2014 IEEE International Conference on Robotics & Automation (ICRA)
Hong Kong Convention and Exhibition Center
May 31 - June 7, 2014. Hong Kong, China

Vision-Based Robust Road Lane Detection in Urban Environments

Michael Beyeler[1,2], Florian Mirus[1], Alexander Verl[1]

***Abstract*— Road and lane detection play an important role in autonomous driving and commercial driver-assistance systems. Vision-based road detection is an essential step towards autonomous driving, yet a challenging task due to illumination and complexity of the visual scenery. Urban scenes may present additional challenges such as intersections, multi-lane scenarios, or clutter due to heavy traffic. This paper presents an integrative approach to ego-lane detection that aims to be as simple as possible to enable real-time computation while being able to adapt to a variety of urban and rural traffic scenarios. The approach at hand combines and extends a road segmentation method in an illumination-invariant color image, lane markings detection using a ridge operator, and road geometry estimation using RANdom SAmple Consensus (RANSAC). Employing the segmented road region as a prior for lane markings extraction significantly improves the execution time and success rate of the RANSAC algorithm, and makes the detection of weakly pronounced ridge structures computationally tractable, thus enabling ego-lane detection even in the absence of lane markings. Segmentation performance is shown to increase when moving from a color-based to a histogram correlation-based model. The power and robustness of this algorithm has been demonstrated in a car simulation system as well as in the challenging KITTI data base of real-world urban traffic scenarios.**

I. Introduction

Road and lane position detection are integral components in autonomous driving and intelligent transportation systems. A large body of research has focused on vision-based road detection methods [1]–[6], and some of these methods are already successfully used in commercial lane estimation systems [4]. Most vision-based road detection algorithms rely on low-level image features (such as color or texture) in combination with a post-processing step for outlier removal (using tracking, filtering, data fusion, or computational models of both road and vehicle) [1]. However, additional constraints usually come at an additional computational cost, and can restrict the algorithm to highly structured roads and uncluttered scenes (such as freeways or interurban traffic scenarios). In urban traffic scenarios, however, a road detection algorithm might be challenged by the presence of substantial clutter due to heavy traffic, worn-out lane markings, intersections, slip lanes, or others; and abrupt changes in illumination strongly impact the quality of visual sensing. Finding practical solutions to such real-world scenarios is currently a major research topic in autonomous and highly automated driving.

[1] Michael Beyeler, Florian Mirus, and Alexander Verl are with the Fraunhofer Institute for Manufacturing Engineering and Automation IPA, Robot and Assistive Systems Department, Nobelstraße 12, 70569 Stuttgart, Germany Email: florian.mirus@ipa.fraunhofer.de

[2]Michael Beyeler is with the University of California, Irvine, Department of Computer Science, Irvine, CA 92697, USA Email: mbeyeler@uci.edu

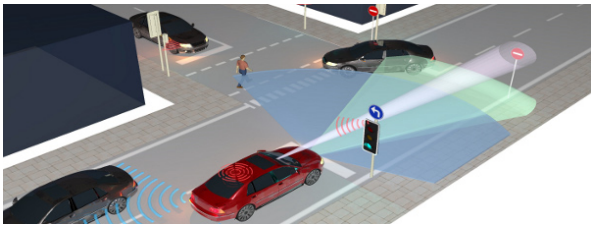

Fig. 1. Autonomous vehicle tackling an urban traffic scenario.

The work at hand proposes a fast and robust ego-lane detection algorithm that aims to be as simple as possible to enable real-time computation while being able to adapt to a variety of driving scenarios. The algorithm combines and extends three independent pieces of research as follows. Road segmentation is performed on a "shadow-free" representation [6] of an RGB image using a histogram correlation method, and serves as a contextual prior for the detection of lane markings, which are extracted by means of ridge or "creaseness" detection [5]. A computational model of the ego-lane geometry is then rapidly constructed by fitting candidate lane markings to the model parameters using RANdom SAmple Consensus (RANSAC) [7]. Thus the resulting algorithm is able to determine on a frame-by-frame basis: (i) the road region, (ii) the ego-lane region and markings, (iii) the relative position and orientation of the car with respect to the ego-lane, and (iv) the local lane width. The effectiveness and robustness of this algorithm is demonstrated in a car simulation system (PreScan) and in the challenging real-world KITTI Vision Benchmark Suite [8]. Road segmentation results from the here employed histogram correlation method show a significant improvement in performance when compared to a previously proposed color-based approach [6]. Furthermore, representative sample frames demonstrate the correct extraction of ego-lane information (i) in multi-lane scenarios, (ii) for a variety of road curvature values, (iii) in the presence of cars on the road, (iv) in the presence of strong shadows, (v) at intersections, and (vi) in the absence of explicit lane markings. Thus the model at hand signifies the first step towards constructing a powerful but efficient vision-based algorithm capable of aiding an autonomous vehicle in navigating complex urban scenes.

II. Implementation

The process flow of the proposed algorithm is shown in Fig. 2. The following subsections will explain the model in detail.

978-1-4799-3684-7/14/$31.00 ©2014 IEEE

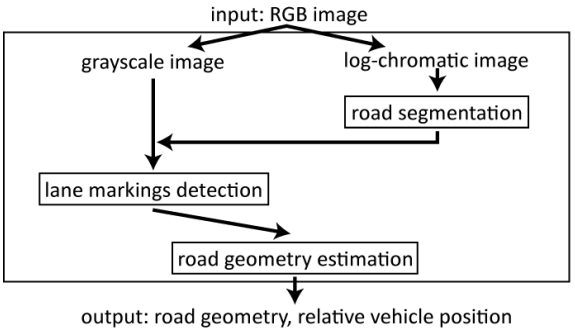

Fig. 2. Process flow of the model.

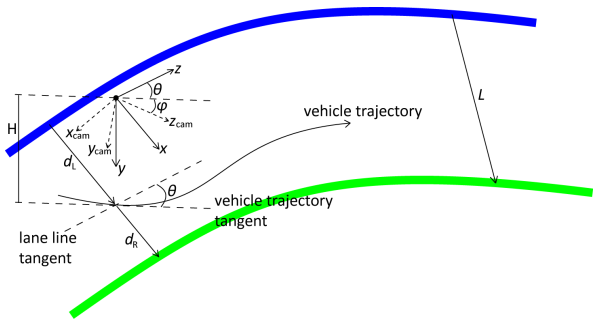

Fig. 3. Camera setup for image acquisition (adapted from [5]).

A. Camera Setup

The camera setup is shown in Fig. 3. For the purpose of this study, it is assumed that a stationary, forward-facing color camera is located at the center of the windshield, a horizontal distance of d_{L} meters from the left lane border (blue) at height H above the ground, and a distance d_{R} from the right lane border (green), such that total lane width $L = d_{\mathrm{L}} + d_{\mathrm{R}}$. The camera coordinate system sustains an angle θ with the road tangent line (yaw angle), and an angle ϕ with the road plane (pitch angle). Whereas L, d_{L}, d_{R}, θ, and ϕ may vary over time, H and the location of the horizon can be assumed to remain constant. Thus, it is possible to project the horizon line into each image frame and consider only pixel rows below this line for further processing.

B. Road Segmentation

In order to remove shadows, the RGB image is converted to an illuminant-invariant or "log-chromaticity" space [6], to perform a texture-based road segmentation. The illuminant-invariant transformation is illustrated in Fig. 4. For a given RGB image with standard color channels R, G, and B the log-chromaticity space has two axes, $r = \log(R/G)$ and $b = \log(B/G)$, using the G channel for normalization. In this space, a set of color surfaces of a given chromaticity value imaged under different lighting conditions is projected onto a straight line (labeled "illumination-dependent" in Fig. 4). Sets of color surfaces with different chromaticity values form parallel straight lines, whose offset corresponds to chromatic difference ("chromaticity-dependent"). Furthermore, these lines define an orthogonal gray-level axis ℓ_α, where a surface under different illuminations is represented by the same point. Movements along ℓ_α imply changing the surface chromaticity, independently of illumination. An example of such a transformation is shown in Fig. 4.

Computation of this image requires the knowledge of the angle α, which is an intrinsic parameter of the camera sensor, and can be estimated using entropy-minimization [6].

In order to classify small image patches as either road surface or background, it is proposed to extend the color-based classifier introduced by Alvarez and Lopez [6] to a classifier based on histogram correlation. Thus, instead of assigning each pixel a probability of belonging to the road based on its illuminant-invariant color value, it is proposed to compare the normalized histogram of a small image patch with that of a road model. A road model $\mathcal{M}$ was built by computing the normalized histogram $h_{\mathcal{M}}$ of illuminant-invariant color values in a neighborhood of a set of road pixels. For this, a set of N_{s} seeds was placed following an equidistant distribution along two rows R_{s} at the bottom part of each image. The normalized histogram was then computed within a squared area of $K_{\mathrm{s}} \times K_{\mathrm{s}}$ pixels surrounding each seed. Finally, the i-th image patch consisting of $K_{\mathrm{s}} \times K_{\mathrm{s}}$ pixels was classified as "road" if its histogram h_i strongly correlated with $h_{\mathcal{M}}$; that is, if $corr(h_i, h_{\mathcal{M}}) \geq \lambda$.

Following the implementation of Alvarez and Lopez, bin width was chosen according to Scott's rule, and a post-processing step was performed based on mathematical morphology operations (a closing with a 5px × 3px structuring element followed by a flood-fill operation) [6]. However, the here described histogram correlation approach has the clear advantage of being capable to differentiate even between image patches of similar chromaticity given that h_i and $h_{\mathcal{M}}$ are only weakly correlated. This is advantageous especially in urban environments, where it is likely to find a multitude of grayish structures (similar color) but with distinct textures. An example is shown in Fig. 5, where the approach by Alvarez and Lopez (top) is not able to differentiate between road, sidewalk and train tracks, but the approach at hand (bottom) is. In order to ensure that Fig. 5 shows the best possible segmentation result, an exhaustive search for parameters K_{s} and λ was performed, and the result compared to the manually segmented road region. The top image used $K_{\mathrm{s}} = 11$ and $\lambda = 0.01$, whereas the bottom image used $K_{\mathrm{s}} = 15$ and $\lambda = 0.84$ (note the different definitions of λ). The segmented road region was then used as a search window for the detection of lane markings, reasoning that lane markings must be located on the road.

C. Lane Markings Detection

The lane markings detection algorithm is based on a low-level image feature called a "ridge", which is a measure of "creaseness" [5], [9]. Ridges of a gray-level image are the center lines of bright, elongated structures. Let $I(\mathbf{x})$ be a grayscale image with spatial coordinates $\mathbf{x} = (x_1, x_2)$. If the image is considered a landscape by plotting $I(\mathbf{x})$ against $\mathbf{x}$, then these center lines correspond to the landscape ridges. It has been shown that this analysis performs better than, for

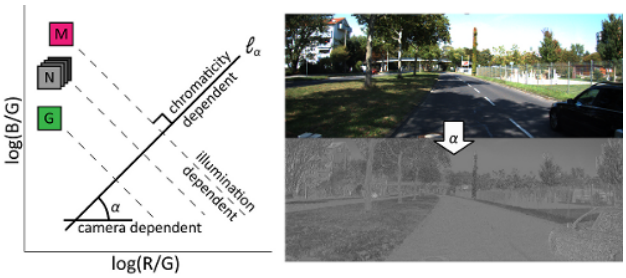

Fig. 4. Illuminant-invariant transformation ($\alpha = 0.85$).

Fig. 5. Segmentation example. Top: color-based, Alvarez and Lopez [6]. Bottom: histogram correlation-based, this paper.

example, lane markings detection based on steerable filters or edge detection, because of its invariance to image translation, rotation, and monotonic gray-level transforms [5], [9]. An enhanced creaseness measure $\tilde{\kappa}$ can be obtained as follows.
(1) Compute the structure tensor field $\mathbf{S}$ based on the gradient vector field w. For the computation of $\mathbf{S}$ a Gaussian neighborhood $G(\mathbf{x};\sigma_I)$ of size σ_I ("integration scale") is used. In the 2D case this gives

$$\mathbf{S}(\mathbf{x};\sigma_I;\sigma_D) = \begin{pmatrix} s_{11} & s_{12} \\ s_{21} & s_{22} \end{pmatrix}, \quad (1)$$

$$s_{ij}(\mathbf{x};\sigma_I;\sigma_D) = G(\mathbf{x};\sigma_I) * \frac{dI}{dx_i}(\mathbf{x};\sigma_D)\frac{dI}{dx_j}(\mathbf{x};\sigma_D)),$$

where $*$ denotes convolution and σ_D ("differentation scale") denotes the standard deviation of the Gaussian kernel involved in the differentation process needed to compute w in a well-posed manner. The differentation scale is tuned to the size of the objects whose orientation has to be determined, whereas the integration scale is tuned to the size of the neighborhood in which an orientation is dominant.
(2) Perform the eigensystem analysis of $\mathbf{S}$ analytically. The eigenvector which corresponds to the highest eigenvalue $w'(\mathbf{x};\sigma_I)$ of $\mathbf{S}(\mathbf{x};\sigma_I)$ yields the dominant gradient orientation at $\mathbf{x}$, where "dominant" means inside the Gaussian neighborhood [9]. Note that the gradient of a function points toward the direction of maximum change. In order to apply the divergence operator in the next step, a proper direction to the dominant gradient orientation w' needs to be assigned. To recover such direction, w' is put in the same 2D-quadrant as w. Thus the new vector field is $\tilde{w} = \text{sign}(w'^t w)w'$.
(3) Compute the new enhanced creaseness measure $\tilde{\kappa}_d = -\text{div}(\tilde{w})$. Positive values of $\tilde{\kappa}_d$ measure the similarity of a neighborhood to a ridge structure. In fact, it has been shown that these values lie in the range $[0.0, 2.0]$, where 0 means not at all ridge, around 1.0 quite and 2.0 perfect local maximum [9]. The algorithm only considers those pixels $\mathbf{x}$ for which $\tilde{\kappa}_d(\mathbf{x}) > 0.25$, a value set experimentally but fixed for all simulations.
Due to perspective, the imaged lane lines' width decreases with distance. In order not to miss them when computing the image gradient, anisotropic Gaussian smoothing is performed, where the variance σ_D of the Gaussian (along the horizontal direction) increases with increasing row number.
(4) Compute a suitable confidence measure C to reduce creaseness in the structures that are not essential for further steps. Ridges should be enhanced if there is a single dominant orientation within a neighborhood, which is indicated by a large dissimilarity of the eigenvalues in the structure tensor. Therefore, denoting the eigenvalues of $\mathbf{S}$ by λ_1, λ_2, a logical choice consists of testing whether the sum of quadratic differences

$$\lambda_\Delta(\mathbf{x};\sigma_I) = \sum_{i=1}^{d}\sum_{j=i+1}^{d}(\lambda_i(\mathbf{x};\sigma_I) - \lambda_j(\mathbf{x};\sigma_I))^2 \quad (2)$$

exceeds a pre-defined threshold c characteristic for λ_Δ in the structure to be enhanced. A suitable function is $C(\mathbf{x};\sigma_I;c) = 1 - \exp(-(\lambda_\Delta(\mathbf{x};\sigma_I))^2/2c^2)$. Finally $\tilde{\kappa}_d C$ is used as a measure of creaseness.

D. Road Geometry Estimation

Detected ridges are fed into the road geometry estimation module as candidate lane markings, which are then fit to a parametric road module using RANSAC.
Under the assumption of a flat road and constant curvature, a lane line is projected onto the image plane as a hyperbola [5], [10]. Furthermore, the curvature is either constant or varies linearly with the arc length s, such that curvature $C = 1/R = C_0 + C_1 s$. This is consistent with a road formed by segments of constant curvature connected by clothoids [11]. The changes in curvature are assumed to be smooth, such that the linear term of the curvature is negligible, i.e. $C_1 \approx 0$.
The road geometry can be expressed as a pair of hyperbolas, modeling the left and right lane lines, that share the horizon as horizontal asymptote and correspond to two parallel lines L meters apart when back-projected to the road plane (see Fig. 3). The corresponding linear system $A\theta_{\mathcal{M}} = b$ that relates two points (u_L, v_L) and (u_R, v_R) on the left resp. right lane line to the road parameters can be written as

$$\begin{bmatrix} 1 & 0 & -v'_\text{L} & 1/v'_\text{L} \\ 1 & v'_\text{R} & -v'_\text{R} & 1/v'_\text{R} \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \\ a_3 \\ a_4 \end{bmatrix} = \begin{bmatrix} u_\text{L} \\ u_\text{R} \end{bmatrix}, \quad (3)$$

where $v'_{\mathrm{L}} = v_{\mathrm{L}}/E_v + \tan\phi$, $v'_{\mathrm{R}} = v_{\mathrm{R}}/E_v + \tan\phi$, and E_u resp. E_v is the focal length in pixels/meter along the horizontal resp. vertical camera axis [5], [10]. The model $\theta_{\mathcal{M}} = [a_1, \ldots, a_4]^t$ can be translated to more meaningful entities

$$\begin{aligned} \theta &= \frac{\cos\phi}{E_u} a_1, & \mathrm{L} &= \frac{H}{E_u \cos\phi} a_2, \\ d_{\mathrm{L}} &= \frac{H}{E_u \cos\phi} a_3, & C_0 &= \frac{4\cos^3\phi}{E_u H} a_4, \end{aligned} \tag{4}$$

where θ is the yaw angle, L is the road width, d_{L} is the camera's horizontal distance from the left lane line, and C_0 is the road curvature (compare Fig. 3). Parameters E_u, E_v, H, and ϕ can be estimated through a camera calibration process [12].

Since the camera is located at the center of the windshield (see Fig. 3) and forward-facing, it is straight-forward to estimate the u-coordinate of the vanishing point u_{vanish}, when the vehicle is centered in a straight line. The assumption that the lane lines must be located on either side of u_{vanish} allows to split the candidate lane markings into two sets. However, if the ridge point is close to the horizon, this simple assumption does not hold. Since the point could belong to either the left or the right lane line, a third set is needed.

In each RANSAC iteration, four points from the three sets of candidate lane markings are picked randomly (not all from the same set), which is the minimum number required to solve Eq. 3. Based on these four points, a road model $\tilde{\theta}_{\mathcal{M}}$ is estimated. If the model yields a reasonable value $L \in [L_{\text{min}}, L_{\text{max}}]$ for the lane width, the model support is computed by taking into consideration all candidate lane markings. Otherwise four new points are picked from the sets.

In order to determine the set of inliers for a hyperbola, a metric or error function is needed. Other studies have employed the geometric distance of a point to a conic or the Sampson's distance [5] (although the latter has been criticized for being conceptually and computationally inaccurate [13]). The work at hand proposes a computationally much simpler yet effective step. First, the absolute error between a candidate lane markings point and the current model estimate $\tilde{\theta}_{\mathcal{M}}$ is computed

$$\begin{aligned} \epsilon_{\mathrm{L},i} &= |(\tilde{a}_1 - v'_{\mathrm{L},i}\tilde{a}_3 + \tilde{a_4}/v'_{\mathrm{L},i}) - u_{\mathrm{L},i}| \\ \epsilon_{\mathrm{R},j} &= |(\tilde{a}_1 + v'_{\mathrm{R},i}\tilde{a}_2 - v'_{\mathrm{R},j}\tilde{a}_3 + \tilde{a}_4/v'_{\mathrm{R},j}) - u_{\mathrm{R},j}|, \end{aligned} \tag{5}$$

where $(u_{\mathrm{L},i}, v_{\mathrm{L},i})$ resp. $(u_{\mathrm{R},j}, v_{\mathrm{R},j})$ is the i-th resp. j-th candidate lane marking point for the left resp. right lane line. If the error ϵ_{L} resp. ϵ_{R} for left resp. right lane candidate is smaller than some threshold ϵ_{max}, the candidate point is considered an inlier and put in the consensus set. ϵ_{max} was set to 5.0.

A score z for the two hyperbolas is then computed by considering all inliers i in the consensus set

$$z = \sum_i \frac{1}{1+\epsilon_{\mathrm{L},i}} + \sum_j \frac{1}{1+\epsilon_{\mathrm{R},j}}. \tag{6}$$

The contribution of an inlier in the total hyperbola score is inversely related to the inlier's approximation error $\epsilon_{l,i}$. The overall result is that hyperbolas with many good inliers have the greatest score. Combining the scores of both left and right hyperbolas has the advantage that a good fit of one hyperbola can compensate a poor fit of the other. If both hyperbolas are fit poorly, then the total score z will be relatively low. If z is below some threshold, it is discarded.

After a number (experiments show 200 to be sufficient) of RANSAC iterations, the pair of hyperbolas with the greatest score is selected as the lane lines.

TABLE I
ROAD SEGMENTATION RESULTS

	$corr_{\text{Bhattacharyya}}$	$corr_{\text{simple}}$	color-based [6]
$\hat{\mathcal{F}}$	0.9333 ± 0.0023	0.9236 ± 0.0033	0.8709 ± 0.0210
$\hat{\sigma}_{\hat{\mathcal{F}}}$	0.0669 ± 0.0163	0.0697 ± 0.0125	0.1745 ± 0.0353

III. EVALUATION

A. Road Segmentation

Table I reports the effectiveness of the here proposed histogram correlation-based road segmentation algorithm (using Bhattacharyya distance resp. simple correlation) in comparison with a segmentation method based on illuminant-invariant color [6] (re-implemented for the purpose of performance comparison). Results were obtained by following a cross-validation procedure to apply the different segmentation algorithms to the "Rainy Day" sequence by Alvarez and Lopez [6]. This image sequence was taken during a rainy day, just after the rain stopped so that the road was wet, and can be downloaded from http://www.cvc.uab.es/adas/databases.

Quantitative evaluations are provided by three pixelwise measures: (i) precision $P = (\sum \mathcal{G} I_r)/\sum I_r$, (ii) recall $R = (\sum \mathcal{G} I_r)/\sum \mathcal{G}$, and (iii) effectiveness $F = (2PR)/(P+R)$, where $\mathcal{G}$ and I_r are the ground-truth mask and the segmentation result of a given color image, respectively. 35 images of the "Rainy Day" sequence were selected randomly for learning the classification threshold λ, whereas the rest of the images was used for testing. This procedure was repeated 20 times, and average values were computed for effectiveness $\hat{\mathcal{F}} = (1/20)\sum_{i=1}^{20} \hat{F}_i$ and standard deviation $\hat{\sigma}_{\hat{\mathcal{F}}} = (1/20)\sum_{i=1}^{20} \sigma_{\hat{F}_i}$. To ensure fair comparison, for all algorithms $\alpha = 44°$, $K_s = 11$, $N_s = 9$. Correlation based on Bhattacharyya distance proved the most effective and stable (lowest standard deviation), whereas the color-based results were comparable to the ones reported by Alvarez and Lopez [6].

B. Ego-Lane Detection

In order to assure the robustness and correctness of the proposed algorithm, testing has been conducted using both a PreScan simulation scenario (see Fig. 6) and the KITTI Vision Benchmark Suite [8] (see Fig. 7). Left and right lane lines are shown in blue and green, respectively (compare

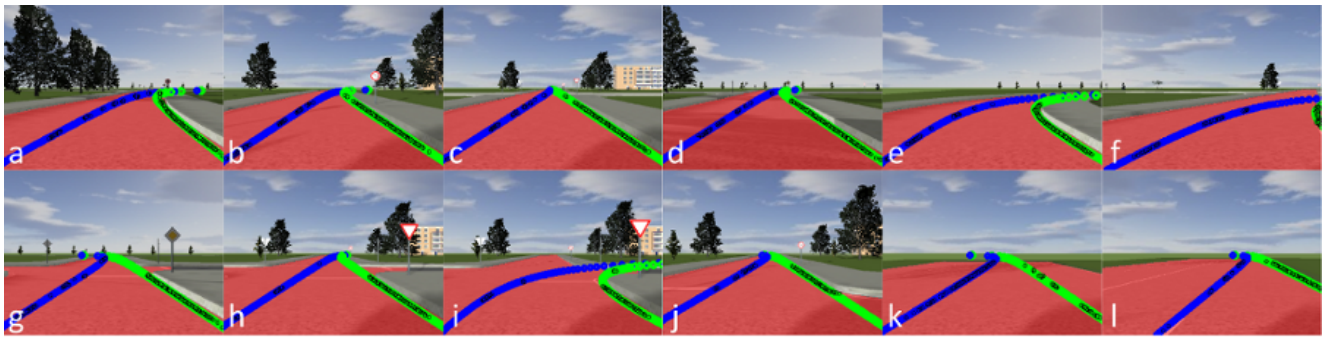

Fig. 6. Simulation results using PreScan.

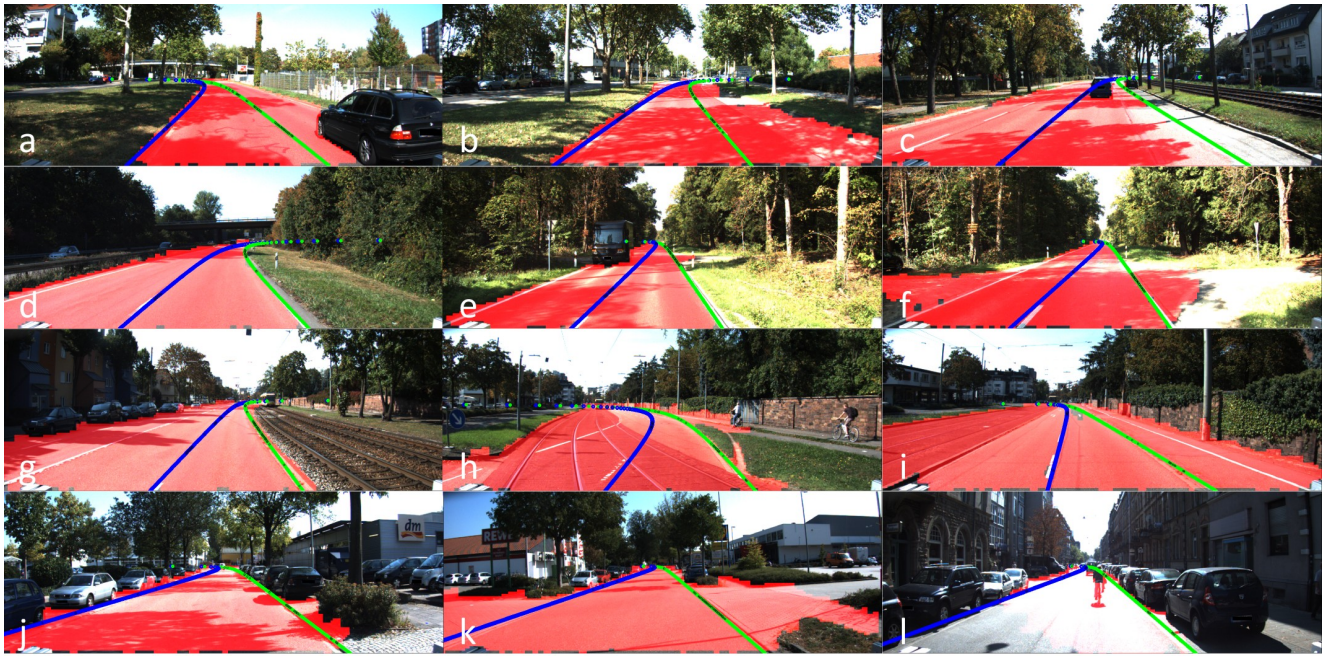

Fig. 7. Real world results using the KITTI Vision Benchmark Suite.

Fig. 3), whereas ridge points in the final RANSAC consensus set are plotted as small black circles. The segmented road region is shown in a semi-transparent red.

A major problem in acquiring quantitative results for meaningful parameters such as the actual road curvature, is the lack of ground truth and precise knowledge of road shape, camera position and the viewing direction at each frame. Thus, qualitative results in the form of video frames are presented, to show examples of challenging situations involving multiple lanes, intersections, shadows, and poor lighting conditions.

1) PreScan (a car simulation software tool): The proposed algorithm was able to segment the road region with high precision, even in the presence of shadows (Panels b–d, h–l). Estimation of lane geometry was robust for a variety of curvature magnitudes (Panels a–f), which is obvious especially in Panel f, where only a few points of the right lane are visible. When navigating an intersection (Panels g–j), the algorithm is confronted with the problem of having multiple lanes present, and when close enough to the intersection (Panel i), the algorithm might in fact favor a turn over going straight ahead. However, this is intended behavior, as two equally likely road models are present in the frame. In this scenario, another intelligence could easily inform the road geometry module to look for a road patch of a certain curvature (e.g., enforce small curvature magnitudes to go straight ahead). After passing the intersection (Panel j), the algorithm recovers quickly.

In the presence of multiple lanes (Panels k–l) the algorithm constantly signals the right lane due to the fact that candidate lane markings were partitioned into either possibly belonging to the left or right lane line (see Section II-D). Thus the algorithm will never confuse a neighboring lane with its own.

2) KITTI Vision Benchmark Suite (captured by autonomous vehicle "AnnieWAY" in the city of Karlsruhe): The here presented algorithm demonstrates robust behavior in a variety of urban environments (Panels a–c, g–l), in rural and interurban environments (Panels d–f), even in the presence of strong shadows (Panels a–c, e, f, and j) or objects on the road (Panel a, c, e, and l).

Panel i depicts a particularly difficult scenario, in which a simple edge-detection based algorithms might confuse the solid white line on the right-hand side for a lane boundary, when it in fact belongs to a bicycle track. The here presented algorithm avoids this pitfall by correctly detecting the curb as candidate lane boundary and by discarding the outer white line because the resulting lane width would be too large.

Panels j–l demonstrate that the here presented algorithm is capable of ego-lane detection even in scenarios where no explicit lane markings are present. One might assume that in such an environment ridge detection would return an empty result, and thus that RANSAC could not be performed. However, by constraining the search area for lane markings to the segmented road region, the search for overall weakly pronounced but locally dominant ridge structures becomes computationally tractable, which in turn enables the use of curbs, surface subsidences, or even markings of a parking spot as candidate lane boundaries for the correct estimation of road geometry.

IV. Discussion

A. Contribution

The work at hand presents an integrative vision-based approach to aid an autonomous vehicle in navigating in urban environments, by combining and extending a set of independent pieces of research.

A previously published road segmentation algorithm was extended to an approach based on histogram correlation, which significantly improved segmentation performance. This alteration proves viable especially in dense urban environments, which show a predominance of grayish textures such as sideways, train tracks, or walls. Thus the algorithm is capable of separating regions of similar color given that their histogram-based texture descriptors differ significantly.

Employing the segmented road region as a prior for lane markings extraction significantly improves the execution time and success rate of the RANSAC algorithm. Additionally, as the here presented results demonstrate, no computationally intensive RANSAC error calculation (such as computing the geometric distance to a conic sector) is necessary to correctly infer road geometry in a variety of driving scenarios. Moreover, limiting the search window for candidate lane markings makes the detection of weakly pronounced ridge structures computationally tractable, thus enabling ego-lane detection even in the absence of explicit lane markings.

B. Model Limitations and Future Work

Although the proposed segmentation algorithm works well in a variety of urban environments, more work and testing needs to be done in order to make the algorithm practical for use in a real-world urban environment.

More attention needs to be directed to difficult scenarios such as intersections and heavy traffic. RANSAC success rate may be impeded if only a limited number of candidate lane markings are present (such as in heavy traffic), if the present lane bifurcates (such as in slip lanes), or when multiple equally valid lane paths are available (such as at intersections and roundabouts). Some of these conditions may require a more sophisticated (e.g., probabilistic) framework [3]. In the latter condition, external human or artificial input may inform the vision-based algorithm of the driver's intended trajectory.

The here proposed algorithm processes each frame independently. However, in order to improve robustness, one might consider the fact that road parameters do not change significantly from frame to frame, given a sufficiently high frame rate. Especially RANSAC execution time might benefit from prohibiting the estimated road model to change significantly on a frame-by-frame basis. For this, a Kalman or particle filter could be employed.

Future work could also be directed in the field of sensor fusion, such as using laser scanners to detect weak surface subsidences or structures that are difficult to detect using computer vision algorithms.

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*, 2005.

[2] S. Thrun, "Winning the DARPA grand challenge," in *Knowledge Discovery in Databases: PKDD 2006*, ser. Lecture Notes in Computer Science, J. Fürnkranz, T. Scheffer, and M. Spiliopoulou, Eds. Springer Berlin / Heidelberg, 2006, vol. 4213, pp. 4–4, 10.1007/11871637-4. [Online]. Available: http://dx.doi.org/10.1007/11871637-4

[3] Z. Kim, "Robust Lane Detection and Tracking in Challenging Scenarios," *IEEE Transactions on Intelligent Transportation Systems*, vol. 9, no. 1, 2008.

[4] A. Amditis, M. Bimpas, G. Thomaidis, M. Tsogas, M. Netto, S. Mammar, A. Beutner, N. Mohler, T. Wirthgen, S. Zipser, A. Etemad, M. Da Lio, and R. Cicilloni, "A Situation-Adaptive Lane-Keeping Support System: Overview of the SAFELANE Approach," *IEEE Transactions on Intelligent Transportation Systems*, vol. 11, no. 3, pp. 617–629, 2010.

[5] A. M. López, Serrat Joan, C. Cañero, F. Lumbreras, and T. Graf, "Robust lane markings detection and road geometry computation," *International Journal of Automotive Technology*, vol. 11, no. 3, pp. 395–407, 2010.

[6] J. M. Á. Alvarez and A. M. Lopez, "Road Detection Based on Illuminant Invariance," *IEEE Transactions on Intelligent Transportation Systems*, vol. 12, no. 1, pp. 184–193, 2011.

[7] M. A. Fischler and R. C. Bolles, "Random sample consensus: A paradigm for model fitting with applications to image analysis and automated cartography," *Communications of the ACM*, vol. 24, no. 6, pp. 381–395, 1981.

[8] A. Geiger, P. Lenz, C. Stiller, and R. Urtasun, "Vision meets robotics: The KITTI dataset," *International Journal of Robotics Research (IJRR)*, to appear.

[9] A. M. López, D. Lloret, J. Serrat, and J. J. Villanueva, "Multilocal Creaseness Based on the Level-Set Extrinsic Curvature," *Computer Vision and Image Understanding*, vol. 77, no. 2, pp. 111–144, 2000.

[10] A. Guiducci, "Parametric model of the perspective projection of a road with applications to lane keeping and 3D road reconstruction," *Computer Vision and Image Understanding*, vol. 73, no. 3, pp. 414–427, Mar. 1999. [Online]. Available: http://dx.doi.org/10.1006/cviu.1998.0737

[11] E. Dickmanns and B. Mysliwetz, "Recursive 3-D road and relative ego-state recognition," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 14, no. 2, pp. 199–213, 1992.

[12] Z. Zhang, "A flexible new technique for camera calibration," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 22, no. 11, pp. 1330–1334, Nov. 2000. [Online]. Available: http://dx.doi.org/10.1109/34.888718

[13] R. I. Hartley and A. Zisserman, *Multiple View Geometry in Computer Vision*, 2nd ed. Cambridge University Press, 2004.