Robust lane lines detection and quantitative assessment

Antonio López, Joan Serrat, Cristina Cañero, Felipe Lumbreras

Computer Vision Center & Computer Science Dept.
Edifici O, Universitat Autònoma de Barcelona,
08193 Cerdanyola, Spain

Abstract. Detection of lane markings based on a camera sensor can be a low cost solution to lane departure and curve over speed warning. A number of methods and implementations have been reported in the literature. However, reliable detection is still an issue due to cast shadows, wearied and occluded markings, variable ambient lighting conditions etc. We focus on increasing the reliability of detection in two ways. Firstly, we employ a different image feature other than the commonly used edges: ridges, which we claim is better suited to this problem. Secondly, we have adapted RANSAC, a generic robust estimation method, to fit a parametric model of a pair or lane lines to the image features, based on both ridgeness and ridge orientation. In addition this fitting is performed for the left and right lane lines simultaneously, thus enforcing a consistent result. We have quantitatively assessed it on synthetic but realistic video sequences for which road geometry and vehicle trajectory ground truth are known.

1 Introduction

A present challenge of the automotive industry is to develop low cost advanced driver assistance systems (ADAS) able to increase traffic safety and driving comfort. Since vision is the most used human sense for driving, some ADAS features rely on visual sensors [2]. Specifically, lane departure warning and lateral control can be addressed by detecting the lane markings on the road by means of a forward–facing camera and computer vision techniques. In this paper we focus on this problem, which is one of the first addressed in the field of ADAS. Many papers have been published on it, since it is a difficult and not yet completely solved problem due to shadows, large contrast variations, vehicles occluding the marks, wearied markings, vehicle ego–motion etc. Recent reviews of detection methods can be found in [1,2].

The main contributions of this paper are three. The first one is to employ a different low–level image feature, namely, *ridgeness*, to obtain a more reliable lane marking points detection under poor contrast conditions (section 2). Aside from this practical consideration, conceptually, a ridge describes better than an edge what a lane line is: the medial axis of a thick, brighter line. Secondly, we have adapted RANSAC, a generic robust estimation method, to fit a parametric model

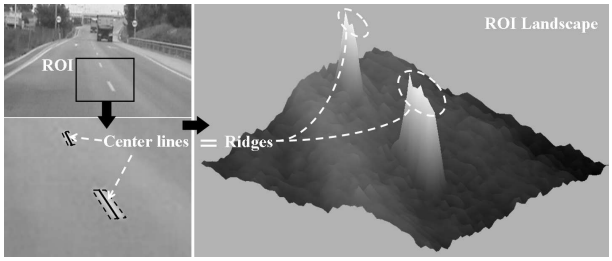

Fig. 1. Left: road image with a region of interest (ROI) outlined. Right: ROI seen as a landscape, where lane markings resemble mountains and ridges correspond correspond to the center of the lane markings.

to the candidate lane marking points, using as input data both ridgeness and ridge orientation (section 3). Thirdly, we quantitatively assess the method with regard to four quantities derived from the segmented lane markings: vehicle yaw angle and lateral offset, lane curvature and width. This is possible on synthetic sequences, for which we know exactly the value for these parameters since they are provided as input to the program which generates them (section 4). Section 5 draws the main conclusions and comments future work. A previous version of this work was presented in [3]. However, it did not include the validation on synthetic sequences and the pair of lane lines model was different.

2 Lane Markings as Ridges

Ridges of a grey–level image are the center lines of elongated, bright structures. In the case of a lane line is its longitudinal center. This terminology comes from considering an image as a landscape, being the intensity the z axis or height, since then these center lines correspond to the landscape's ridges (figure 1). Accordingly, ridgeness stands for a measure of how much a pixel neighborhood resembles a ridge. Therefore, a ridgeness measure must have high values along the center of the line and decrease as the boundary is approached. A binary ridge image, corresponding to the centerline, can be obtained by simple thresholding, provided we have a well–contrasted and homogeneous ridgeness measure.

This notion of ridge or medial axis is a simpler and, as we will see in short, computationally better characterization of lane lines than that provided by edges. Instead of defining (and trying to find out) a lane line as points between two parallel edge segments with opposite gradient direction, a ridge is the center of the line itself, once a certain amount of smoothing has been performed. And this amount is chosen as the scale at witch ridges are sought.

There are different mathematical characterizations of ridges. In [4] a new one is proposed which compares favorably to others and that we have adapted for the problem at hand. Let $\mathbf{x} = (u, v)$ be the spatial coordinates (u columns, v rows). Then, ridgeness is calculated as the positive values of the divergence of

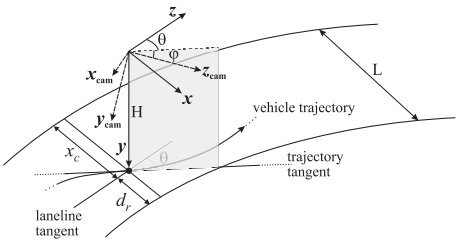

Fig. 2. Image acquisition geometry.

the normalized gradient vector field $\tilde{\mathbf{w}}$ of the image:

$$\tilde{\kappa}_{\sigma_\mathrm{d},\sigma_\mathrm{i}}(\mathbf{x}) = -\mathrm{div}(\tilde{\mathbf{w}}_{\sigma_\mathrm{d},\sigma_\mathrm{i}}(\mathbf{x})) \tag{1}$$

The parameter σ_d is the *differentiation scale*, in opposition to σ_i which is the *integration scale*. The former must be tuned to the size of the target structures, while the later determines the size of the neighborhood we want to use in order to compute the dominant orientation.

We only take into account those pixels $\mathbf{x}$ for which $\tilde{\kappa}_{\sigma_\mathrm{d},\sigma_\mathrm{i}}(\mathbf{x}) > 0.25$, a value fixed experimentally but with a large margin before the selected pixels change significantly. Due to perspective, the imaged lane lines width decreases with distance. In order not to miss them, σ_d also decreases with the row number so that upper rows are less smoothed than lower rows.

Interesting properties of $\tilde{\kappa}_{\sigma_\mathrm{d},\sigma_\mathrm{i}}(\mathbf{x})$ are invariance to image translation and rotation, as one would expect, but also to monotonic grey–level transforms. The later greatly helps in lane detection in presence of shadows and low contrast conditions, opposite to gradient–based measures. Shadows cast by vehicles and road infrastructure (like fences, tunnel entries, lamp posts) give rise to long and straight contour lines which can fool edge–based lane detection methods.

3 Lane model and fitting

3.1 Lane lines model

A number of geometrical models for the projected lane lines have been proposed, from simple straight lines to quadratic, spline and other polynomial curves, with the aim of performing a good image segmentation. However, few are built on a sound geometrical base like in [6]. There it is shown that, under several reasonable assumptions (flat road, constant curvature), a lane line on a planar road is projected onto the image plane as an hyperbola. Admittedly, this is not a new model, but what that work reveals are the relationships among model parameters and meaningful and interesting geometrical entities such as lane width, curvature and the vehicle's lateral position, which we want to compute in order to validate our method, aside of their own evident applicability in ADAS.

As illustrated in figure 2, the camera coordinate system has Y axis coincident with the vehicle's direction and sustains an angle $\theta \ll 1$ rad. with the

road tangent line (also referred as yaw angle). It also forms an angle φ with the road plane (pitch angle). The lane has width L and the camera is located at a horizontal distance of d_r meters from the right border and at height H above the ground. Of course, L, d_r, θ and φ may vary over time, but H is supposed constant. Finally, let be E_u and E_v the focal lengths in pixels/meter along the horizontal and vertical camera axes, and the image origin centered in the principal point (intersection of the optical axis with the image plane). Then, the following equation relates (u_r, v_r), the pixel coordinates where the right lane line is imaged, to the road parameters it belongs to [6]:

$$u_r = E_u \left(\frac{\theta}{\cos\varphi} + \frac{d_r \cos\varphi}{H E_v}(v_r + E_v \tan\varphi) + \frac{E_v H C_0 / \cos^3\varphi}{4(v_r + E_v \tan\varphi)} \right) \tag{2}$$

The former equation clearly follows the formulation of a hyperbola with a horizontal asymptote. In order to enforce parallelism of lane borders, we introduce a new variable x_c, which is the signed distance along the X axis between the camera projection on the road plane and the central axis of the left lane line (figure 2). It follows that $d_r = x_c - L$, $d_l = x_c$ and we have the following couple of equations, for points $(u_l, v_l), (u_r, v_r)$ on the left and right border, respectively:

$$\begin{aligned} u_l &= E_u \left(\frac{\theta}{\cos\varphi} + \frac{\cos\varphi}{H E_v} x_c (v_l + E_v \tan\varphi) + \frac{E_v H C_0 / \cos^3\varphi}{4(v_l + E_v \tan\varphi)} \right) \\ u_r &= E_u \left(\frac{\theta}{\cos\varphi} + \frac{\cos\varphi}{H E_v}(x_c - L)(v_r + E_v \tan\varphi) + \frac{E_v H C_0 / \cos^3\varphi}{4(v_r + E_v \tan\varphi)} \right) \end{aligned} \tag{3}$$

Since parameters E_u, E_v, H and φ can be estimated through a camera calibration process [7], equation (4) is linear with respect to the four unknowns θ, x_c, L and C_0. It can be compactly rewritten as

$$\begin{bmatrix} 1 & 0 & v'_l & 1/v'_l \\ 1 & -v'_r & v'_r & 1/v'_r \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \\ a_3 \\ a_4 \end{bmatrix} = \begin{bmatrix} u_l \\ u_r \end{bmatrix} \tag{4}$$

with $v'_r = v_r/E_v + \tan\varphi, v'_l = v_l/E_v + \tan\varphi$ and

$$\theta = \frac{\cos\varphi}{E_u} a_1, \quad L = \frac{H}{E_u \cos\varphi} a_2, \quad x_c = \frac{H}{E_u \cos\varphi} a_3, \quad C_0 = \frac{4\cos^3\varphi}{E_u H} a_4 \tag{5}$$

Note that according to this model, four points, not all on the same line, define *a pair* of hyperbolas sharing the same horizontal asymptote. In addition, they correspond to two parallel curves L meters apart, when back projected to the road plane. This implies that we are going to fit *both left and right* lane lines at the same time and enforcing parallelism, that is, consistency in the solution. Besides, the sparsity of candidates in one lane side due to shadows, occlusions or dashed lane marks can be compensated by those in the other side.

3.2 Model fitting

A minimum of four points are necessary in order to solve equation (4), provided there is at least one point on each curve. If more points are known, we get an overconstrained system that is solved in the least–squares sense. The problem, of course, is the selection of the right points among all candidates from the previous detection step. We need a robust technique in the sense of, simultaneously, classify candidate points into lane points (inliers) and not lane points (outliers, at least for that side), and perform the fitting only to the former ones. RANSAC, which stands for Random Sample Consensus [9], is a general estimation technique for such purpose based on the principle of hypotheses generation and verification.

An observation must be made concerning the lane model of equation (4). In it we supposed the pitch angle φ known from the calibration process, but actually it suffers variations around its nominal value due to non–planar roads, acceleration, brake actioning etc. To account for this fact, quite influential in the instantiated model because it changes its horizontal asymptote, we test several possible values for φ, taking n_φ equispaced samples within $\varphi \pm \Delta\varphi$.

4 Results

As pointed out in a recent survey on video–based lane departure warning [10], results in the literature are often presented only in the form of several frames, where the reader can check the correspondence between detected lane lines and real lane markings. We also present results in this qualitative way, but just to show examples of challenging situations (see figure 3). However, since our fitted model has a direct relation to geometrically meaningful parameters of interest in the context of ADAS, we base the evaluation on the comparison of these computed values with the actual ones. The construction of digital road maps at lane line resolucion is a research issue in itself. Therefore, we have resorted to build a simulator which generates sequences of synthetic but realistic images of 3D roads of known geometry.

The simulator, implemented in Matlab, models the road geometry, photometry and the camera focal length, trajectory and pose with regard a world coordinate system. Whereas the camera roll angle has been fixed and yaw depends on the trajectory, the pitch angle φ has not, since it is responsible for the horizon line vertical motion, which is not static in real sequences. Besides, it turns out that this parameter is quite influential on the results. Thus, we have randomly varied the pitch angle so as to mimic the effects of 1) uneven road surface and 2) acceleration and brake actioning, both observed in real sequences. Specifically, pitch variation is generated by adding two random noises: the first one of high frequency and small amplitude ($\leq 0.2°$) and the second one of low frequency but larger amplitude (between $0.5°$ and $1°$), which account respectively for the two former pitch variation sources.

We have performed several tests on synthetic sequences in order to calculate the error in the estimation of C_0, L, θ, x_c and also φ on a 5 Km long road, one

Fig. 3. Segmented lane line curves. From left to right and top to bottom : dashed lines, occlusion, tunnel exit, special road marks, shadows, night image with reflections. Complete sequences from which these frames have been extracted can be viewed at `www.cvc.uab.es/adas/projects/lanemarkings/IbPRIA07.html`.

frame per meter. Sudden slope changes introduce large errors, though logically localized in time, whereas almost no error is observed at curvature changes. At frames where the pitch variation has its largests peaks ($t = 380, 720, 1200, 1450$), the error is small for x_c, moderate for L but large for C_0 and θ (figure 4, due to space limitations, we do not include the figures for all the parameters). The reason is that x_c and L are local road measures very close to the camera position and thus not affected by the global lane line shape, specially its shape at a large distance, close to the horizon line. On the contrary, C_0 and θ do depend on the global shape (according to the road model, the curvature is supposed to be constant) which is in turn dependent on the shared lane lines horizontal asymptote. In addition, a small amplitude noise appears elsewhere mainly due to the small amplitude pitch variation.

We have tried to minimize the effect of pitch changes (both large and small) by considering φ another parameter to estimate. The second row of figure 4 shows the result for $n_\varphi = 7$ and $\Delta\varphi = 1°$. Close examination of the estimated pitch angle allows us to conclude that the best pitch search is often able to correctly estimate it (the four largest pitch variations are well detected), but not always. The most prominent errors are localized around the four slope changes, where this simple approach of guessing the best pitch fails. Elsewhere, a sort of impulsive error is observed, caused by a small number of inliers. In spite of it, a causal median filter (median of a number of pitch estimations before the current frame) produces an acceptable result, even for a small n_φ. Finally, figure 5 shows the root–mean square error between computed and ground truth for four different values of n_φ, and also for their median filtered versions. Whereas there is only a slight improvement, or even no improvement at all, when n_φ increases, the error of the filtered parameters clearly decreases. Therefore, it seems that it does not pay to look for the best pitch if no filtering is performed afterwards.

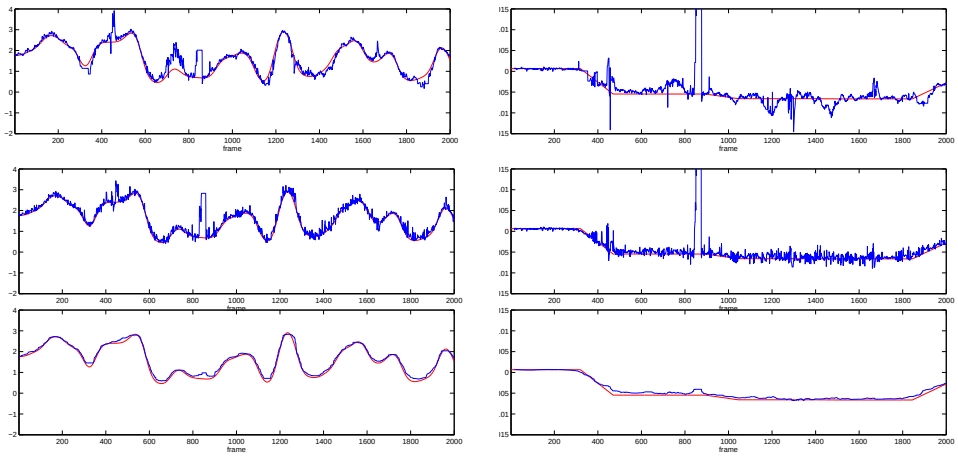

Fig. 4. Ground truth and computed x_c (left) and C (right) for (top to bottom): n_φ=1 and 7 pitch angles around nominal camera pitch, and median filtering of this latter result.

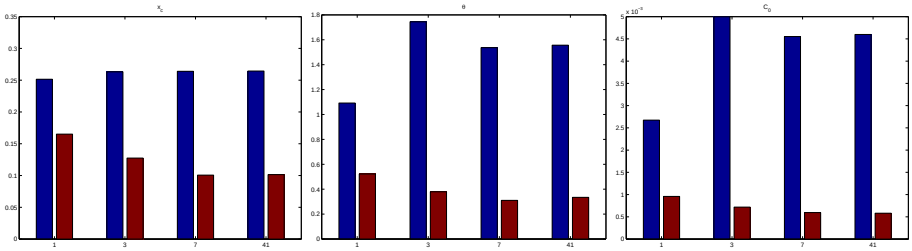

Fig. 5. Root-mean square error of computed x_c, θ and C_0 (angles in degrees) n_φ =1, 3, 7 and 41. In each pair, left bar corresponds to the computed value and right bar to its causal median filtered version.

5 Conclusions

We have developed a new method for the extraction of lane lines from video sequences. Robustness is achieved both in the feature detection phase, where we employ an image feature well suited to this problem, and in the model fitting phase, which we have addressed with the RANSAC approach. This method relies just on images, that is, we do not take into account data from other vehicle sensors, like the steering angle or yaw rate. Also, we have avoided any kind of result post processing: each frame is processed independently of the others. The reason is that our aim was to build a 'baseline' system to which add later filtering and data fusion to improve the result. Our lane line extraction method has the advantage of computing four road and vehicle trajectory parameters which are of interest in the context of ADAS. We have compared the computed values with ground truth from a synthetic but realistic road. The present implementation of feature detection, model fitting and parameter computation runs in real–time,

at 40 ms/frame on a 2 Ghz Pentium IV, no matter if the image is synthetic or real. From these experiments we conclude that we can compute reasonable estimations of road curvature, width and vehicle lateral position and direction, even in the case where the road does not follow the assumed model of flatness, constant curvature and known camera pitch. However, the weak point of our method is the estimation of the pitch angle, which we expect to predict on the basis of previous frames.

Acknowledgments. This research has been partially funded by Spanish MEC grant TRA2004-06702/AUT.

References

1. C.R. Jung, C.R. Kelber. *Lane following and lane departure using a linear–parabolic model*, Image and Vision Computing 23, pg. 1192–1202, 2005.
2. M. Bertozzi, A. Broggi, A. Fascioli. *Vision–based Intelligent Vehicles: State of the Art and Perspectives*, Robotics and Autonomous Systems, vol. 32, pg. 1–16, 2000.
3. A. López, C. Cañero, J. Serrat, J. Saludes, F. Lumbreras, T. Graf. *Detection of Lane Markings based on Ridgeness and RANSAC*, IEEE Conf. on Intelligent Transportation Systems, pg. 733–738, Vienna, 2005.
4. A. López, D. Lloret, J. Serrat, J. Villanueva. *Multilocal Creaseness Based on the Level–Set Extrinsic Curvature*, Computer Vision and Image Understanding, vol. 77, pg. 111–144, 2000.
5. E.D. Dickmanns and B.D. Mysliwetz. *Recursive 3D road and relative ego–state recognition*, IEEE Trans. on Pattern Analysis and Machine Intelligence, vol. 14, pg. 199–213, 1992.
6. A. Guiducci. *Parametric Model of the Perspective Projection of a Road with Application to Lane Keeping and 3D Road Reconstruction*, Computer Vision and Image Understanding, vol. 73, pg. 414–427, 1999.
7. Z. Zhang. *A flexible new technique for camera calibration*, IEEE Transactions on Pattern Analysis and Machine Intelligence, 22(11), pg. 1330–1334, 2000.
8. R. Aufrère and R. Chapuis and F. Chausse. *A Model–Driven Approach for Real–Time Road Recognition*, Machine Vision and Applications, vol. 13, pg. 95–107, 2001.
9. M.A. Fischler and R.C. Bolles. *Random Sample Consensus: a Paradigm for Model Fitting with Applications to Image Analysis and Automated Cartography*, Commun. ACM, vol. 24, pg. 381–395, 1981.
10. J.C. McCall and M.M. Trivedi. *Video–based lane estimation and tracking for driver assistance: survey, system and evaluation*, IEEE Trans. on Intelligent Transportation Systems, Vol. 7, No. 1, March 2006, pp.20–37.