ROBUST MULTIPLE LANE ROAD MODELING BASED ON PERSPECTIVE ANALYSIS

Marcos Nieto, Luis Salgado, Fernando Jaureguizar, and Jon Arróspide

Grupo de Tratamiento de Imágenes - E. T. S. Ing. Telecomunicación
Universidad Politécnica de Madrid - Madrid (Spain)
{mnd, lsa, fjn, jal}@gti.ssr.upm.es

ABSTRACT

Road modeling is the first step towards environment perception within driver assistance video-based systems. Typically, lane modeling allows applications such as lane departure warning or lane invasion by other vehicles. In this paper, a new monocular image processing strategy that achieves a robust multiple lane model is proposed. The identification of multiple lanes is done by firstly detecting the own lane and estimating its geometry under perspective distortion. The perspective analysis and curve fitting allows to hypothesize adjacent lanes assuming some a priori knowledge about the road. The verification of these hypotheses is carried out by a confidence level analysis. Several types of sequences have been tested, with different illumination conditions, presence of shadows and significant curvature, all performing in real-time. Results show the robustness of the system, delivering accurate multiple lane road models in most situations.

Index Terms— Multiple lane detection, vanishing point, homography, curve fitting, confidence measures.

1. INTRODUCTION

Focusing on the field of driver assistance systems (DAS), two major objectives are road modeling and vehicle detection within in-vehicle vision systems. Usually, the road model is firstly computed to obtain a reliable environment description which afterwards allows to accurately detect vehicles. For this purpose, there are typically two main processing stages, *feature extraction*, which extracts features from images, and *model fitting*, that uses those features to obtain the number of lanes, their width, curvature, etc. Most works found in literature detect, as features, the lane markings within the road [1][2], although other features, like the color of the road, are also used [3].

Model fitting is the processing stage that introduces *a priori* knowledge to constrain the problem and achieve adequate

results. For instance, [4] assumes that the road is flat, without slope changes and with parallel lane markings, while other works use much more accurate models such as B-snakes based models [2].

Nonetheless, most works identify only one lane due to the difficulty of finding reliable features for adjacent lanes in the image [3][5][6]. Although there are authors that detect multiple lanes [1][4], no validation process is done which ensures that detected lanes are correct with some degree of confidence, and therefore there is a lack of robustness, particularly in presence of non-homogenous road color, shadows, vehicles, etc.

In this work a novel strategy that overcomes these limitations is presented. It is based on the estimation of the vanishing point, which allows to reconstruct the geometry of the road and perform hypotheses on the number of lanes, followed by an hypotheses verification algorithm which assigns confidence levels to hypothesized lanes.

2. STRATEGY OVERVIEW

The aim of this work is to show that it is possible to model the road in an innovative and robust way by detecting not only the lane where the car is being driven (the *own lane*) but the *adjacent lanes*, that are of great interest for DAS.

Four separate modules compose the system: (i) feature extraction; (ii) perspective estimation; (iii) own lane estimation; and (iv) adjacent lanes estimation.

Each image of the sequence, I_k, is processed by the feature extraction module, which extracts a set of feature points $\{z_k^i\}$ likely belonging to lane markings in the image. These points are used to estimate the perspective of the image, i.e. obtaining the homography, H_k that allows to transform the image into a non-perspective *bird-view* of the road (image plane to road plane homography). Feature points are transformed into $\{z'_k^i\}$ which are used to estimate the own lane in this non-perspective domain by robustly fitting circumference curves. The final step is to form hypotheses of adjacent lanes assuming the same width and curvature for all lanes. A confidence criterium is applied to verify the hypothesized lanes.

3. FEATURE EXTRACTION

The best way to describe a road is to identify the lane markings that define lanes in almost any well-painted road. The first step is to generate an image where discontinuous lane markings become continuous, and thus easier to detect by further analysis. Therefore, a temporal filtered image, I'_k, in discrete time k, is generated as follows:

$$I'_k(x,y) = \max\{I_k(x,y), \ldots, I_{k-K}(x,y)\} \quad (1)$$

where (x,y) are pixel coordinates inside the image and K the number of considered previous images. An example of this transformation is shown in Fig. 1; where (a) shows the lower part of the original image, and (b) the resulting filtered image.

Low gradient pixels at the lower part of the image likely belong to the road asphalt. Hence the gradient image is derived and the modified histogram of the lower part of I'_k is computed, taking into account only low gradient pixels.

This histogram is typically unimodal, with mean and standard deviation values (μ, σ), which accurately parameterize the gray level of the road in the current image. The contrast between lane markings and asphalt pixels is enhanced if all pixels with intensity values lower than $\{\mu + 2.5\sigma\}$ are set to zero. The result of this extraction is image I''_k shown in Fig. 1 (c) where all its pixels are set to zero but those likely belonging to lane markings.

For each row and each column of I''_k, the intensity level evolution is analyzed in order to locate lane markings based on their properties: (i) they appear as bursts of high intensity values, and (ii) are mostly surrounded by zeros. Therefore they are detected applying derivative filters, which enhance steep intensity changes, followed by an adaptive thresholding technique that classifies as belonging to lane markings only those pixels with high intensity values and response to the filter.

For each row and column, the filter $D = \frac{1}{2}[-1 \; \underline{0} \; 1]$ is applied, obtaining signed values enhancing steep changes. For each pixel, the adaptive threshold is computed as:

$$t(x,y) = s \cdot \exp\left(-\frac{\ln(s)}{s}I''_k(x,y)\right) \quad (2)$$

where s is the scale factor (255 for 8-bit gray images). The higher the pixel value is, the lower the threshold, so that pixels with high values (likely belonging to lane markings) easily exceed the threshold, whereas other pixels, with lower values, will obtain higher thresholds.

As a result, each pixel obtains one of these three values: $\{-1, 0, 1\}$. The pattern that describes a lane marking within a column or a row is a 1 followed by an arbitrary number of zeros and ended by a -1. When found, the central pixel of this pattern defines a lane marking point, $\mathbf{z}_k = (x,y)^\top$. Fig. 1 (d) shows an example of the set of points $\{\mathbf{z}_k^i\}_{i=1}^N$, where N is the total number of points in the image.
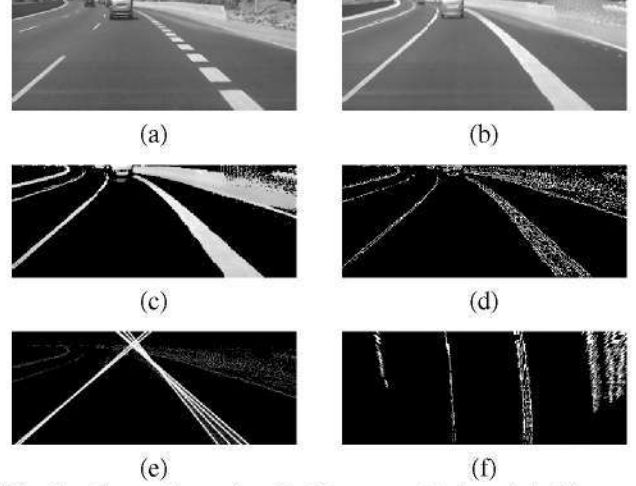

Fig. 1. Feature Extraction Module steps: (a) I_k, original image; (b) I'_k, maximum image for $K = 10$; (c) I''_k, asphalt pixels are set to zero; (d) resulting set of points $\{\mathbf{z}_k^i\}$ in white; (e) vanishing point estimation through Hough transform; and (f) resulting non-perspective bird-view $\{\mathbf{z}'^i_k\}$.

4. PERSPECTIVE ESTIMATION

The vanishing point, which is, in a perspective image, the point to which parallel lines seem to converge, is used to compute the homography between the image plane and the road plane, assumed to be flat [4]. Lanes estimation is simplified in this non perspective domain as lane markings can be approximated as parallel straight lines in the lower part of the image, even when there is a significant curvature ahead. Outlier points not belonging to the road, such as trees, the horizon, etc., are also removed.

4.1. Vanishing point estimation

The vanishing point is computed as the intersection point of the straight lines prolonged from the lower part of the image. The well known Hough transform [6] is used for that purpose over the set of points $\{\mathbf{z}_k^i\}$ as it is robust against outliers offering multiple line fitting. Each line is parameterized with an angle θ and a distance ρ as follows:

$$x \cdot \sin\theta + y \cdot \cos\theta = \rho \quad (3)$$

The Hough transform may lead to more than two lines for these lane markings, as shown in Fig. 1 (e), which result in multiple intersection points. Hence, the vanishing point $\mathbf{v}_k = (v_x, v_y)^\top$ is selected as the solution of the system of equations built with the equations of each detected line:

$$\begin{bmatrix} \mathbf{s} & | & \mathbf{c} \end{bmatrix} \cdot \mathbf{v}_k = \mathbf{p} \quad (4)$$

where $\mathbf{s} = (\sin\theta_0, \ldots, \sin\theta_{l-1})^\top$, $\mathbf{c} = (\cos\theta_0, \ldots, \cos\theta_{l-1})^\top$, $\mathbf{p} = (\rho_0, \ldots, \rho_{l-1})^\top$, and l is the number of lines. This equation is solved with singular value decomposition (SVD), giving the least squares error solution $\mathbf{v}_k$ to the system.

The computed vanishing point for the k-th image is stabilized with a Kalman filter that estimates the state-vector $s_k = (v_x, v_y, \dot{v}_x, \dot{v}_y)^\mathsf{T}$ taking as measure the abovementioned least squares error vanishing point.

4.2. Homography estimation

The perspective distortion of the image is governed by an homography, H_k, between the image plane and the road plane, understood as a *bird-view* model of the road, as shown in Fig. 1 (f). The estimated vanishing point is used to define a set of correspondences between points of these planes so that the homography is computed through the Direct Linear Transformation (DLT) [7].

If it is assumed that the yaw angle of the camera with respect to the road plane is close to zero, the correspondences $\{x_i \leftrightarrow x'_i\}$ (to simplify notations k index is removed in this section) are set as shown in Fig. 2 (a) and (b): in (a) the original domain is depicted, with the two main straight lines modeling the own lane. The selected points are obtained as the intersection of these lines with arbitrary horizontal lines in the lower part of the image. Within the transformed domain, shown in (b), these points form a rectangle in the lower part of the image, taking into account the assumption of Section 4.1, even when there might exist curvature ahead.

Once given the set of correspondences, the homography is computed with the direct linear transformation (DLT) algorithm. This transformation is given by the cross product $x'_i \times Hx_i = 0$. If the j-th row of the matrix H is denoted by $h^{j\mathsf{T}}$, then we have

$$Hx_i = \begin{pmatrix} h^{1\mathsf{T}}x_i \\ h^{2\mathsf{T}}x_i \\ h^{3\mathsf{T}}x_i \end{pmatrix} \qquad (5)$$

considering $x'_i = (x'_i, y'_i, z'_i)^\mathsf{T}$, the cross product leads to [7]:

$$\begin{bmatrix} 0^\mathsf{T} & -z'_i x_i{}^\mathsf{T} & y'_i x_i{}^\mathsf{T} \\ z'_i x_i{}^\mathsf{T} & 0^\mathsf{T} & -x'_i x_i{}^\mathsf{T} \end{bmatrix} \begin{pmatrix} h^1 \\ h^2 \\ h^3 \end{pmatrix} = 0 \qquad (6)$$

which has the form $A_i h = 0$, where A_i is a 2×9 matrix, and h is a 9-vector made up of the entries of the matrix H. By assembling the A_i matrices into a single $2n \times 9$ matrix A, where n is the number of correspondences, we obtain the equation $Ah = 0$ which is solved by obtaining the SVD of A.

5. OWN LANE ESTIMATION

The own lane is typically the best seen in the images, and so are its lane markings. That way, first, the own lane is estimated and then it is used to "extrapolate" adjacent lanes by assuming that lanes in the same road have equal width and curvature.

Once the homography that removes perspective from the original images is known, feature points $\{z^i_k\}$ are transformed
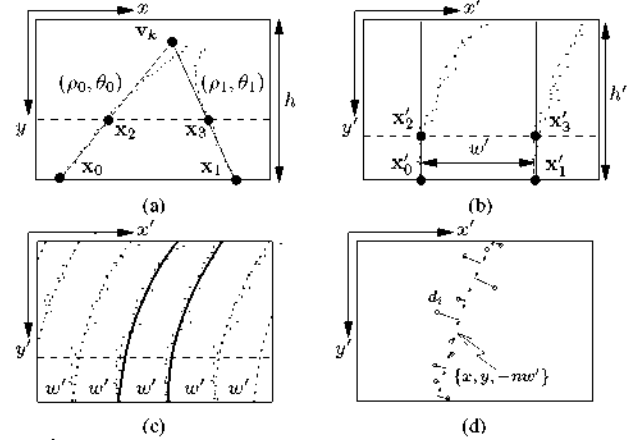

Fig. 2. Perspective analysis: (a) image plane points; (b) road plane points. Black dots represent the set of feature points; (c) fitted curves for own lane in solid lines, hypothesized curves in dotted lines; and (d) verification through distance measurements.

into $\{z'^i_k\}$ in the road plane, with $z'^i_k = H_k z^i_k$. Parabolic curves are typically used to fit to these points [1], offering enough accuracy for both transformed and image domain. However, for the transformed domain, generic circumference arc models show better performance for most situations.

First, the set of points is clustered into two sub-sets, according to the lane marking that is nearest to them. Then, each sub-set is fitted into an independent circumference. The parametric expression of this curve is given by:

$$x^2 + y^2 + 2Ax + 2By + C = 0 \qquad (7)$$

where the center is the point $(x, y) = (-A, -B)$, and the radius is $r = \sqrt{A^2 + B^2 - C}$.

If each transformed point is denoted as $z'^i_k = (x'_i, y'_i)^\mathsf{T}$, the curve is estimated as the least squares solution to the following system of equations:

$$\begin{bmatrix} 2x'_1 & 2y'_1 & 1 \\ \cdots & \cdots & \cdots \\ 2x'_N & 2y'_N & 1 \end{bmatrix} \begin{bmatrix} A \\ B \\ C \end{bmatrix} = \begin{bmatrix} -(x'^2_1 + y'^2_1) \\ \cdots \\ -(x'^2_N + y'^2_N) \end{bmatrix} \qquad (8)$$

where N is the number of points. The accuracy of the model is significantly improved by applying RANSAC [8] which identifies and removes outliers.

6. ADJACENT LANES ESTIMATION

Let $\{x, y, r\}$ be the parameters that define one of the estimated curves of the own lane. Then the parameters of adjacent curves are extrapolated as $\{x, y, r \pm nw'\}$, where w' is the width of the own lane and n indexes that the extrapolated curve is the n-th curve at the left $(-n)$ or right $(+n)$ of the own lane. A set of hypothesized lanes is shown in Fig. 2 (c).

For each lane marking, the verification is carried out by checking the distances, d_i, between the hypothesized points

Fig. 3. Examples of multiple lane road model.

(those points belonging to the hypothesized curves) and the transformed feature points $\{\mathbf{z'}_k^i\}$ as shown in Fig. 2 (d). The average distance of all points belonging to each curve is computed, $\bar{d}_i$, and used to obtain a confidence level as in (9), which is weighted by a confidence factor which grows with the number of points analyzed, N_j.

$$\Psi_j = (1 - \exp(-N_j)) \cdot \exp\left(\left(-\tau \bar{d}_i\right)\right) \qquad (9)$$

where τ is a decreasing factor of the exponential.

As a consequence, hypothesized lane markings at left and right of the own lane obtain confidence levels according to: (i) how well they fit with actual features obtained in the features extraction phase, $\{\mathbf{z'}_k^i\}$, and (ii) how many feature points are near them. Those lane markings having a confidence value Ψ_j higher than a certain threshold are considered as detected.

Finally, each point of the detected curves is retro-projected into the image domain by using $\mathbf{H}^{-1}$, delivering the actual estimation of the lanes in the image plane.

7. RESULTS

Several video sequences have been captured with an 8-bit gray-level camera with resolution 360×288 pixels installed inside a car, while driving on motorways in Madrid (Spain). The analyzed video sequences include varying illumination conditions, multiple and single lane roads, well and bad asphalted and painted roads, curvature, occluding vehicles, etc.

The proposed algorithm shows excellent results due to its robustness, not only in the estimation of the vanishing point, but also in the feature extraction processing module. Fig. 3 shows multiple lane detection in different example images. Among them, there are examples where curvature is accurately detected for all lanes (upper-left). In others, the presence of shadows and non-homogeneous road color does not disturb the road model, nor does the presence of vehicles partially occluding lane markings. In all these cases the strategy is robust and offers good results performing in real-time, derived from the application of robust techniques such as RANSAC, or Kalman filtering. Fig. 4 depicts this robustness by showing the confidence level along time for the two adjacent lane markings of example sequence shown in Fig. 1. These confidence levels are maintained through time, showing that the hypothesized lane markings do actually fit well
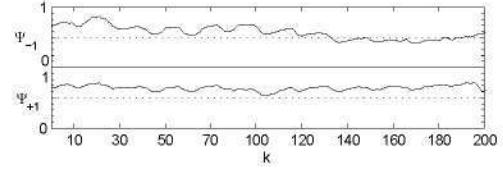
with the features corresponding to these adjacent lanes. From $k = 120$, Ψ_{-1} is slightly lower than Ψ_{+1}: the reason is the movement of the vehicle within the own lane, making more visible the right lane marking.

8. CONCLUSIONS

In this paper we have described a novel strategy for multiple lane road modeling with a single camera based on efficient vanishing point estimations which allow to reconstruct the geometry of the road. This reconstruction is undertaken to generate a transformed domain which corresponds to a *bird-view* of the road ahead. Robust curve fitting is applied to model the own lane and to hypothesize adjacent lanes, which are evaluated with a confidence function that determines how well features, corresponding to these lanes, fit to the estimated geometry of the own lane. Results show very accurate multiple lane road models for different illumination conditions, presence of vehicles, curvature, and non-homogeneous road color.

9. REFERENCES

[1] J. C. McCall and M. M. Trivedi, "Video-based lane estimation and tracking for driver assistance: Survey, system, and evaluation," *IEEE Transactions on Intelligent Transportation Systems*, vol. 7, No. 1, pp. 20–37, March 2006.

[2] Y. Wang, E. K. Teoh, and D. Shen, "Lane detection and tracking using b-snakes," *Image and Vision Computing*, vol. 22, pp. 269–289, 2004.

[3] J. P. González and U. Ozguner, "Lane detection using histogram-based segmentation and decision tree," in *Proc. Int. Trans. Systems*. IEEE, 2000, pp. 346–351.

[4] M. Bertozzi and A. Broggi, "Gold: A parallel real-time stereo vision system for generic obstacle and lane detection," *IEEE Trans. on Image Processing*, vol. 7, No. 1, pp. 62–81, 1998.

[5] S. Beucher and X. Yu, "Road recognition in complex traffic situations," in *7th Symp. on Trans. Systems*, Tianjin, China, August 24-26, 1994, pp. 423–428.

[6] K. Macek, B. Williams, S. Kolski, and R. Siegwart, "A lane detection vision module for driver assistance," in *Proc. Conf. on Mech. and Robotics*. IEEE/APS, Aachen, Germany, 2004.

[7] R. Hartley and A. Zisserman, *Multiple View Geometry in computer vision*, Cambridge University Press, 2001.

[8] Peter Meer, Charles V. Stewart, and David E. Tyler, "Robust computer vision: An interdisciplinary challenge," *Computer Vision and Image Understanding*, vol. 78, no. 1, pp. 1–7, 2000.