Stereo Vision-Based Pedestrian Detection Using Dense Disparity Map-Based Detection and Segmentation

Chung-Hee Lee*, Dongyoung Kim
Daegu Gyeongbuk Institute of Science & Technology(DGIST), Daegu, KOREA

ABSTRACT

In this paper, we propose a stereo vision-based pedestrian detection method using a dense disparity map-based detection and segmentation algorithm. To enhance a pedestrian detection performance, we use a dense disparity map extracted from a global stereo matching algorithm. First, we extract a road feature information from the dense disparity map, which is a decision basis of presence or absence of obstacles on the road. It is very important to extract the road feature from the disparity for detecting obstacles robustly regardless of external traffic situations. The obstacle detection is performed with the road feature information to detect only obstacles from entire image. In other words, pedestrian candidates including various upright objects are detected in the obstacle detection stage. Each obstacle area tends to include multiple objects. Thus, a disparity map-based segmentation is performed to separate the obstacle area into each obstacle accurately. And then, accurate pedestrian areas are extracted from segmented obstacle areas using road contact and pedestrian height information. This stage enables to reduce false alarms and to enhance computing speed. To recognize pedestrians, classifier is performed in each verified pedestrian candidate. Finally, we perform a verification stage to examine the recognized pedestrian in detail. Our algorithms are verified by conducting experiments using ETH database.

Keywords: dense disparity map, stereo vision, pedestrian, segmentation, detection

1. INTRODUCTION

Recently, many intelligent functions have been applied to vehicle to provide drivers with safety and convenience. Several functions including lane detection, adaptive cruise control, forward warning system are already installed in various vehicles. However, the applied technologies are needed to enhance to offer better functions for drivers. Particularly, there are still a lot of problems about reliability in obstacle recognition. Thus, many researchers are trying to overcome these problems by using some methods, for example, usage of various sensors or complicated algorithms. Using various sensors can be a good solution of solving the problems. But there are restrictions caused by some reasons such as cost, installation and power problem, etc. Vision sensor has a lot of advantages, for example, low cost, small size, providing of many information, etc. Thus, vision sensor is considered to be essential in intelligent vehicle. From now on, the purpose of camera is a simple object detection or environment monitoring. However, the purpose has become more diverse; vehicle or pedestrian detection, traffic sign recognition, etc[1-4]. It is very difficult to detect objects in various traffic situations robustly using only single camera. Thus, stereo vision system which can offer 3D information is utilized to enhance a detection performance. In addition, using a dense disparity with precise and accurate range information is indispensable to ensure reliability of intelligent function in automotive application. There are a lot of detection methods of using the disparity map, for example, u/v-disparity, column detection[5-7]. And the disparity also can be utilized as one kind of features for classifier or very important information for visual tracking. Many studies about how to extract the disparity map easily or how to use the disparity effectively for enhancing a detection performance have still been in progress. In this paper, we propose a stereo vision-based pedestrian detection method using a dense disparity map-based detection and segmentation algorithm. We use a dense disparity map to enhance a pedestrian detection performance. Our algorithm consists of several stages; road feature extraction from v-disparity, obstacle detection with road feature, disparity map-based obstacle segmentation, accurate pedestrian area detection from road contact and pedestrian height information, classification and verification. First, we calculate a v-disparity map from a dense disparity map by accumulating frequency of disparity value in the horizontal direction. And then, we extract a road feature from the v-disparity map using the most frequent values. The obstacle detection is performed with the extracted road feature to detect all of obstacles on the road including various upright obstacles. There are still many

*corresponding author-chlee@dgist.ac.kr

Sixth International Conference on Graphic and Image Processing (ICGIP 2014),
edited by Yulin Wang, Xudong Jiang, David Zhang, Proc. of SPIE Vol. 9443, 94430D
© 2015 SPIE · CCC code: 0277-786X/15/$18 · doi: 10.1117/12.2178695

obstacles in single detection area. Thus, the disparity map-based segmentation is performed to divide or merge the detected obstacles accurately. We extract accurate pedestrian areas in the segmented areas using road contact and pedestrian height information. The classification is performed in each pedestrian candidate to recognize pedestrians using SVM classifier with gradient and color features. Finally, the verification is performed to verify the recognized pedestrians using texture and shape information.

2. DENSE DISPARITY MAP-BASED PEDESTRIAN DETECTION

2.1 Stereo Vision System Modeling

A stereo vision system modeling is shown in Fig. 1. Two cameras are located "*h*" from the road and tilted toward ground by "Θ" and the base line is "*b*". We define that the world coordinate system is $R_w(X_w, Y_w, Z_w)$ and also define that the left and right camera coordinate systems are R_{cl}, R_{cr}, respectively. And the center of the image plane(u, v) is defined as (u_0, v_0). The mapping on the image plane of a point, $P(X, Y, Z, 1)^T$ in the world coordinate is achieved by several transformations[5-7]. The relationship between the image plane and world coordinate is expressed by

$$\begin{aligned} u_l &= u_0 + \alpha \frac{X - b/2}{(Y+h)\sin\theta + Z\cos\theta}, \\ u_r &= u_0 + \alpha \frac{X + b/2}{(Y+h)\sin\theta + Z\cos\theta}, \\ v &= v_0 + \alpha \frac{(Y+h)\cos\theta - Z\sin\theta}{(Y+h)\sin\theta + Z\cos\theta}, \end{aligned} \tag{1}$$

where $\alpha = fm$ is the focal length of the camera in terms of the pixel dimensions in the u and v directions. f is the focal length and m is the number of pixels per unit distance in the image.

2.2 Road Feature Extraction and Obstacle Detection

Fig. 2 represents the flow chart of the road feature extraction and obstacle detection algorithm. First, we calculate the v-disparity map from the dense disparity map by accumulating frequency of disparity value in the horizontal direction. In the v-disparity map, obstacles are represented as vertical lines and a ground surface is represented as a diagonal line[7]. Thus, if we separate the vertical line and the diagonal line in the v-disparity map, we can detect only obstacles in the disparity map. However, it is not easy to separate two types of lines robustly in multiple obstacles or non-flat road situations. Above all, it is difficult to extract the diagonal line accurately by a general line detector. Thus, we have two steps in detection; road feature extraction from v-disparity and obstacle detection with the road feature. In other words, we extract only the diagonal line from the v-disparity and detect only obstacles with the road feature in the disparity map. To extract the diagonal line, namely the road feature information, we binarize the v-disparity map using the most frequent value in row and column direction. Generally, the road feature's values in v-disparity map is higher than other values. Thus, if we choose the most frequent values in each row and column, the values are those of consisting of the

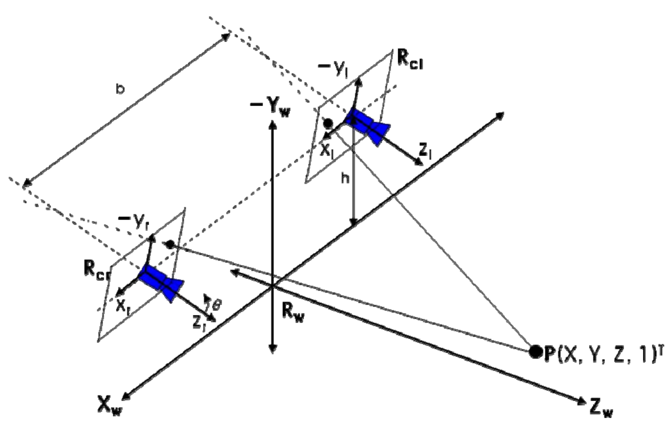


Figure 1. The stereo vision system modeling.

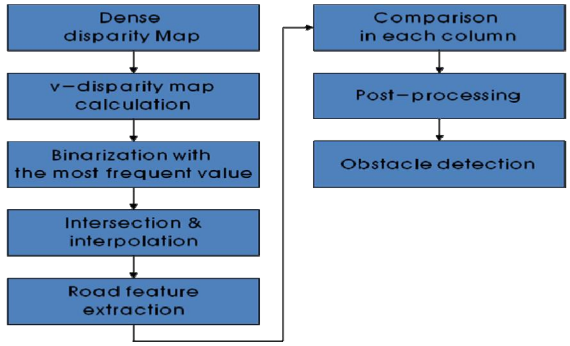
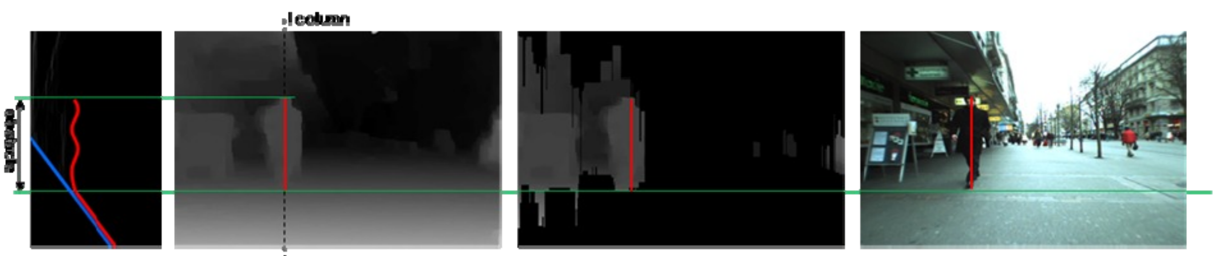


Figure 2. The flow chart of the road feature extraction and obstacle detection algorithm.

road feature. After intersecting binary images to select definitive values, an interpolation is conducted to estimate empty values. To find all obstacles on the road, all columns in the disparity map are compared with the extracted road feature information. Fig. 3 shows the concept of the obstacle detection with the road feature information. If there is a successive interval whose disparity values are higher than those of the road feature in the v-disparity map, the interval is regarded as the obstacle. Fig. 3(a) shows the v-disparity map which represents the road feature information(blue line) and disparity values of arbitrary column(red line). Fig. 3(b) shows the disparity map with one obstacle line(red line) and Fig. 3(c) is the result of obstacle detection from comparison all columns with the road feature. One obstacle line is also marked in the input image(Fig. 3(d)).



(a) (b) (c) (d)

Figure 3. The obstacle detection with the road feature information.

2.3 Disparity Map-Based Segmentation

Even though we use the dense disparity map to detect obstacles accurately, each obstacle area tends to include multiple obstacles because of disparity map resolution or noises. Thus, we need to another processing to divide or merge obstacles in each obstacle area. The disparity map-based bird's-eye-view mapping segmentation is utilized to segment obstacles accurately. First, the areas are converted to bird's-eye-view map to segment obstacles easily. The mapping equation is as follows:

$$
\begin{aligned}
X &= (u_l + u_r - 2u_0)\frac{\{(Y+h)\sin\theta + Z\cos\theta\}}{2\alpha}, \\
Z &= \frac{(v - v_0)b\cos\theta + \alpha b\sin\theta - dh}{d}, \\
Y &= 0.
\end{aligned} \tag{2}
$$

The bird's-eye view map represents the obstacles on a planar X-Z plane. Thus, it is very easy and simple to divide the obstacles using histogram-based clustering algorithm. After clustering, the bird's-eye-view map is converted again to the disparity map using inverse-birds'-eye-view mapping with index matrix. Because the mapping equation is irreversible, the index matrix calculated previously in the bird's- eye- view mapping should be utilized in the inverse mapping.

2.4 Pedestrian Area Detection and Recognition

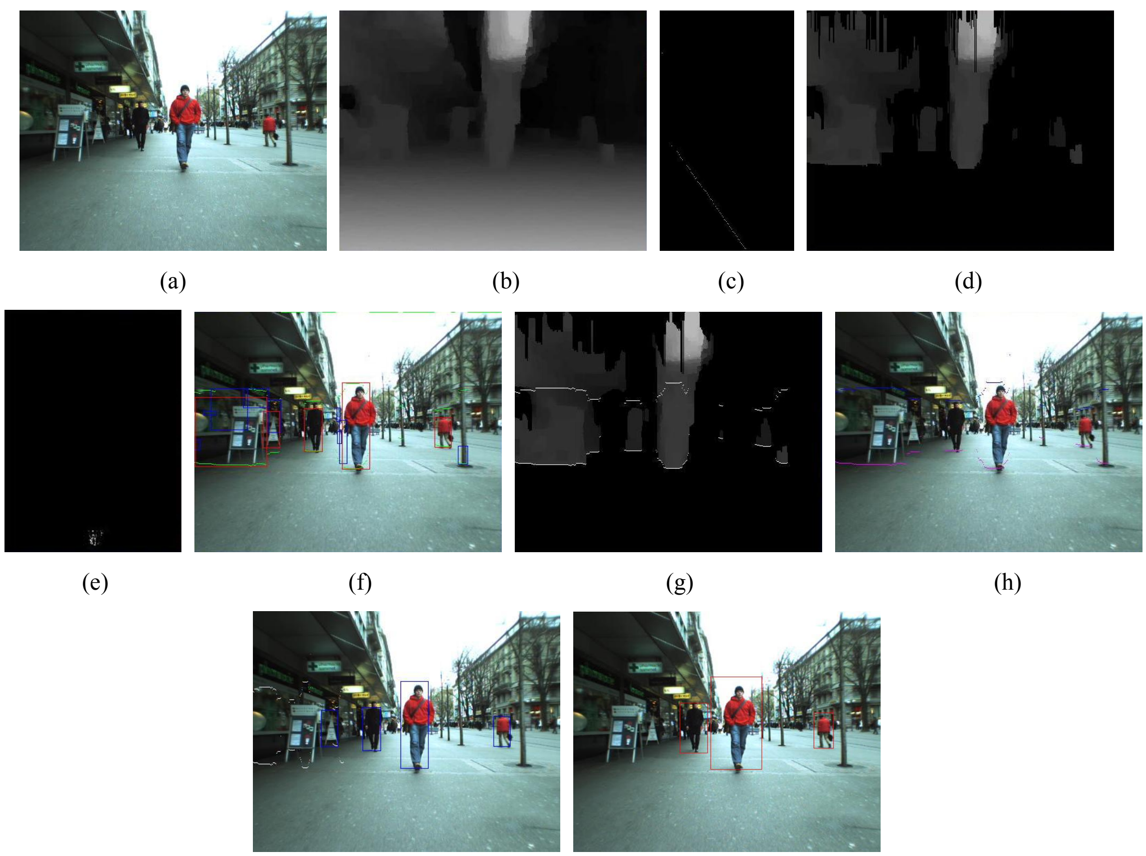
Accurate pedestrian areas are extracted from segmented obstacle area using road contact and pedestrian height information. First, the ground surface is estimated using the road feature information. And the pedestrian areas are calculated using the pedestrian height information. The pedestrian height in the image, $v_1 - v_2$ is calculated by

$$v_1 - v_2 = H \frac{d}{b \cos \theta}, \tag{3}$$

where "H" is the real pedestrian height and "d" is disparity value. "b" and "Θ" are the base line and pitch, respectively. Thus, the segmented obstacle areas are refined by road contact and pedestrian height information. This processing enables to reduce false alarms and to enhance computing speed by handing over certain pedestrian areas to classifier stage. The classification is performed in each candidate to recognize pedestrians using SVM classifier with HOG(Histogram of Oriented Gradients)[8] and CSS(Color Self Similarity) features[9]. Finally, the verification is performed to verify the recognized pedestrians using texture and shape information.

3. EXPERIMENTS

Fig. 4 represents experimental results of pedestrian detection using ETH database. Fig. 4(a) represents a left image and Fig. 4(b) represents a dense disparity map calculated by a belief propagation stereo matching algorithm. Fig. 4(c) represents a road feature information from a v-disparity map binarized by the most frequent values. Only diagonal line is extracted in the v-disparity map properly. The obstacle detection result obtained by comparing the road feature information with all columns of the disparity map is shown in Fig. 4(d). And then, the bird's-eye-view mapping of the obstacle detection result is shown in Fig. 4(e). The segmentation result from the bird's-eye-view mapping is shown in Fig. 4(f). The pedestrian area detection result obtained by the road contact and pedestrian height information is shown in Fig. 4(g) and 4(h). The results of pedestrian recognition and verification are shown in Fig. 4(i) and Fig. 4(j), respectively.



(a) (b) (c) (d)

(e) (f) (g) (h)

(i) (j)

Figure 4. The experimental results of pedestrian detection.

4. CONCLUSION

In this paper, we proposed a stereo vision-based pedestrian detection method using a dense disparity map-based detection and segmentation algorithm. Our algorithm consisted of several stages; road feature extraction from v-disparity, obstacle detection with road feature, disparity map-based obstacle segmentation, accurate pedestrian area detection, classification and verification. Thanks to the dense disparity map-based detection and segmentation processing, the pedestrian detection performance was improved, which can be suitable for automotive applications. Particularly, this processing enabled to reduce false alarms and to enhance computing speed by handing over certain pedestrian areas to classifier stage. However, our method is needed to be verified in various database and we will enhance our classifier by using more features such as motion[9], disparity, symmetry and additional color information[10], etc in the future.

ACKNOWLEDGMENT

This work was supported by the DGIST R&D Program of the Ministry of Science, ICT and Technology of Korea(14-IT-01).

REFERENCES

[1] S. Zehang, B. George, and M. Ronald, "On-Road Vehicle Detection: A Review," IEEE Transactions on Pattern Analysis and Machine Intelligence 28(5), 694-711, (2006).

[2] P. Dollar, C. Wojek, B. Schiele, and P. Perona, “Pedestrian detection: An Evaluation of The State of The Art”, IEEE Transactions on Pattern Analysis and Machine Intelligence 34(4), 743-761, (2012).

[3] P. Dollar, S. Belongie, and P. Perona, “The Fastest Pedestrian Detector in The West”, British Machine Vision Conference, (2010).

[4] Benenson R., Mathias M., Timofte R., and Van Gool L, “Pedestrian Detection at 100 Frames Per Second”, Computer Vision and Pattern Recognition, 2903-2910, (2012).

[5] R. Labayrade, D. Aubert, and J.-P. Tarel, "Real Time Obstacle Detection in Stereovision on Non Flat Road Geometry Through V-disparity Representation," Proc. IEEE Int. Vehicles Symposium, 646-651, (2002).

[6] CH Lee, YC Lim, S Kwon and J Kim, "Stereo Vision-Based Obstacle Detection Using Dense Disparity Map," ICGIP2011, (2011).

[7] CH Lee, YC Lim, S Kwon and JH Lee, "Stereo Vision-Based Vehicle Detection Using A Road Feature and Disparity Histogram," Optical Engineering 50(2), 027004-027004-23, (2011).

[8] Dalal, N. and Triggs, B., "Histograms of Oriented Gradients for Human Detection," Computer Vision and Pattern Recognition, 886-893, (2005).

[9] S. Walk, N. Majer, K. Schindler, and B. Schiele, “New Features and Insights for Pedestrian Detection,” Proc. IEEE Conf. Computer Vision and Pattern Recognition, 1030-1037, (2010).

[10] P. Dollar, Z. Tu, P. Perona, and S. Belongie, “Integral Channel Features”, British Machine Vision Conference, (2009).