_j) = 1$.

We can compute the $P(S = \Lambda \mid L = l_i, R = r_j)$ as

$$\begin{aligned} &P(S = \Lambda \mid L = l_i, R = r_j) \\ &\quad = \beta[\alpha \times \text{ratio}(l_i, r_j) + (1 - \alpha)\text{ratio\_non}(l_i, r_j)] \\ &\qquad + (1 - \beta)\text{vote}(l_i, r_j) \end{aligned} \tag{16}$$

where α and β are always chosen to be 0.7 and 0.9, respectively.

The result of this process is the determination of the hopefully more precise left and right boundaries of the road corresponding to the value $P(L = l_i, R = r_j \mid S = \Lambda)$. If two values computed by (16) are closed to each other, we chose the left and right boundaries pair that is close to the road center.

V. Experiments

Experiments in urban traffic with the THMR-V(Tsinghua Mobile Robot V), an autonomous vehicle developed by Tsinghua University, Beijing, China, demonstrated the feasibility and robustness of our method. The above proposed algorithms have been implemented in both Matlab and C++ on a 1.5-GMHz PC without any code optimization.

A. Offline Computation

In all experiments, an area whose range is 50 m ahead and 20 m wide is mapped to images whose size is 128 × 128 pixels. Before running the road-detection system, a user needs to specify some camera-specific and time-invariant parameters offline. These parameters are as follows:

- *VanishY*: vertical position of the vanishing point that can be computed offline by the result of camera calibration;
- *I2G and G2I*: two "lookup tables"; $\text{I2G}(\text{u}, \text{v}) = (\text{x}, \text{y})$ means that point (u, v) in the image plane corresponds to point (x, y) in the road plane and $\text{G2I}(\text{x}, \text{y}) = (\text{u}, \text{v})$ is the inverse map;
- *Cur1–Cur9*: nine lookup tables; the value stored in $\text{Curi}(\text{x}, \text{y})$ is the "shifting offset" of point (x, y) under assumed curvature i.

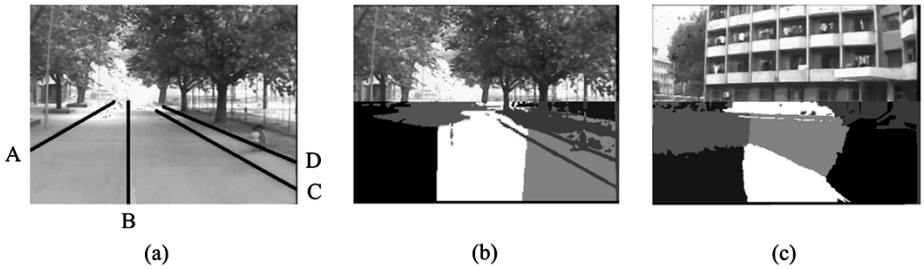


Fig. 11. Result of clustering. (a) Possible road boundaries in image, (b) clustering result of a straight road, and (c) clustering result of a curved road.

B. Online Computation

Here, we show two scenes to demonstrate the effectiveness of the method presented in this paper, one for a case of a straight road and the other for a curved road, as shown in Figs. 9(a) and 10(a).

For Scene 1, an original image is shown in Fig. 9(a). Fig. 9(b) shows the result of projecting image Fig. 9(a) onto the road plane. Fig. 9(c) is the edge image of Fig. 9(b), as extracted by erosion and dilation morphological operators. We then use nine curvatures, shown in Fig. 4(a), to shift the image of Fig. 9(c). The three most likely candidates for road detection are shown in Fig. 9(d). Fig. 9(e) exhibits the voting results according to the columns of images shown in Fig. 9(d). Then, Table I gives us the position of three candidates for the left and right boundaries; values in the table are the positions of the columns in the projection images. From Fig. 9 and Table I, we can know that it is difficult to extract the road boundaries only from the intensity image. Subsequently, Fig. 9(f) shows the area we use to compute the mean value and variance in the Gaussian distribution parameter estimation module. Then, Fig. 9(g) is the result of road-area extraction using the algorithm presented in Section IV-A. Next, Table II shows us the values of probability of (16), which correspond to nine combination of candidate left and right boundaries. Fig. 9(h) is the result of road boundaries finding using the algorithms present in Section IV-B. Fig. 10 is the result in curve situation.

Final results are then given in Figs. 9(h) and 10(h), which show excellent agreement with our visual assessments of the situation.

TABLE I
BOUNDARIES CANDIDATES OF SCENES 1 AND 2

Sn	Left candidate (position)			Right candidate (position)		
	l_1	l_2	l_3	r_1	r_2	r_3
#1	58	81	0	0	106	120
#2	0	64	75	90	102	107

Sn represents scene; value 0 represent the selected position of this candidate is unlikely to be the boundary.

C. Comparison

We will compare the algorithm presented in this paper with three kinds of methods: edge based, area based, and combination edge and area based. The images size captured, which are used to extract the road area, is 192×144 pixels. An area whose range is 50 m ahead and 20 m wide is mapped to images whose size is 128×128 pixels. The computational time in the tables is based on the MATLAB code.

1) Compare With the Edge-Based Method: Figs. 9 and 10 show us that perhaps there may be no clear left–middle–right pairs in the images captured in urban or campus traffic scenes. Sometimes the road surface has clear middle road lane, just as shown in Fig. 9, but sometimes there has not, as shown in Fig. 10. Therefore, there have many confused lines in edge image which may affect the judgement of true road boundaries; for example, the lines A, B, C, and D, shown in Fig. 11(a). Because the urban and campus road surface is not wide, the too narrow detection result based on edge image always is not acceptable, as shown in Figs. 9(f) and 10(f). Therefore, it is difficult to use the algorithms described in [3]–[11] to extract reliable road boundaries from the edge image.

2) Compare With the Area-Based Method: The methods that label every pixel in an image as road or nonroad can be regarded as the area-based method. Some methods exist that are based supervised learning, which always take a lot of time in the training and testing steps [27], [28] and make little use of the existing features of the road in urban and campus environments. We now only consider the image segmentation based on clustering. As described in UNSCARF [13], we cluster the pixel under the vanish point into six classes, using a five-dimensional feature image where each pixel x is

$$x = [\text{red}, \text{green}, \text{blue}, \text{row}, \text{column}]. \tag{17}$$

The clustering method adopted by UNSCARF is iterative self-organizing data-analysis technique algorithm (ISODATA). We also show that the computational time of the k-mean method. In addition, in order to simulate the SCARF algorithm, we first predict the positions of the road area. Then, we cluster the road and nonroad areas into four classes, respectively, using a 3-D feature image where each pixel x is

$$x = [\text{red}, \text{green}, \text{blue}]. \tag{18}$$

The values in Table III show us that only the clustering step is too slow to meet the real-time requirement. In addition, the basic assumption of the SCARF and UNSCARF is that the color of the surrounding area can be distinguished from the color of the road area. Fig. 11(b) and (c), which are the clustering results of Figs. 9(a) and 10(a), show us that it is difficult to distinguish the road from its surroundings only by color for the similar color of road area and the area that is beyond road shoulder. Furthermore, the road-model-matching step of SCARF and UNSCARF

TABLE II
PROBABILITY OF NINE COMBINATIONS

Sn	$P(S=\Lambda \mid L=l_i, R=r_j)$								
	$l_1\&r_1$	$l_1\&r_2$	$l_1\&r_3$	$l_2\&r_1$	$l_2\&r_2$	$l_2\&r_3$	$l_3\&r_1$	$l_3\&r_2$	$l_3\&r_3$
#1	0	0.56	0.54	0	0.46	0.43	0	0	0
#2	0	0	0	0.55	0.63	0.60	0.46	0.53	0.50

Sn represents scene.

TABLE III
COMPUTATIONAL TIME COMPARISON

Methods	Clustering step	Time(s)
UNSCARF	k-means clustering (6 classes)	5.02
	ISODATA (6 classes)	110.8
SCARF	Road Area clustering (k-means 4 classes)	1.32
	Non-road Area clustering (k-means 4 classes)	0.9210
Our method	Boundary Module	0.7310
	Area Module	0.7110

also is time consuming. Because they do not use any edge information of the image, they must compute the confident values for a lot of models in order to find the true boundaries of the road. The delimited area segmentation method also cannot acquire a good segmentation when it does not use the edge information in image segmentation [18].

3) Compare With the Edge-and-Area Combination-Based Method: Most of edge- and area-based methods were composed of independent edge-detection and image-segmentation modules [21]–[24]. Therefore, it also is time consuming. In addition, the fusion of these two modules may give the wrong detection result because they did not consider several road boundaries candidates. For example, the area between C and D in Fig. 11(a) can be regarded as road surface by the methods presented in [21] and [22] if the threshold value is chosen improperly. The reason is that the area delimited by C and D has a similar to color the road surface although it is separated from the road area by road shoulders. In our method, we overcome this problem by choosing the road boundaries pair that is close to the road center if the probability values of two combinations are close to each other. In their methods, they also can regard the area between A and B in Fig. 11(a) as the only road surface because the searching of maximum voting value is for the entire image without any previous knowledge. However, in our algorithm, we search the left and right candidates in different ranges of the image.

Based on the previous discussion, we can conclude that it is difficult to design an algorithm that can work well on all kinds of roads. The methods based on intensity image can give a good result for road detection on the highway and well-structure roads; the image-segmentation methods are suitable for the roads in park, where the difference between road and nonroad areas is often characterized by color. The methods presented in this paper can work well on the road in urban traffic or on a campus, where the road is semistructured.

VI. CONCLUSION

In this paper, we describe and implement a framework for efficient exacting road area and boundaries. This method combines the result of boundaries estimation of graylevel image and road-area extraction based on color image. It succeeds in addressing the road-detection problem in urban traffic.

Future work categories may include:

- *Predict the position of boundaries and color of road surface frame to frame:* This work can perhaps be done using a Kalman filter.
- *Removing the affect of shadow and water on road surface:* We can achieve this goal by analyzing the color component of hue-saturated value (HSV) color space.

ACKNOWLEDGMENT

The authors would like to thank Dr. H. Araujo for valuable discussions. They would also like to thank Dr. A. M. Bruckstein for much help regarding the language and content of this paper.

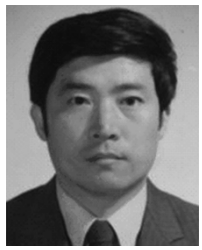
REFERENCES

[1] M. Bertozzi, A. Broggi, and A. Fascioli, "Vision-based intelligent vehicles: State of the art and perspectives," *Robot. Autonom. Syst.*, vol. 32, pp. 1–16, 2000.
[2] D. Dufourd and A. Dalgalarrondo, "Quantitative evaluation of image processing algorithms for unstructured road edge detection," in *Proc. SPIE Unmanned Ground Vehicle Technology V*, vol. 5083, Orlando, FL, Apr. 2003, pp. 440–451.
[3] M. Bertozzi and A. Broggi, "GOLD: A parallel real-time stereo vision system for generic obstacle and lane detection," *IEEE Trans. Image Processing*, vol. 7, pp. 62–81, Jan. 1998.
[4] A. Broggi, "Robust real-time lane and road detection in critical shadow conditions," in *Proc. IEEE Int. Symp. Computer Vision*, Coral Gables, FL, Nov. 1995, pp. 353–359.
[5] A. Broggi and S. Bertè, "Vision-based road detection in automotive systems: A real-time expectation-driven approach," *J. Artificial Intell. Res.*, vol. 3, pp. 325–348, 1995.
[6] D. Pomerleau, "RALPH: Rapidly adapting lateral position handler," in *Proc. IEEE Intelligent Vehicles Symp.*, Detroit, MI, Sept. 1995, pp. 506–511.
[7] B. Southall and C. J. Taylor, "Stochastic road shape estimation," in *Proc. IEEE Int. Conf. Computer Vision (ICCV'01)*, vol. 1, Vancouver, BC, Canada, July 2001, pp. 205–212.
[8] J. P. Tarel and F. Guichard, "Combined dynamic tracking and recognition of curves with application to road detection," in *Proc. IEEE Int. Conf. Image Processing*, vol. 1, Vancouver, BC, Canada, Sept. 2000, pp. 216–219.
[9] K. Kanatani and K. Watanabe, "Reconstruction of 3-D road geometry from images for autonomous land vehicles," *IEEE Trans. Robot. Automat.*, vol. 6, pp. 127–132, Feb. 1990.
[10] K. Kluge and C. Thorpe, "Representation and recovery of road geometry in YARF," in *Proc. IEEE Symp. Intelligent Vehicles*, July 1992, pp. 114–119.

[11] P. Katsande and P. Liatsis, "Adaptive order explicit polynomials for road edge tracking," in *Proc. IEEE Int. Conf. Advanced Driver Assistance Systems, (ADAS'01)*, Sept. 2001, pp. 63–67.
[12] J. D. Crisman and C. E. Thorpe, "SCARF: A color vision system that tracks roads and intersections," *IEEE Trans. Robot. Automat.*, vol. 9, pp. 49–58, Feb. 1993.
[13] J. Crisman and C. E. Thorpe, "UNSCARF: A color vision system for the detection of unstructure roads," in *Proc. IEEE Int. Conf. Robotics and Automation*, Apr. 1991, pp. 2496–2501.
[14] B. Serge and B. Michel, "Road segmentation and obstacle detection by a fast watershed transform," in *Proc. IEEE Intelligent Vehicles Symp.*, Oct. 1994, pp. 296–301.
[15] C. Rasmussen, "Grouping dominant orientations for ill-structured road following," in *Proc. IEEE Comp. Soc. Conf. Computer Vision and Pattern Recognition*, vol. 1, July 2004, pp. 470–477.
[16] ——, "Combining laser range, color, and texture cues for autonomous road following," in *Proc. IEEE Int. Conf. Robotics and Automation*, vol. 4, May 2002, pp. 4320–4325.
[17] K. Onoguchi, N. Takeda, and M. Watanabe, "Planar projection stereopsis method for road extraction," *IEICE Trans. Inform. Syst.*, vol. E81-D, no. 9, pp. 1006–1018, 1998.
[18] R. Chapuis, R. Aufrere, and F. Chausse, "Accurate road following and reconstruction by computer vision," *IEEE Trans. Intell. Transport. Syst.*, vol. 3, pp. 261–270, Dec. 2002.
[19] S. Y. Oh, J. H. Lee, and D. H. Choi, "A new reinforcement learning vehicle control architecture for vision-based road following," *IEEE Trans. Veh. Technol.*, vol. 49, pp. 997–1005, May 2000.
[20] F. Paezold and U. Franke, "Road recognition in urban environment," in *Proc. IEEE Int. Conf. Intelligent Vehicles*, 1998, pp. 87–91.
[21] F. W. J. Gibbs and B. T. Thomas, "The fusion of multiple image analysis algorithms for robot road following," in *Proc. IEEE 5th Int. Conf. Image Processing and Its Applications*, July 1995, pp. 394–398.
[22] N. Apostoloff and A. Zelinsky, "Robust vision based lane tracking using multiple cues and particle filtering," in *Proc. IEEE Intelligent Vehicles Symp.*, June 2003, pp. 558–563.
[23] M. Lutzeler and S. Baten, "Road recognition for a tracked vehicle," in *Proc. SPIE Enhanced and Synthetic Vision*, 2000, pp. 171–180.
[24] Y. Xiuqing, L. Jilin, and G. Weikang, "An integrated vision system for alv navigation," *Int. J. Pattern Recogn. Artif. Intell.*, vol. 14, pp. 929–940, 2000.
[25] P. Golland and A. M. Bruckstein, "Motion from color," *Comp. Vision Image Understanding*, vol. 68, no. 3, pp. 346–362, 1997.
[26] S. Birchfield, "Elliptical head tracking using intensity gradients and color histograms," in *Proc. IEEE Conf. Computer Vision and Pattern Recognition*, Santa Barbara, CA, June 1998, pp. 232–237.
[27] S. Baluja, "Evolution of an artificial neural network based autonomous land vehicle controller," *IEEE Trans. Syst., Man, Cybernetics B*, vol. 26, pp. 450–463, June 1996.
[28] P. Conrad and M. Foedisch, "Performance evaluation of color based road detection using neural nets and support vector machines," in *Proc. IEEE Applied Imagery Pattern Recognition Workshop*, Oct. 2003, pp. 157–160.

Yinghua He received the B.S. and M.S. degrees in computer science from Huazhong University of Science and Technology, Wuhan, China, in 1997 and 2000. She is currently working toward the Ph.D. degree in computer science and technology at Tsinghua University, Beijing, China.

Her research interests include computer vision, intelligent robotics, pattern recognition, and machine learning.



Hong Wang received the Ph.D. degree from the Department of Computer Science and Technology, Tsinghua University, Beijing, China, in 1993.

He is an Associate Professor in the Department of Computer Science and Technology, Tsinghua University. His main research interests include mobile robots, computer vision, data fusion, and artificial intelligence. He has published more than 40 papers in these fields.

Bo Zhang was born in Fujian, China. He graduated from Tsinghua University, Beijing, China, in 1958.

He now a Professor in the Computer Science and Technology Department, Tsinghua University, and an Academician of the Chinese Academy of Science. His main research interests include artificial intelligence, neural networks, robotics, and pattern recognition. He has published approximately 130 papers and three books in these fields.