Computer Vision and Image Understanding 149 (2016) 186–196

Contents lists available at ScienceDirect

Computer Vision and Image Understanding

journal homepage: www.elsevier.com/locate/cviu

Pedestrian lane detection in unstructured scenes for assistive navigation

Son Lam Phung*, Manh Cuong Le, Abdesselam Bouzerdoum

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

ARTICLE INFO

Article history:
Received 15 April 2015
Accepted 23 January 2016

Keywords:
Pedestrian lane detection
Vanishing point estimation
Assistive and autonomous navigation
Benchmark dataset

ABSTRACT

Automatic detection of the pedestrian lane in a scene is an important task in assistive and autonomous navigation. This paper presents a vision-based algorithm for pedestrian lane detection in unstructured scenes, where lanes vary significantly in color, texture, and shape and are not indicated by any painted markers. In the proposed method, a lane appearance model is constructed adaptively from a sample image region, which is identified automatically from the image vanishing point. This paper also introduces a fast and robust vanishing point estimation method based on the color tensor and dominant orientations of color edge pixels. The proposed pedestrian lane detection method is evaluated on a new benchmark dataset that contains images from various indoor and outdoor scenes with different types of unmarked lanes. Experimental results are presented which demonstrate its efficiency and robustness in comparison with several existing methods.

1. Introduction

Locating the pedestrian lane in a given scene is a key component in many assistive and autonomous navigation systems. It enables a vision-impaired person to find the walkable path and maintain his or her balance while walking – a challenging task that at present is performed mostly using a white cane or a guided dog [1]. It also allows a smart wheelchair to traverse a pedestrian lane with little guidance from the disabled user [2]. Pedestrian lane detection is also useful for autonomous vehicles in sensing off-limit regions or pedestrians in a scene [3]. Furthermore, algorithms for finding the pedestrian lane can be extended to locate open roads for self-driven cars or robots. Pedestrian lane detection in fact complements other features, e.g. obstacle detection [4,5] and GPS-based guidance [6], of electronic navigation devices.

Despite its significance, there are only a few methods proposed for pedestrian lane detection, which are mostly concerned with pedestrian lanes having white markers [7–10]. To address this gap, this paper focuses on camera-based detection of pedestrian lanes in unstructured environments. In this paper, a pedestrian lane is assumed to exist in the scene. However, the scene is under varying lighting conditions and could be indoor or outdoor. Furthermore, the pedestrian lanes can have arbitrary surfaces with no painted markers.

Existing algorithms for unmarked lane detection (including pedestrian lanes) rely on color and texture of lane surfaces to differentiate the lane pixels from the background [11–14]. These algorithms require off-line training, and hence their detection accuracy decreases when the lane appearance differs from the training data. In practice, the lane appearance varies significantly due to different lane surfaces or illumination conditions. Other existing algorithms locate the lane boundaries among the edges that point to the vanishing point of the image [15,16]. However, algorithms based on finding the lane borders are sensitive to background clutter. In this paper, we propose a new method to detect unmarked pedestrian lanes using both color, edge, and shape features. In contrast to the existing methods, the proposed approach constructs a lane model from the input image, and is therefore more adaptive to different illumination conditions and lane surfaces. The main contributions of the paper can be briefly described as follows:

- Firstly, we propose an improved vanishing point estimation method using local orientations of color edge pixels. Estimating the vanishing point using edge pixels is more efficient than using all pixels as done in the existing methods [15,16]. In addition, to estimate local orientations and edge pixels more robustly, we apply the color tensor on multiple color channels, instead of relying on only the intensity channel.
- Secondly, we present a method to define automatically a sample region, from which a lane appearance model is constructed adaptively for each input image. This sample region is determined using the vanishing point and the geometric and color

* Corresponding author. Tel. +61 242213407.
E-mail addresses: phung@uow.edu.au (S.L. Phung), clm635@uowmail.edu.au (M.C. Le), bouzer@uow.edu.au (A. Bouzerdoum).

http://dx.doi.org/10.1016/j.cviu.2016.01.011

features of lane borders and surfaces. The lane model is therefore adaptive to various types of lane surfaces. To make the lane model more robust to the lighting conditions, this paper employs the illumination-invariant color space (IIS). In addition, we propose a novel lane score that combines color, edge, and shape features for detecting unmarked pedestrian lanes.

- Lastly, we create a new dataset with manually annotated ground-truth for objective evaluation of algorithms for vanishing-point estimation and pedestrian-lane detection. Although several datasets exist for road/lane detection for vehicles, our dataset, to the best of our knowledge, is the first for pedestrian lanes. This dataset, collected from realistic indoor and outdoor scenes, with various shapes, textures, and surface colors, is expected to facilitate research in vanishing-point estimation and pedestrian-lane detection. It is available at www.uow.edu.au/~phung/plvp_dataset.html.

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

2. Related work

Current vision-based approaches for detecting pedestrian lanes in unstructured scenes can be divided into two categories: (i) lane segmentation, and (ii) lane-border detection. In the *lane segmentation* approach, off-line color models are used to differentiate the lane pixels from the background [11,12,17,18]. Different color spaces and classifiers have been used. For example, Crisman and Thorpe use Gaussian models in the red–green–blue (RGB) color space to represent the on-road and off-road classes [11]. Also using the RGB space, Tan et al. capture the variability of the road surface with multiple color histograms, and the background with a single color histogram [12]. Instead of using the RGB space, Ramstrom and Christensen employ UV, normalized red and green, and luminance components and construct Gaussian mixture models for the road-surface and background classes [18]. Sotelo et al. employ the hue-saturation-intensity (HSI) color space [17]. In their method, achromatic pixels (i.e. with extreme intensities or low saturations) are classified using intensity only, whereas other pixels are classified by thresholding their chromatic distance to the training colors. Because the color models are trained off-line, these methods do not cope well with the appearance variations in lane surfaces.

To address this problem, several methods model the lane pixels directly from sample regions in the input image [19–22]. These methods determine the sample lane regions in different ways. For example, Alvarez et al. select small random areas at the bottom and in the middle of the input image [22,23]. Miksik et al. initialize the sample lane region as a trapezoid at the bottom and center of the image, and then refine the sample region using the vanishing point [21]. He et al. form a sample lane region from the candidate lane boundaries, which are detected using the vanishing point and an assumption about the lane width [19]. The performance of these methods depends on the quality of the sample regions, which in turn relies on prior knowledge about the walking lane.

In the *lane-border detection* approach, the lane boundaries are determined using the vanishing point [15,16] or templates of the lane boundaries [24]. In [15], the lane borders are detected among the edges pointing to the vanishing point. The optimal left and right edges are judged using an objective function that measures the color and texture differences between lane and non-lane regions. This method is effective only when the lane region is homogeneous and differs significantly from non-lane regions in terms of color and texture. Kong et al. also find the lane borders from the edges pointing to the vanishing point, except that their method ranks edges using texture orientation and color features [16]. Because this method relies only on edges for lane-border detection, it is sensitive to background edges. In another method, the lane boundaries are found from the edges of homogeneous color regions by matching with lane templates [24]. Recently, Chang et al. propose combining lane-border detection and road segmentation for detecting lanes [25]. Similarly to [16], 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 of the input image. In [26], the two left and right borders of the lane are found among the rays that point to the vanishing point; this approach is suitable mainly for pedestrian lanes with straight-line borders. This paper extends this approach to detect pedestrian lanes with curved borders and varied surfaces.

3. Proposed pedestrian lane detection method

In this section, we present the new method for detecting unstructured pedestrian lanes, which comprises three main stages: (i) vanishing point estimation, (ii) sample region selection, and (iii) lane segmentation.

3.1. Vanishing point estimation

The vanishing point in an image is often located using either line segments [27–29] or local orientations [15,16,30]. For unstructured scenes with non-straight edges, using local orientations is more suitable than using line segments for vanishing point estimation. However, most existing methods based on local orientations have high computational complexity and are sensitive to background clutter. Furthermore, they rely on the intensity channel only, even though color channels provide photometric information that can lead to more robust detection of edges and local orientations. In this paper, we propose to improve the accuracy and efficiency of vanishing point detection, by employing color tensor to capture image structure and focusing on edge pixels only.

The color tensor is a tool for analyzing the local differential structure of a color image [31]. Consider an image with three color channels: $F = \{F_k;\, k = 1, 2, 3\}$. Let $D_{k,\,x}$ and $D_{k,\,y}$ denote the derivatives of F_k along the horizontal and vertical direction, respectively. Let $\mathbf{w}$ be the convolution kernel of a smoothing filter. The color tensor of the image is represented as

$$\begin{pmatrix} G_{xx} & G_{xy} \\ G_{yx} & G_{yy} \end{pmatrix} \quad \text{where} \quad \begin{cases} G_{xx} & = \mathbf{w} * \left[\sum_{k=1}^{3} D_{k,x} \circ D_{k,x} \right] \\ G_{yy} & = \mathbf{w} * \left[\sum_{k=1}^{3} D_{k,y} \circ D_{k,y} \right] \\ G_{xy} & = \mathbf{w} * \left[\sum_{k=1}^{3} D_{k,x} \circ D_{k,y} \right] \end{cases}. \tag{1}$$

Here, $*$ denotes the 2-D convolution, and $\circ$ denotes the element-wise multiplication (Hadamard product). Based on eigenvalue analysis of the color tensor [31], we estimate the dominant local orientation $\boldsymbol{\theta}$ and the edge strength $\boldsymbol{\lambda}$ for all image pixels as

$$\boldsymbol{\theta} = \frac{1}{2} \arctan\left(\frac{2G_{xy}}{G_{xx} - G_{yy}} \right) + \frac{\pi}{2}, \tag{2}$$

$$\boldsymbol{\lambda} = \frac{1}{2}\left(G_{xx} + G_{yy} + \sqrt{(G_{xx} - G_{yy})^2 + 4G_{xy}^2} \right), \tag{3}$$

where the arithmetic operations are performed element-wise. Next, the edge pixels in the image are identified via non-maximum suppression and hysteresis thresholding, as done in the intensity-based Canny edge detector. The main difference in this paper is