Proceedings of the 11th International IEEE
Conference on Intelligent Transportation Systems
Beijing, China, October 12-15, 2008

Curb Detection Based on a Multi-Frame Persistence Map for Urban Driving Scenarios

Florin Oniga, Sergiu Nedevschi, and Marc Michael Meinecke

***Abstract*—An approach for the detection of straight and curved curbs (border of relevant traffic isles, sidewalks, etc) is presented, in the context of urban driving assistance systems. A rectangular elevation map is built from 3D dense stereo data. Edge detection is applied to the elevation map in order to highlight height variations. We propose a method to reduce significantly the 3D noise from dense stereo, using a multi-frame persistence map: temporal filtering is performed for edge points, based on the ego car motion, and only persistent points are validated. The Hough accumulator for lines is built with the persistent edge points. A scheme for extracting straight curbs (as curb segments) and curved curbs (as chains of curb segments) is proposed. Each curb segment is refined using a RANSAC approach to fit optimally the 3D data of the curb. The algorithm was evaluated in an urban scenario. It works in real-time and provides robust detection of curbs.**

I. Introduction

CURBS are an important driving area delimiter in urban scenarios. Compared to lane markings, curbs have various "looks" and are harder to detect from monocular intensity images.

Existing stereo-based systems focus mainly on detecting 3D obstacles (vehicles, pedestrians etc.), and curbs are usually considered road inliers. Most of the existing algorithms try to compute the road/lane surface, and then use it to discriminate between road and obstacle points.

Disparity space-based algorithms take as input the standard output of stereo matching: the disparity map. The "v-disparity" [1] approach is well known and used to detect the road surface in a variety of applications [2].

3D space-based algorithms are often used for ego-pose estimation [3], [4], but also for lane and obstacle detection [5], [6], road, traffic isles and obstacle detection [7], or curb detection [8], [9].

The 3D space is represented as an elevation map and a quadratic road surface is computed in [7]. Obstacle and traffic isles are detected based on their position and density relative to the road surface. Curb points can be extracted as the borders of traffic isles. However, if the road surface cannot be fitted (lack of texture or non-quadratic surface) then traffic isles are not detected.

Approaches presented in [8] and [9] deal explicitly with curb detection. Both methods are based on the extraction of straight curbs with the Hough transform.

Image edge points are detected in [9]. A weight is computed, for each edge point, as a function of the image brightness gradient and the 3D elevation gradient. These weights are used for voting in the Hough accumulator. One dominant straight curb (per scene) is extracted from the Hough accumulator.

The 3D space, from dense stereo, is transformed into an elevation map in [8]. Edges are detected on the elevation map and the Hough accumulator is built. At most two straight curbs are extracted from the accumulator. A curb is considered valid if it has a specific height variation profile.

Both methods [8] and [9] detect curbs as lines in the space of interest, without extracting the actual curb segments along these lines. In many scenes, real curbs are segments (with visible ends, for instance in intersections) in the space of interest.

The algorithm that will be presented in this paper is related to the approach from [8], but it has several key improvements:

- It uses (multi-frame) temporal filtering of curb points, thus fewer false curbs are detected.
- It locates curb segments along linear curbs extracted with Hough, thus providing a better description of curbs.
- It extracts curved curbs (often present in urban scenes) as chains of segments.

Curbs are detected based on local derivatives of the elevation map, so the shape of the road surface or varying pitch and roll angles are irrelevant (no additional processing time is spent).

The curb detection algorithm will be presented next in section II, results and an extended evaluation in section III and conclusions in IV.

II. The Curb Detection Algorithm

A. Overview

The proposed algorithm is robust if the following assumptions are fulfilled:

- 3D points are reconstructed around the curb and noise is not predominant,
- A significant height variation (elevation gradient) is present around the curb (at least 5 cm) and the height variation is sharp (like a step function).

F. Oniga and S. Nedevschi are with the Computer Science Department, Technical University of Cluj-Napoca, Romania, email: florin.oniga, sergiu.nedevschi@cs.utcluj.ro

M. M. Meinecke is with the Research Department of Volkswagen AG, Wolfsburg, Germany, e-mail: marc-michael.meinecke@volkswagen.de.

The algorithm has the following main steps:

1. Building the Digital Elevation Map (DEM).
2. Detection of DEM edges: points with specific 3D height variation are detected.
3. Temporal filtering of DEM edges: only edge points static relative to the road are validated, based on the motion model of the ego vehicle. A multi-frame persistence map is proposed.
4. Extraction of straight and curved curbs from the Hough accumulator.

Compared to the approach presented in [8], only the first two steps are similar, while steps 3 and 4 are novel.

B. *Building the Digital Elevation Map (DEM)*

We used the DEM representation presented in [7]. Some modifications were needed for the problem of curb detection.

A region of interest, from the set of 3D points (Fig. 1), is transformed into a rectangular DEM (Fig. 2). Each map cell stores a single height value. Contrary to [7], the lowest height of the 3D points contained by a cell is stored (instead of the highest). This is necessary to avoid the situation when 3D points from floating objects (traffic signs, trees etc) overwrite the height of the underlying curb points.

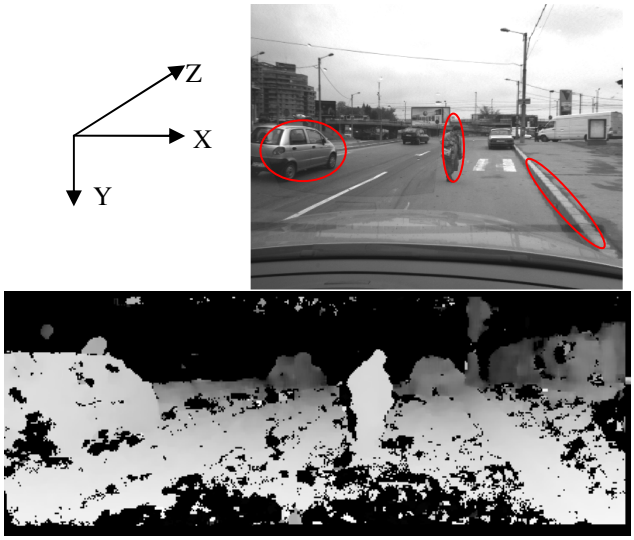
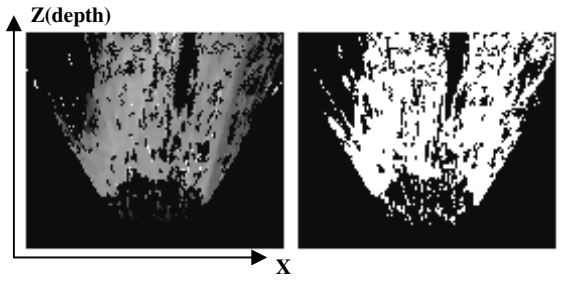


Fig. 1. The left image (top) and a perspective view of the 3D set of points from dense stereo (bottom), brighter means closer.

The region of interest (ROI) is smaller than in [7]. Considering the road plane from calibration, the ROI vertical limits (Y-axis) are -2 to +2 meters, lateral (X-axis) -6 to +6 meters, and along the depth (Z-axis) from 0 up to 10 meters. We selected these limits for several reasons. The vertical span guaranties that curb points are stored in the DEM even for extreme pitch and roll angle of road relative to ego car (extreme uphill or downhill of up to 20% slope), or various road geometries (as long as they fit in the vertical interval). The depth limit of the ROI has to fulfill two constraints: height variations of 5 centimeters should be detectable on the DEM, and enough 3D points are reconstructed so the DEM is not sparse along the depth (the height uncertainty model and depth resolution model from [7] were applied to our stereo configuration).

A DEM cell has a size of 10 cm x 10 cm in the XZ world plane. This size provides a good tradeoff between the overall processing time and the accuracy of 3D curbs localization. The 3D height was translated and scaled for a better visualization of the DEM (road level from system calibration has a gray value 128 in the DEM). Only heights around the road (1 meter band) are displayed correctly (for the images in this paper), due to a limited number of gray values (256). This has no influence upon the detection of curbs because the whole range of heights is used for processing.



a. b.

Fig. 2. The DEM in a: darker means higher in 3D. The validity map in b shows which cells (white) are valid (contain 3D data).

C. *Detection of DEM edges*

The main feature of a curb point is that the 3D height (elevation) changes sharply around the point. Curb points should be included in the set of edge points of the DEM.

We used a robust method (Canny, [10]) for detecting edges (Fig. 3). Only edges having the gradient in a desired interval (equivalent to a 3D variation between 5 and 35 centimeters) are detected.

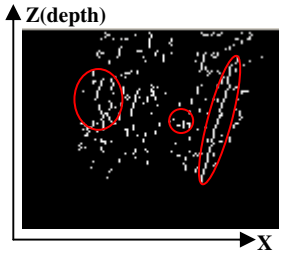


Fig. 3. Edges detected on the DEM from figure 2. The only areas where height variations exist are marked with red ellipses (in the 3D scene: the left-side car, the pedestrian and the right-side curb). The other edge points represent false height variations.

Some false bumps are reconstructed on the road surface, causing false edges (fig. 3). This noise appears due to low-accuracy reconstruction of the road surface (weak texture combined with image noise). It can make the process of curbs extraction quite difficult, because the noise has height variations similar to small curbs. This is the main reason why, in [8], at most two curbs were detected (to avoid an increased rate of false detections).

D. Temporal filtering of DEM edges based on multi-frame persistence

We propose a temporal filtering process to decrease the presence of noise, based on the two assumptions:

- Curbs are static related to the road surface: if the ego motion between successive frames is compensated, then curb points from the current frame should overlap curb points from the previous frame.
- Weak road texture combined with image noise will cause randomly located DEM edge points. These points should have different locations between consecutive frames.

Coordinates from the current reference frame O(t) can be transformed into the previous reference frame O(t-1), assuming the translation d and rotation angle α are known (Fig. 4). The ego car's standard speed and yaw-rate sensors can be used to estimate these parameters. The following motion model was used: the ego has a circular trajectory between successive frames, and the arc length and radius are computed based on the ego car speed v, yaw-rate value γ and frame relative timestamp Δt. The yaw-rate sensor provides the rotation angle, and the translation is computed geometrically:

$$\alpha = \gamma \cdot \Delta t \,, \tag{1}$$

$$d = 2 \cdot \frac{v \cdot \Delta t}{\alpha} \cdot \sin\left(\frac{\alpha}{2}\right). \tag{2}$$

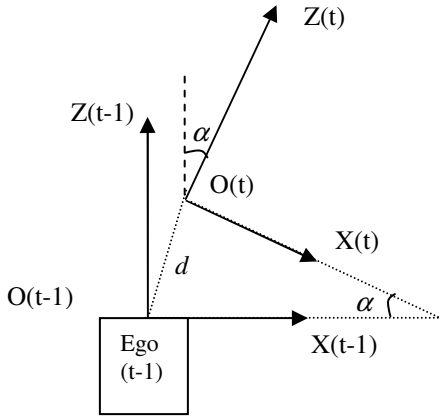


Fig. 4. Ego motion is expressed as the Euclidian distance between origins and relative angle between axes.

First, let us define the concept of multi-frame persistence map (PM). The PM is a rectangular map of the same size as the DEM. A cell (i,j) (i for the Z direction and j for X) of PM shows the lifetime (in consecutive frames) of DEM cell (i,j): for how many consecutive frames was the cell detected as an edge point, in a global reference frame (the same edge point relative to the road, along the sequence of frames).

The persistence map PM_T for the current frame is built from the PM_{T-1} of the previous frame and the set of edge points detected on the DEM_T of the current frame.

For each location (i, j) of PM_T:

- If (i,j) is not an edge point then $PM_T(i,j)=0$;
- Otherwise, if (i,j) is edge then:
  1. Compute the coordinates of current frame point (i, j) in the previous frame, as real numbers (i', j').
  2. $PM_T(i,j)$=MAXIMUM (W) + 1, where the set W contains the persistence values of the previous frame PM_{T-1} for the 4 closest neighbors of (i',j'),

$$W = \left\{ \begin{matrix} PM_{T-1}(\lfloor i' \rfloor, \lfloor j' \rfloor), PM_{T-1}(\lfloor i' \rfloor, \lceil j' \rceil), \\ PM_{T-1}(\lceil i' \rceil, \lfloor j' \rfloor), PM_{T-1}(\lceil i' \rceil, \lceil j' \rceil) \end{matrix} \right\}.$$

Using the maximum persistence of the four points closest (instead of a single point) to the real coordinates from the previous frame is required to compensate some sources of errors:

- Possible lack of accuracy from the ego sensors,
- Most of the time, integer coordinates (i,j) from DEM_T do not have an correspondent DEM_{T-1} location with integer coordinates (i', j'),
- Pitch angle variation might occur between frames (due to road bumps etc), causing small depth shifts of the coordinates between successive frames. Assuming an extreme pitch variation α of 5 degrees, the shift along the depth Z, at 10 meters, is $10*(1-\cos(\alpha))=0.04$ meters, less than half the DEM cell size.

Once the persistence map PM is computed, candidate curb points should be selected as points having the persistency higher than a threshold T. A value too low for T will cause many false curb points to occur, while a value too high will cause unjustified delay for detecting true curb points. We performed the following experiment for selecting the optimal T: a sequence of 200 frames was acquired while driving through an empty parking lot, without any objects or curbs in the analyzed ROI (Fig. 5). The road surface had normal texture, with common features such as braking traces, different color patches etc.

Fig. 5. The first frame of the sequence used.

All detected edge points are false curb points, due to the poor accuracy of the 3D reconstructed road. A total number of 30421 edge points were detected (an average of 152 points per frame). The total number of edge points having the same persistence value was evaluated (Fig. 6) relative to

the total number of edge points. 22.6% persist for two frames, 7.8% persist for three frames, 3.56% persist for four frames, and less than 1.5% points persist more than four frames. A value of 3 or 4 frames is optimal for the threshold T, greatly reducing the number of false edge points (Fig. 7). The only downside is that a true curb point is validated only after T frames since it entered the analyzed ROI. However, this has minor influence upon detection (no curbs are missed). Even for a speed of 50km/h, it will take about 18 frames for the ego car to reach a curb point placed at the maximum depth.

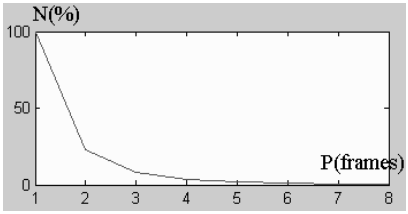


Fig. 6. The percentage of points N as a function of the persistence value P.

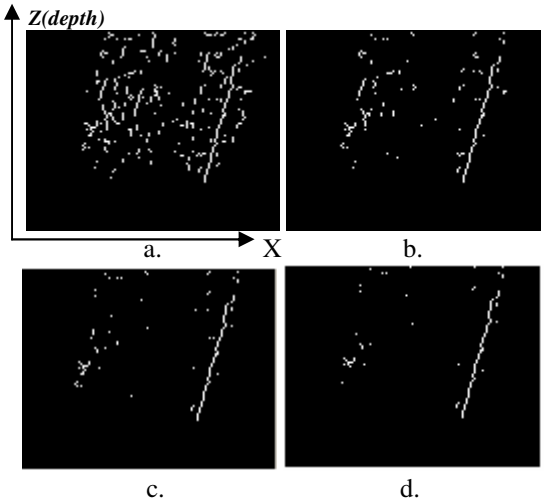


Fig. 7. The edges from the DEM shown in figure 2 validated using a persistency threshold T of one frame (no filtering, in a), two frames (b), three frames (c), and four frames (d).

Another positive effect of the proposed temporal filtering is the rejection of most edges from dynamic scene components (cars, pedestrians, etc). This simplifies the process of curbs extraction from the Hough accumulator, since most of outliers are rejected from the set of curb points.

E. Extraction of curbs from the Hough accumulator

The Hough accumulator [11] is built. Lines are represented in the (r,θ) space and we used a resolution of 1 degree for the angle θ (0...360 degrees) and a resolution of 1 pixel for r. Local maxima of the accumulator are relevant lines that might contain curb segments. However, the Hough accumulator does not provide any clues about how curb points are placed along the relevant lines. These points can be scattered along the line (most likely noise) or can form continuous chains (possible curb segments). Curb segments can be extracted from the relevant lines by analyzing the height variation profile of each line.

First, a robust way to estimate the height variation for a line will be presented. After this, we will present a scheme to extract curb segments and chains of curb segments.

1) The line variation profile

Each line point must be checked to see if it is a curb point. We assume a curb point has two main features:

- Its height variation is in a specific interval (5..35 centimeters),
- The height variation is positive from the roadside of the curb to the opposite side (sidewalk or traffic isle).

The roadside of the curb can be detected geometrically since the road is placed on the same side as the ego vehicle (under normal circumstances) relative to the curb.

Computing the height variation for a line point can be unstable if only points adjacent to the line are used (a simple gradient, Sobel mask). Although the gradient was robust enough to detect height variations as edges, it is less accurate for estimating the magnitude of the variation. It measures the overall variation: if the supporting road surface of the curb has a large local slope, this slope will contribute to the estimated variation, making it larger than the actual variation.

A larger area must be used in order to have a robust result. The shape of the area cannot be rectangular since it would not provide symmetry for all possible line orientations. We used a circular mask (5x5 or 7x7 pixels, similar to corner detection algorithms). The height variation for one point is computed as a difference function between points placed on the same side as the ego car (relative to the line) and the opposite points, inside the circular nucleus.

An example with real data is presented in figure 8.b: pixels inside the mask with their values, the line pixels drawn with black and, in the center, the current point x for which the computation is performed.

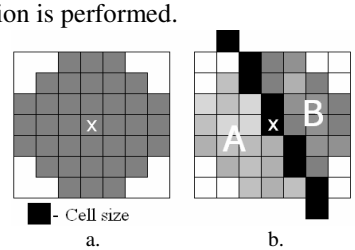


Fig. 8. The circular mask (7x7 pixels, in a) used for computing the height variation for the central point x. An example with real data from the DEM is shown in b, line pixels are drawn with black.

The median values $H_m(A)$ and $H_m(B)$ are computed for regions A, respectively B. The height variation for the line point x is computed as $H_m(A)-H_m(B)$, if A is on the same side as the ego car, or $H_m(B)-H_m(A)$ otherwise (B is on the ego's side).

Finally, the height variation profile of each relevant line is evaluated by computing the height variation for each line point. Only positive height variations (between 5 and 35 centimeters) of feasible curb points are stored in the profile.

2) Extracting chains of curb segments

One requirement for this process is to avoid multiple responses for the same curb. To fulfill this, the votes of extracted curb segments are subtracted from the accumulator. Another requirement is to detect chains of curb segments belonging to the same curved curb.

The scheme of curb segments extraction is:

1. Select the global maximum L of the Hough accumulator H.
   a. If $H(L)$ is above a threshold T_L then:
      I. Compute the height variation profile HP of L.
      II. Extract the dominant curb segment C represented by the largest continuous interval of curb points in the HP (of at least T_L points). Recursively add curb segments at each end of C (detailed next).
      III. Compute all pairs of (r,θ) for each edge point of the detected chain of curb segments, and subtract one vote from correspondent accumulator locations.
   b. Otherwise stop.
2. Repeat step 1.

Step II is the core of the scheme. If C is placed on a curved curb then it is most likely that its extremities are the starts of adjacent segments. All half-lines are extracted, for each extremity, having the following features: they contain the extremity, have a Hough score above T_L, and form an obtuse angle with the segment C. Starting from the extremity, the best (based on length and the height profile variation) interval of curb points is searched along the half lines. If it contains at least T_L points then the segment associated to this interval is selected as the adjacent curb segment of the extremity. Then, this step is applied recursively to the newly added segments, for their unconnected extremities.

The value of T_L was 10 edge points, equivalent of at least 1-meter 3D length. A different value can be used (lower/greater), depending on the quality of stereo reconstruction. Some particular urban scene might have, beside real curbs, curb-like items in the background of real curbs. Such artifacts can be eliminated by comparing their 3D location with the location of the closest curbs to the ego car (simple geometric constraints).

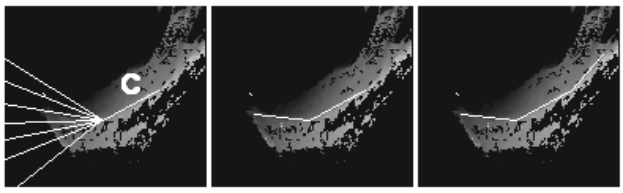


Fig. 9. A curved curb is approximated as a chain of curb segments. The first segment is extended with the longest curb segment from the set of candidate half-lines (only few are drawn), at each end.

Each curb, or chain of curb segments, is mapped back from the DEM space into the 3D space, and projected onto the left image for visualization. The steps presented so far provide a good lateral and longitudinal localization of curbs. Without the detection of the vertical orientation/location of each segment, the results are less robust (Fig. 11).

Fig. 10. The curb segment projected back onto the left image, for the scene shown in figure 2.

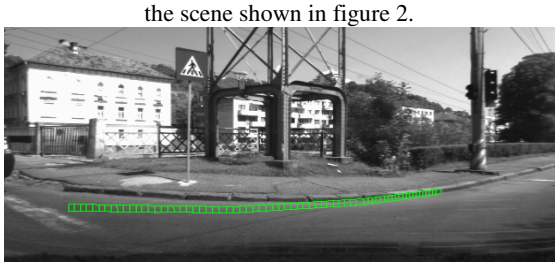


Fig. 11. The chain of curb segments projected back onto the left image (the vertical coordinate of curb points is the road level from calibration), for the scene analyzed in figure 9.

We consider two DEM lines for each curb segment (one on each side of the curb segment as the adjacent surfaces of the curb, presumably road and sidewalk/traffic isle). RANSAC is applied, on the height data around the curb segment, in the DEM, to compute these lines, based on several constraints: two points per sample for each line, lines from adjacent segments must have the same height for junctions, lines associated with a segment are parallel and the vertical distance between them is the average height variation of the curb. The average height variation of the curb is the median of the height variation profile from the dominant curb segment. The configuration with the best score is selected as the best vertical localization of the curb. A good 3D alignment is obtained (Fig. 12).

Fig. 12. Better results are obtained using vertical alignment of each curb segment.

III. RESULTS

The algorithm was implemented in C++. The dense 3D data was acquired using a calibrated stereo rig with grayscale cameras and a commercial dense stereo board. We used cameras with a large horizontal field of view (about 67 degrees), appropriate for urban scenes.

A processing time around 8-9 ms was achieved for the algorithm (on Pentium 4 Dual Core processor).

Four different scenes with curbs of various heights were used for the first criterion. The computed height variation was accurate (Table 1), with a maximum error of 5%.

Table 1

Scene number	1	2	3	4
Real height (cm)	5.0	7.0	11.0	14.0
Computed height (cm)	4.8	7.3	10.6	13.8

The algorithm was evaluated on a sequence of 2050 images and has a detection rate of about 96% of the total number of curbs. Two false curbs appeared on the whole sequence (Fig. 14). They were caused by false elevations (large areas) in the set of 3D points provided by dense stereo. The rate of false curbs is related to the quality of the stereo reconstruction.

Detection of non-sharp curbs (so called "traversable") is not stable because a 3x3 mask is used for edge detection. A possible solution with multiple mask sizes will be studied in the future.

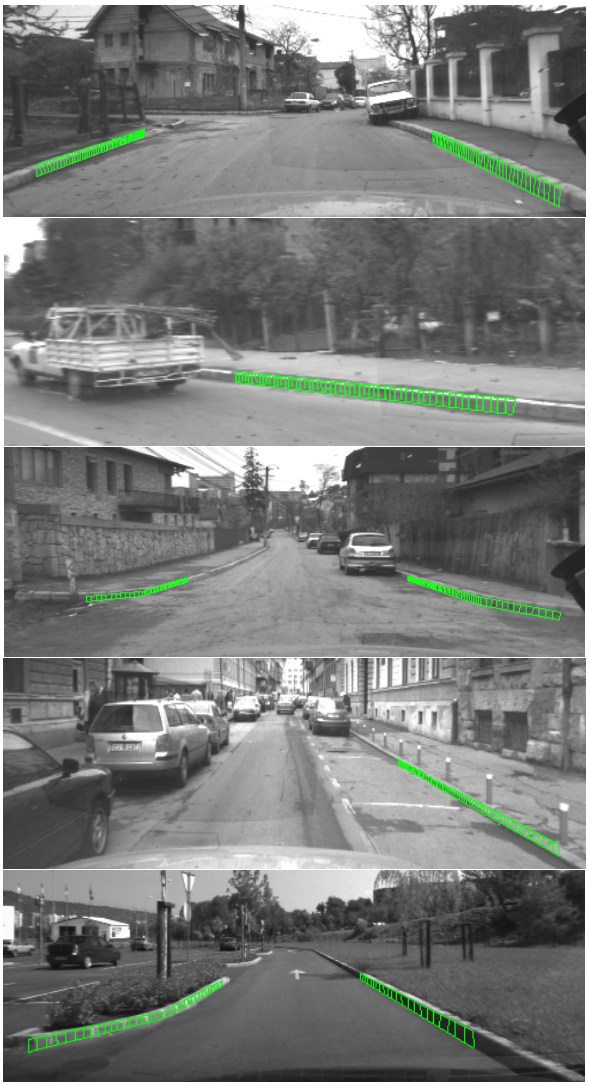


Fig. 13. Various 3D curbs projected onto the left image.

Fig. 14. A scene with a false curb.

IV. Conclusions

A novel curb detection algorithm was presented. It deals with a problem less approached, the detection of straight and curved curbs in the urban environment. The proposed algorithm works in real-time. It can be used in a variety of applications that require detection of road-elevated delimiters such as curbs: from simple lateral ego-vehicle control to complex path planning.

Future research is needed for improving the generality of the algorithm, by providing solutions for scenarios with poor 3D data.

References

[1] R. Labayrade, D. Aubert, and J.-P. Tarel, "Real time obstacle detection on non flat road geometry through V-disparity representation," in IEEE Intelligent Vehicles Symposium, Versailles, June 2002, pp. 646–651.

[2] Broggi, C. Caraffi, P. Paolo Porta, and P. Zani, "The Single Frame Stereo Vision System for Reliable Obstacle Detection used during the 2005 DARPA Grand Challenge on TerraMax", in IEEE Intelligent Transportation Systems Conference, Toronto, Canada, September 17-20, 2006, pp. 745-752.

[3] Sappa, D. Gerónimo, F. Dornaika, and A. López, "Real Time Vehicle Pose Using On-Board Stereo Vision System", Int. Conf. on Image Analysis and Recognition, LNCS Vol. 4142, Portugal, September 18-20, 2006, pp. 205-216.

[4] M. Cech, W. Niem, S. Abraham, and C. Stiller, "Dynamic ego-pose estimation for driver assistance in urban environments". In IEEE Intelligent Vehicles Symposium, Parma, Italy, 2004, pp. 43-48.

[5] S. Nedevschi, R. Danescu, D. Frentiu, T. Marita, F. Oniga, C. Pocol, R. Schmidt, T. Graf, "High Accuracy Stereo Vision System for Far Distance Obstacle Detection", IEEE Intelligent Vehicles Symposium (IV2004), Parma, Italy, pp. 292-297, 2004.

[6] S. Nedevschi, R..Schmidt, T. Graf, R. Danescu, D. Frentiu, T. Marita, F. Oniga, C. Pocol, "3D Lane Detection System Based on Stereovision", IEEE Intelligent Transportation Systems Conference (ITSC), Washington, USA, pp.161-166, 2004.

[7] F. Oniga, S. Nedevschi, M-M. Meinecke, T-B. To, "Road Surface and Obstacle Detection Based on Elevation Maps from Dense Stereo", The 10th International IEEE Conference on Intelligent Transportation Systems, Sept. 30 - Oct. 3, 2007, Seattle, Washington, USA.

[8] F. Oniga, S. Nedevschi, M-M. Meinecke, "Curb Detection Based on Elevation Maps from Dense Stereo", The 3rd International IEEE Conference on Intelligent Computer Communication and Processing, pp.119-125, 6-8 Sept. 2007, Cluj-Napoca, Romania.

[9] R. Turchetto, R. Manduchi, "Visual Curb Localization for Autonomous Navigation", Proceedings of IEEE/RSJ International Conference on Intelligent Robots and Systems, pp. 1336-1342, Las Vegas, October 2003.

[10] J. Canny, "A computational approach to edge detection", IEEE Trans. Pattern Analysis and Machine Intelligence, pp. 679-698, Nov. 1983.

[11] D. H. Ballard, "Generalizing the Hough Transform to Detect Arbitrary Shapes", Pattern Recognition, 13(2):111-122, 1981.