Disparity Map Refinement using Occlusion Handling for 3D Scene Reconstruction

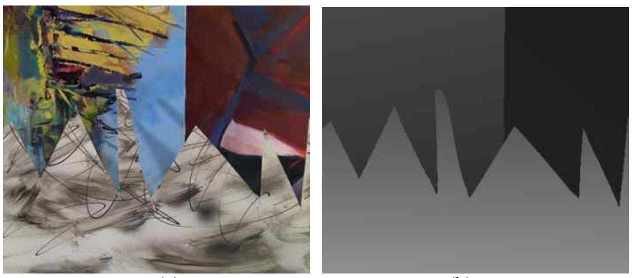
Woo-Seok Jang and Yo-Sung Ho

Gwangju Institute of Science and Technology (GIST)
261 Cheomdan-gwagiro, Buk-gu, Gwangju, 500-712, Korea
Telephone: +82-62-715-2263, Fax: +82-62-715-3164
E-mail: {jws, hoyo}@gist.ac.kr

***Abstract*-An occurrence of occlusion is the challenge problem in stereo matching. In this paper, we propose a refinement algorithm of disparity maps with occlusion handling. In order to refine the disparity map, we obtain an initial disparity map with an optimization algorithm based on the modified constant-space belief propagation (CSBP) which has low complexity. An occlusion map is computed using unique constraint from initial left and right disparity map. We classify occlusion into two types from obtained occlusion map and apply suitable occlusion handling process, respectively. Proposed handling process based on potential energy function extends the disparity values of the visible pixels to the occluded pixels. Experimental results show that the proposed algorithm is confirmed to be valid and can be widely used for 3D scene reconstruction.**

I. INTRODUCTION

Stereo matching is a widely researched topic in computer vision. Stereo matching finds corresponding pixels in two images. The final result of stereo matching represents disparity map which is a set of the displacement vectors between the corresponding pixels. Two images which are captured from different positions are given and it is assumed that these images are rectified for simplicity and accuracy of the stereo matching. From this assumption, corresponding points are found in same horizontal line of two images. A disparity map acquired by stereo matching can be represented by a gray scale image. Fig. 1 shows one example of a disparity map. Depth of each pixel is perceived from the disparity map. The object is close to viewpoint as intensity value of a pixel in the disparity map is high.



(a) (b)

Figure 1. Example of disparity map. (a) Color image, (b) Disparity map

In general, stereo matching algorithms can be categorized into two approaches: local methods and global methods. The local methods are generally efficient for complexity [1]. However, it makes blurred object border and the removal of small detail at the depth discontinuity. In order to solve this problem, global methods have been proposed [2]. Global methods define energy function by Markov random field and optimize this function using several optimization algorithms such as belief propagation [3] and graph cut [4] [5]. In this paper, we used global method optimized by a belief propagation based method. However, a belief propagation algorithm is too computationally complex even for low resolution image and small number of disparity level. No matter how good performance is, it is difficult to use too complex algorithm in real life. Recently, several methods have been proposed to reduce the complexity of belief propagation [6]-[8]. We choose constant-space belief propagation (CSBP) [8] which is the most recently presented and reduces complexity significantly for our method.

Occlusion is important in stereo matching. Since two images are captured from different position, occlusion is generated. Occluded parts in the image are only visible in an image. Therefore, accurate estimation of the disparity value is difficult in this part. However, accurate assignment of disparity values in an occlusion part makes good result for final disparity map.

In this paper, we propose a disparity map refinement method using occlusion handling. We use cross-checking method for occlusion detection. We use probability model to refine the disparity map. In fact, even if proposed method has low complexity, the performance is reasonable for 3D scene reconstruction.

II. OCCLUSION PROBLEM

Occlusion is the challenging part in stereo matching. Fig. 2 illustrates the case of occlusion occurrence.

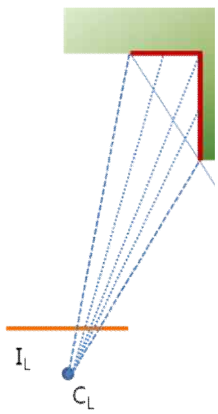


Figure 2. Occurrence of occlusion

Two kinds of constraints: the ordering constraint and the uniqueness constraint have been typically used for occlusion handling in stereo matching. The ordering constraint preserves the order of matching along the scanline in two input images [9]. The ordering constraint has limitation. It is violated in image that contains thin objects or narrow holes. Fig. 3 shows an example of violation for ordering constraint. The circled stick in the left image is to the left of the letter ‘u’. However, the circled stick in the right image is to the right of the letter ‘u’. Therefore we use the cross-checking method which is a kind of the uniqueness constraint for occlusion detection. The uniqueness constraint uses the fact that the corresponding points between two input images are one-to-one mapping. The uniqueness constraint can be iteratively applied to global optimization method.

(a) (b)

Figure 3. Violation of ordering constraint. (a) Left image, (b) Right image

III. PROPOSED METHOD

A. *Overall Framework*

Fig. 4 represents the overall framework of the proposed algorithm. First, initial disparity map is obtained for the left image and the right image. CSBP that has low complexity is used for optimization of initial disparity map. Occlusion is detected using cross-checking method [10] from both disparity maps. Estimation of disparity value is performed at the occluded pixels. Finally, disparity map with occlusion handling is generated.

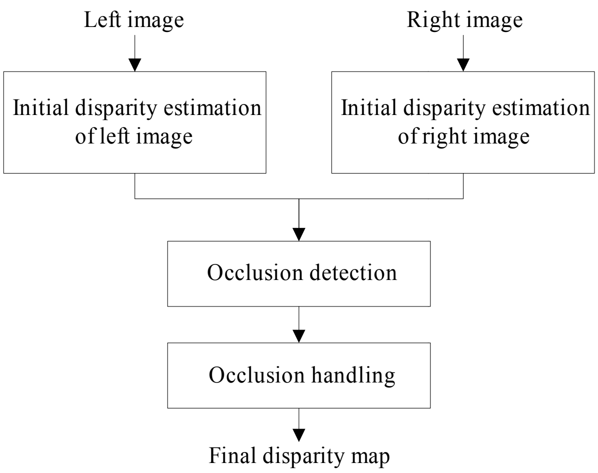


Figure 4. Overall framework of our proposed algorithm

B. *Occlusion Detection*

In order to detect the occlusion, firstly initial disparity is obtained. An initial disparity map is calculated by CSBP. CSBP hierarchically reduces the disparity search range and fixes the number of disparity levels on the original resolution [8]. The complexity of CSBP depends on only constant space, that is O(1). Therefore CSBP is efficient for real application. However, CSBP cannot make enough quality for result. We need to refine the initial disparity map by CSBP.

We apply occlusion handling for disparity map refinement. Occlusion maps of the left and right images are obtained from initial disparity map. Cross-checking method evaluates the mutual consistency from both initial disparity maps. If a particular pixel in the left image is not occluded pixel, the disparity values from the left and the right disparity maps should be consistent. Cross-checking method can be represented by (1).

$$D_L(x_s) = D_R(x_s - D_L(x_s)) \tag{1}$$

where D_L and D_R are the left disparity map and right disparity map respectively. x_s is a pixel in left image. If (1) is not satisfied, we assume that that pixel is in the occlusion part. Fig. 5 illustrates the cross-checking method. Corresponding points have the same disparity value in both images.

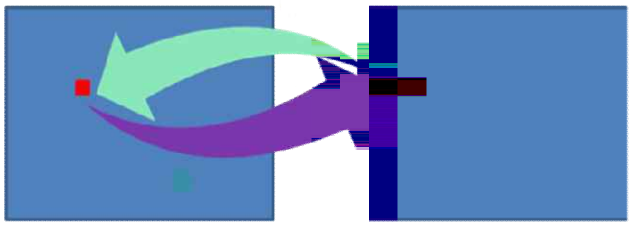


Figure 5. Cross-checking method

C. *Occlusion Handling*

After detecting the occlusion, the reasonable disparity value should be assigned to the occluded pixel. Occlusion is only visible in one image. Thus, it is impossible to determine the accurate disparity value. If we use the disparity values from neighboring pixels in the non-occlusion region, estimation of the disparity value in the occlusion region is possible. Generally, disparity values in occluded pixels are similar to the disparity values of visible pixels in the background. In our proposed method, we try to propagate the visible pixels to occluded pixels.

(a) (b)

Figure 5. Two kinds of occlusion, (a) Color image (b) Occlusion map

First, we classify occlusion regions into two parts called the left-side part and the general part. Fig. 5 shows the left image and the corresponding occlusion map. The circled part in Fig. 5 (b) is the left-side part and the rest of the occlusion is the general part. The left-side part disappears into the left of the image in the right image. Because of this, occlusion in this part is generated. In this part, it is useless to estimate the disparity value using the disparity values of the neighboring pixels, since our algorithm does not use iterative optimization. Thus, we extend the disparity of the leftmost visible pixels to the left-side part for each horizontal line. For the general part, we define the potential energy function. Let L(s) be the neighboring pixels which distance from occluded pixel s is smaller than predefined distance and $C=\{s,t\,|\,s>t, t\in L(s)\}$ be the set of all nearby pixels which affect pixel s. The potential energy function for occlusion handling is defined as

$$E_{OH}(s,\, d_s)=\sum_{t\in C\backslash B}(1-o_t)\frac{1}{dist(s,t)}\exp\left(-\frac{diff_{s,t}}{\sigma^2}\right) \qquad (2)$$

where $B=\{s,t\,|\,d_s\neq d_t, t\in C\}$, o_t is the occlusion value from the obtained occlusion map. $dist(s,t)$ is the distance between occluded pixel s and visible pixel t. $diff_{s,t}$ is the color difference defined in (3).

$$diff_{s,t}=\sum_{c\in\{R,G,B\}}|\,I_c(s)-I_c(t)\,| \qquad (3)$$

The disparity value which has the maximum value of (2) is determined as the disparity value for the pixel s. This process works only occluded pixels which are near visible pixels. Thus, it completely handles the thin or small parts of the occlusion. However, wide and large parts of the occlusion are processed only near the visible pixels. In order to solve this problem, we apply potential energy function for occlusion handling one more time. At this time, we do not consider only visible pixels to assign the disparity value to occluded pixel, since visible pixels are sufficiently propagated to the occluded pixels until previous process. Second potential energy function for occlusion handling is as follows.

$$E_{OH}(s,\, d_s)=\sum_{t\in C\backslash B}\frac{1}{dist(s,t)}\exp\left(-\frac{diff_{s,t}}{\sigma^2}\right) \qquad (4)$$

This function is very similar to (2), however (4) does not consider whether source for occlusion handling is a visible pixel or not. Proposed occlusion handling process assigns disparity values which have maximum value of the potential energy function for occlusion handling without optimization of an energy function. Therefore, its complexity is much lower than the other algorithms which optimize the energy function iteratively. Fig. 6 shows the flowchart of the final occlusion handling process.

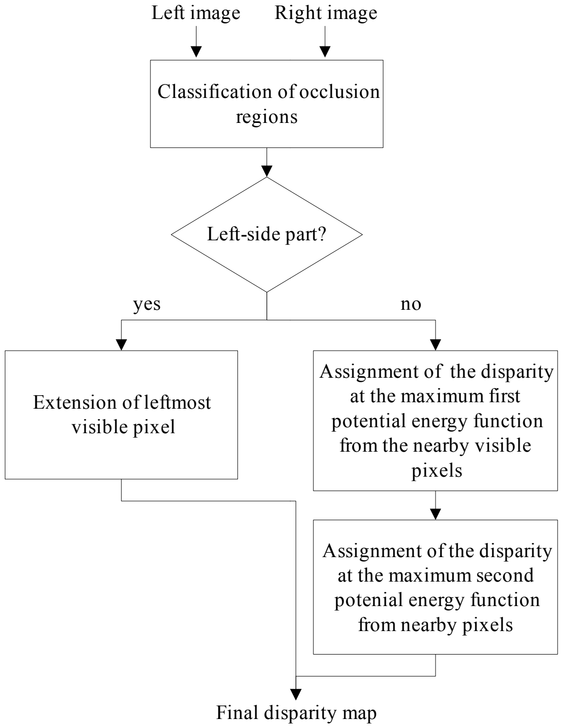
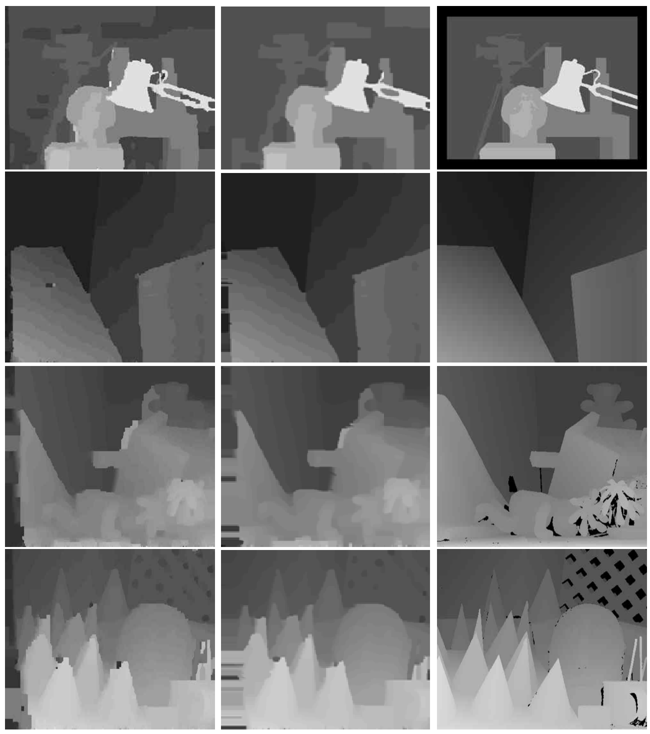


Figure 6. Occlusion handling process

IV. EXPERIMENTAL RESULT

In order to evaluate the performance of our proposed method, we follow the methodology which measures the percentages of bad matching pixels [11]. Fig. 7 illustrates the visual comparison of our disparity maps with disparity maps of CSBP and ground truth.



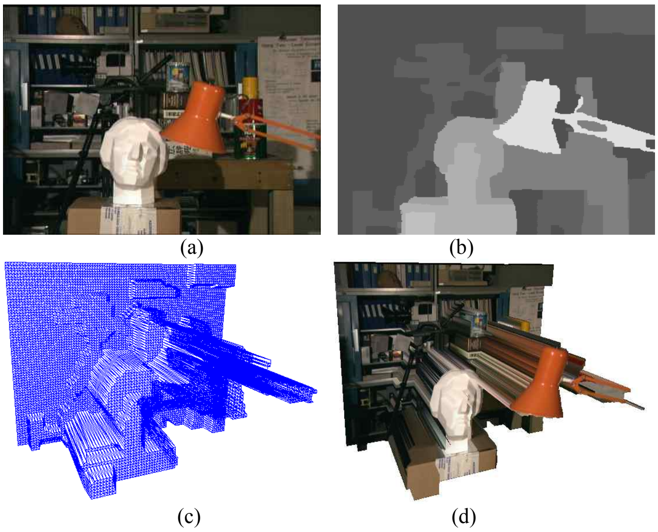
(a) (b) (c)

Figure 7. Final experimental result. (a) CSBP (b) Proposed method (c) Ground truth

Objective evaluation is presented in table I. Table I shows the percentage of the mismatching pixels between the proposed method and ground truth. When the absolute disparity error is greater than 1 pixel, the pixel is regard as the bad matching pixel. The result shows that proposed method improve quality of a CSBP based disparity map.

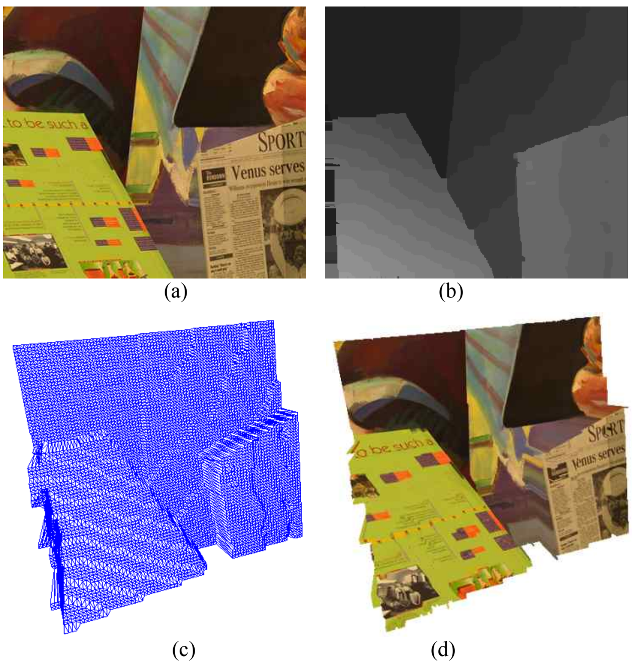
TABLE I
EVALUATION FOR DISPARITY MAP

Image	CSBP	Proposed method
Tsukuba	4.17	3.10
Venus	3.11	2.79
Teddy	20.20	18.02
Cone	16.50	15.49



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

Figure 8. 3D scene reconstruction for Tsukuba. (a) Original color image (b) Disparity map of proposed method (c) Mesh representation for 3D scene (d) 3D scene modeling



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

Figure 9. 3D scene reconstruction for Venus. (a) Original color image (b) Disparity map of proposed method (c) Mesh representation for 3D scene (d) 3D scene modeling

Fig. 8 and fig. 9 shows the original color image, the disparity map from proposed method, and reconstructed model of the 3D scene for Tsukuba and Venus. In order to represent the 3D scene, we used mesh modeling. Laplacian smoothing was applied to all vertices of the mesh. Mesh smoothing by Laplacian smoothing reduces the curvature variation and removes the noise. From the result of 3D scene reconstruction, we can feel sense of distance easily.

V. CONCLUSION

In this paper, we proposed the disparity refinement method considering occlusion. CSBP is the fastest and takes the smallest memory among fast belief propagation algorithms. CSBP was applied for optimization of an initial disparity map basically. Occlusion detection was performed by cross-checking method. Proposed occlusion handling method assigned reasonable disparity values to the occluded pixels. The proposed method is computationally efficient since it does not use iterative global optimization technique at occlusion handling. Experimental results show that the proposed method improves the performance of stereo matching and provides natural 3D scene. The 3D scene model from accurate disparity map can be widely used for many application related to 3D image processing.

ACKNOWLEDGMENT

This research was supported in part by the MCST (Ministry of Culture, Sports and Tourism) and KOCCA (Korea Creative Content Agency) in the CT (Culture Technology) Research & Development Program 2010.

REFERENCES

[1] H. Hirschmuller, P. R. Innocent, and J. M. Garibaldi, “Real-time correlation-based stereo vision with reduced border errors,” International Journal of Computer Vision, vol. 47 no.1/2/3, pp. 229-246, April 2002.
[2] O. Veksler, “Fast variable window for stereo correspondence using integral images,” IEEE Computer Society Conference on Computer Vision and Pattern Recognition, vol 1, pp. 556-561, June 2003.
[3] J. Sun, N. N. Zheng, and H. Y, Shum, "Stereo matching using belief propagation," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 25, no. 7, pp. 787-800, July 2003.
[4] Y. Boykov, O. Veksler, and R. Zabih, ”Fast approximate energy minimization via graph cuts," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 23, no. 11, pp.1222-1239, Nov. 2001.
[5] V. kolmogorov and R. Zabih, “Computing visual correspondence with occlusions using graph cuts,” IEEE International Conference on Computer Vision, pp. 508-515, 2001.
[6] P. Felzenszwalb and D. Huttenlocher, “Efficient belief propagation for early vision,” IEEE Computer Society Conference on Computer Vision and Pattern Recognition, pp.261-268, 2004.
[7] Q. Yang, L. Wang, R. Yang, S. Wang, M. Liao, and D. Nister, “Real-time global stereo matching using hierarchical belief propagation,” British Machine Vision Conference, pp. 989-998, 2006.
[8] Q. Yang, L. Wang, N. Ahuja, "A constant-space belief propagation algorithm for stereo matching," IEEE Computer Society Conference on Computer Vision and Pattern Recognition, pp. 1458-1465, 2010.
[9] A. Bobick and S. Intille, “Large occlusion stereo,” International Journal of Computer Vision, vol.33, no. 3, pp. 181-200, Sep. 1999.
[10] G. Egnal and R. Wildes, “Detecting binocular halfocclusions: empirical comparisons of five approaches,” IEEE Transactions on Pattern Analysis and Machine Intelligence, vol.24 no. 8, pp. 1127-1133, Aug. 2002.
[11] D. Scharstein and R.Szeliski, “A taxonomy and evaluation of dense two-frame stereo correspondence algorithms,” International Journal of Computer Vision, vol. 47, no. 1, pp. 7-42, April 2002.