Lane Detection in Unstructured Environments for Autonomous Navigation Systems

Manh Cuong Le, Son Lam Phung and Abdesselam Bouzerdoum

School of Electrical, Computer and Telecommunications Engineering
University of Wollongong, Australia

Abstract. Automatic lane detection is an essential component for autonomous navigation systems. It is a challenging task in unstructured environments where lanes vary significantly in appearance and are not indicated by painted markers. This paper proposes a new method to detect pedestrian lanes that have no painted markers in indoor and outdoor scenes, under different illumination conditions. Our method detects the walking lane using appearance and shape information. To cope with variations in lane surfaces, an appearance model of the lane region is learned on-the-fly. A sample region for learning the appearance model is automatically selected in the input image using the vanishing point. This paper also proposes an improved method for vanishing point estimation, which employs local dominant orientations of edge pixels. The proposed method is evaluated on a new data set of 1600 images collected from various indoor and outdoor scenes that contain unmarked pedestrian lanes with different types and surface patterns. Experimental results and comparisons with other existing methods on the new data set have demonstrated the efficiency and robustness of the proposed method.

1 Introduction

Lane detection plays a vital role in assistive navigation for blind people, autonomous vehicles and mobile robots. The aim of a lane detection system is to locate the lane region in each scene in front of the traveler. The system must cope with variations in the scene, the illumination condition, and the lane type. Automatically finding lanes using cameras is a popular approach for assistive navigation systems [1–3]. For autonomous cars, numerous vision-based algorithms of vehicle lane detection have been proposed [2, 4–8]. However, there has been little work on pedestrian lane detection for assistive navigation of visually impaired people [1, 9, 10]. Furthermore, most existing pedestrian lane detection methods are designed to find pedestrian crossings, which are identified by painted markers [1, 9–12]. To address this gap, this paper concentrates on vision-based detection of pedestrian lanes that have no painted markers for indoor and outdoor scenes, under varying illumination conditions and lane surfaces.

Most existing methods for unstructured (i.e. unmarked) lane detection exploit the appearance properties (e.g. color and texture) of lane surfaces to classify the lane pixels from the background [6, 13–15]. In these methods, the classifiers

are trained off-line, and hence the detection performance is degraded when lane appearance differs from the training data (e.g. due to change in lane surface types or illumination conditions). In another approach, several algorithms detect the lane boundaries based on edge features (e.g. color and orientation) [2, 16]. However, using only edges to identify the lane borders is sensitive to background clutter, and the detection performance is significantly affected by the robustness of edge detection.

This paper proposes a method for detecting unmarked pedestrian lanes, using appearance and shape information. In contrast to existing methods, the appearance model of the lane region is constructed on-the-fly employing the vanishing point, and therefore is invariant to varying illumination conditions and different lane surfaces. Furthermore, shape context [17] is used to model the shape of pedestrian lanes. The main contributions of the paper can be briefly described as follows:

- We propose an improved vanishing point estimation method, which is based on the votes of local orientations from color edge pixels. Using only edge pixels for voting the vanishing point is more efficient than using all pixels as in the existing methods [2, 16]. Furthermore, to estimate robustly local orientations and edge pixels under severe illumination conditions, our method employs multiple color channels, instead of only the intensity channel.
- We propose using the vanishing point to identify a sample region on the input image for learning the appearance model of the pedestrian lane surface. The appearance model is thus adaptive to various types of lane surfaces. To make the appearance model invariant to different illumination conditions, a so-called illumination invariant space (IIS) is adopted.
- We propose a probabilistic model that combines both appearance and shape information for detecting unstructured pedestrian lanes. To evaluate pedestrian lane detection methods, we also create a new data set, collected from various indoor and outdoor environments with different types of unmarked lanes.

The remainder of the paper is organized as follows. Existing methods for lane detection in unstructured environments are reviewed in Section 2. The proposed method is described in Section 3. Experimental results are presented in Section 4. Finally, conclusions are given in Section 5.

2 Related work

This section presents briefly vision-based approaches for unstructured lane detection. There are two major approaches: lane segmentation and lane border detection.

In the *lane segmentation* approach, off-line color models are used for classifying the lane pixels from the background [6, 18, 19, 13]. The color models are first constructed from manually-selected sample regions, and then updated from the

detected regions in the sequence frames. Different color spaces and color classifiers have been used. For example, Crisman and Thorpe use Gaussian models of red-green-blue (RGB) color components to represent the appearances of the road surface and background [6]. Tan *et al.* also use RGB components, but model the variability of the road surface by multiple histograms and the background by a single histogram [13]. Instead of using RGB components, Ramstrom and Christensen employ UV, normalized red and green components, and luminance to construct Gaussian mixture models for road surface and background classes. Sotelo *et al.* employ the hue-saturation-intensity (HSI) color space [18]. Because the color models are trained off-line, these methods do not cope well with the appearance variation of lane surfaces.

To address this problem, several algorithms construct the appearance model of the lane pixels directly from sample regions in the input image [20–22, 7]. These algorithms determine the sample lane regions by different ways. For example, in [7, 8], the sample lane regions are selected as small random areas at the bottom and middle parts of the input image. Miksik *et al.* initialize the sample lane region as a trapezoid area at the bottom and central part of the image, and then refine the sample region using the vanishing point [22]. He *et al.* determine the sample lane region from the potential lane boundaries, which are detected using the vanishing point and the width prior of lanes [20]. The performance of these algorithms depends on the quality of the sample regions, which in turns relies on the prior knowledge of the lanes.

In the *border detection* approach, the lane boundaries are determined using the vanishing point [16, 2] or the predefined models of the lane boundaries [23]. In [16], the lane borders are detected simultaneously from edges directing to the vanishing point, employing the color difference between the lane region and non-lane regions. This method is effective only when the lane region is homogeneous and differs significantly in color from non-lane regions. Kong *et al.* also find the lane borders from edges directing to the vanishing point, except that their method uses the orientation and color cues of the edges [2]. Since this method is only based on edges for lane border detection, it is sensitive to background edges. In another method, the lane boundaries are located from the edges of homogeneous color regions, using the predefined models [23]. Recently, Chang *et al.* propose combining lane border detection and road segmentation for detecting the lane region [3]. Similarly to [2], their method detects lane borders using the vanishing point. The lane region is segmented using the color model learned from a homogeneous region at the bottom and middle part of the input image. Chang *et al.* 's method also relies on the prior knowledge of the lane location.

3 Proposed method

The proposed method for detecting unstructured pedestrian lanes is based on the appearance and shape of the pedestrian lane. To make the detection method adaptive to different road surface structures, the appearance model of the lane

region is learned automatically from a sample lane region, which is selected using the vanishing point. Shape context descriptor [17] is employed to model the shape of the lane region. The proposed method includes three main stages: vanishing point estimation, sample region selection, and lane detection. Each stage is described in the following subsections.

3.1 Vanishing point estimation

The vanishing point in an image is often found based on either line segments [4, 24, 25] or local orientations [26, 2]. The algorithms using line segments are only suitable for structured environments where there exist straight edges. For unstructured environments, most existing vanishing point detection methods use local orientations [26, 2, 27]. These methods compute the local orientations of pixels using Gabor filters on the intensity channel, and therefore are not robust under challenging illumination conditions. Furthermore, the methods have high complexity and are sensitive to background clutter. To cope with the problems, we estimate the local orientations using the color tensors and detect the vanishing point employing the local orientations of edge pixels.

Given a color image $\mathbf{f}$, the tensor components are calculated on three color channels as in [28]:

$$\begin{cases} \mathbf{g}_{xx} &= \sum_{k=1}^{3} \mathbf{g}_{xk}\mathbf{g}_{xk}, \\ \mathbf{g}_{yy} &= \sum_{k=1}^{3} \mathbf{g}_{yk}\mathbf{g}_{yk}, \\ \mathbf{g}_{xy} &= \sum_{k=1}^{3} \mathbf{g}_{xk}\mathbf{g}_{yk}, \end{cases} \tag{1}$$

where $\mathbf{g}_{xk} = \mathbf{w} * \mathbf{d}_x^k$ and $\mathbf{g}_{yk} = \mathbf{w} * \mathbf{d}_y^k$; $*$ denotes the convolution operator; $\mathbf{w}$ is the convolution kernel of a Gaussian filter; $\mathbf{d}_x^k$ and $\mathbf{d}_y^k$ denote the spatial derivatives of the color channel k. The local dominant orientation field is estimated as

$$\theta = \frac{1}{2} \arctan\left(\frac{2\mathbf{g}_{xy}}{\mathbf{g}_{xx} - \mathbf{g}_{yy}}\right). \tag{2}$$

Figure 1(b) shows the local orientations of sampled pixels for the input image in Fig. 1(a).

To estimate the edge map of $\mathbf{f}$, we apply the color Canny edge detector, which is proposed in [28]. This edge detector computes the magnitude and orientation of pixels using the color tensor, and therefore is more robust than the conventional Canny edge detector using intensity gradients. Figure 1(c) shows the edge map detected from the input image in Fig. 1(a).

The vanishing point is determined by a voting scheme as follows. Each pixel location $v = (x_v, y_v)$ is considered as a vanishing point candidate, and voted by edge pixels $p = (x_p, y_p)$ that are below v. The voting score is computed as in [2]:

$$s(v,p) = \begin{cases} \frac{1}{1+(\delta\mu)^2} & \text{if } \delta \leq \frac{5}{1+2\mu}, \\ 0 & \text{otherwise.} \end{cases} \tag{3}$$

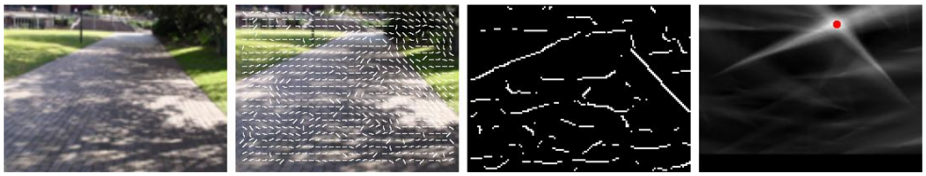


Fig. 1. Illustration of vanishing point estimation: (a) input image, (b) local orientations estimated by the color tensor for sampled pixels, (c) edge map, (d) voting map and the vanishing point (in red).

Here, δ is the difference between the local orientation $\theta(x_p, y_p)$ and the angle of the line L_{vp} connecting v and p; μ be the ratio between the length of L_{vp} and the diagonal length of the image. Equation (3) means that $s(v, p)$ is high when edge pixel p is close to v and is consistent in orientation with the line L_{vp}. The vanishing point is finally found as the point that has the highest sum of the voting scores. Figure 1(d) demonstrates the voting map and the vanishing point computed for the image in Fig. 1(a).

3.2 Sample region selection

Since the appearance (e.g. texture and color) of pedestrian lane regions is varied and strongly affected by illumination conditions, it is difficult to obtain an robust appearance model via off-line training. In our method, the appearance model is computed directly on the input image. Based on the vanishing point estimated in the previous step, a sample region of the pedestrian lane is automatically selected, and the appearance model is then constructed from the pixels in the sample region.

In existing methods (e.g. [8, 21, 22]), the sample region is chosen as a small region at the bottom and middle of the input image. However, the sample region selected in such a manner may include the background when the lane region is not located at the middle of the image. In our method, the sample region is automatically detected using the vanishing point and the geometric and appearance characteristics of the lane region.

Given the vanishing point estimated in the previous step, a set of N imaginary rays $\{r_1, r_2, ..., r_N\}$ is created as shown in Fig. 2(a). These rays are uniformly distributed in a fixed angle range $[\alpha_{\min}, \alpha_{\max}]$ relative to the horizontal direction. The training region is identified by finding a ray pair (r_i, r_j) that best represents the characteristics of the lane region. These characteristics include: 1) the lane direction φ estimated as the direction of the bisector between r_i and r_j; 2) the uniformity u of color pixels in the lane region R_{ij} formed by a ray pair (r_i, r_j) as shown in Fig. 2(b). The uniformity u is computed using the color histogram. This is similar to [29], except that our method uses three color channels instead

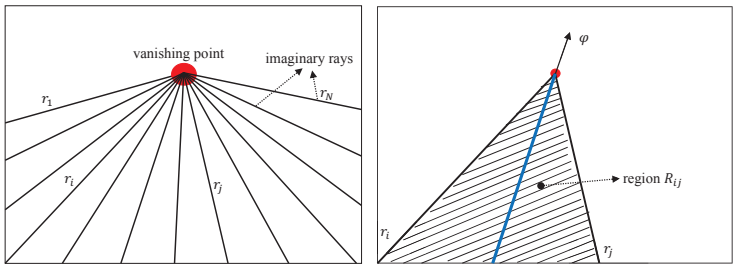


Fig. 2. Imaginary rays: (a) imaginary rays; (b) a pair of rays.

of intensity:

$$u = \sum_{m=1}^{M} \sum_{n=1}^{M} \sum_{k=1}^{M} h(m,n,k)^2. \tag{4}$$

Here, h is the normalized 3-D color histogram of R_{ij}, and M is the number of bins for each color channel.

Let $\mathcal{L}$ denote the lane class. The conditional probability of $\mathcal{L}$ for two features φ and u can be expressed as

$$P(\mathcal{L}|\varphi,u) \propto p(\varphi,u|\mathcal{L})P(\mathcal{L}) \propto p(\varphi|\mathcal{L})p(u|\mathcal{L}), \tag{5}$$

where $P(\mathcal{L})$ is prior probability, $p(\varphi|\mathcal{L})$ and $p(u|\mathcal{L})$ are the class-conditional probability density functions (*pdf*). Here, we assume that $p(\varphi,u)$ is uniform, $P(\mathcal{L})$ is constant, φ and u are statistically independent for a given $\mathcal{L}$.

We have found that the distribution of the lane directions φ in the training set is similar to a normal distribution. Therefore, $p(\varphi|\mathcal{L})$ is modeled as

$$p(\varphi_{ij}|\mathcal{L}) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(\varphi-\overline{\varphi})^2}{2\sigma^2}}, \tag{6}$$

where $\overline{\varphi}$ and σ are the mean value and standard deviation of φ that are computed using the training data.

Our experiments have shown that lane regions have high uniformity and u varies in a range from 0 to 1. Based on these characteristics, we model $p(u|\mathcal{L})$ using the beta function as

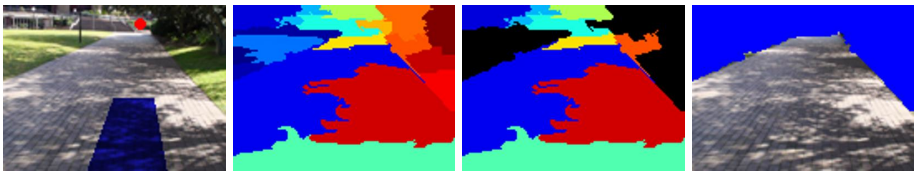
$$p(u|\mathcal{L}) = \frac{1}{B(\alpha,\beta)} u^{\alpha-1}(1-u)^{\beta-1}, \tag{7}$$

where $B(\alpha,\beta)$ is the beta function, α and β are positive parameters. These parameters are selected so that $p(u|\mathcal{L})$ is high when u is high, and vice versa.

Finally, the sample region R^* is obtained from a pair of rays (r_i^*, r_j^*) as

$$(r_i^*, r_j^*) = \arg\max_{r_i, r_j} P(\mathcal{L}|\varphi,u). \tag{8}$$

Since the vanishing point could be located outside the lane region, we use only the bottom half of R^* for training. Figure 3(a) shows an example of sample region selection.



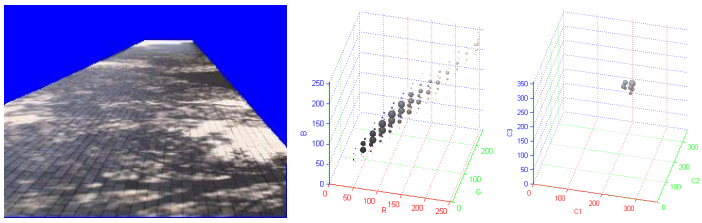
(a) (b) (c) (d)

Fig. 3. Illustration of the proposed method for pedestrian lane detection: (a) training region (blue area) extracted from the vanishing point (red dot); (b) segmented regions using graph-based segmentation method [30]; (c) candidate regions considered in $\mathcal{R}'$ (background regions are marked in black); (d) detected walking lane.

Given the sample region R^*, the appearance model of the lane region is represented as the class-conditional *pdf* $p(\mathbf{c}|\mathcal{L})$ of colors $\mathbf{c}$ for the lane class $\mathcal{L}$. The *pdf* $p(\mathbf{c}|\mathcal{L})$ is estimated using the color histogram of pixels in R^*. However, instead of using the RGB space which is sensitive to illumination conditions and shading, we convert colors from the RGB space to an illumination invariant space (IIS) [31] as

$$\begin{cases} C_1 &= \arctan\{R/\max(G,B)\}, \\ C_2 &= \arctan\{G/\max(R,B)\}, \\ C_3 &= \arctan\{B/\max(R,G)\}. \end{cases} \tag{9}$$

Figure 4 shows the color distribution of a lane region in the RBG and IIS space. The lane pixels have less variations in the IIS space than the RGB space.



(a) (b) (c)

Fig. 4. Color distributions of lane pixels: (a) a sample lane region; (b) color distribution of lane pixels in the RGB space; (c) color distribution of lane pixels in the IIS space.

3.3 Lane detection

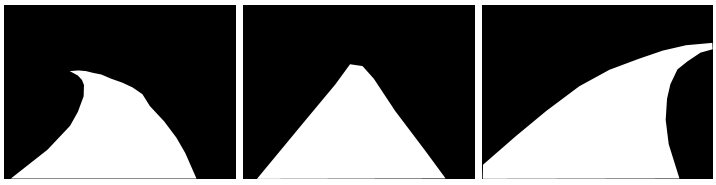
This subsection presents a method to detect the walking lane in the input image using both appearance and shape information. In our method, the input image is first segmented into homogeneous regions. The walking lane region is then determined by merging those image regions using appearance and shape criteria.

Let $\mathcal{R} = \{R_1, R_2, ...\}$ be the set of homogeneous regions obtained using the graph-based segmentation method in [30]. Figure 3(b) illustrates the homogeneous regions segmented from the input image in Fig. 3(a). The lane region is a combination of connected regions of $\mathcal{R}$. The two regions R_i and R_j are considered as connected if there exist two pixels $p_i \in R_i$ and $p_j \in R_j$ that are connected (e.g. 4-connected pixels). The lane region is the subset Z^* of $\mathcal{R}$ such that

$$\begin{aligned} Z^* &= \arg\max_{Z \subset \mathcal{R}} p(Z|\mathcal{L}) \\ &= \arg\max_{Z \subset \mathcal{R}} p(\bar{\mathbf{c}}_z, \mathbf{s}_z|\mathcal{L}) \\ &= \arg\max_{Z \subset \mathcal{R}} p(\bar{\mathbf{c}}_z|\mathcal{L}) p(\mathbf{s}_z|\mathcal{L}). \end{aligned} \tag{10}$$

Here, $\bar{\mathbf{c}}_z$ is an appearance feature and $\mathbf{s}_z$ is a shape feature of region Z. It is also assumed that $\bar{\mathbf{c}}_z$ and $\mathbf{s}_z$ are statistically independent.

In (10), $\bar{\mathbf{c}}_z$ is defined as the mean color of all pixels in Z, and $p(\bar{\mathbf{c}}_z|\mathcal{L})$ is the *pdf* of the lane color and is learned from the sample region. For shape feature $\mathbf{s}_z$, we adopt the shape context descriptor proposed in [17]. Shape context descriptor is known for its robustness to local shape deformation and partial occlusion, and its invariance to scale and rotation. The shape context of a point p is the histogram of locations of points other than p in relative to p. The similarity between two shapes is computed as the matching cost between the corresponding sets of points on the two shapes.



(a) (b) (c)

Fig. 5. Several shape templates of the lane region: (a) left-curved lane, (b) straight lane, (c) right-curved lane.

Let $\mathcal{T} = \{T_1, T_2, ...\}$ be a set of shape templates for the pedestrian lane . In our implementation, 10 templates are used, and some of the templates are shown in Fig. 5. Each template is sampled so that the distance between two adjacent

points on the template is approximately 5 pixels. The sampling is performed similarly for regions Z in the test image. The likelihood $p(\mathbf{s}_z|\mathcal{L})$ is defined as

$$p(\mathbf{s}_z|\mathcal{L}) = \exp\left[-\lambda \min_{T\in\mathcal{T}} D(\mathbf{s}_z, T)\right], \tag{11}$$

where $D(\mathbf{s}_z, T)$ is the matching cost between the approximated outer contour $\mathbf{s}_z$ and a template T as in [17]. Note that the smaller is the matching cost $D(\mathbf{s}_z, T)$, the higher is the similarity between $\mathbf{s}_z$ and T, and therefore the likelihood $p(\mathbf{s}_z|\mathcal{L})$ is higher. In (11), λ is a positive scalar determined through training data.

Given the appearance likelihood $p(\bar{\mathbf{c}}_z|\mathcal{L})$ calculated from the sample region and shape likelihood $p(\mathbf{s}_z|\mathcal{L})$ defined in (11), the optimal set Z^* can be obtained with a computational complexity of $O(2^{|\mathcal{R}|})$ by exhaustively searching every possible sub-set of $\mathcal{R}$. To reduce the computational burden, we propose a greedy algorithm that generates Z^* by iteratively adding and removing regions. To further accelerate the algorithm, we only consider regions $R_i \in \mathcal{R}$ with $p(\bar{\mathbf{c}}_i|\mathcal{L})$ greater than or equal to τ_c, where $\bar{\mathbf{c}}_i$ is the mean color of all pixels in R_i and τ_c is a predefined threshold. The greedy algorithm is described in *Algorithm* 1.

Algorithm 1 Adding and moving regions for lane detection.

$\mathcal{R}' \leftarrow \{R_i \in \mathcal{R} \mid p(\bar{\mathbf{c}}_i|\mathcal{L}) \geq \tau_c\}$
$Z^* \leftarrow \arg\max_{R_i\in\mathcal{R}'} p(\bar{\mathbf{c}}_i|\mathcal{L})$
continue $\leftarrow$ TRUE
while (*continue*) **do**
  $\mathcal{R}_a \leftarrow \{R_i \in \{\mathcal{R}' - Z^*\}$ so that $Z^* \cup R_i$ is a connected set$\}$
  $R^+ = \arg\max_{R_i\in\mathcal{R}_a} p(\{Z^* \cup R_i\}|\mathcal{L})$
  $\mathcal{R}_r \leftarrow \{R_i \in Z^*$ so that $\{Z^* - R_i\}$ is a connected set$\}$
  $R^- = \arg\max_{R_i\in\mathcal{R}_r} p(\{Z^* - R_i\}|\mathcal{L})$
  if $p(\{Z^* \cup R^+\}|\mathcal{L}) \geq p(\{Z^* - R^-\}|\mathcal{L})$ **and** $p(\{\ Z^* \cup R^+\}|\mathcal{L}) > p(Z^*|\mathcal{L})$ **then**
    $Z^* \leftarrow Z^* \cup R^+$
  else if $p(\{Z^* - R^-\}|\mathcal{L}) > p(Z^*|\mathcal{L})$ **then**
    $Z^* \leftarrow \{Z^* - R^-\}$
  else
    continue $\leftarrow$ FALSE
  end if
end while

In Algorithm 1, at each iteration when an image region R_i is added to Z^* or removed from Z^*, the connectivity of $\{Z^* \cup R_i\}$ and $\{Z^* - R_i\}$ is checked. A set of regions is considered to be connected if any two regions in the set are connected. For example, supposed that set Z^* in Fig. 6 consists of regions 2, 6, 7, 8, and 10. Region 8 will not be removed from Z^* because doing so will break the connectivity of Z^*. Similarly, region 3 will not be added to Z^*.

Because the number of regions is finite and operators in Algorithm 1 are deterministic, the algorithm will converge. Figure 3(d) illustrates the result of lane detection using Algorithm 1 for the input image shown in Fig. 3(a).

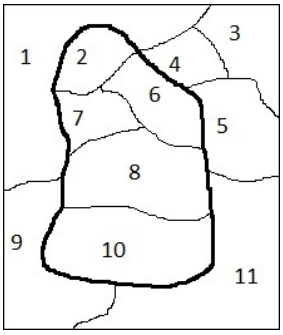


Fig. 6. Illustration of adding and removing regions.

4 Experimental Results

This section describes the image data, evaluation measures and parameters employed in the proposed method. It also presents experimental results for vanishing point estimation and pedestrian lane detection.

4.1 Experimental methods

To evaluate the proposed method, we collected a data set of 1600 images in different indoor and outdoor scenes. The data set includes unmarked pedestrian lanes with various surface structures and shapes. In many cases, lane regions are affected by extreme illumination conditions (e.g. very low or high illumination). We manually annotated lane regions and determined the vanishing point in each image. In the experiments, we used 500 images for training, and 1100 images for testing. Note that the training set was employed only for estimating the orientation angles of lane regions and finding the parameters of the proposed method, it was not used for learning the appearance model.

To measure the detection performance, detected regions are compared with annotated regions. Let R_d denote a detected region and R_g denote a ground-truth region. The matching score between R_d and R_g is computed as

$$\chi(R_g, R_d) = \frac{|R_g \cap R_d|}{|R_g \cup R_d|}, \tag{12}$$

where $|R|$ is the area of region R, $\cap$ and $\cup$ denote the intersection and union of R_d and R_g, respectively. A detected region R_d is considered as *correct* if there exists a ground-truth region R_g where $\chi(R_g, R_d)$ is greater than or equal to an evaluation threshold τ_e. Similarly to the evaluation of other object detection systems [32], τ_e is set to 0.5.

The pedestrian lane detection performance is evaluated by two measures: recall and precision. *Recall* is the percentage of the ground-truth lanes that are

detected correctly. *Precision* is the percentage of the detected lanes that are considered to be correct.

We also evaluate the accuracy of vanishing point estimation. Let P_d be the detected vanishing point, and P_g be the ground-truth vanishing point. Similarly to [27], the estimation error is measured by the ratio of the Euclidean distance from P_d to P_g and the diagonal length D_I of the image:

$$\text{ERR} = \frac{|P_d - P_g|}{D_I}. \tag{13}$$

In our experiments, the parameters of the proposed method were chosen based on analyzing the performance of the pedestrian lane detection on the training set. The window size $L \times L$ of $\mathbf{w}$ in Section 3.1 and the number of imaginary rays N in Section 3.2 are chosen as $N = 29$ and $L = 13$. Each color component in the IIS space is quantized into 180 bins. The parameters α and β in (7), λ in (11) and threshold τ_c in Section 3.3 are set as $\alpha = 2$, $\beta = 1$, $\lambda = 25$, and $\tau_c = 0.02$.

4.2 Experimental results

The proposed vanishing point estimation (VPE) method was evaluated and compared with two existing algorithms: Hough-based method [4] and Gabor-based method [2]. The Hough-based method applies the Hough transform on the edge map to find line segments, and then computes the vanishing point by voting the intersections of line pairs in the Hough space [4]. The Gabor-based method employs Gabor filters for computing local orientations, and a local adaptive scheme for estimating the vanishing point [2].

Table 1. Accuracy and speed of algorithms for vanishing point estimation.

Methods	*Average error*	*Computational time (s)*
Hough-based method [4]	0.250	0.06
Gabor-based method [2]	0.086	3.00
Proposed method	**0.057**	**0.60**

Table 1 shows the performance of different VPE algorithms. The average error of the proposed method (0.057) was significantly lower than the Gabor-based method (0.086) and the Hough-based method (0.250). Furthermore, the average processing time per image (of size 100×140 pixels) of the proposed method (0.60 s) was significantly shorter than the Gabor-based method (3.00 s). The Hough-based method had the shortest processing time, but it also had the lowest accuracy. Figure 7 shows several visual results of different VPE methods.

For pedestrian lane detection, we evaluated the proposed method using the IIS and RGB color space. In this comparison, the proposed VPE method was

Fig. 7. Visual results of vanishing point (VP) detection: red dot is the ground-truth VP; green dot is the VP detected by the proposed method; yellow dot is the VP detected by Hough-based method [4]; blue dot is the VP detected by Gabor-based method in [2]. See electronic color image.

Fig. 8. Visual results of different methods for pedestrian lane detection. *Row 1*: input images. *Row 2*: pedestrian lanes detected by the method in [2]. *Row 3*: pedestrian lanes detected by the proposed method using the RGB color space. *Row 4*: pedestrian lanes detected by the proposed method using the IIS color space.

used for both color spaces. As shown in the last two rows of Table 2, using the IIS space, the proposed method achieved a recall rate of 94.8% and a precision rate of 95.1%. Using the RGB space, the recall and precision rate decreased to 91.2% and 92.6%, respectively.

We also compared the proposed method with the lane border detection method of Kong *et al.* [2]. The method in [2] employs the vanishing point to determine the lane borders. As shown Table 2, the proposed method (recall rate of 94.8% and precision rate of 95.1%) outperformed significantly the method in [2] (recall rate of 63.9% and precision rate of 66.0%). Figure 8 shows several comparative results of different methods for pedestrian lane detection. These results demonstrate that the proposed method using the IIS color space is more

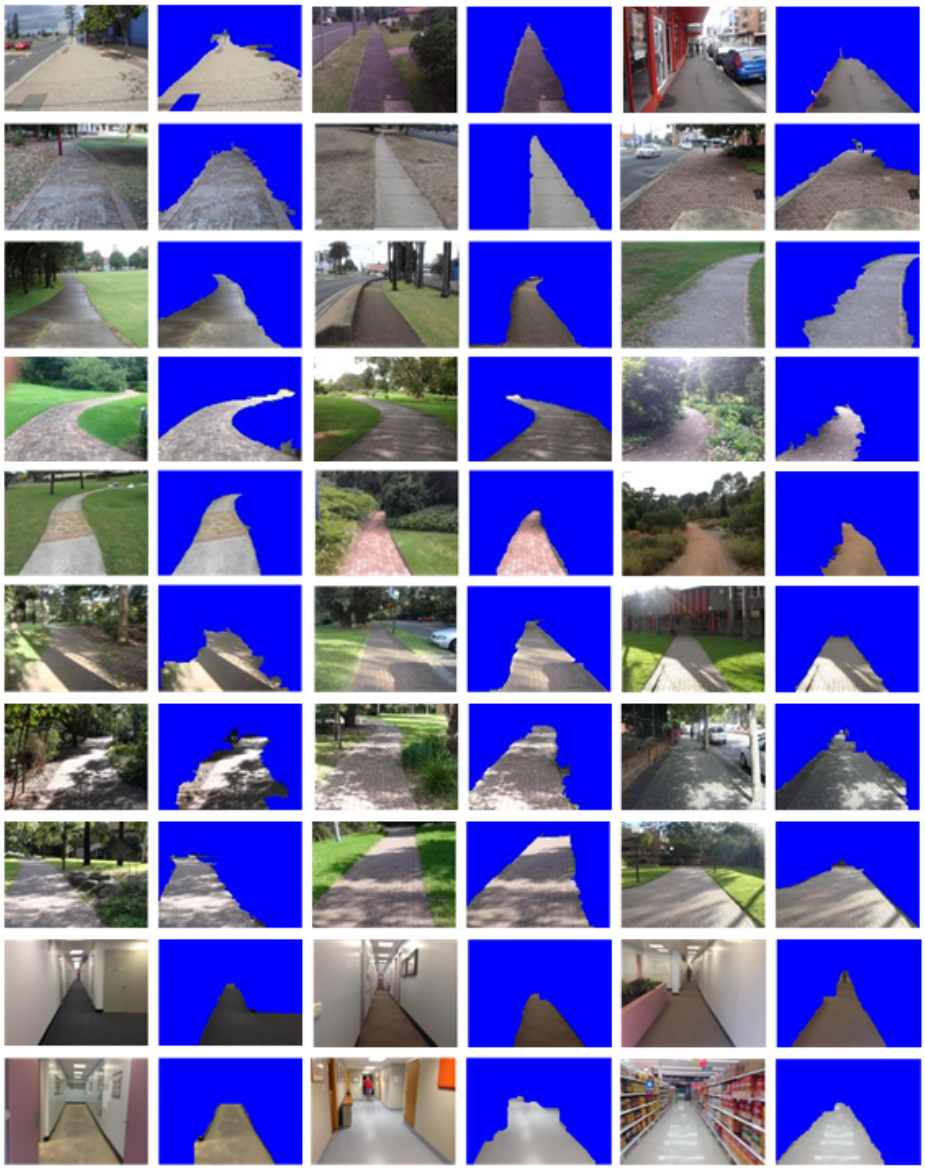


Fig. 9. Visual results of pedestrian lane detection. Column 1,3 and 5:input images. Column 2,4 and 6: detected lanes.

robust than the proposed method using the RGB color space and the lane border detection method in [2].

Table 2. Comparison of algorithms for pedestrian lane detection.

Methods	*Recall (%)*	*Precision (%)*	*Processing time (s)*
Lane border detection [2]	63.9	66.0	3.7
Proposed VPE + RGB	91.2	92.6	1.8
Proposed VPE + IIS	**94.8**	**95.1**	**1.8**

Our MATLAB implementation took 1.8 seconds, on the average, to process an image of size 100×140 pixels. We consider speeding-up the method as future work. Several detection results of the proposed method are demonstrated in Fig. 9. The experimental results show that the proposed method detects robustly pedestrian lanes with various lane surface structures and shapes and under different illumination and weather conditions.

5 Conclusion

This paper presents a method for pedestrian lane detection in unstructured environments. The proposed method makes use the vanishing point to automatically determine a sample region from which an appearance model of the lane surface is constructed. Shape context descriptor is used to describe the shape of the lane region. The walking lane is then detected using both the appearance and shape features in a probabilistic approach. The proposed method is evaluated and compared with existing methods on a large data set collected from various environments. The experimental results have demonstrated that the proposed method is able to detect various types of unstructured pedestrian lanes under challenging environmental conditions. It also compares favorably with the existing methods.

References

1. Ivanchenko, V., Coughlan, J., Huiying, S.: Detecting and locating crosswalks using a camera phone. In: IEEE Conference on Computer Vision and Pattern Recognition. (2008) 1–8
2. Hui, K., Audibert, J.Y., Ponce, J.: General road detection from a single image. IEEE Transactions on mage Processing **19** (2010) 2211–2220
3. Chang, C.K., Siagian, C., Itti, L.: Mobile robot monocular vision navigation based on road region and boundary estimation. In: IEEE/RSJ International Conference on on Intelligent Robots and Systems. (2012) 1043–1050
4. Wang, Y., Teoh, E.K., Shen, D.: Lane detection and tracking using B-snake. Image and Vision Computing **22** (2004) 269–280
5. Kim, Z.W.: Robust lane detection and tracking in challenging scenarios. IEEE Transactions on Intelligent Transportation Systems **9** (2008) 16–26
6. Crisman, J.D., Thorpe, C.E.: Scarf: A color vision system that tracks roads and intersections. IEEE Transactions on Robotics and Automation **9** (1993) 49–58

7. Alvarez, J.M., Lopez, A.M.: Road detection based on illuminant invariance. IEEE Transactions on Intelligent Transportation Systems **12** (2011) 184–193
8. Alvarez, J.M., Gevers, T., LeCun, Y., Lopez, A.M.: Road scene segmentation from a single image. In: European Conference on Computer Vision. (2012) 376–389
9. Uddin, M.S., Shioyama, T.: Bipolarity and projective invariant-based zebra-crossing detection for the visually impaired. In: IEEE Conference on Computer Vision and Pattern Recognition. (2005) 22–30
10. Le, M.C., Phung, S.L., Bouzerdoum, A.: Pedestrian lane detection for assistive navigation of blind people. In: International Conference on Pattern Recognition. (2012) 2594–2597
11. Le, M.C., Phung, S.L., Bouzerdoum, A.: Pedestrian lane detection for the visually impaired. In: International Conference on Digital Image Computing Techniques and Applications. (2012) 1–6
12. Se, S.: Zebra-crossing detection for the partially sighted. In: IEEE Conference on Computer Vision and Pattern Recognition. (2000) 211–217
13. Tan, C., Tsai, H., Chang, T., Shneier, M.: Color model-based real-time learning for road following. In: IEEE Conference on Intelligent Transportation Systems. (2006) 939–944
14. Sha, Y., Zhang, G.y., Yang, Y.: A road detection algorithm by boosting using feature combination. In: IEEE Intelligent Vehicles Symposium. (2007) 364–368
15. Alvarez, J.M., Gevers, T., Lopez, A.M.: Vision-based road detection using road models. In: IEEE International Conference on Image Processing. (2009) 2073–2076
16. Rasmussen, C.: Texture-based vanishing point voting for road shape estimation. In: British Machine Vision Conference. (2004) 470–477
17. Belongie, S., Malik, J., Puzicha, J.: Shape matching and object recognition using shape contexts. IEEE Transactions on Pattern Analysis and Machine Intelligence **24** (2002) 509–522
18. Sotelo, M., Rodriguez, F., Magdalena, L., Bergasa, L., Boquete, L.: A color vision-based lane tracking system for autonomous driving on unmarked roads. Autonomous Robots **16** (2004) 95–116
19. Ramstrom, O., Christensen, H.: A method for following unmarked roads. In: IEEE Intelligent Vehicles Symposium. (2005) 650–655
20. He, Y., Wang, H., Zhang, B.: Color-based road detection in urban traffic scenes. IEEE Transactions on Intelligent Transportation Systems **5** (2004) 309–318
21. Oh, C., Son, J., Sohn, K.: Illumination robust road detection using geometric information. In: International IEEE Conference on Intelligent Transportation Systems. (2012) 1566–1571
22. Miksik, O., Petyovsky, P., Zalud, L., Jura, P.: Robust detection of shady and highlighted roads for monocular camera based navigation of ugv. In: IEEE International Conference on Robotics and Automation. (2011) 64–71
23. Crisman, J., Thorpe, C.: Unscarf, a color vision system for the detection of unstructured roads. In: IEEE International Conference on Robotics and Automation. (1991) 2496 – 2501
24. Tardif, J.P.: Non-iterative approach for fast and accurate vanishing point detection. In: International Conference on Computer Vision. (2009) 1250–1257
25. Andal, F.A., Taubin, G., Goldenstein, S.: Vanishing point detection by segment clustering on the projective space. In: European Conference on Computer Vision. (2012) 324–337
26. Rasmussen, C.: Grouping dominant orientations for ill-structured road following. In: IEEE Conference on Computer Vision and Pattern Recognition. (2004) 470–477

27. Moghadam, P., Starzyk, J.A., Wijesoma, W.S.: Fast vanishing-point detection in unstructured environments. IEEE Transactions on Image Processing **21** (2012) 425–430
28. Weijer, J.v.d., Gevers, T., Smeulders, A.W.M.: Robust photometric invariant features from the color tensor. IEEE Transactions on Image Processing **15** (2006) 118–127
29. Gonzalez, R., Woods, R. In: Digital Image Processing using MATLAB. Prentice Hall (2004)
30. Felzenszwalb, P., Huttenlocher, D.: Efficient graph-based image segmentation. International Journal of Computer Vision **59** (2004) 167–181
31. Gevers, T., A.W.M., S., Stokman, H.: Photometric invariant region detection. In: British Machine Vision Conference. (1998) 659–669
32. Everingham, M., Gool, L., Williams, C.K., Winn, J., Zisserman, A.: The pascal visual object classes (VOC) challenge. International Journal of Computer Vision **88** (2010) 303–338