Curb Segments Detection with Temporal Filtering for Urban Driving Scenarios

Florin Oniga,   Sergiu Nedevschi,
Technical University of Cluj-Napoca, Romania
florin.oniga, sergiu.nedevschi@cs.utcluj.ro

Marc Michael Meinecke
Volkswagen A.G., Germany
marc-michael.meinecke@volkswagen.de

Abstract

An improved approach for the detection of 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. Temporal filtering is performed for edge points, based on the ego car motion, and only persistent points are validated. Relevant lines are extracted using a Hough accumulator. Curb segments are extracted from each relevant line by applying a set of proposed conditions. The algorithm was evaluated in an urban scenario. It works in real-time and provides robust detection of curbs.

1. Introduction

Curbs are an important driving area delimiter in urban scenarios. Unlike lane markings, curbs have various "looks" and are harder to detect from monocular intensity images. Most of the existing stereo-based systems focus mainly on detecting 3D obstacles (vehicles, pedestrians etc.), and curbs are usually considered road inliers.

Approaches [1] and [2] deal explicitly with curb detection. Both are based on the extraction of linear curbs with the Hough transform.

Image edge points are detected in [2]. 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 building the Hough accumulator. One dominant curb (per scene) is extracted as the line with the maximum score in the accumulator.

The 3D space, from dense stereo, is transformed into an elevation map in [1]. Edges are detected on the elevation map and the Hough accumulator is built. At most two curbs are extracted from the accumulator. A curb is considered valid if it has a specific height variation profile.

Both methods [1] and [2] detect curb 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) in the space of interest.

The algorithm that will be presented in this paper is related to the approach from [1], but it has several advantages. It uses 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 the scene. The number of curbs segments extracted is not limited. 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.

The curb detection with temporal filtering algorithm will be presented next in section 2, results and an extended evaluation in section 3 and conclusions in 4.

2. The curb detection with temporal filtering algorithm

The proposed algorithm is optimal if the following assumptions are fulfilled: 3D points are reconstructed around the curb and noise is not predominant, curbs are linear or quasi-linear (small curvatures), and a significant height variation is present around the curb (at least 5 cm) and the height variation is sharp (like a step function).

The algorithm has the following main steps:
1. Building the Elevation Map (EM).
2. Detection of EM edges.
3. Temporal filtering of EM edges.
4. Hough based extraction of relevant lines.
5. Extraction of curb segments.

Compared to the approach presented in [1], only the first two steps are similar. Step 4 is fundamentally improved, and steps 3 and 5 are new.

2.1. Building the Elevation Map (EM)

The elevation map representation used was presented in [3]. We performed some modifications needed for the problem of curb detection.

A region of interest, from the set of 3D points (Fig. 1), is transformed into a rectangular elevation map (Fig. 2). Each map cell stores a single height value. Contrary to [3], 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.

Figure 1. The left image (top) and a perspective view of the 3D set of points from dense stereo (bottom).

The region of interest (ROI) is smaller than in [3]. 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 two reasons: curb points are stored in the EM even for extreme pitch and roll angle, and enough 3D accuracy is available from stereo at the maximum depth.

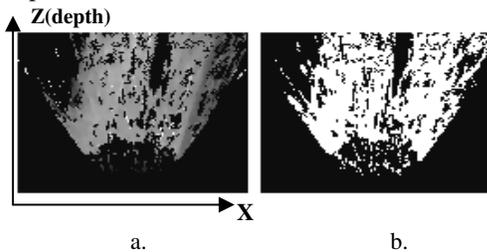An EM cell has a size of 10 cm x 10 cm in the XZ world plane.

Figure 2. Elevation map in a: darker means higher in 3D. The validity map in b shows which cells (white) are valid (contain at least one 3D point).

2.2. Detection of EM edges

The main feature of a curb point is that the 3D height (around the point) changes sharply. Curb points are included in the set of edge points from the elevation image. The elevation map, built from dense stereo points, presents some spike-like noise (rough errors in the dense 3D reconstruction). A median 3x3 filter is applied, before edge detection, in order to reduce this noise.

We used an optimal method [4] for detecting edges (Fig. 3) with the gradient in a desired interval (equivalent to a 3D variation between 5 and 35 centimeters). Using this interval was necessary in order to filter most of the edges from other scene items (cars, pedestrians etc.).
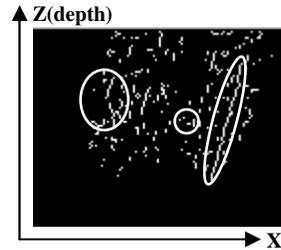

Figure 3. Edges detected on the elevation map from figure 2. The only areas where height variations exist are marked with ellipses (in the 3D scene: the left-side car, the pedestrian and the right-side curb).

The median filter applied eliminates the spike-like noise, but some fine noise (false bumps) remains on the road surface. This noise appears due to low-accuracy reconstruction of the road surface, (weak texture combined with image noise). This is the main reason why, in [1], at most two curbs were detected (to avoid an increased rate of false detections).

2.3. Temporal filtering of EM edges

We propose a temporal filtering process to decrease the level 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 produce randomly located edge points. These points should have random locations between consecutive frames.

Coordinates from the previous reference frame O(t-1) can be transformed into the current reference frame O(t), assuming the translation d and rotation angle α are known (Fig. 4). 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 speed v, yaw-rate angle γ and frame relative timestamp Δt. The yaw-rate sensor provides the rotation angle, and the translation is computed geometrically:

$$\alpha = \gamma \cdot \Delta t \,, \quad d = 2 \cdot \frac{v \cdot \Delta t}{\alpha} \cdot \sin\left(\frac{\alpha}{2}\right). \qquad (1)$$
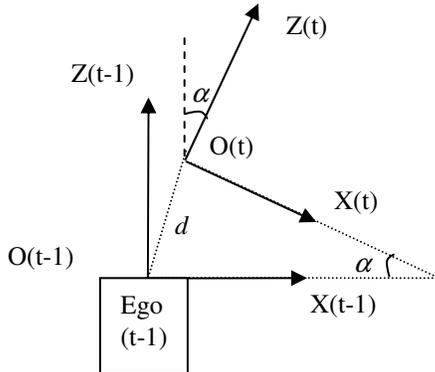

Figure 4. Ego motion is expressed as the Euclidian distance between origins and relative angle between axes.

EM edges detected in frame O(t-1) were mapped (rotation and translation) and dilated into O(t) reference frame (Fig. 5.a). Morphological dilation was necessary to compensate two sources of errors:

- Possible lack of accuracy from the ego sensors,
- Integer coordinates of the EM map are transformed between the frames.

The edges detected in the current frame O(t) were superimposed (Fig. 5.c, white points) to the dilated edges, mapped to O(t) (Fig. 5.c, gray points), from the previous frame O(t-1). Only persistent edges (from previous to current frame) were validated (Fig. 5.d).
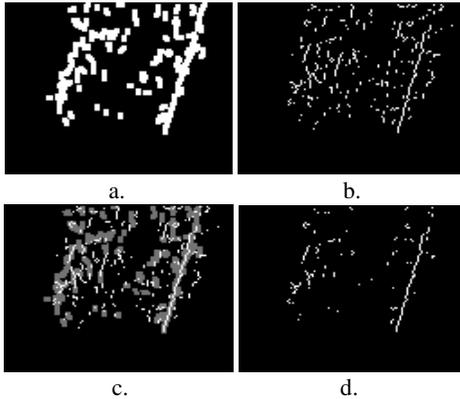

Figure 5. The dilated edges from the previous frame O(t-1) are mapped into O(t), in a. Edges detected in the current frame O(t) are shown in b. Superimposing (c) helps obtaining the filtered set of edge points in the current frame (d), with less noise.

About 70-80% of noisy edges are eliminated. Another positive effect is the elimination of edges from dynamic scene components (cars, pedestrians, etc). This simplifies the process of extracting linear curbs, since most of outliers are rejected from the set of curb points.

2.4. Hough based extraction of relevant lines

The Hough accumulator [5] is built. It counts how many edge points are placed on each possible line in the elevation map. Lines are represented in the (r,θ) space, which allows modeling a finite number of lines in the image (line parameters are bounded). We used a resolution of 1 degree for the angle θ (0...360 degrees) and a resolution of 1 pixel for r.

Due to the reduction of noisy edges through temporal filtering, a larger number of relevant lines (15 lines) will be extracted for further analysis, compared to [1] (only 5 lines, and at most 2 valid curbs).

The main challenge of this process is to avoid multiple responses for the same line. We used a optimal approach, by taking out the votes of extracted lines from the accumulator. The process of selecting relevant lines from the Hough accumulator H is:

1. Select the global maximum L of the Hough accumulator.
   a. If H(L) is above a threshold T_L(=10 edge points, equivalent of at least 1 meter cumulated length) then:
      - Store L as a relevant line,
      - Compute all pairs of (r,θ) for each edge point of L and subtract one vote from correspondent accumulator locations.
   b. Otherwise stop.
2. If the number of relevant lines is below a threshold (15 lines) then go to step 1, otherwise stop.

The relevant lines extracted from the Hough accumulator have only one attribute: they contain at least T_L(=10) edge points. These edge points can be scattered along the line (most likely noise) or can form continuous chains (possible curb segments).

2.5. Extraction of curb segments

In [1] each relevant line was analyzed by counting how many curb points it contains. Each line containing at least 40% curb points is considered curb. The main disadvantage is that the whole line is considered curb. Curbs might span into the drivable space for some scenarios (ex. in intersections). We propose the detection of curb segments along the relevant lines, by analyzing the height variation profile of the lines.

Curb segments are searched in the height profile (computed as described in [1]) as continuous intervals of curb points. A curb segment is considered valid if it has a 3D length of at least 1 meter. Each curb segment

has an average height variation computed from its profile. Curb segments are mapped back into the 3D space and projected onto the left image for visualization.

Figure 6. The curb segment projected back onto the left image.

3. Results

The algorithm was implemented in C++. The dense 3D information was generated using a calibrated stereo rig with grayscale cameras and a commercial dense stereo board. A processing time around 6 ms was achieved for the algorithm itself (on Pentium 4 Dual Core processor). Building the elevation map is the most time consuming part, taking between 3 to 4 ms. The detection of curbs on the elevation map is very fast, about 2 ms.

Two criteria were used to evaluate the proposed algorithm: the height variation accuracy, and the rate of false positives and negatives.

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	7	11	14
Computed height(cm)	4.8	7.3	10.6	13.8

Regarding the second criterion we performed the following evaluation: 250 stereo images of different (random) scenes were used for testing (out of two hours of stored stereo images, recorded while driving the ego-car). Two variants of the algorithm were evaluated: with temporal filtering (TF) and without. The following detection rates were obtained:

Table 2

Algorithm	Detected curbs (%)	Missed curbs (%)	False curbs (/frames)
Without TF	96.4	3.6	42/250
With TF	95.8	4.2	3/250

The false detection rate dropped significantly by using both edge and curb filtering based on persistency between two consecutive frames (various results in figure 7).

4. Conclusions

A robust curb detection algorithm was presented. It deals with a problem less approached, the real-time detection of curbs in the urban environment. 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 detecting non-sharp and strongly curved curbs.

Figure 7. Various 3D curbs projected onto the left image.

5. References

[1] 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.

[2] 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.

[3] 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.

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

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