ARTICLE IN PRESS

IMAVIS-03632; No of Pages 12

Image and Vision Computing xxx (2017) xxx–xxx

Contents lists available at ScienceDirect

Image and Vision Computing

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

Ego-Lane Analysis System (ELAS): Dataset and algorithms☆

Rodrigo F. Berriel*, Edilson de Aguiar, Alberto F. de Souza, Thiago Oliveira-Santos

Universidade Federal do Espírito Santo, Brazil

ARTICLE INFO

Article history:
Received 5 April 2016
Received in revised form 18 April 2017
Accepted 21 July 2017
Available online xxxx

Keywords:
Ego-lane analysis
Lane estimation
Kalman filter
Particle filter
Dataset
Image processing

ABSTRACT

Decreasing costs of vision sensors and advances in embedded hardware boosted lane related research – detection, estimation, tracking, etc. – in the past two decades. The interest in this topic has increased even more with the demand for advanced driver assistance systems (ADAS) and self-driving cars. Although extensively studied independently, there is still need for studies that propose a combined solution for the multiple problems related to the ego-lane, such as lane departure warning (LDW), lane change detection, lane marking type (LMT) classification, road markings detection and classification, and detection of adjacent lanes (i.e., immediate left and right lanes) presence. In this paper, we propose a real-time Ego-Lane Analysis System (ELAS) capable of estimating ego-lane position, classifying LMTs and road markings, performing LDW and detecting lane change events. The proposed vision-based system works on a temporal sequence of images. Lane marking features are extracted in perspective and Inverse Perspective Mapping (IPM) images that are combined to increase robustness. The final estimated lane is modeled as a spline using a combination of methods (Hough lines with Kalman filter and spline with particle filter). Based on the estimated lane, all other events are detected. To validate ELAS and cover the lack of lane datasets in the literature, a new dataset with more than 20 different scenes (in more than 15,000 frames) and considering a variety of scenarios (urban road, highways, traffic, shadows, etc.) was created. The dataset was manually annotated and made publicly available to enable evaluation of several events that are of interest for the research community (i.e., lane estimation, change, and centering; road markings; intersections; LMTs; crosswalks and adjacent lanes). Moreover, the system was also validated quantitatively and qualitatively on other public datasets. ELAS achieved high detection rates in all real-world events and proved to be ready for real-time applications.

1. Introduction

Decreasing costs of vision sensors and advances in embedded hardware boosted traffic lane related research (detection, estimation, tracking, etc.) in the past two decades. The interest increased even more with the demand for advanced driver assistance systems (ADAS) and self-driving cars. The need for these solutions is also supported by the fact that humans are the main cause of car accidents [1]. Lane detection is an essential task in this context and has been extensively studied [2–4]. Nevertheless, drivers rely not only on the position of the lanes for driving safely, but also use visual cues (e.g., pavement markings) to understand what is and what is not allowed (direction, lane change, etc.) in a given lane.

Ego-lane and host lane are names given to the lane where the vehicle is positioned. Ego-lane analysis comprises multiple tasks related to the host lane, such as lane estimation (LE) [5], lane departure warning (LDW) [6], lane change detection [7], lane marking type (LMT) classification [8], road markings detection and classification [9], and detection of adjacent lanes, also known as multiple lanes detection [10,11]. In general, different sensors have been used to address these issues: monocular [12] and stereo [13] cameras, LiDAR [14], and Sensor-fusion [15]. Each sensor has its own drawbacks (e.g., sudden illumination changes for cameras; sparsity and range limit for laser).

Many camera-based solutions [2,10,16] focus only on lane detection and estimation. Given the associated cost in generating datasets for these tasks, the ones used in many of the solutions in the literature are kept private. Alongside the lack of public datasets, there is a lack of public implementations which hinders fair comparisons. Although extensively studied independently, there is still need for studies that propose a combined real-time solution for the multiple problems related to the ego-lane including validation, public

☆ This paper has been recommended for acceptance by Branislav Kisacanin.
* Corresponding author.
E-mail address: rfberriel@inf.ufes.br (R.F. Berriel).

http://dx.doi.org/10.1016/j.imavis.2017.07.005
0262-8856/© 2017 Elsevier B.V. All rights reserved.

or open source implementation and public datasets, enabling fair comparisons.

In this context, this paper presents a real-time vision-based Ego-Lane Analysis System (ELAS) capable of estimating ego-lane position, classifying LMTs and road markings, performing LDW, detecting adjacent lanes (i.e., immediate left and right lanes), and detecting lane change events. The proposed monocular vision-based system works on a temporal sequence of images. The final estimated lane is modeled as a spline using a combination of methods (Hough lines with Kalman filter and spline with particle filter). Based on the estimated lane, all other events are detected using image processing techniques. A novel dataset was created and annotated. Alongside this dataset, the implementation of ELAS is publicly available for future fair comparisons. The dataset contains more than 20 different scenes (over 15,000 frames) and a variety of scenarios (urban road, highways, traffic, shadows, rain, etc.). Moreover, ELAS was evaluated on this dataset, achieving high detection rates in all real scenarios, proving to be ready for real-time real-world applications.

2. Ego-Lane Analysis System (ELAS)

ELAS processes a temporal sequence of images, analyzing each of them individually. These images are expected to come from a monocular forward-looking camera mounted on a vehicle. The general work-flow is described in Fig. 1. Firstly, general road markings features are extracted and stored in feature maps. Based on these features, pavement markings (i.e., crosswalks, arrow markings) are detected and removed from the maps. Subsequently, lanes are estimated by using a combination of Hough lines with Kalman filter and spline with particle filter. These two processes are performed separately: the Hough lines with Kalman filter are used to estimate the base of the lane (i.e., base point, lane width, and lane direction); and, the spline-based particle filter is used to estimate the curvature of the lane. Essentially, the lane base estimated by the Hough lines and Kalman filter is used as a starting point to the particle filter, reducing the freedom of the spline near the car and guiding the spline curvature direction in the far depth of view. This combination was used to take advantage of the lane stability and linearity near the car (Hough and Kalman), while reducing the number of parameters of the particle filter responsible for the curvature. Finally, based on the estimated lane position, the remaining tasks are performed: LMT classification, adjacent lanes detection, and deviation from the lane center. All these tasks are performed in real-time (i.e., more than 30 frames per second) and are explained in detail in the following subsections. As a result, for each image the system outputs information describing the lane position, lane markings type, crosswalks, road signs, presence of adjacent lanes, and deviation of the car related to the lane center. It is worth mentioning that some of the algorithms hereby presented were designed considering the driving rules from Brazil (e.g., lane marking type, road signs). Many of these rules extend to other countries in the world, and therefore the assumptions and limitations would still be valid. In case of different rules, some adaptations of the presented method would be required.

2.1. General feature extraction

A general set of feature maps is extracted from each frame and the maps are used by each module according to its task. An overview of this process is shown in Fig. 2.

2.1.1. Preprocessing

Before generating the feature maps, the input image is converted to grayscale, since only pixel intensities are analyzed for generating the feature maps. Secondly, a region of interest (RoI) is set in order to remove irrelevant parts of the image. Finally, an Inverse Perspective Mapping (IPM) [17] is applied in order to reduce perspective distortion, resulting in a top-view image (also called bird's-eye view). To apply the IPM, ELAS assumes a constant ground plane along the input frames, using a static homography matrix. In this bird's-eye view image, lane markings tend to have constant width. However, due to road inclination and casual bumps, lane markings width constancy can be temporarily lost because fixed ground plane assumption may fail. As explained later, this issue is addressed by the lane model used.

2.1.2. Feature maps generation

Four feature maps are generated using threshold-based techniques. Each feature map is a binary image. White pixels in these maps are called evidences. An overview of this process is described in Fig. 2.

2.1.2.1. Step Row Filter Map (I^{SRF}). Lane markings are usually areas brighter than its surroundings (asphalt). In this feature map, lane marking evidences are detected on the original grayscaled image using the step row filter defined by Eq. (1), as presented in [18].

$$y_i = 2x_i - (x_{i-\tau} + x_{i+\tau}) - |x_{i-\tau} - x_{i+\tau}| \quad (1)$$

Here, τ is assumed large enough to surpass the size of double lane markings. Additionally, this filter was applied in the original image, linearly adjusting τ according to the vertical axis.

2.1.2.2. Horizontal Difference of Gaussians Map (I^{DOG}). Based on the same assumption (lane markings are brighter than asphalt), in this map, evidences are found using a horizontal Difference of Gaussians [19] (DoG), equivalent to a horizontal Mexican hat with central opening with the average size of a lane width. DoG is applied in the IPM grayscale image in order to take advantage of lane width invariability, and is followed by a threshold.

2.1.2.3. Vertical Absolute Derivate Map (I^{VAD}). Some objects of interest (e.g., vehicles, stop lines) contain horizontal edges in the bird's-eye view image. To extract these features vertical changes are calculated using the absolute value of the y-image derivative followed by a threshold.

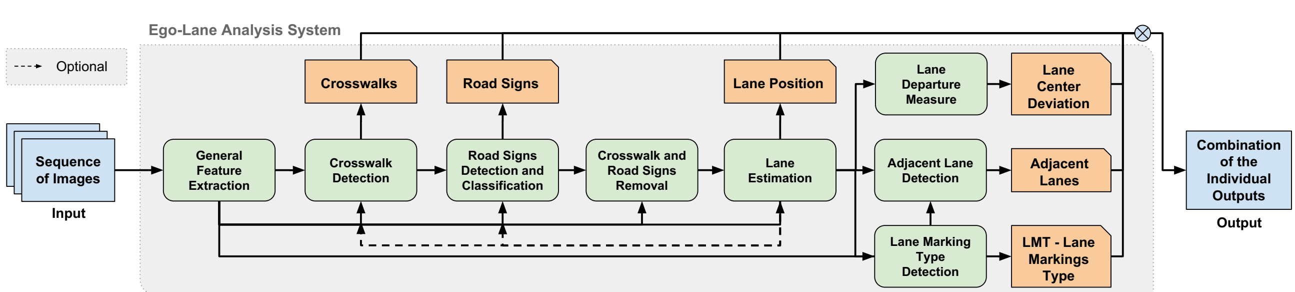

Fig. 1. Overview of the Ego-Lane Analysis System (ELAS).