PROCEEDINGS OF THE FIRST SYMPOSIUM SEGVAUTO-TRIES-CM. TECHNOLOGIES FOR A SAFE, ACCESSIBLE AND SUSTAINABLE MOBILITY. R&D+I IN AUTOMOTIVE: RESULTS

17th-18th November 2016, Madrid

EDITORS
Francisco Aparicio Izquierdo
Luis Martínez Sáez
Juan José Herrero Villamor

EDITORIAL
FUNDACIÓN GENERAL DE LA UPM
c/ Pastor 3
28003 Madrid

ISBN:978-84-16397-47-1
Legal Deposit: M-39741-2016

A Comparison of Road Curb Detection

Using Curvature Values from LIDAR and Stereo Vision

Carlos Fernández, Javier Lorenzo, Javier Alonso, Álvaro García, David F. Llorca, Miguel A. Sotelo

University of Alcalá, Spain, carlos.fernandezl@uah.es

Introduction

Road detection has been a topic of the utmost relevance in the fields of Advanced Driver Assistance Systems (ADAS) and autonomous driving. In the last years, the quality and performance of ADAS has improved significantly, providing a crucial contribution for making autonomous driving a reality. Nowadays, some production vehicles are equipped with systems that take the control of the vehicle in order to reduce the damage to the occupants in case of accident.

All these systems require accurate scene understanding capability. For such purpose, vision sensors have been extensively used by the automotive industry in diverse systems. Lane departure warning (LDW) and lane keeping systems are clear examples of that. These types of systems are consolidated in today's cars. However, their use is limited to highways and roads with clearly visible lane markers. Robust operation of LDW systems on unmarked roads and urban environments is still a challenge. In most non-urban roads, lanes are limited by road markings. However, metropolitan areas are more complex and the free space can be restricted by road markings, parked cars, traffic signs, lampposts and curbs of quite different heights. The accuracy and reliability of vision based systems are strongly affected by the large variety of street configurations, different materials and textures, illumination changes, etc. In order to deal with this adversities, research resources must be oriented to developing algorithms for the reliable detection of curbs and road edges, see Figure 1.

This paper focuses on curb detection on urban scenarios. We propose a versatile curvature-based method which can operate using dense 3D point clouds regardless the specific sensor. Thus, our approach has been validated using 3D data captured from stereo vision and LIDAR. Free space is crucial for understanding the road scene, providing relevant information to the path planning module in an autonomous vehicle.

Materials and methods

In this paper we propose an algorithm for online road curb detection based on curvature surface

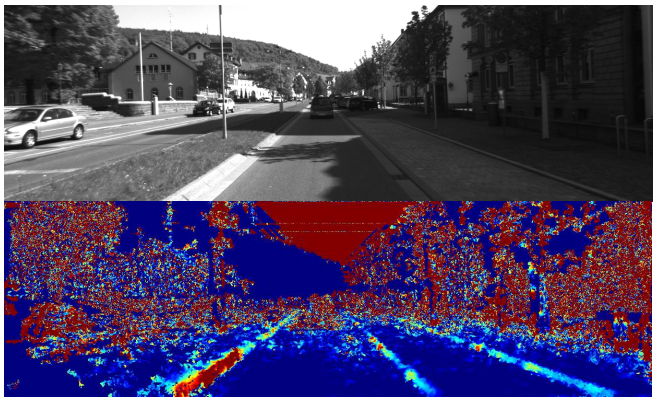


Fig. 1: Top: example of a complex urban scenario. Bottom: the curvature representation. On the left side, the scene has a high curb in the closer distance and a lower one further. On the right side, there are two small curbs, the first one is the boundary for the parking place and the second one is the limit of the sidewalk.

estimation in 3D point clouds. The clouds can be captured from different types of sensors, for example LIDAR or stereo cameras. The point cloud from LIDAR is sparse because of the vertical resolution. In order to increase the cloud density, an iterative algorithm is applied to align the latest scans. The proposed method is included in a more complex system for urban scene understanding with particular emphasis on unmarked roads.

In this paper, a curb detection method based on surface curvature is compared using two different inputs. The first one is the 3D reconstruction point cloud using the Semi Global Matching (SGM) algorithm based on stereo vision [1]. The second one is the 3D reconstruction point cloud obtained from Velodyne.
Since we focus on urban environments, the performance of the system is evaluated using the public dataset: KITTI Vision Benchmark Suite [2]. The dataset provides images and information of urban scenarios from different types of sensors, such as monochrome and color cameras, multilayer LIDAR, GPS and IMU. The height of the cameras is 1.65 meters and the height of the LIDAR is 1.73 meters. Furthermore, the LIDAR is installed 27 cm behind the cameras. Consequently, the relationship between camera and LIDAR poses must be calibrated. The accuracy of the stereo-based 3D point cloud is reasonable high even for long distances due to the long stereo baseline (0.54 m) and the high

resolution of the cameras. On the one hand, road scene reconstruction from stereo is more dense than LIDAR reconstruction but it is also affected by mismatching errors. On the other hand, LIDAR provides low noise measurements even at long range while at large distances the 3D reconstruction is sparse and dependent on the vertical resolution.

Surface Curvature

The proposed curb detection method is based on surface curvature estimation presented in [3]. This feature has been also used in [4] for free space detection and for curb detection in [5]. The curvature describes the variation along the surface normal and it varies between 0 and 1, where low values correspond to flat surfaces. The curvature feature is more robust and stable than tangent plane normal vectors. For each point *p*, the nearest neighbors (NN) p_i in a surrounding area defined by a radius *R* are selected. These points are used to create a weighted covariance matrix, where *k* denotes the number of *NN*.

$$\bar{p} = \frac{1}{k}\sum_{i=1}^{k} p_i \quad (1)$$

$$\mu = \frac{1}{k}\sum_{i=1}^{k} |p - p_i| \quad (2)$$

$$w_i = \begin{cases} exp\left(-\frac{(p-p_i)^2}{\mu^2}\right) & if\ |p-p_i| \geq \mu \\ 1 & otherwise \end{cases} \quad (3)$$

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

The eigenvector *V* and eigenvalues λ of *C* are computed as *C*·*V*= λ·*V*. A curvature measure γ_z^p is defined by the following equation, where $\lambda_x \leq \lambda_y \leq \lambda_z$ are the eigenvalues of the covariance matrix *C*.

$$\gamma_z^p = \frac{\lambda_z}{\lambda_x + \lambda_y + \lambda_z} \quad (5)$$

Curb Detection

In our reference system, the *Z* axis is orthogonal to the road, therefore the curvature γ_z provides a discriminative descriptor of the road shape. Curb height and curvature γ_z are highly correlated. Consequently, after thorough observation of urban scenes in the KITTI dataset, a set of thresholds $\alpha_i = \{1 \dots N\}$ is used to label the type of curb, see Table 1:

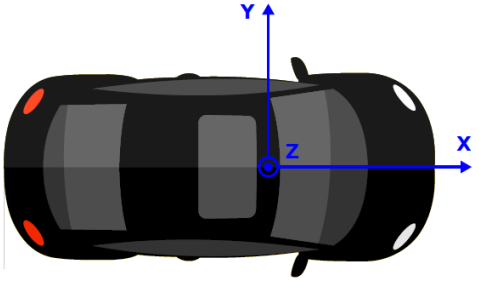
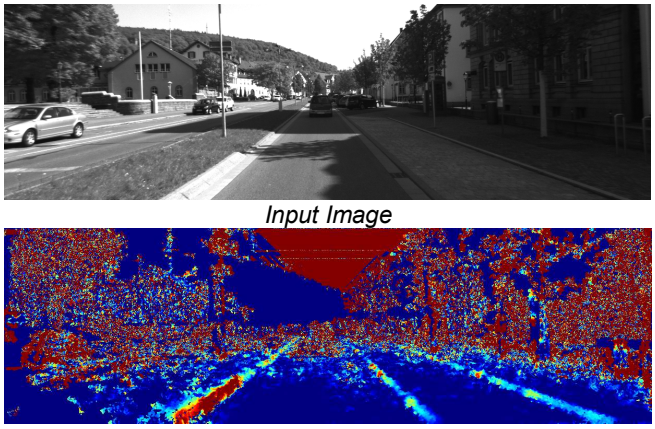


Fig. 2: Reference system of the vehicle.

Table 1: Curb curvature values

DESCRIPTION	CURVATURE	COLOR
Flat Surface	$0 \leq \gamma_z < \alpha_0$	Not Painted
Very Small Curbs (~3 cm)	$\alpha_0 \leq \gamma_z < \alpha_1$	Yellow
Small Curbs (~5 cm)	$\alpha_1 \leq \gamma_z < \alpha_2$	Orange
Regular Curbs (~10 cm)	$\alpha_2 \leq \gamma_z < \alpha_3$	Red
Big Obstacles (Walls, cars)	$\alpha_3 \leq \gamma_z < \alpha_4$	Purple

Curb curvature is different in each scene. For example, if the curb is a regular one, most of the points exhibit curvature values $\gamma_z \epsilon [\alpha_2, \alpha_3)$, but there are also some curb points yielding significantly different values, see Figure 3.

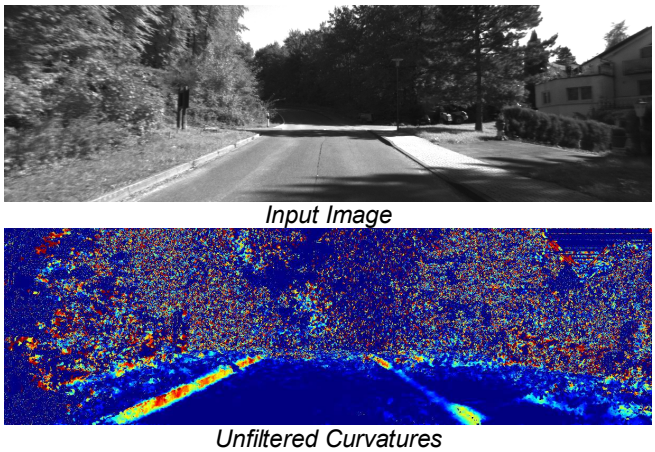


Input Image

Fig. 3: Curvature values codified in a color scale, where cold colors correspond with flat surfaces.

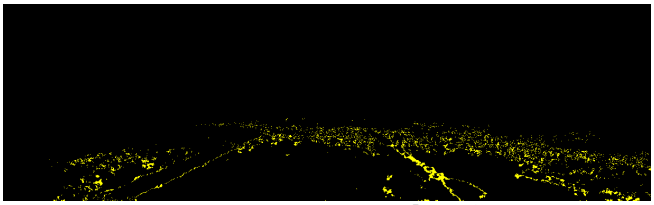
These measurement outliers are removed by means of a filtering process. A binary mask is applied for each curvature range using morphological operations and contour analysis.

The resulting masks are merged and re-filtered in order to get an image like the one shown in Figure 4. The use of fixed or empirical thresholds is then avoided given that the proposed function is adapted automatically for different scenes depending on the predominant curvature value.

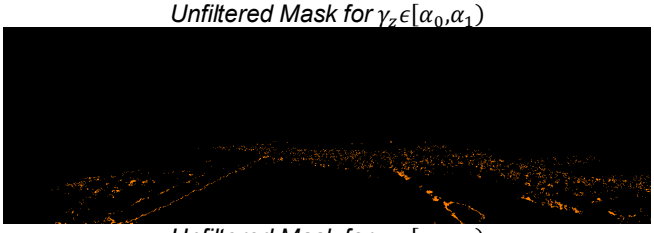


Input Image

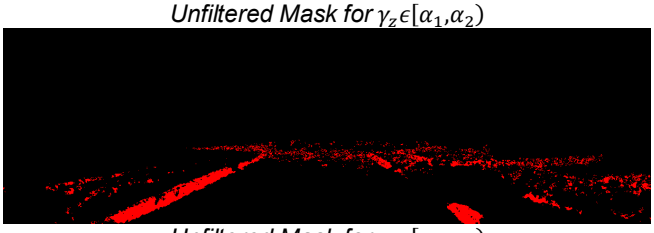
Unfiltered Curvatures



Unfiltered Mask for $\gamma_z \epsilon [\alpha_0, \alpha_1)$



Unfiltered Mask for $\gamma_z \epsilon [\alpha_1, \alpha_2)$



Unfiltered Mask for $\gamma_z \epsilon [\alpha_2, \alpha_3)$

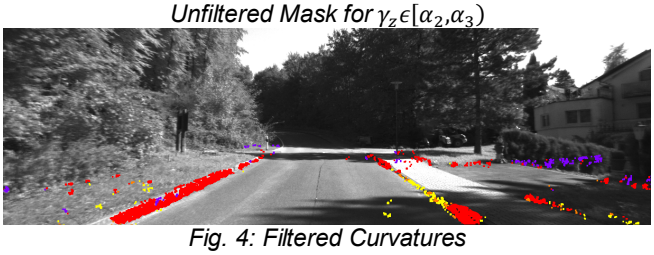


Fig. 4: Filtered Curvatures

Increase Cloud Density

The point cloud obtained from the stereo sensor is acceptably dense, while the LIDAR cloud is sparse given that the vertical resolution is limited to 0.4 degrees. As a consequence of that, the algorithm described in the previous section cannot be applied directly. The cloud requires being dense enough. For that purpose, for every point the nearest neighbors (NN) are fitted to a surface. Then, new points are generated by drawing samples from the estimated surface. If the road surface is estimated using a polynomial, the resulting shape of small curbs is smoothed up and, consequently, those curbs are partially removed. In order to avoid this problem, an iterative algorithm is applied to align the sparse clouds from instant *t-n* to *t*. As a result of that alignment, a dense cloud is created. This algorithm is denoted as Iterative Closest Point (ICP) [6,7].

ICP minimizes the difference between two point clouds iteratively. The first one *A* is kept fixed and the other one *B* is warped to match the reference. If the clouds are close enough to each other the initial values for *T* can be set to identity. For each iteration, every point in *B* is transformed using the current transformation matrix *T* and matched with the corresponding point in *A*. If the distance between the points is greater than d_{max}, the points are rejected. As depicted in Figure 5, the ICP algorithm does not smooth the original data and small curbs can still be detected using the algorithm explained in the previous section.

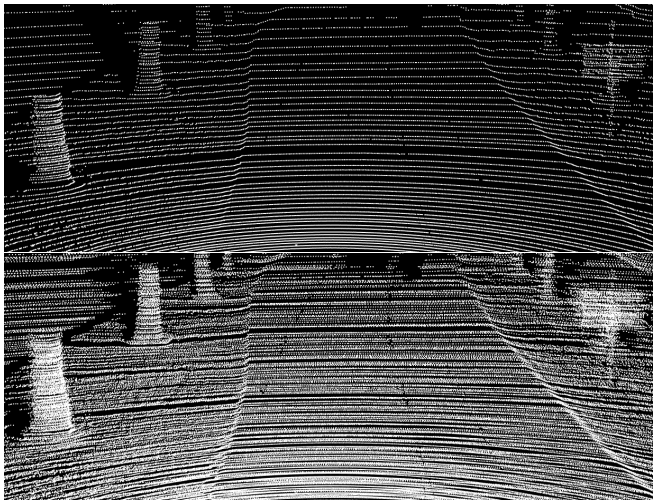


Fig. 5: Top: Single LIDAR scan. Bottom: Point cloud obtained with the Iterative Closest Point algorithm using the last 5 scans of the LIDAR.

Results and discussion

The lateral accuracy was evaluated in road sequences from the KITTI dataset. In such dataset no ground-truth is provided for curbs detection. A contribution of this paper is the creation of a public dataset containing a number of sequences with manually labeled curbs. The ground-truth was manually annotated in the stereo image and also in the LIDAR data. The number of labeled sequences in our dataset is still low, although it will be gradually increased.

For the evaluation, straight lines are fitted to the detected curb points and also to the labeled ground-truth points. The estimated line and the ground truth are defined by the points *p* and *q* and the direction vectors $\vec{u}$ and $\vec{v}$ respectively.

The accuracy is evaluated by computing the equation 8, where *a* and *b* are the evaluated range, from 6 meters up to 20 meters. In urban scenarios the driving speed is usually low. Thus, we consider that an estimation range of 20 meters is enough for safe maneuvering in most cases.

$$l = \frac{u_y}{u_x}(x - p_x) + p_y \quad (6)$$

$$\hat{l} = \frac{v_y}{v_x}(x - q_x) + q_y \quad (7)$$

$$RMSE = \sqrt{\frac{1}{b-a}\int_a^b (l - \hat{l})^2 dx} \quad (8)$$

In the case of the stereo data, the disparity images are computed using Semi Global Matching (SGM). By accumulating the latest 5 scans of the LIDAR, the density of the point cloud is enhanced, making it good enough for the curvature estimation algorithm. The density of the LIDAR cloud is increased using ICP, therefore some errors during the matching stage can yield point clouds with slightly worse quality than a dense stereo point cloud.

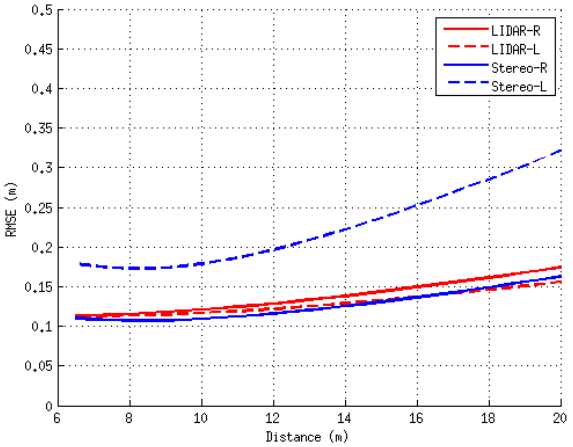
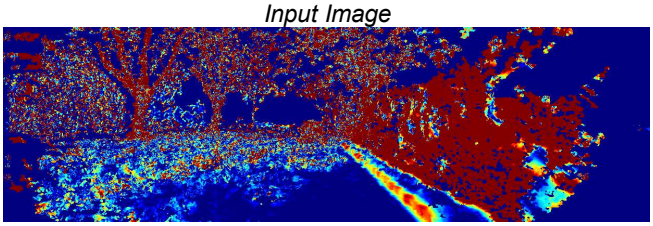


Fig. 6: Lateral RMSE obtained comparing ground truth and the result of the algorithm. The results for left and right curbs are separated to evaluate the performance individually.

In Figure 6, the lateral error is depicted. The LIDAR error is plotted in red and the error from the stereo is plotted in blue. The performance of our approach using the LIDAR data is quite similar both for the left and right curbs. In addition, the performance using the 3D data obtained from stereo vision is also similar to LIDAR for the right most curb. However, the lateral error on the left curb is significantly higher.

Input Image



Curvatures From Stereo

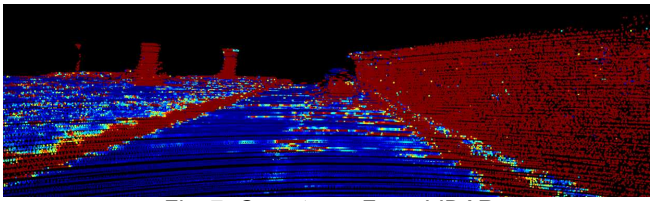


Fig. 7: Curvatures From LIDAR

By analyzing the disparity images in several sequences, we noted that the left part of the image provides noisy and unrealistic 3D information (see Figure 7, which can be caused by several factors (e.g., inaccurate calibration). As shown in Figure 6, when accurate disparity is available (right curb) the performance of the curvature estimation based on stereo vision is even better than the LIDAR because stereo-based point cloud is more dense than that obtained from the LIDAR. Therefore, we can state that the use of stereo cameras for ADAS in urban environments is a good-enough technology for accurate road curb detection. The curb detection algorithm presents a lateral RMSE of 12cm in a range from 6 meters to 20 meters for the right most curb, while the value of RMSE on the left side is 22cm.

Conclusion

In this paper a curb detection estimation algorithm based on 3D point clouds has been presented. The use of fixed or empirical thresholds is avoided given that the proposed function is adapted automatically for different road scenes depending on the predominant curvature value. The algorithm presented in this paper has successfully been applied for the detection of straight and curved curbs. However, a geometric detection method is not enough to get a robust free space detection system due to the fact that some road limits have the same height as the road. For this reason, texture and color provide key information that complements and enhance the geometric reconstruction. As a future work, a new machine learning approach will include the geometric reconstruction presented in this paper together with texture and color information in order to detect the drivable area.

The proposed algorithm is compared using different sources for the 3D cloud data. Although the LIDAR cloud is more accurate than the stereo cloud, it is sparser. In order to improve the LIDAR cloud density, the Iterative Closest Point (ICP) has been applied. The density of the resulting cloud is similar to that of stereo vision systems, while preserving the accuracy of LIDAR sensors.

Acknowledgements

This work was supported by the Regional Government of Madrid under research grant SEGVAUTO-TRIES (S2013/MIT-2713) and by the General Traffic Division of Spain under research grant CUAHC SPIP2015-01774.

References

[1] H. Hirschmuller, “Accurate and efficient stereo processing by semiglobal 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.

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

[3] M. Pauly, M. Gross, and L. Kobbelt, “Efficient simplification of pointsampled surfaces,” in Visualization, 2002. VIS 2002. IEEE, Nov 2002,pp. 163–170.

[4] C. Fernandez, R. Izquierdo, D. Llorca, and M. Sotelo, “Road curb and lanes detection for autonomous driving on urban scenarios,” in Intelligent Transportation Systems (ITSC), 2014 IEEE 17th International Conference on, Oct 2014, pp. 1964–1969.

[5] A. Hata, F. Osorio, and D. Wolf, “Robust curb detection and vehicle localization in urban environments,” in Intelligent Vehicles Symposium Proceedings, 2014 IEEE, June 2014, pp. 1257–1262.

[6] S. Rusinkiewicz and M. Levoy, “Efficient variants of the icp algorithm,” in 3-D Digital Imaging and Modeling, 2001. Proceedings. Third International Conference on, 2001, pp. 145–152.

[7] M. Greenspan and M. Yurick, “Approximate k-d tree search for efficient icp,” in 3-D Digital Imaging and Modeling, 2003. 3DIM 2003. Proceedings. Fourth International Conference on, Oct 2003, pp. 442–448.