

Stereo-based Road Obstacle Detection and Tracking

Intae Na*, Sang Hyun Han*, Hong Jeong*

*Department of Electronic and Electrical Engineering, Pohang University of science and technology,
790-784 san 31, Hyoja-dong, Nam-gu, Pohang Gyungbuk, Korea.

bluewing@postech.ac.kr, han567@postech.ac.kr, hjeong@postech.ac.kr

Abstract— We present a novel visual obstacle detection and tracking system based on stereo vision. A robust stereo matcher, an obstacle detector and a tracker module are implemented and tested under actual driving conditions. Implemented system shows reliable range estimation, detection and tracking performance. Our system has showed 82.4 percent correctly detected rate in complex traffic situations.

Keywords— Smart car, Stereo matching, Obstacle detection, Obstacle tracking, Hierarchical belief propagation

I. INTRODUCTION

Stereo vision is widely used in existing vision-based smart cars and driver assistance systems [1], [2]. Range information from stereo vision is indispensable for many vehicle applications such as autonomous navigation, stop-and-go, obstacle localization and collision avoidance.

In moving vehicle applications, vision systems are required to have both high accuracy and fast processing speed, which is a very difficult trade-off problem, especially for stereo vision-based systems. Global matching approaches are usually avoided in vehicle applications because they require heavy computational burden, making real-time system intractable. However, current stereo vision studies based on parallel processing offer accurate depth maps in real-time or semi real-time [3], [4].

In this paper, we introduce global stereo vision-based road obstacle detection and tracking system which shows robust and reliable performance under general traffic environments.

II. METHOD DESCRIPTION

Overall system consists of stereo matcher, object detector and object tracker, as shown in Figure 1. First, stereo images are input to stereo matcher to produce a disparity map. The disparity map is analyzed in object detector to detect obstacles.

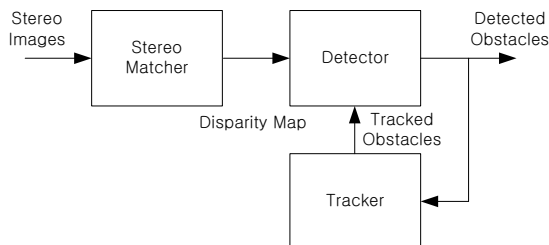


Figure 1. Block diagram of overall system

When obstacles are detected, tracker tracks those obstacles. In the following subsections, we describe each module in detail.

A. Stereo Matcher

Stereo matcher consists of two main parts - matching cost calculator and optimizer. Matching cost calculator is based on Normalized Cross Correlation (NCC) to enhance accuracy and reliability of the system. Matching cost $D(d_p)$ is formulated as follows,

$$D(d_p) = \min(C_d \cdot (1 - |NCC|), K_d), \quad (1)$$

$$NCC = \frac{\sum_{\Delta p} (I_R(p' + d_p) - \mu_R)(I_L(p') - \mu_L)}{\sqrt{\sum_{\Delta p} (I_R(p' + d_p) - \mu_R)^2 \sum_{\Delta p} (I_L(p') - \mu_L)^2}}, \quad (2)$$

where I_R, I_L are intensity values of right and left images, d_p is disparity, μ_L, μ_R are average intensity within a pixel window. Final matching cost is computed by (1) and fed to an optimizer.

Our optimizer used a global matching method based on the Hierarchical belief propagation (HBP) [5]. Message update and belief computation equation is defined as follows.

$$m_{pq}^l(d_q) = \min_{d_p} \left(V(d_p, d_q) + D(d_p) + \sum_{r \in N(p) \setminus q} (m_{rp}^{l-1}(d_p) - \alpha) \right), \quad (3)$$

$$\hat{d}_q = \arg \min_{d_q} \left(D(d_q) + \sum_{p \in N(q)} m_{pq}^l(d_q) \right). \quad (4)$$

To prevent a smoothing effect across the object boundary, we applied constraints on message propagation using an image gradient. Gradient orientation is quantized into 8 directions, and represented by integer index $g_i \in [1, 8]$ in a counterclockwise manner, as shown in Figure 2. Messages

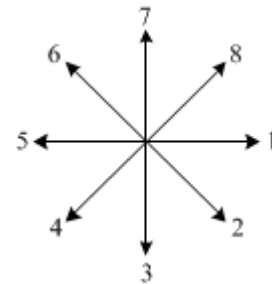


Figure 2. Indexed gradient direction

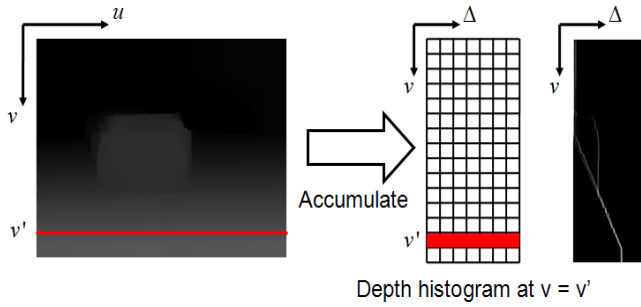


Figure 3. Creating v-disparity image

are constrained by the following procedure:

```

for  $i=1$  to  $MN$ 
  truncate upward message by  $T_u$  if  $g_i=2,3,$  or  $4$ 
  truncate downward message by  $T_d$  if  $g_i=6,7,$  or  $8$ 
  truncate leftward message by  $T_l$  if  $g_i=4,5,$  or  $6$ 
  truncate rightward message by  $T_r$  if  $g_i=1,2,$  or  $4$ 
end

```

In the above procedure, messages across different objects are limited by Truncation parameters, $T_u=T_d=9.7$ and $T_l=T_r=12.7$. M, N are height and width of the image.

B. Obstacle Detector

A v-disparity-based method [6] is used for initial detection. Our stereo vision-based detector has four steps for obstacle detection. First, v-disparity and u-disparity images are constructed using a disparity image from a stereo matcher, as shown in Figure 4. Horizontal and vertical axes of v-disparity images indicate disparity and vertical axis of the disparity map, respectively. Intensity of v-disparity image is the number of pixels which have corresponding disparity values. Figure 3 shows how v-disparity is created in graphics. u-disparity image is a horizontal version of v-disparity image. In the next

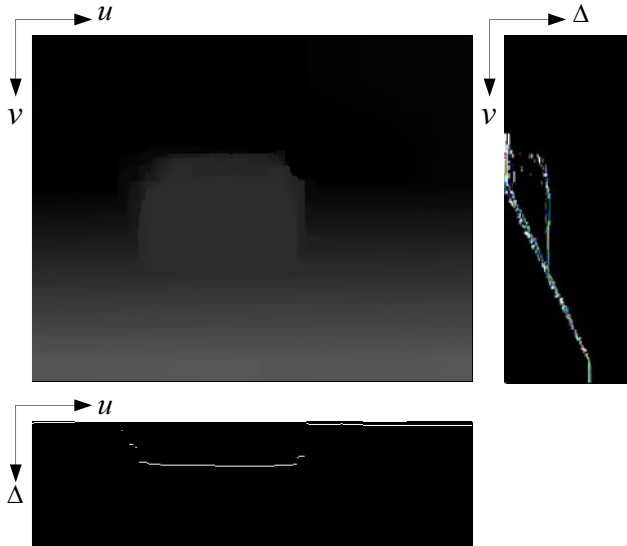


Figure 4. u-v disparity images

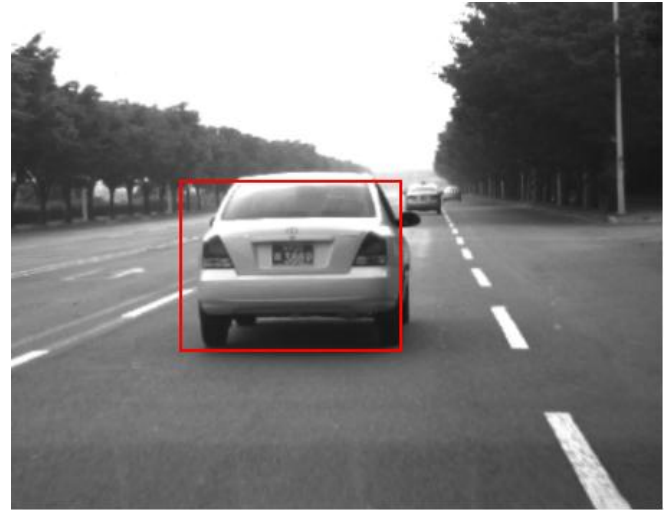


Figure 5. A detection result image

step, we searched road profile by extracting dominant line segment in v-disparity image using Hough transform. Detected road is parameterized by vehicle height h and pitch angle θ . Next, road region is segmented and removed from region of interest and vertical line segments in v-disparity image are extracted using Hough transform. An intersection between the vertical line segment and road profile locates the contact position of the obstacle and the road surface. Height of the road obstacle is assumed to be constant (three meters). Using the assumption and extracted road profile, the height of the bounding box is computed using,

$$v = v_0 - f \tan \theta + \frac{1}{b} \left(h - \frac{C}{\cos \theta} \right) \Delta, \quad (5)$$

$$\Delta v = -\frac{C \Delta}{\cos \theta}, \quad (6)$$

where f is focal length, v_0 is vertical coordinate of zero disparity point in v-disparity image, b is baseline length of stereo vision system, and C is model height. In the final step, the width of the bounding box is found by extracting horizontal line segments from the u-disparity image, as shown in Figure 5.

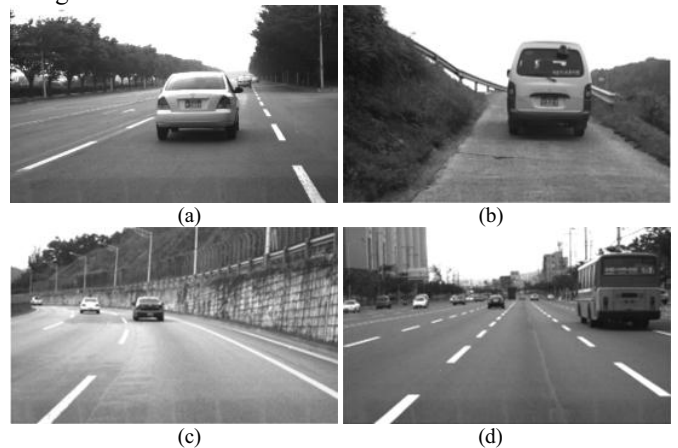


Figure 6. Sample of image sets for experiment. (a)-(d) images are from set 1-4, respectively.



Figure 7. Original image of resulting range images in Figure 8 and 9. Image from (a) sample set 3 and (b) sample set 4

C. Obstacle Tracker

To reduce redundant initial searching for all image regions, we applied detect-then-track strategy [1] to our obstacle detection system.

Pre-detected N obstacles at time t are represented by rectangular boxes in the image, as shown in Figure 5, which have width $w_k(t)$, height $h_k(t)$, and dominant disparity $\Delta_k(t)$ as state parameters, where $k \in [1, N]$. Given state parameters of each detected obstacle, a search region is set up using predefined offset values for width, height and disparity. Offset values adaptively adjusted with respect to $\Delta_k(t)$. Within the cropped search region or the region of interest (ROI), we construct a disparity histogram within the region. Most frequently appearing disparity value is selected as the box disparity, and the result is verified by local u-v disparity analysis. Local u-v disparity images are constructed using

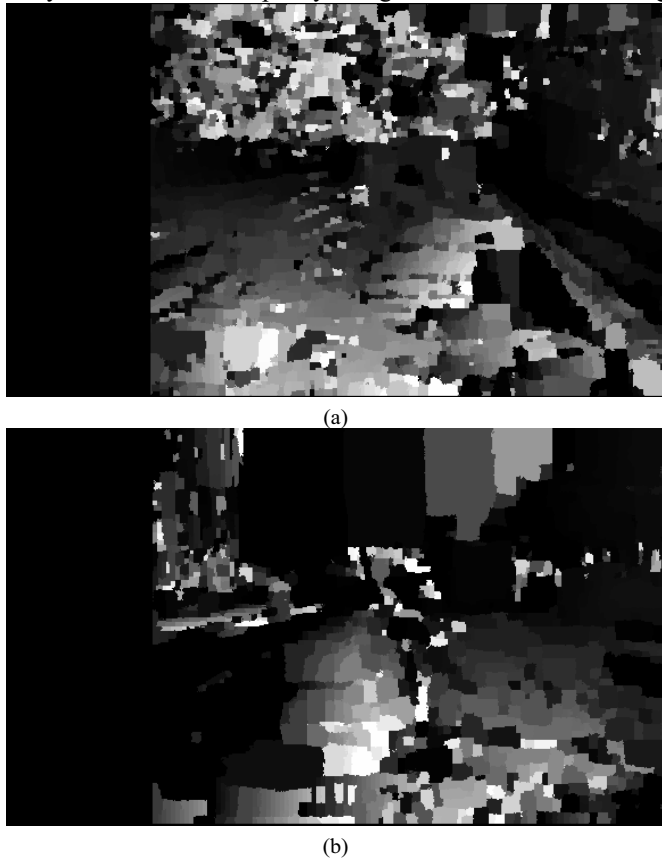


Figure 8. Resulting range image of (a) Figure 7a and (b) 5b from conventional BP stereo matcher

only disparity values within ROI, which aims to reduce the overall search time. State parameters of tracked obstacles are stored in memory and corresponding image regions are eliminated to prevent repeated detection by the obstacle detector.

III. EXPERIMENTAL RESULTS

Our system is realized using Microsoft's Visual Studio. Two cameras and a laptop with 2GHz dual core processor are used for image acquisition and processing. Experimental images are carefully selected to measure the performance in more realistic situations. Four sets of image sequence are selected for evaluation, as shown in Figure 6.

Figure 8 and Figure 9 show sample of range images from conventional system and our system, respectively, applied on images in Figure 7. Even state-of-art global matching system fails in realistic situation due to bad illumination condition, while our system shows far more reliable result.

Table 1 shows evaluation of detection performance. The performance is measured using correctly detected rate and number of type I (false positive) and type II (false negative) errors. Obstacle instances are generated manually for all image sets. In set 1, scenes consist of large illumination and high contrast. Also, reflecting regions exist in the frontal road and vehicles, which possibly break down the stereo matching result. Trees near the driving vehicle caused most of type I errors, since they have strong vertical profiles as

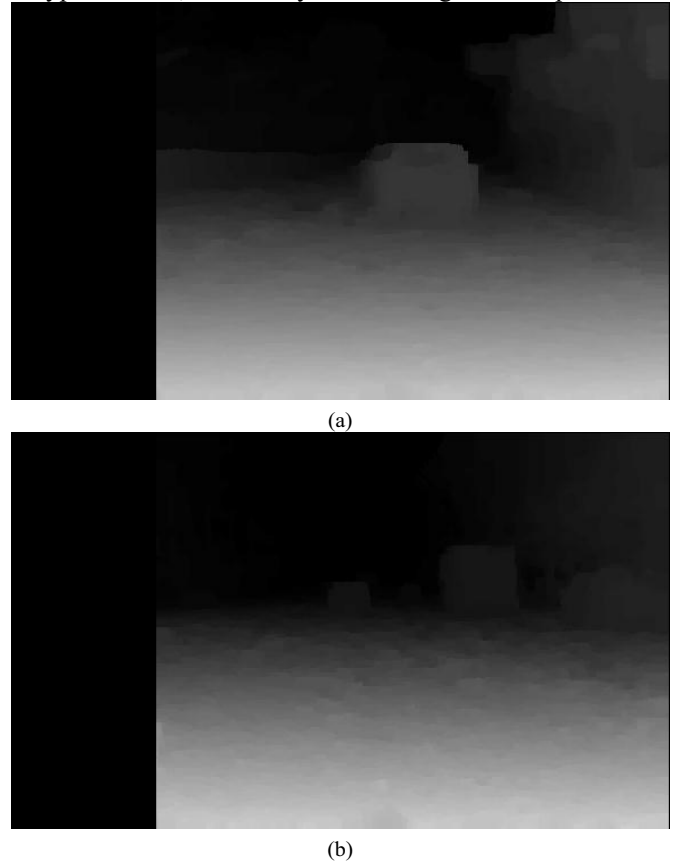


Figure 9. Resulting range image of (a) Figure 7a and (b) 5b from suggested robust stereo matcher

TABLE 1. EVALUATION OF OBSTACLE DETECTION SYSTEM

Set	Frames	Correct (%)	Type I	Type II
1	70	75.3	59	23
2	41	100.0	37	0
3	299	73.0	185	154
4	110	81.1	5	110

our obstacle model. Set 2 consists of unpaved road and also has illumination and contrast problems similar to set 1. Only a single road obstacle is presented, so detection performance has no meaning. But we also observed many type I errors by artificial structures in the image such as guiding rails, bushes, road signs, and small hills. The road in set 3 has median strip and elevation in the vicinity of road. Detection rate is fair enough, but type I and type II errors occurred because roadside structures interrupted detection of road obstacles due to their strong and coherent vertical profiles. Set 4 is the most challenging sequence, because the scenes are very cluttered and complicated. Important obstacle may have various shape and size with respect to their location and pose variation. Roadside structures and buildings in the background region are less coherent than set 3, hence the result shows only few type I errors. On the other hand, object variety and large number of instances caused many type II errors. Average correctly detected rate of our detection and tracking system was 82.4 percent.

IV. CONCLUSIONS

New stereo-based road obstacle detection and tracking system is proposed. The depth information in outdoor environment was successfully extracted with high accuracy and reliability using robust global matching technique.

The high correctly detected rate of detection and tracking system led us one step closer to the ultimate autonomous smart cars. To achieve more reliable system, further effort to reduce false alarms from obstacle-like structures and background objects is needed.

ACKNOWLEDGMENT

This work was supported by the following funds: the Brain Korea 21 project and the Ministry of Knowledge Economy under Human Resources Development Program for Convergence Robot Specialists.

REFERENCES

- [1] Z. Sun, G. Bebis, and R. Miller, "On-Road Vehicle Detection : A Review," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 28, no. 5, May, 2006, pp. 694-711.
- [2] M. Bertozzi, and A. Broggi, "GOLD : A Parallel Real-Time Stereo Vision System for Generic Obstacle and Lane Detection," *IEEE Trans. on Image Processing*, vol. 7, no. 1, Jan. 1998, pp. 62-81.
- [3] Q. X. Yang, L. Wang, and R. G. Yang, "Real-Time Global Stereo Matching Using Hierarchical Belief Propagation," *The British Machine Vision Conference*, 2006.
- [4] S. Park and H. Jeong, "Real-Time Stereo Vision FPGA Chip with Low Error Rate," *In Intl' Conference on Multimedia and Ubiquitous Engineering*, pp. 751-756, 2007.
- [5] P. F. Felzenszwalb and D. R. Huttenlocher, "Efficient Belief Propagation for Early Vision," *In CVPR'04*, no. 1, pp. 1261 - 1268, 2004.
- [6] R. Labayrade, D. Aubert, and J. P. Tarel, "Real Time Obstacle Detection in Stereovision on Non Flat Road Geometry Through "V-disparity" Representation," *In IVS'02*, Versailles, France, pp. 646-651, 2002.