2014 IEEE 17th International Conference on
Intelligent Transportation Systems (ITSC)
October 8-11, 2014. Qingdao, China

Road curb and lanes detection for autonomous driving on urban scenarios

C. Fernández, R. Izquierdo, D. F. Llorca, M. A. Sotelo

Abstract— **This paper addresses a framework for road curb and lanes detection in the context of urban autonomous driving, with particular emphasis on unmarked roads. Based on a 3D point cloud, the 3D parameters of several curb models are computed using curvature features and Conditional Random Fields (CRF). Information regarding obstacles is also computed based on the 3D point cloud, including vehicles and urban elements such as lampposts, fences, walls, etc. In addition, a gray-scale image provides the input for computing lane markings whenever they are present and visible in the scene. A high level decision-making system yields accurate information regarding the number and location of drivable lanes, based on curbs, lane markings, and obstacles. Our algorithm can deal with curbs of different curvature and heights, from as low as 3 cm, in a range up to 20 m. The system has been successfully tested on images from the KITTI data-set in real traffic conditions, containing different number of lanes, marked and unmarked roads, as well as curbs of quite different height. Although preliminary results are promising, further research is needed in order to deal with intersection scenes where no curbs are present and lane markings are absent or misleading.**

I. INTRODUCTION AND RELATED WORK

Road detection and tracking has traditionally been an exhaustive topic of research in the fields of Advanced Driver Assistance Systems (ADAS) and Autonomous Driving. On the one hand, ADAS have mainly focused on increasing the safety of drivers and road users by means of warnings to drivers and assisted interventions. On the other hand, it is undebatable that autonomous driving has become a high priority issue on the research and commercial agendas of major car makers in the latest years, aiming at producing fully autonomous vehicles by 2020. The deployment of autonomous cars will bring a number of clear benefits in terms of increased traffic efficiency and reduced accident toll, deriving in unquestionable higher energy efficiency and enhanced road safety. For such purpose, available technologies, such as vision sensors, have been exhaustively used by the automotive industry in their mass-produced units. As an example, the use of cameras has made possible the development of lane departure warning and lane keeping systems. These are largely consolidated ADAS functions that can be easily found in a good deal of today's cars. However, the use of these functions is frequently limited to highways and roads with clearly visible lane markers. Safe operation on roads where lane markings are poorly visible or where no lane markers exist at all remains still an open challenge for researchers. Urban scenarios are particularly challenging in this aspect given the large variety of configurations and different situations that can be encountered, such as illumination conditions strongly affected by surrounding vehicles and buildings, interaction with other vehicles and vulnerable road users, etc. All these factors cause a great impact on the accuracy and reliability of vision-based systems.

In order to overcome these difficulties, research efforts must be oriented to developing algorithmic solutions for the reliable detection of road edges and curbs. A review of related literature reveals that color and texture are potential features to characterize the road. In [1] [2], the hue-saturation-intensity (HSI) color space is used together with road shape restrictions and ad hoc post processing in order to recover undetected shadowed areas in a network of unmarked roads. Other authors [3] use the red-green-blue (RGB) color space and put more emphasis on the classification process using a mixture of Gaussians (MoG). In [4] a shadow-invariant feature space is combined with a model-based classifier in an attempt to develop a system robust to shadows. For such purpose, the authors propose to use the illuminant invariant image introduced in [5]. A hierarchical two-stage approach for learning the spatial layout of road scenes is developed in [6]. In the first stage, base classifiers analyze the local visual properties of patches extracted from monocular camera images and provide metric confidence maps. In a second stage, the so-called SPatial RAY (SPRAY) features are computed from each metric confidence map. The ego-lane is extracted afterwards following a semantic segmentation approach. In other cases [7], vision is used for road boundary detection but it is further enhanced by radar in order to detect static barriers along the road side such as guardrails. The use of stereo-vision for recovering the road structure is generally more robust than monocular-based solutions, at least in the presence of road curbs. Accordingly, a stereo-based homography associated with a semantic graph is proposed in [8], where the Viterbi algorithm is used to find the most likely road edges. As an extension of the previous work [9], the authors propose to solve the homography as a maximum a posteriori (MAP) problem in a Markov Random Field (MRF) that alternates between computing the binary labels for the road/non road regions and learning the optimal parameters for the probabilistic algorithm. In [10] a 2D/3D approach is applied to road detection in urban environments. The 2D layer provides a number of pixel's clusters based on the Watershed transform while the 3D layer classifies the clusters using the V-Disparity technique. A remarkable approach for classifying the different road

C. Fernández, R. Izquierdo, D. F. Llorca, and M. A. Sotelo are with the Computer Engineering Department, Polytechnic School, University of Alcalá, Madrid, Spain. email: carlos.fernandez, llorca, sotelo@aut.uah.es.

surfaces and obstacles in urban environments is proposed in [11] using digital elevation maps (DEM). Two classifiers are used for obstacles and road classification, based on density and road surface, respectively. In [12], the authors propose a generic framework for curb detection and reconstruction using Conditional Random Fields (CRF). Points extracted from the 3D point cloud are assigned to curb adjacent surfaces, i.e. street and sidewalk using Loopy Belief Propagation. As a result, curbs can be reconstructed even of low height up to a distance of 20 meters. The proposed system was further enhanced in [13] by including a temporal filter that improves the robustness and accuracy, particularly in case of low-height curbs and missing measurements.

In this paper we propose a framework for road curb and lanes detection in the context of urban autonomous driving, with particular emphasis on unmarked roads. Based on a 3D point cloud, curbs are computed using curvature features and Condition Random Fields (CRF). Information regarding obstacles is also computed based on the 3D point cloud, including vehicles and urban elements such as lampposts, fences, walls, etc. In addition, a gray-scale image provides the input for computing lane markings whenever they are present and visible in the scene. A high level decision-making system yields accurate information regarding the number and location of drivable lanes, based on curbs, lane markings, and obstacles. Our algorithm can deal with curbs of different curvature and heights, from as low as 3 cm, in a range up to 20 m. The rest of the paper is structured as follows: section II presents a general description of the system, including the methods for curb reconstruction, obstacle detection, road marking extraction, free space estimation, curb height measurement, and number of lanes computation. Results and discussion are presented in section III. Finally, we analyze our conclusions and future work in section IV.

II. SYSTEM DESCRIPTION

A. Sensors Description

All input data to our system is provided by the KITTI Vision Benchmark Suite [14]. This public database contains road scenes information obtained with different types of sensors: INS (OXTS RT 3003), LIDAR (Velodyne HDL 64E), 2 grayscale and 2 color cameras 1.4 Mpx (Point Grey Flea 2) and 4 Varifocal lenses 4-8 mm (Edmund Optics NT59-917). LIDARs are very useful because they provide low noise measurements at long range distance. However, the goal of this paper is to understand the road scene by using only stereo vision.

B. General Description

As mentioned in section I, precise understanding of urban scenarios is crucial for autonomous driving. In the proposed method, we estimate the number of drivable lanes as well as its location w.r.t. the ego-vehicle. As shown in Figure 1, the free space or drivable area is computed based on the outputs provided by three independent modules in charge of detecting road markings, curbs, and obstacles, respectively. In addition, the curb detection module also provides an estimate of the curb height.

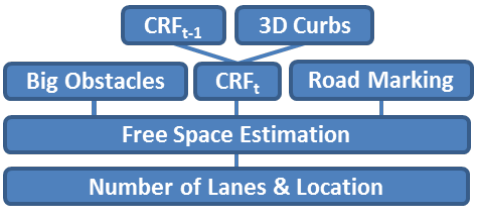


Fig. 1. Functional relationship between the different modules

The proposed curb and obstacle detection methods are based on surface curvatures and surface normals. In a first step, a 3D point cloud is obtained using the well known Semi-Global Matching (SGM) method [15]. This method is based on the idea of pixelwise matching of Mutual Information and approximation of a global, 2D smoothness constraint by combining many 1D constraints. The camera axes is translated to the ground so that axis are renamed with respect to the world coordinate axes as follows:

$$\begin{aligned} x_{world} &= z_{cam} \\ y_{world} &= -x_{cam} \\ z_{world} &= -y_{cam} + 1.65 \end{aligned} \tag{1}$$

C. Curb Detection Method

The proposed curb detection method is based on surface curvature estimation [16]. For each point p, we select the nearest neighbors (NN) p_i in a surrounding area defined by a radius of 0.25 meters. These points are used to create a weighted covariance matrix, where k denotes the number of NN.

$$\bar{p} = \frac{1}{k}\sum_{i=1}^{k} p_i \; ; \; \mu = \frac{1}{k}\sum_{i=1}^{k} |\bar{p} - p_i| \tag{2}$$

$$w_i = exp\left(-\frac{(p-p_i)^2}{\mu^2}\right) \tag{3}$$

$$C = \sum_{i=1}^{k} w_i \cdot (p_i - \bar{p})^T \cdot (p_i - \bar{p}) \tag{4}$$

The eigenvector V and eigenvalue λ are computed as $C \cdot V = \lambda \cdot V$. Curvature γ_p is defined by equation 5, where $\lambda_0 \leq \lambda_1 \leq \lambda_2$ are the eigenvalues of the covariance matrix C.

$$\gamma_p = \frac{\lambda_0}{\lambda_0 + \lambda_1 + \lambda_2} \tag{5}$$

Only curvature in the Z axis (in the 3D world coordinates) is taken into account given that it provides a very discriminant description of road curvature changes. Curvature values vary between 0 and 1. After thorough observation of urban scenes in the KITTI dataset, road curbs can be clustered into three groups, see Table I:

Curb curvature is different in each scene. For example, if the curb is a regular one, most of the points exhibit curvature

TABLE I
CURB CURVATURE VALUES

DESCRIPTION	CURVATURE	COLOR
Very Small Curbs (∼3 cm)	0.12 … 0.17	yellow
Small Curbs (∼5 cm)	0.17 … 0.22	orange
Regular Curbs (∼10 cm)	0.22 … 0.50	red
Big Obstacles	0.50 … 1.0	purple

values between 0.22 and 0.50, but there are also some curb points yielding significantly different values. These measurement outliers are removed by means of a filtering process. A binary mask is processed for each curvature range using morphological operations and contour analysis. The resulting masks are merged and refiltered in order to get an image like the one shown in Figure 2. The use of fixed or empirical thresholds is then avoided given that the proposed function is adapted automatically for different scenes. The result obtained is used as an input to a Conditional Random Field (CRF) in order to enhance the result. Ordinary classifiers predict a label for a single sample without taking into account its neighbors. Contrary to that, the use of a 4-connected neighborhood and efficient loopy belief propagation in a CRF [17] [18] enhances significantly the quality of the curb detection stage. The CRF has 2 labels (curb/not curb). At time t, the prior is considered as the result of the CRF at time $t-1$, while the likelihood is modeled from the output of the curb detection function. An example of the process is shown in Figure 2.

(a) Input image

(b) Curvature values

(c) Curb detection result filtered with the CRF

Fig. 2. Curvatures are computed and drawn in different colors on the input image. Pixels detected as curbs are depicted in blue in 2(c).

D. Obstacles detection

There are many types of obstacles in urban environments: vehicles, curbs, traffic lights, traffic signs, walls, etc. An specialized algorithm has been described in Section II-C for road curb detection. Other obstacles are labeled as "big obstacles" because their height is significantly higher than that of curbs. Accordingly, all these obstacles are treated using the same method. For each point p in the 3D point cloud, the nearest neighbors are computed and the normal vectors are estimated by fitting the points to a plane.

In the world reference frame, a point p is considered an obstacle if $0.50 \leq \alpha \leq 1.0$, where $\alpha = \{n_x, n_y, c_z\}$ and n_x, n_y are the components of the normal vector and c_z is the curvature in Z axes. The precision of the method is conditioned by the quality of the stereo sensor. In some occasions, the quality of the disparity image is not good enough in general, as can be observed in the reconstructed 3D cloud, but in areas with correct disparity values the results are acceptable for obstacle detection. See Figure 3.

Fig. 3. Points labeled as obstacles using normal vectors

E. Road Marking Detection

The proposed road marking detection method is based on state of the art techniques. However, we provide a brief description for completeness purpose. As explained in [19], a median filter is applied to the input image. The window size of the median filter needs to be twice larger than the road marking. If the road marking is larger than the window, for example in a zebra crossing, the border is well detected but the areas inside the zebra crossing are not. In order to keep the window size constant, a bird-eye view of the scene is reconstructed. An adaptable threshold is then applied to the input image. After that, both images are subtracted and the final result is filtered based on area features, rectangularity and 3D height.

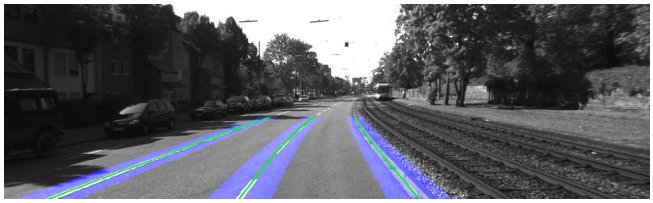
A temporal filter is then applied using an EKF and a clothoid model for the road lines [20]. A region of interest of 0.5 m around the clothoid model has been used for validation in the tracking stage. The state vector of the clothoid model is composed of $\begin{bmatrix} C_0 & C_1 & x_0 & \psi & w \end{bmatrix}^T$, where C_0 is the initial curvature, C_1 is the curvature variation, x_0 is the lateral displacement with respect to the world coordinate system, ψ is the angular displacement with respect to the lane orientation, and w is the lane width.

F. Free Space Estimation

The drivable road area is obtained based on the outputs of the curb, obstacles, and road marking detection modules.

(a) Input Image

(b) Unfiltered result



(c) Road marking detected

Fig. 4. In 4(b) the unfiltered result is shown for the method described above. In 4(c), the area of interest is displayed in transparent blue and clothoid models are shown in solid green.

Depending on the scene type, road limits can be defined by curbs, by a curb and a parked car, or by multiple combinations of them. In most cases, the free space is limited by the first obstacle on the left and the first obstacle on the right. In some cases, there are not any obstacles on the road, and the road limits are defined by road markings.

Besides, a curb mask (Figure 2(c)) is used to find candidate lines in the point cloud using RANSAC. In addition to those lines, the same method is applied to the obstacles mask (Figure 3). The algorithm is applied to the point cloud so that the result is a line in 3D. A line L is defined by a 3D point p and a direction vector $\vec{d}$. Using the line with the largest number of inliers as the main line L_0, other parallel lines to L_0 are computed. The k lines obtained L_k are divided into two groups depending on their position, left (L_k^l) or right (L_k^r). The lines obtained from curbs and obstacles are merged so that the road limits satisfy equations 6 and 7 in world coordinates.

$$road_l = min_y(p_k^l) \tag{6}$$

$$road_r = max_y(p_k^r) \tag{7}$$

If lines from obstacles or curbs are not detected, the lines obtained in Section II-E are used as input to the free space algorithm. In Figure 5, lines detected from curb points and obstacle points are displayed in blue and red respectively.

Fig. 5. Lines detected using RANSAC in the point cloud. The yellow line indicates the direction for other parallel lines, the blue ones are curbs candidates, and the red ones represent bigger obstacles.

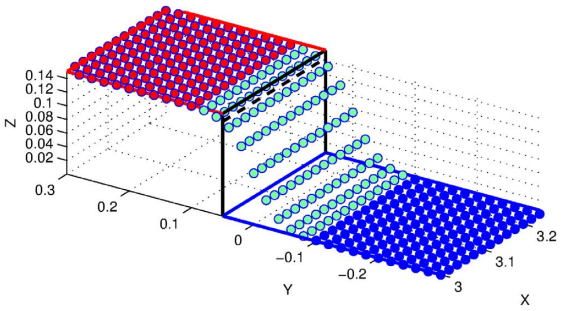
G. Curb Height Estimation

The curb height is calculated using the lines extracted from curb points in Section II-F. Only lines labeled as road limits are taken into account. These lines are divided in slots of 0.3 m. The NN of the middle point p of each slot are computed and separated into two groups (NN_l, NN_r) using the vertical plane Π_v defined by the line. For each group, the estimated planes (Π_l, Π_r) are intersected with Π_v, yielding two different lines (r_l, r_r). The point height is obtained by evaluating both lines in the p_x and p_y coordinates.

$$r_l = \Pi_l \cap \Pi_v \tag{8}$$

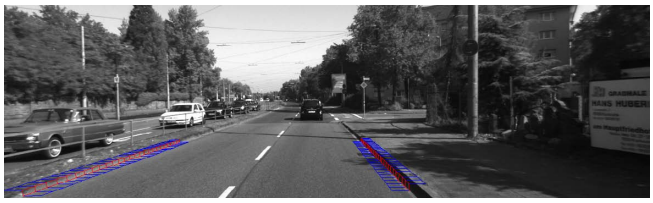
$$r_r = \Pi_r \cap \Pi_v \tag{9}$$

Road curb location is accurately detected in most cases. However, curb height estimation is more sensitive to mismatching errors in the disparity image. In general, height is constant along the curb. Therefore height variations between slots are smoothed out by a filtering process. Curb points with height ranging in $\left[\bar{h}-2\sigma, \bar{h}+2\sigma\right]$ are considered inliers. The curb height is computed as the mean value of all inliers.



(a) Planes estimation

(b) Unfiltered result



(c) Filtered result

Fig. 6. Curb height estimation steps.

H. Estimation of the Number of Lanes

A neural network-based module is trained to provide the probability that the road has 1, 2 or 3 lanes based on road

curbs, road markings and obstacles. In the training stage, a total of 138 different road scenes have been processed manually to label the road width and the number of lanes. The neural network provides a hit rate of 93.47%. The values to separate the free space into lanes are the following: if the free space is less than 4.06 meters, it is considered as a one lane road. If the road width is between 4.06 and 8.57 it is regarded as a two lanes road. For values greater than 8.57 meters it is considered as a three lanes road.

III. RESULTS

Figure 7, shows some representative results in different urban scenarios. The first row shows a road in which the drivable area is restricted by lane markings and regular curbs. The road is affected by shadows and is surrounded by grass. The grass surface is irregular and is not parallel to the road plane. Therefore, accurate curb height estimation becomes a challenge. The free space is limited by the left and right curbs and the number of lanes is detected using road markings. Detection of small curbs is complex given that, in some occasions, the measurement noise is even greater than the curb height. Despite these difficulties, our method is able to detect very small curbs (3 cm), as shown in the second row. The third row shows a very complex scene: the road is limited by a small curb on the left and a regular curb on the right. In addition, a car is parked on the drivable area and there is not any road marking. Although the scene is affected by shadows, the contrast is good enough to obtain valid disparity values. The fourth row shows a scene with an intersection on the right. As a consequence, there is not any road curb limiting the free space. Instead, the road limits are given by the road markings. Road markings are also used to estimate the number of lanes. On the left hand side, the road has a regular curb. Finally, the fifth row shows an urban street made of bricks and affected by shadows. The road is limited by regular curbs on the left and right sides but only the right one is properly detected because the shadows decrease the quality of the disparity image in a significant way. In addition, the street has no road markings. Therefore the number of lanes is estimated using the free space width. In this case, the system successfully estimates two lanes in the scene, although the accuracy of the curb detector should be further improved. For results evaluation, 38 curbs have been manually labelled in different scenes in order to be compared against the automatic values obtained from our method. The final estimated curb height in our experiments exhibits a RMSE of 0.014 meters.

The method described in [12] fits the measurements to a sigmoid function in order to estimate both the curb position and the curb height. Our method proposes a solution from a different viewpoint. We accomplish curb detection using curvature values in a first step. After that, surface reconstruction is applied only for height computation. In [21] and [13] the authors use temporal integration of curb measurements using egomotion data to improve the estimation, reaching a detection range of 20 meters. By using the method proposed in this paper, we can detect curbs up to 20 meters without temporal integration. However, the use of temporal integration is proposed as part of our future work in order to further enhance the robustness of the curb detection module in cluttered environments.

IV. CONCLUSIONS AND FUTURE WORKS

We have developed a system for curb detection and free space estimation for autonomous driving in urban scenarios. The curb detection system is able to detect a wide range of curbs with varying heights from 2-3 cm up to 10 cm or higher in a range of up to 20 meters. The accuracy of the method is very much dependent on the quality of the disparity images. In road scenes exhibiting dense disparity image, road curb detection is robust and curb height is estimated accurately. Curb height is estimated separating the curb points into two groups: sidewalk and road. The height difference between both sides define the curb height. Obstacles like cars, traffic lights, pedestrians or cyclists are detected using surface normal vectors estimation. Furthermore, the road marking detection system used in this work is based on state of the art techniques. Lines are detected are filtered using an EKF and a clothoid model. The free space computation module estimates the road width and the number of drivable lanes based on inputs received from the curb, obstacles and road marking detection methods.

The current method for curb segmentation is based on lines fitting using RANSAC. Therefore, although the method is valid for straight roads, which is the usual case, a more advanced method that takes into account corners and lateral curb curvature will be developed as part of our future work. Cutting-edge matching algorithms will be tested in order to improve the quality of the 3D reconstructed point cloud. The presence of shadows on the road surface affects the performance of the road marking and road curb detection methods. Accordingly, a vision-based method to remove shadows and obtain an homogeneous road surface will be developed.

V. ACKNOWLEDGMENTS

This work is supported by the Spanish Ministry of Education under Research Grant ONDA-FP TRA2011-27712-C02-02.

REFERENCES

[1] M. A. Sotelo, F. J. Rodriguez, and L. Magdalena, "Virtuous: vision-based road transportation for unmanned operation on urban-like scenarios," *IEEE Transactions on Intelligent Transportation Systems*, vol. 5, no. 2, pp. 69–83, June 2004.

[2] M. A. Sotelo, F. J. Rodriguez, L. Magdalena, L. M. Bergasa, and L. Boquete, "A color vision-based lane tracking system for autonomous driving on unmarked roads," *Autonomous Robots*, vol. 16, no. 1, pp. 95–116, 2004.

[3] H. Dahlkamp, A. Kaehler, D. Stavens, S. Thrun, and G. Bradski, "Self-supervised monocular road detection in desert terrain," in *Proceedings of Robotics: Science and Systems*, Philadelphia, USA, August 2006.

[4] J. Alvarez and A. Lopez, "Road detection based on illuminant invariance," *IEEE Transactions on Intelligent Transportation Systems*, vol. 12, no. 1, pp. 184–193, March 2011.

[5] G. Finlayson, S. Hordley, C. Lu, and M. Drew, "On the removal of shadows from images," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 28, no. 1, pp. 59–68, Jan 2006.

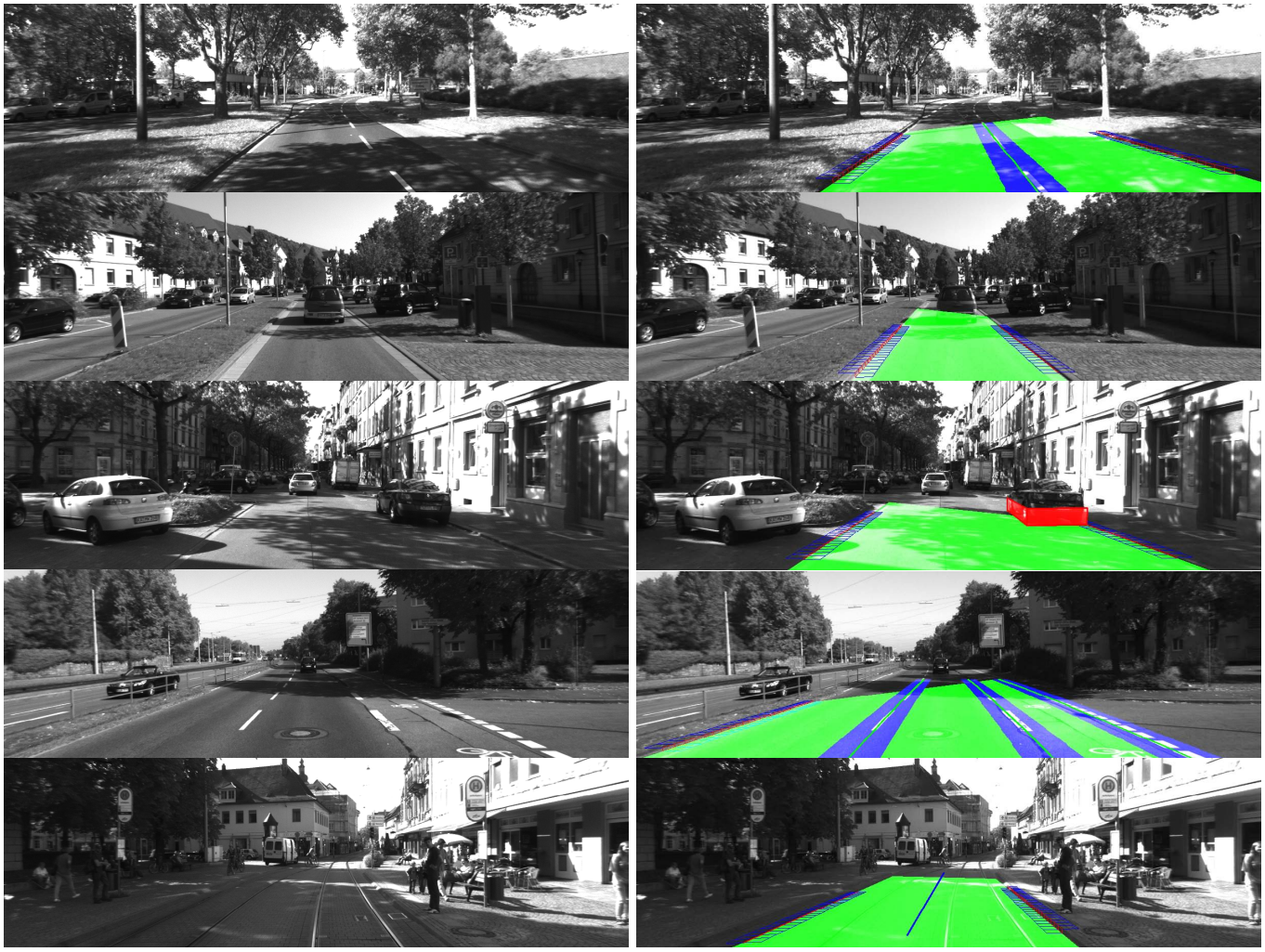


Fig. 7. Final results in different urban scenarios.

[6] J. Fritsch, T. Kuhnl, and F. Kummert, "Monocular road terrain detection by combining visual and spatial information," *IEEE Transactions on Intelligent Transportation Systems*, vol. PP, no. 99, pp. 1–11, 2014.

[7] F. Janda, S. Pangerl, E. Lang, and E. Fuchs, "Road boundary detection for run-off road prevention based on the fusion of video and radar," in *IEEE Intelligent Vehicles Symposium (IV)*, June 2013, pp. 1173–1178.

[8] C. Guo, T. Yamabe, and S. Mita, "Robust road boundary estimation for intelligent vehicles in challenging scenarios based on a semantic graph," in *IEEE Intelligent Vehicles Symposium (IV)*, June 2012, pp. 37–44.

[9] C. Guo, S. Mita, and D. McAllester, "Robust road detection and tracking in challenging scenarios based on markov random fields with unsupervised learning," *IEEE Transactions on Intelligent Transportation Systems*, vol. 13, no. 3, pp. 1338–1354, Sept 2012.

[10] G. Vitor, D. Lima, A. Victorino, and J. Ferreira, "A 2d/3d vision based approach applied to road detection in urban environments," in *IEEE Intelligent Vehicles Symposium (IV)*, June 2013, pp. 952–957.

[11] F. Oniga and S. Nedevschi, "Processing dense stereo data using elevation maps: Road surface, traffic isle, and obstacle detection," *IEEE Transactions on Vehicular Technology*, vol. 59, no. 3, pp. 1172–1182, March 2010.

[12] J. Siegemund, D. Pfeiffer, U. Franke, and W. Frstner, "Curb reconstruction using conditional random fields," in *IEEE Intelligent Vehicles Symposium (IV)*, June 2010, pp. 203–210.

[13] J. Siegemund, U. Franke, and W. Frstner, "A temporal filter approach for detection and reconstruction of curbs and road surfaces based on conditional random fields," in *IEEE Intelligent Vehicles Symposium (IV)*, June 2011, pp. 637–642.

[14] A. Geiger, P. Lenz, C. Stiller, and R. Urtasun, "Vision meets robotics: The kitti dataset," *International Journal of Robotics Research (IJRR)*, 2013.

[15] H. Hirschmuller, "Accurate and efficient stereo processing by semi-global matching and mutual information," in *Proceedings of the 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05)*. Washington, DC, USA: IEEE Computer Society, 2005, pp. 807–814.

[16] R. B. Rusu, Z. C. Marton, N. Blodow, M. Dolha, and M. Beetz, "Towards 3d point cloud based object maps for household environments," *Robot. Auton. Syst.*, vol. 56, no. 11, pp. 927–941, Nov. 2008.

[17] P. F. Felzenszwalb and D. P. Huttenlocher, "Efficient belief propagation for early vision," *Int. J. Comput. Vision*, vol. 70, no. 1, pp. 41–54, Oct. 2006.

[18] J. Sun, N.-N. Zheng, and H.-Y. Shum, "Stereo matching using belief propagation," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 25, no. 7, pp. 787–800, July 2003.

[19] P. Foucher, Y. Sebsadji, J. P. Tarel, P. Charbonnier, and P. Nicolle, "Detection and recognition of urban road markings using images," in *14th International IEEE Conference on Intelligent Transportation Systems (ITSC)*, Oct 2011, pp. 1747–1752.

[20] 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, Feb 1992.

[21] F. Oniga and S. Nedevschi, "Curb detection for driving assistance systems: A cubic spline-based approach," in *IEEE Intelligent Vehicles Symposium (IV)*, June 2011, pp. 945–950.