Aalborg Universitet

Stereoscopic Roadside Curb Height Measurement using V-Disparity

Matu, Florin-Octavian; Vlaykov, Iskren; Thøgersen, Mikkel; Nasrollahi, Kamal; Moeslund, Thomas B.

Published in:
Proceedings from SPIE 9445, Seventh International Conference on Machine Vision (ICMV 2014)

DOI (link to publication from Publisher):
10.1117/12.2180562

Publication date:
2014

Document Version
Early version, also known as pre-print

Link to publication from Aalborg University

Citation for published version (APA):
Matu, F.-O., Vlaykov, I., Thøgersen, M., Nasrollahi, K., & Moeslund, T. B. (2014). Stereoscopic Roadside Curb Height Measurement using V-Disparity. In A. Verikas, B. Vuksanovic, P. Radeva, & J. Zhou (Eds.), *Proceedings from SPIE 9445, Seventh International Conference on Machine Vision (ICMV 2014)* SPIE - International Society for Optical Engineering. https://doi.org/10.1117/12.2180562

General rights
Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal -

Take down policy
If you believe that this document breaches copyright please contact us at vbn@aub.aau.dk providing details, and we will remove access to the work immediately and investigate your claim.

Downloaded from vbn.aau.dk on: June 18, 2025

Stereoscopic Roadside Curb Height Measurement using V-Disparity

Florin Octavian Matu, Iskren Vlaykov, Mikkel Thogersen, Kamal Nasrollahi, and Thomas Moeslund
Visual Analysis of People Laboratory, Aalborg University, Denmark

ABSTRACT

Managing road assets, such as roadside curbs, is one of the interests of municipalities. As an interesting application of computer vision, this paper proposes a system for automated measurement of the height of the roadside curbs. The developed system uses the spatial information available in the disparity image obtained from a stereo setup. Data about the geometry of the scene is extracted in the form of a row-wise histogram of the disparity map. From parameterizing the two strongest lines, each pixel can be labeled as belonging to one plane, either ground, sidewalk or curb candidates. Experimental results show that the system can measure the height of the roadside curb with good accuracy and precision.

Keywords: Roadside curb, height measurement, V-Disparity, Hough Transform.

1. INTRODUCTION

Municipalities control a high number of road assets, such as street lamps, roadside curbs and gullies. To support their investment decisions, these assets have to be regularly inspected. Depending on the viewpoint, curbs may exhibit various curvatures and heights. This assessment is usually conducted manually which makes it a rather expensive and time consuming job. In this paper we devise an automatic method for the robust measurement of the height of the roadside curbs. The information about the geometry of the scene is obtained as line parameters and used to label each pixel from the image as *ground*, *sidewalk,* or *curb*. World coordinates of the pixels labeled as *ground* or *sidewalk* are computed from the intrinsic parameters of the input device and from the disparity data. These coordinates are then used to construct the equations of the two planes. The approximate curb region is detected in the thresholded gradient image as the strongest line using the Hough transform. By selecting a pixel in the detected curb region, the distance from the ground plane to the sidewalk plane is recorded as the height of the curb. This technique is simple, yet a robust and direct height measurement of the roadside curb. An output of the application is illustrated in Figure 1.

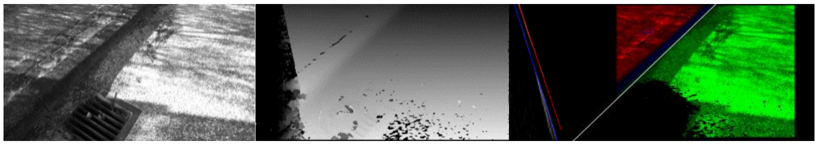


Figure 1. (a) Left input image, (b) Disparity image, (c) V-Disparity image, (d) Detected sidewalk, roadside curb, ground.

The rest of this paper is organized as follows: Section 2 reviews the related previous works in the literature, Section 3 explains the proposed system, Section presents the experimental results and finally Section 5 concludes the paper.

2. THE RELATED WORK

In the field of stereo-vision, many methods have been proposed for the specific task of road boundary detection. However, the challenge of curb (or *step*) detection has mainly been tackled in the Intelligent Transportation System area due to the multiple benefits it presents for robot navigation in urban environments. Xiaoye [1] presented a method in which edges are extracted from the brightness information of an input image. Then 3D data is used for a number of purposes: to determine the ground plane, to eliminate points which are unlikely to be part of the curb, to compute a suitable curvature index of the curb and to allow precise estimation of the height of the curb. The output of their algorithm is a set of either concave, or convex, 3D edge segments which define the curb. Oniga [2-4] and Balta [5] constructed a Digital Elevation Map (DEM) from a 3D point cloud captured using a dense stereo vision system. Oniga [3, 4] applied RANSAC to fit the candidate curb points, also represented by DEM edges, resulting in cubic polynomial curves which nominate the curb. More relevant to our approach is the work of Balta [5], in which the authors proposed to use a 3D Time-of-Flight (ToF) camera to acquire real-time high fidelity depth images leading to more precise V-disparity images. The authors also used an RGB camera in a secondary process that groups pixels belonging to the same physical object. The algorithm for estimating the terrain traversability classifies all image pixels in the TOF image as traversable or not based upon the analysis of the 3D (depth data) and 2D (IR data) content of the data from the TOF camera. In the same manner, Labayrade [6] and Cubber [7] used the output from a stereo vision system, represented by a disparity map. From it, the V-Disparity map is extracted and used to model the basic geometry of the scene. Furthermore, Labayrade [6] discusses a non-flat ground hypothesis, but will not be covered in this paper. A study on the U-V-Disparity can be found in Hu [8], where the authors present thorough and intuitive usages of the information stored in the U-V-disparity images. Their methods are in one sense, an extension to the work presented in Labayrade [6]. Bao [9] combines 2D image processing with 3D-road reconstruction from stereo images in order to collect and evaluate road distress data. The 3D geometrical model is used to detect deficiencies based on shape and depth of the effects.

3. THE PROPOSED SYSTEM

This paper purposes a computer vision system for measuring the height of the roadside curbs. The block diagram of the proposed system is shown in Figure 2. The sub-blocks of this system are explained in the following subsections.

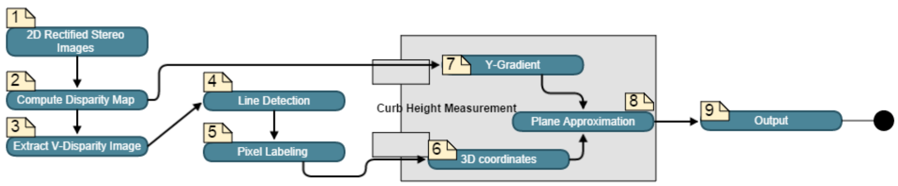


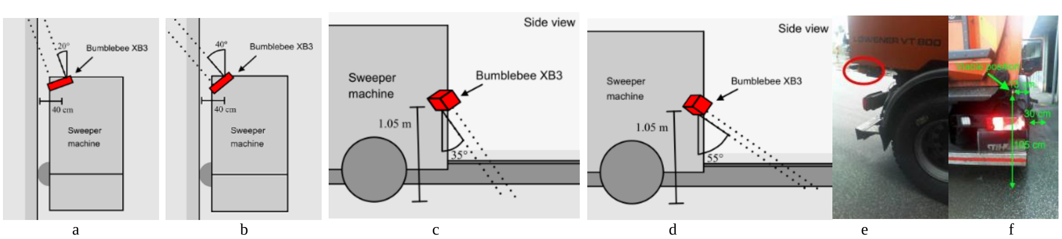
Figure 2. The block diagram of the proposed system.

3.1 Capturing device

Stereo vision technique was chosen for this system because both color and geometry information can be found in the same input data and because of its inherent tolerance to sunlight pollution. These systems rely on two standard camera sensors to capture data. The depth can be measured from such a device by computing the disparity image. To ease depth computation, it is recommended that the sensors are positioned next to each other with the baseline parallel to the scan-line and, preferably, to the horizon-line, while facing the same direction. Thereby, two slightly offset images of the scene are obtained. The device used for capturing the input is a Bumblebee XB3, with a frame-rate that can go up to 16 fps.

3.2 Positioning

The already developed systems for height measurements are nowadays using their own mobile platform. However, we propose to attach such a system to municipalities' vehicles, like a street sweeping machine, which can remove the need for an independent platform. Placing the device at the rear of a street sweeping machine offers multiple advantages. The obvious benefit is that the captured scene will feature clean streets, which greatly improves the quality of the gathered data. Furthermore, since these machines regularly patrol a high proportion of the streets, they provide a steady flow of input. The investigation of a sweeping machine lead to a viable position for the capturing device, see Figure 3 (e), (f). It was found to be on the rear of the sweeping machine, at 1.05 meters height, having a back-down-facing angle due to other components of the machine. Also we set the stereo system at a distance of 35 to 45 centimeters away from the curb. The downward facing angle will be between 35˚ and 55˚ compared to the vertical direction. The backward angle is between 20˚ and 40˚ towards the curb from the parallel direction to the curb. Figure 3(a)-(d) show these additional constraints.

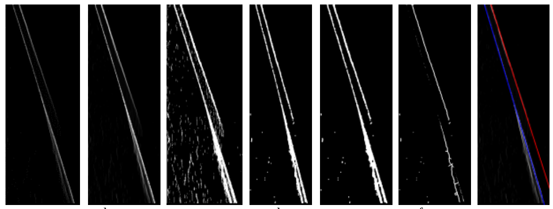


a b c d e f

Figure 3. The constraints for the capturing device: (a) and (b) Top view; (c) and (d) Side view, (e) Viable position on the sweeping machine, and (f) Suggested positioning of the capturing device.

3.3 Line Detection and V-Disparity

In order to increase the precision of the algorithm, the input to the system has captured the images with the constraints explained in the previous section. The raw input images are rectified and then a disparity map is computed. The V-Disparity map is constructed as a row-wise histogram of the disparity image and used to model the basic geometry of the scene. The major planar surfaces from the scene are represented as lines, parameterized using the Hough Transform. Having computed this image, we enlarge it by a factor of two and apply a binary threshold with a value of 15. Then, to remove the noise and keep the structure of the data, we apply a morphological opening followed by a morphological closing (both with a *3x3* Structural Element (SE)). The shapes of the employed SEs are rectangular and elliptical, respectively. The Zhang-Suen skeletonization algorithm is then used to thin the detected regions (lines). Figure 4 depicts these steps in the processing part of the V-Disparity image.



a b c d e f g

Figure 4. a) Computing V-Disparity, b) enlarging by 2, c) thresholding, d) opening, e) closing, f) thinning, and g) strongest lines.

Since the input consists of images of the scene in the vicinity of the vehicle, it is further assumed that the road has a flat geometry. This means that each dominant plane will yield a strong straight line in the V-Disparity image which discards the need for a piece-wise line-fitting. The latter would lead to noticeably more precise results in the case of uphill or downhill scene geometry.

The lines in the Hough-space are represented by ρ (displacement from the origin of the coordinate system) and θ (rotation according to the abscissa of the coordinate system). These two parameters are returned for each detected strong line. An example of the two strongest lines is illustrated in Figure1 (d). For each horizontal coordinate, its corresponding vertical counterpart is calculated. The pixel with this position is labeled in Figure1 (d) as Ground or Sidewalk. The function used for constructing the line is:

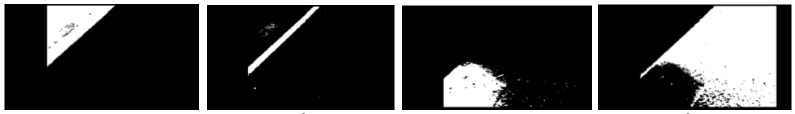
$$y = \frac{cos\theta}{sin\theta} x + \frac{\rho}{sin\theta} \qquad (1)$$

3.4 Pixel Labeling

The Hough Transform is applied with better efficiency on the skeletonized version of the V-Disparity image, leading to more robust results. It is used to grab the parameters of the two strongest lines in the image. Having computed the parameters of a line, the next step is to calculate the disparity value of the plane that is represented by these paramteres. This is done for each row v_r by:

$$v_r = \frac{\rho - sin\theta}{cos\theta} \qquad (2)$$

Comparing the required disparity values to those of each pixel determines the labels that will be assigned: *Ground*, the value for the pixel is within an small offset range to that defined by the strongest line, *Roadside Curb*, the pixel has a value in the range of the two strongest lines or *Sidewalk*, the value of the pixel is within a small offset range to that defined by the second strongest line. Pixels with other disparity values are labeled obstacles and will not be covered in this paper. We use the candidates for each label to create masks. The output masks of the pixel labeling process for the input disparity and V-Disparity images from Figure 1, are displayed in Figure 5.



a b c d

Figure 5: Labels used in the pixel mapping process: (a) Sidewalk, (b) Curb, (c) Gully, (d) Ground.

3.5 Curb Height Measurement

For the purpose of measuring the height of the roadside curb, we compute the two plane equations from the 3D coordinates of the pixels labeled as *Ground* or Sidewalk. At the current stage, the data has already been filtered in the *Pixel Labeling* step. This leads to a simple solution for approximating the planes, using principal component analysis. Because the data is representing a plane, the direction of least variation of the data can be used as the normal vector. To obtain the direction of least variation, the covariance of the data is used, as in:

$$\sum = \begin{bmatrix} cov(X,X) & cov(X,Y) & cov(X,Z) \\ cov(Y,X) & cov(Y,Y) & cov(Y,Z) \\ cov(Z,X) & cov(Z,Y) & cov(Z,Z) \end{bmatrix} \qquad (3)$$

The matrix obtained from Eq. 3, *Sigma*, describes the covariance between the data. To find the direction of least variation between the dimensions, the eigenvectors of the covariance matrix can be used. This will result in three vectors, two of these will be in the direction of the most variation and the last will be in the direction of least variance. The vector is chosen based on the corresponding eigenvalues. Next a point on the plane has to be determined. This is chosen to be the mean of the data, which will, once again, be a good approximation, due to the discrimination performed using the V-disparity. The result of this estimation is visible in Figure 6. To find the final height of the curb, the intersection of the unit normal vector of the ground plane with the curb plane needs to be found. This intersection can be computed as the Euclidean distance from one plane to the other.

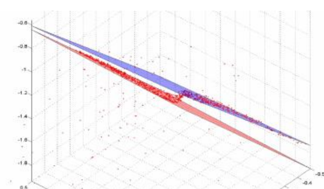


Figure 6: Sample of plane approximation using principal component analysis algorithm from the 3D points.

Experiments have shown that noise on the sidewalk, in the form of grass or bushes, can affect the distance measurement. This caused larger imprecision in the height estimation process. Therefore, the approximate position of the curb is found in the gradient image. Figure 7 shows the steps needed to estimate an approximate curb position from the vertical gradient of the disparity image. The gradient image is computed using the Sobel derivative. A low binary threshold is applied after median filtering with an aperture of *7x7*. On the output from Canny Edge Detector we apply the Hough transform which yields the parameters of either the top or the base of the curb. Since it is the area where the candidates of the two major planes meet, we consider this to be the area with the highest precision for measuring the height of the roadside curb, on the entire length of the curb line.

a b c

Figure 7: Estimation of the position of the curb: a) Vertical gradient image, b) Thresholded, c) Canny Detector.

4. EXPERIMENTAL RESULTS

The chosen configuration for this system uses the center and left sensors of the device, leading to a baseline of *120* millimeters, with disparity values ranging between [*100*,*260*]. To test the validity of the developed system, ground truth data, in form of the real height of the curbs, was collected on a *140* meter stretch. The measurements where conducted at the end of each curb stone, which gives 140 measurements, as one general curb stone is *1* meter long. The measured data is plotted in Figure 8 as the blue graph. Running the program over the stereo images of the stretch of *440* frames gives the results plotted in red. Approximately 3 measurements where generated per curbstone, therefore the data has been averaged with a moving average of window size 3 to obtain the results. Overall, the Sum Absolute Error (SAE) of the height estimation is *1.0298* centimeters, while the Root Mean Square (RMS) is *1.63* centimeters.

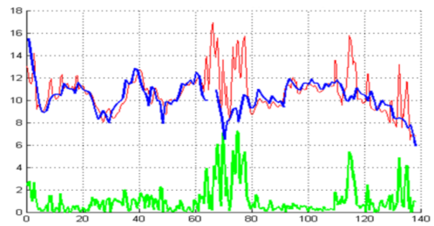


Figure 8: Comparing the height ground truth data with the results of our algorithm, blue, red, and green, represent, ground truth data, output, and absolute error, respectively. On x and y axis are positions in meters, and heights in cm, respectively.

The observed spikes in Figure 8 are caused by improper line selection. The Hough Transform applied to the V-Disparity is a very sensitive step in the algorithm. If the lines are not precisely selected, the computed plane equations will yield results which vary greatly from the ground truth. This is caused because a rather basic selection of the output from the Hough Transform is employed. Since the two strongest lines only give information about the ground and sidewalk planes, the imprecision from their selection will directly affect the curb candidate selection.

5. CONCLUSION

This paper has introduced a system using stereoscopic vision techniques for measuring roadsides' curbs that is of great importance for many municipalities. The proposed system utilizes the geometry of the scene to label each pixel from the image as *ground, sidewalk,* or *curb*. The pixels labeled as *ground* or *sidewalk* are computed from the intrinsic parameters of the stereo setup and from the disparity data. These coordinates are then used to find the two planes representing the ground and the sidewalk. The approximate curb region is found by thresholding the gradient image and employing the Hough transform for finding the two strongest lines. The distance from the ground plane to the sidewalk plane is then used to estimate the height of the curb. The experimental results on real data that have been collected during this work show that the proposed systems can efficiently detect the roadside curbs and estimate their height.

6. ACKNOWLEDGEMENT

We hereby acknowledge the support of Niels Heeser Nielsen from *iLocator*, in Arhus, Denmark.

REFERENCES

[1] 1. Xiaoye, L., Manduchi, R.: Detection and localization of curbs and stairways using stereo vision. International Conference on Robotics and Automation, 2005.

[2] Oniga, F., Nedevschi, S., Meinecke, M.: Curb detection based on a multi-frame persistence map for urban driving scenarios. 11th International IEEE Conference on Intelligent Transportation Systems, pp. 67–72, 2008.

[3] Oniga, F., Nedevschi, S.: Polynomial curb detection based on dense stereovision for driving assistance. 13th International IEEE Conference on Intelligent Transportation Systems, pp. 1110–1115, 2010.

[4] Oniga, F., Nedevschi, S.: Curb detection for driving assistance systems: A cubic spline-based approach. Intelligent Vehicles Symposium, pp. 945–950, 2011.

[5] Balta, H., Cubber, G.D., Doroftei, D., Baudoin, Y., Sahli, H.: Terrain traversability analysis for off- road robots using time-of- flight 3d sensing. 7th IARP International Workshop on Robotics for Risky Environment-Extreme Robotics, 2013.

[6] Labayrade, R., Aubert, D., Tarel, J.P.: Real time obstacle detection in stereovision on non flat road geometry through v-disparity representation. Intelligent Vehicle Symposium, pp. 646–651, 2002.

[7] Cubber, G.D., Doroftei, D.: Multimodal terrain analysis for an all-terrain crisis management robot. In: IARP HUDEM, 2011.

[8] Hu, Z., F, L., K, U.: A complete u-v-disparity study for stereovision based 3d driving environment analysis. 5th International Conference on 3-D Digital Imaging and Modeling, pp. 204–211, 2005.

[9] Bao, G.: Road Distress Analysis using 2D and 3D Information. PhD thesis, University of Toledo, 2010.