Efficient Disparity Map Estimation Using Occlusion Handling for Various 3D Multimedia Applications

Woo-Seok Jang and Yo-Sung Ho, *Senior Member*, IEEE

Abstract — *Disparity acquisition is beneficial to various 3D multimedia applications, i.e., free viewpoint TV (FTV), 3DTV, and 3D scene model reconstruction. In this paper, we propose an efficient disparity map estimation method with occlusion handling taken into account. In order to detect occlusion, we obtain left and right initial disparity maps via optimization based on modified constant-space belief propagation, a low complexity algorithm. The initial disparity maps provide clues for occlusion detection. From the clues, an energy function is defined and optimized by energy minimization framework. We define two occlusion types and apply suitable occlusion handling processes, accordingly. The proposed occlusion handling method extends disparity values of visible pixels to occluded pixels. Experimental results show that our method provides near-ground truth occlusion and a superior disparity map compared to other state-of-the-art methods with occlusion handling. Our contribution provides solid advantages regarding disparity map estimation, which is useful for various 3D multimedia applications*[1].

Index Terms — 3DTV, FTV, occlusion handling, stereo matching.

I. INTRODUCTION

Recently, 3D entertainment systems have received a lot of attention mainly due to the financial success of many 3D films. Generally, 3D experience is achieved by the process of left and right eyes seeing slightly different scenes. In other words, 3D perception is derived from two separate views. In order to represent 3D, several data formats have been proposed. Texture-plus-depth approach, one of these formats, uses an ordinary 2D image accompanied by a depth map. As discussed above, two views are necessary for 3D perception. The non-existing view can be synthesized by employing depth image based rendering (DIBR) [1]. Some of the benefits include the flexibility to render views with variable baseline and increased compressibility of depth data [2]. Thus, this format is practical for many 3D multimedia applications.

Depth information can be acquired by several methods: active sensor based, passive sensor based, and hybrid sensor fusion depth estimation. Active sensor based method [3] employs a physical sensor to measure depth data directly, while passive sensor based method extract depth data from 2D images indirectly. Hybrid sensor fusion method combines both methods to make up for weakness of each method [4], [5]. Active sensor based and hybrid sensor fusion methods require expensive physical sensors such as depth cameras. Thus, we adopt a passive sensor based depth estimation that acquires depth data with low price, although occlusion handling is a bit more problematic.

Disparity information can be converted into depth information by using a stereo image pair in combination with triangulation [6]. Stereo matching, widely researched topic in computer vision, is one of the most useful methods to acquire disparity information from stereo images. Stereo matching acquires disparity information by finding the corresponding points in the other image. The correspondence problem is to compute the disparity map which is a set of displacement vectors between corresponding pixels. Essentially, two images of the same scene taken from different viewpoints are given and the images are assumed to be rectified for the sake of simplicity and accuracy. Consecutively, corresponding points are found in the same horizontal line of two images. A disparity map acquired by stereo matching can be represented by a gray scale image. Depth of each pixel is perceived from the disparity map. If an object is close to the viewpoint, the intensity in the disparity map is high.

In general, stereo matching can be categorized into two approaches: local and global methods. Local methods are generally efficient due to its low complexity [7]. However, these methods blur object borders and removes small details at depth discontinuities depending on the size of correlation window. In order to solve this problem, global methods have been proposed [8]. Global methods define an energy function using Markov Random Field (MRF) and optimize this via several optimization algorithms such as belief propagation [9], dynamic programming [10] and graph cut [11], [12]. However, global methods are too computationally complex even for low resolution images and a small number of disparity levels. Thus, they are not suitable for practical use. Recently, several methods have been proposed to reduce the complexity of belief propagation [13]-[15]. We choose constant-space belief propagation (CSBP) [15] which is recently presented and significantly reduces complexity for practical use.

Since stereo images are captured from different positions, occluded pixels which are only visible in one image occur. Therefore, accurate disparity estimation in these pixels is challenging. Moreover, occluded pixel detection and reasonable disparity assignment are crucial. If stereo matching

[1]This research was supported in part by the MKE under the ITRC support program supervised by NIPA (NIPA-2011-(C1090-1111-0003)), and 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 2011.

W.S. Jang and Y.S. Ho are with the School of Information and Communications, Gwangju Institute of Science and Technology (GIST), Gwangju, Korea (e-mail: jws@gist.ac.kr, hoyo@gist.ac.kr).

Contributed Paper
Manuscript received 10/15/11
Current version published 12/27/11
Electronic version published 12/27/11.

is not equipped with occlusion handling, inadequate results will be achieved, not practical for many applications.

In this paper, we propose a stereo matching method with occlusion handling to estimate disparity map efficiently. We consider the low complexity optimization algorithm for practical use. We use cross check and warping constraints to detect occlusion accurately. The obtained occlusion is handled by potential energy functions, stochastically.

The remainder of this paper is organized as follows. In Section II, we briefly explain the occlusion problem in stereo matching and conventional constraints for occlusion handling. In Section III, we explain our proposed occlusion detection and occlusion handling methods. Section IV analyzes experimental results of the proposed method. Finally, our conclusions are presented in Section V.

II. Occlusion Problem in Stereo Matching

Occlusion is an important and challenging problem in stereo matching. Fig. 1 illustrates how occlusion occurs. The bold lines are only visible in the left image, becoming occlusion.

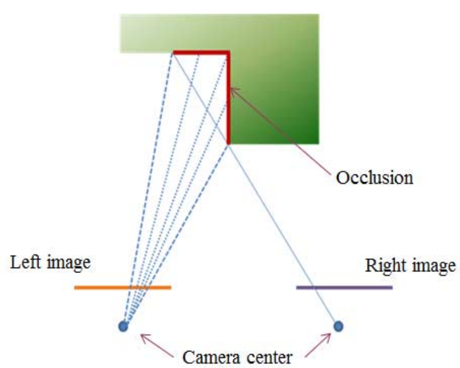


Fig. 1. Occurrence of occlusion in stereo images

The simplest method for occluded pixel detection and disparity estimation uses cross-checking [16] and extrapolation. For each pixel, cross-checking tests the consistency of disparity values from left and right disparity maps, determining occluded pixels for occlusion handling. Subsequently, the disparities of visible pixels are extended into the occluded pixels by extrapolation.

In general, ordering constraint and uniqueness constraint have been typically used for occlusion handling in stereo matching. The ordering constraint provides clues by preserving the order of matching along the scan line in two input images [17]. However, violation occurs around thin objects or narrow holes. Fig. 2 shows an example of such a violation. The uniqueness constraint [18] uses one-to-one mapping of the corresponding points between two input images. Several stereo matching methods which use uniqueness constraint alternate between occlusion estimation and disparity estimation. Disparity map and occlusion map are used for occlusion estimation and disparity estimation, respectively [12]. Yet, in these cases, the complex rate increases due to the large amount of iterations required.

(a) (b)

Fig. 2. Violation of ordering constraint. (a) Left image (b) Right image

There are several stereo matching methods with effective occlusion handling. Kolmogorov *et al.* properly addressed occlusion term in defining the energy function and optimize it using graph cut [12]. Regarding the occlusion term, solely depending on the uniqueness constraint remains as a disadvantage of this method. Liu *et al.* separately handled narrow occlusion and large occlusion, generating results by means of a local method [19]. Thus, utilizing the similarity of neighboring disparity values is challenging. Ben-Ari *et al.* proposed a new energy function considering both discontinuity preserving and occlusion handling [20]. Yet, repetitive optimization significantly increases complexity. In the next section, we present an accurate disparity map generation method with avoiding the limits of the above methods.

III. Proposed Disparity Estimation Method

The goal of the proposed method is to improve the conventional methods using the discussed constraints. We do not consider the ordering constraint due to its ambiguity.

Fig. 3 represents the overall framework of the proposed algorithm. First, initial disparity maps are obtained for the left and right images with optimization via CSBP. Afterward, using the disparity maps, we perform occlusion detection followed by disparity estimation for occluded pixels. Lastly, we generate the final disparity map which is processed with occlusion handling.

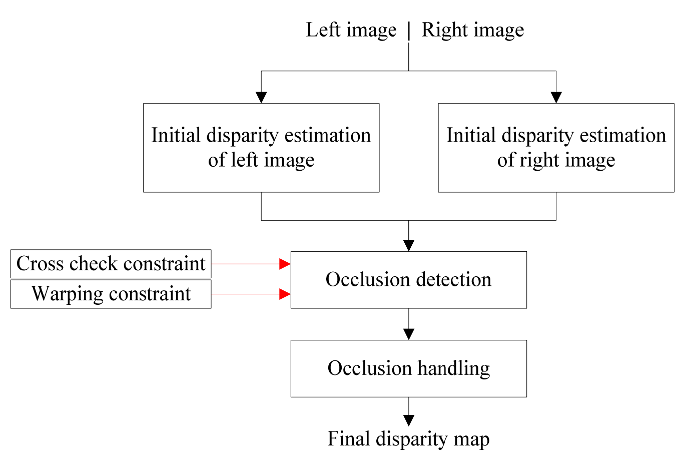


Fig. 3. Overall framework of our method

A. Initial Disparity Based on Modified CSBP

Many stereo matching algorithms define an energy function and solve it through several optimization techniques such as graph cut [11] and belief propagation [9]. The energy function in MRF is defined in (1).

$$E(f)=\sum_{s} D_s(f_s)+\sum_{s,t\in N(s)} S_{s,t}(f_s, f_t) \quad (1)$$

$D_s(\cdot)$ is the data term of node s. $S_{s,t}(\cdot)$ is the smoothness term between nodes s and t. f_s represents the state of each node s. $N(s)$ is the neighbors of the node s. In stereo matching, a node represents a pixel and the data term is generally defined by intensity consistency of pixel correspondences for hypothesized disparity. We use the luminance difference between two pixels as the matching cost. The matching cost as a data term of MRF is defined in (2).

$$D_s(d_s)=\min(|I_L(x_s,y_s)-I_R(x_s+d_s,y_s)|,T_d) \quad (2)$$

I_L and I_R are the left and the right images, respectively. x_s and y_s are horizontal and vertical coordinates of pixel s in the image. d_s is disparity of pixel s. T_d controls the data cost limit. The smoothness term is based on the degree of difference among disparities of neighboring pixels. In our method, the smoothness term is defined in (3).

$$S_{s,t}(d_s,d_t)=\min(\lambda\,|d_s-d_t|,T_s) \quad (3)$$

T_s is the constant controlling to deny cost increase. The smoothness strength λ is generally represented by a scalar constant. However, this smoothness strength is very sensitive. Thus, we adaptively refine the smoothness strength to let the method be more practical. First, color differences between pixel s and its neighboring pixels are calculated. If color difference is high, it may be regarded as an edge in the color image. We assume that the color edge in the color image is remarkably consistent with the depth edge in the depth image. The smoothness strength should be small in the depth edge. On the other hand, the smoothness strength can be high in non-edge. The color difference is defined as

$$diff_{s,t}=\sum_{c\in\{R,G,B\}}|I_c(s)-I_c(t)|. \quad (4)$$

The sum of absolute difference (SAD) is used as a difference measure and each color channel of R, G, and B components is used for color difference computation. After obtaining the color difference, the scale is controlled to set the average of the color difference to '1'. The new color difference scale is defined as follows.

$$diff_{scale}=1-(diff_{s,t}-diff_{mean})/diff_{\max} \quad (5)$$

$diff_{mean}$ is the mean of the color difference and $diff_{\max}$ is the maximum value of the color difference in the whole image. We replace λ value in (3) with λ' value to acquire a suitable smoothness strength.

$$\lambda'=\lambda\cdot diff_{scale} \quad (6)$$

The energy function is now completed. A global optimization method is used to find a disparity value which minimizes energy at each pixel. Belief propagation as an optimization method provides high quality results. However, the general belief propagation is too complex since the cost is converged after numerous iterations. In practice, when image size is N, the number of disparity levels is L, and the number of iterations is T, the computational complexity is originally $O(TNL^2)$ in standard belief propagation [9]. Hence, applying standard belief propagation is not sound for real applications. A low complexity algorithm is required even if the resulting quality is not as superior as that of belief propagation. Fast belief propagation algorithms reduces the complexity to $O(TNL)$ by using hierarchical coarse-to-fine manner [13], [14]. Such an algorithm facilitates real-time computation if it is implemented by Graphic Processing Unit (GPU). However, a lower complexity algorithm is preferred for practical use. CSBP [15] is one of the fastest algorithms. Its complexity depends solely on constant space, O(1). However, the quality is not high enough. In the proposed method, we use CSBP with remarkable fast speed and refine the results sufficiently.

B. Occlusion Detection

We present two constraints for occlusion detection. These are warping constraint and cross check constraint which is a kind of uniqueness constraint. Concerning the warping constraint, all pixels in the left image are projected to the right image using the left disparity map to generate the left occlusion map. In the case of many-to-one mapping, if the disparity map is reliable, among the matching pixels in the left image, the one which possess the largest disparity value is selected as the visible pixel. The rest of the matching pixels become occluded pixels. However, since our initial disparity map based on the modified CSBP is not sufficient enough, we consider all the matching pixels as occluded pixels. Fig. 4 illustrates the warping constraint. The dark-colored pixels are regarded as candidates of occluded pixels.

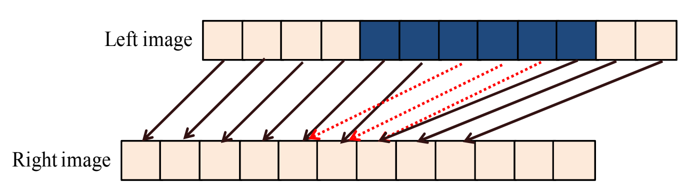


Fig. 4. Warping constraint

We define an energy function for warping constraint considering above characteristics.

$$E_G(D_L)=\sum_s w_b\,|\,o_s - G_L(s,D_L)\,| \quad (7)$$

$G_L(s, D_L)$ is a binary map constructed by the warping constraint. Multiple matching pixels in the left image are set to '1'. o_s is the occlusion value. When pixel s is supposed to the occluded pixel, the occlusion value o_s is set to '1'. w_b is the weighting factor which is applied to the pixel of the largest disparity value and the other pixels, differently.

The cross check constraint evaluates the mutual consistency from both disparity maps. If a particular pixel in the image is not an occluded pixel, the disparity values from the left and the right disparity maps should be consistent as shown in Fig. 5. The corresponding points have the same disparity value in both images.

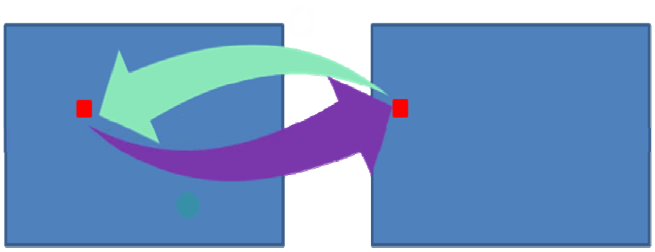


Fig. 5. Cross check constraint

The energy function for cross check constraint is defined as follows.

$$E_C(D_L,D_R)=\sum_s |\,o_s - C_L(s;D_L,D_R)\,| \quad (8)$$

$$\begin{Bmatrix} C_L=0, & if\ D_L(x_s)=D_R(x_s-D_L(x_s)) \\ C_L=1, & otherwise \end{Bmatrix} \quad (9)$$

D_L and D_R are the left and right disparity maps respectively. x_s is a pixel in left image. When C_L =1, it has a high possibility that a current pixel is occlusion.

The final energy function for occlusion detection is defined as

$$\begin{aligned} E_O=&\sum_s (1-o_s)D_s(d_s)+\lambda_O o_s \\ &+\lambda_G E_G(D_L)+\lambda_C E_C(D_L,D_R)+\sum_{s,t\in N(s)}\lambda_s\,|\,o_s-o_t\,| \end{aligned}. \quad (10)$$

(10) includes the difference of luminance component for the data term in addition to the warping and cross check constraint. This comes from the assumption that the large difference of luminance generates wrong matching even if a particular pixel is regarded as a visible pixel by two constraints. The last term represents the smoothness term for the energy function of occlusion detection and it uses SAD among the neighboring pixels of pixel s. This final function is optimized by belief propagation.

C. Disparity Assignment for Occlusion

After occlusion detection, the reasonable disparity value should be assigned to the occluded pixel. Since occlusion is only visible in one image, it is impossible to determine the accurate disparity value by means of conventional stereo matching. If we use the disparity values from neighboring pixels in the non-occluded region, disparity estimation in the occlusion region is possible. Generally, disparity values in occluded pixels are similar to those of visible pixels in the background. In the proposed method, we propagate the disparity values of visible pixels to occluded pixels.

First, we classify occlusion regions into two parts: left-side and general. Fig. 6 shows the left image and the corresponding occlusion map. The circled part in Fig. 6 (b) is the left-side and the rest of the occlusion is the general part. Occlusion in the left-side is generated due to the non-existence of left-side occlusion region in the leftmost of a right image. In this part, it is useless to estimate the disparity value using disparity values of the neighboring pixels, since our algorithm does not employ iterative optimization. Thus, we extend the disparity of the leftmost visible pixels to the left-side part for each horizontal line.

(a) (b)

Fig. 6. Two kinds of occlusion. (a) Color image (b) Occlusion map

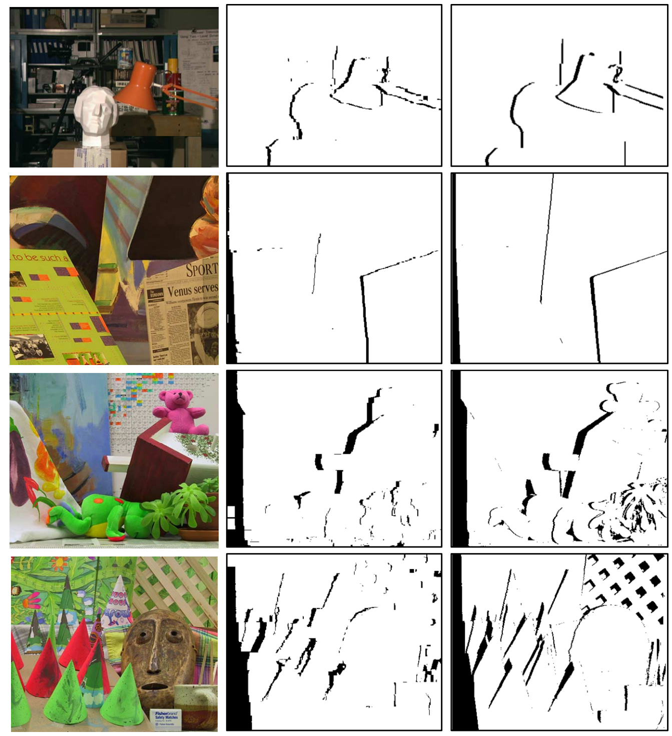
For the general part, we define a potential energy function. Let *L(s)* be the neighboring pixels whose distance from occluded pixel s is smaller than predefined distance and $C=\{s,t\,|\,s>t, t\in L(s)\}$ be the set of all nearby pixels which affect pixel s. The potential energy function for occlusion handling is defined in (11).

$$E_{OH}(s,\ d_s)=\sum_{t\in C\setminus B}(1-o_t)\frac{1}{dist(s,t)}\exp\left(-\frac{diff_{s,t}}{\sigma^2}\right) \quad (11)$$

$B=\{s,t\,|\,d_s\neq d_t, t\in C\}$ and 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 (4). The disparity value, which has the maximum value of (11), is determined as the disparity value for the pixel s. This process works at only occluded pixels which are near visible pixels. Thus, it completely handles thin or small occlusion. However, wide and large occlusion is processed at only near visible pixels. In order to solve this problem, we apply the 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 the previous process. The second potential energy function for occlusion handling is defined as follows.

$$E_{OH}(s, d_s) = \sum_{t \in C \setminus B} \frac{1}{dist(s,t)} \exp(-\frac{diff_{s,t}}{\sigma^2}) \tag{12}$$

This function is very similar to (11). However, (12) does not consider whether the source for occlusion handling is a visible pixel or not. The proposed occlusion handling process assigns disparity values which have the maximum value of the potential energy function for occlusion handling without complex optimization methods.



(a) (b) (c)

Fig. 7. Result of occlusion detection. (a) Original image (b) Proposed method (c) Ground truth

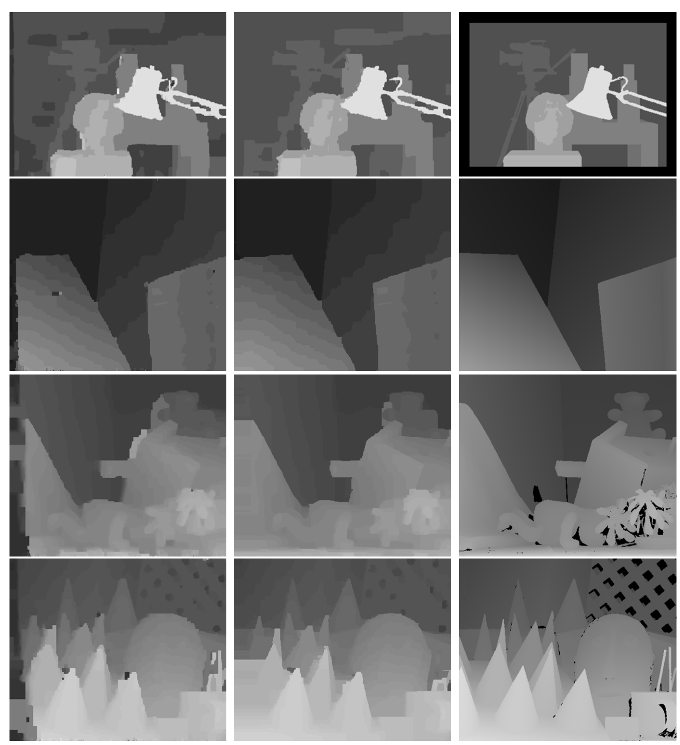
IV. Experimental Results

In order to evaluate the performance of our proposed method, we follow the methodology which measures the percentages of bad matching pixels [21]. First, we evaluate the occlusion map. The occlusion map in Fig. 7 illustrates the visual comparison of our occlusion maps with ground truth. Table I shows the percentage of the mismatching pixels between the proposed method and ground truth. These results verify that our occlusion detection method is a high performance method.

In order to evaluate our final disparity map, we compare our proposed method with the other methods which have good performance with occlusion handling. Fig. 8 demonstrates that the proposed method based on CSBP improves the quality considerably. Objective evaluation is presented in Table II. This measure is computed for three subsets of the image: non-occluded, whole, and discontinuity regions, denoted as "nonocc", "all", and "disc", respectively. When the absolute disparity error is greater than one pixel, the pixel is regarded as the bad matching pixel. The subscript of error rate in Table II represents rankings among the presented methods.

TABLE I
Evaluation for Occlusion Map

Image	Tsukuba	Venus	Teddy	Cone
Error Rate(%)	1.74	1.16	4.75	6.78



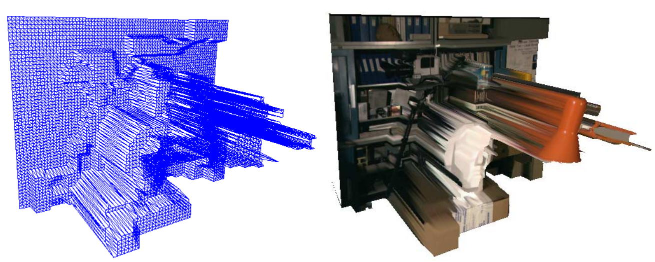
(a) (b) (c)

Fig. 8. Final disparity maps. (a) CSBP (b) Proposed method (c) Ground truth

TABLE II
PERFORMANCE COMPARISON

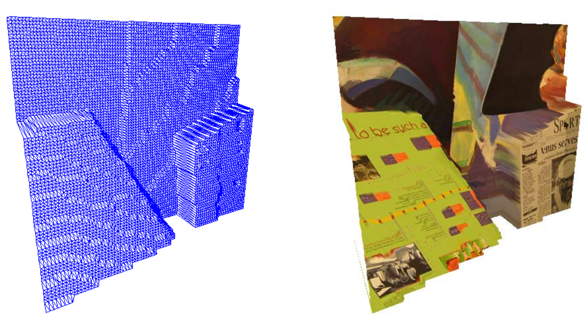
Algorithm		CSBP [8]	**Proposed method**	GC+occ [8]	CCH+SegAggr [12]	VarMSOH [13]
Average ranking		4.42	**2.33**	2.75	2.92	2.58
Tsukuba	nonocc	2.00_{4}	**1.42_{2}**	1.19_{1}	1.74_{3}	3.97_{5}
	all	4.17_{4}	**2.30_{3}**	2.01_{1}	2.11_{2}	5.23_{5}
	disc	10.50_{4}	**7.94_{2}**	6.24_{1}	9.23_{3}	14.90_{5}
Venus	nonocc	1.48_{4}	**0.91_{3}**	1.64_{5}	0.41_{2}	0.28_{1}
	all	3.11_{5}	**1.54_{3}**	2.19_{4}	0.94_{2}	0.76_{1}
	disc	17.70_{5}	**12.71_{4}**	6.75_{3}	3.97_{2}	3.78_{1}
Teddy	nonocc	11.10_{4}	**6.34_{1}**	11.20_{5}	8.08_{2}	9.34_{3}
	all	20.20_{5}	**13.62_{1}**	17.40_{4}	14.30_{2}	14.30_{3}
	disc	27.50_{5}	**17.59_{1}**	19.80_{2}	19.80_{3}	20.00_{4}
Cone	nonocc	5.98_{4}	**4.96_{2}**	5.36_{3}	7.07_{5}	4.14_{1}
	all	16.50_{5}	**12.70_{3}**	12.40_{2}	12.90_{4}	9.91_{1}
	disc	16.00_{4}	**14.44_{3}**	13.00_{2}	16.30_{5}	11.40_{1}
Average bad pixels		11.34_{5}	**8.04_{1}**	8.27_{4}	8.07_{2}	8.17_{3}

Fig. 9 and Fig. 10 show reconstructed models of 3D scenes for Tsukuba and Venus, respectively. In order to represent 3D scenes, we used mesh modeling. Laplacian smoothing was applied to all vertices of the mesh. Mesh smoothing by Laplacian smoothing reduces curvature variation and removes noise. The results of 3D scene reconstruction from an accurate disparity map make us easily feel the sense of distance.



(a) (b)

Fig. 9. 3D scene reconstruction for Tsukuba. (a) Mesh representation for 3D scene (b) 3D scene modeling



(a) (b)

Fig. 10. 3D scene reconstruction for Venus. (a) Mesh representation for 3D scene (b) 3D scene modeling

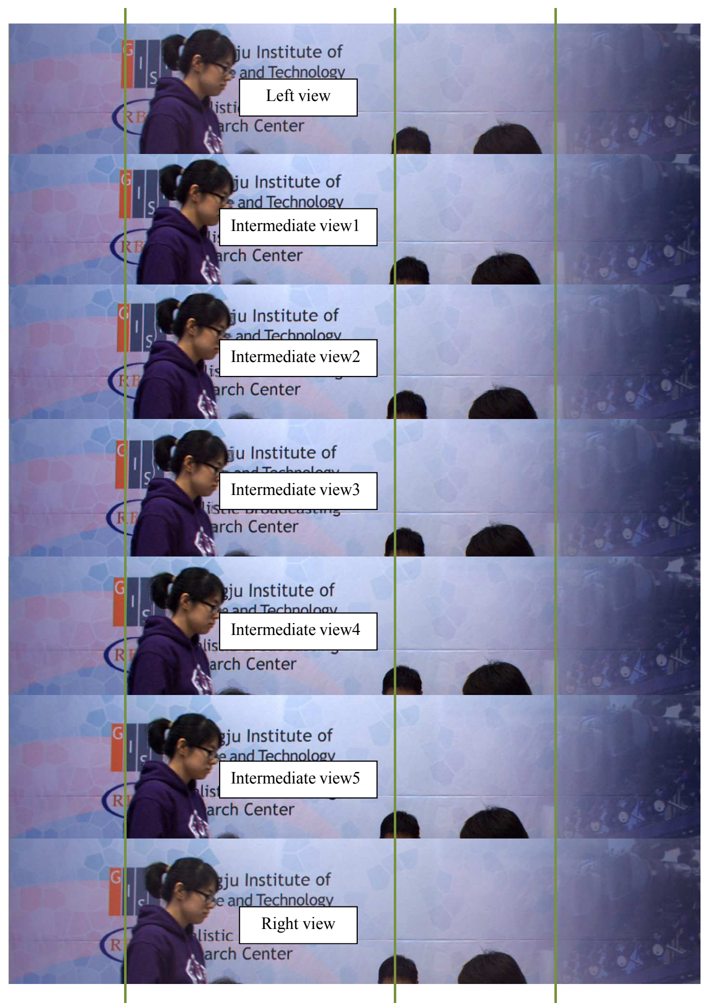


Fig. 11. Intermediate view generation

An accurate disparity map allows users to interactively control the viewpoint and generate new views of a scene. Especially, this is a basis for free viewpoint television (FTV). With FTV, the focus of attention can be controlled by the viewer rather than the director. Fig. 11 shows several intermediate views obtained from two views and their disparity maps. The results demonstrate that our system is useful for various 3D multimedia applications.

V. CONCLUSION

In this paper, we proposed a disparity estimation method considering occlusion to generate 3D information. Initially, we applied modified CSBP for initial disparity map optimization. Consecutively, occlusion was detected by exploiting the warping constraint and cross check constraint. Further, in regards to occlusion handling, we assigned reasonable disparity values to occluded pixels. Experimental results show that our method produces more accurate disparity maps compared to other state-of-the-art methods. Therefore, our method will be a viable attraction to consumer product engineers.

REFERENCES

[1] L. Zhang and W. J. Tam, "Stereoscopic image generation based on depth images for 3DTV," *IEEE Trans. on Broadcasting*, vol. 51, no. 2, pp. 191-199, June 2005.

[2] G. Tech, K. Muller, and T. Wiegand, "Evaluation of view synthesis algorithms for mobile 3DTV," *In Proc. 3DTV Conference*, pp. 132(1-4), May 2011.

[3] S. Y. Kim, J. H. Cho, and A. Koschan, "3D video generation and service based on a TOF depth sensor in MPEG-4 multimedia framework," *IEEE Trans. on Consumer Electronics*, vol. 56, no. 3, pp. 1730-1738, Aug. 2010.

[4] E. K. Lee and Y. S. Ho, "Generation of multi-view video using a fusion camera system for 3D displays," *IEEE Trans. on Consumer Electronics*, vol. 56, no. 4, pp. 2797-2805, Nov. 2010.

[5] S. Y. Kim, S. B. Lee, and Y. S. Ho, "Three-dimensional natural video system based on layered representation of depth maps," *IEEE Trans. on Consumer Electronics*, vol. 52, no. 3, pp. 1035-1042, Aug. 2006.

[6] R. Hartley and A. Zisserman, *Multiple View Geometry in Computer Vision*, 2nd ed., Cambridge University Press, 2003, pp. 262-278.

[7] H. Hirschmüller, P. R. Innocent, and J. Garibaldi, "Real-time correlation-based stereo vision with reduced border errors," *Int. Journal of Computer Vision*, vol. 47, no. 1/2/3, pp. 229-246, Apr. 2002.

[8] O. Veksler, "Fast variable window for stereo correspondence using integral images," *In Proc. IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, vol. 1, pp. 556-561, June 2003.

[9] J. Sun, N. N. Zheng, and H. Y, Shum, "Stereo matching using belief propagation," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 25, no. 7, pp. 787-800, July 2003.

[10] Z. W. Gao, W. K. Lin, Y. S. Shen, C. Y. Lin, and W. C. Kao, "Design of signal processing pipeline for stereoscopic cameras," *IEEE Trans. on Consumer Electronics*, vol. 56, no. 2, pp.324-331, May 2010.

[11] Y. Boykov, O. Veksler, and R. Zabih, "Fast approximate energy minimization via graph cuts," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 23, no. 11, pp.1222-1239, Nov. 2001.

[12] V. kolmogorov and R. Zabih, "Computing visual correspondence with occlusions using graph cuts," *In Proc. IEEE International Conference on Computer Vision*, pp. 508-515, July 2001.

[13] P. Felzenszwalb and D. Huttenlocher, "Efficient belief propagation for early vision," *In Proc. IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pp.261-268, June 2004.

[14] Q. Yang, L. Wang, R. Yang, S. Wang, M. Liao, and D. Nistér, "Real-time global stereo matching using hierarchical belief propagation," *In Proc. British Machine Vision Conference*, pp. 989-998, Sept. 2006.

[15] Q. Yang, L. Wang, N. Ahuja, "A constant-space belief propagation algorithm for stereo matching," *In Proc. IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pp. 1458-1465, June 2010.

[16] G. Egnal and R. Wildes, "Detecting binocular halfocclusions: empirical comparisons of five approaches," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol.24 no. 8, pp. 1127-1133, Aug. 2002.

[17] A. Bobick and S. Intille, "Large occlusion stereo," *Int. Journal of Computer Vision*, vol.33, no. 3, pp. 181-200, Sep. 1999.

[18] D. Marr and T. A. Poggio, "Cooperative computation of stereo disparity," *Science*, vol. 194, no. 4262, pp. 283-287, Oct. 1976.

[19] T. Liu, P. Zhang, and L. Luo, "Dense stereo correspondence with contrast context histogram, segmentation-based two-pass aggregation and occlusion handling," *Lecture Notes in Computer Science*, vol. 5414, pp. 449-461, Jan. 2009.

[20] R. Ben-Ari and N. Sochen, "Stereo matching with Mumford-shah regularization and occlusion handling," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 32, no. 11, pp. 2071-2084, Nov. 2010.

[21] D. Scharstein and R.Szeliski, "A taxonomy and evaluation of dense two-frame stereo correspondence algorithms," *Int. Journal of Computer Vision*, vol. 47, no. 1, pp. 7-42, Apr. 2002.

BIOGRAPHIES

Woo-Seok Jang received his B.S. degree in Electronic engineering from Chonnam National University, Korea, in 2007 and M.S. degree in Information and Communication Engineering from Gwangju Institute of Science and Technology (GIST), Korea, in 2009. He is currently a Ph.D. student in the Department of Information and Communications at GIST, Korea. His research interests are digital image processing, 3D object reconstruction, depth estimation, and realistic broadcasting.

Yo-Sung Ho received the B.S. and M.S. degrees in electronic engineering from Seoul National University, Seoul, Korea, in 1981 and 1983, respectively, and the Ph.D. degree in electrical and computer engineering from the University of California, Santa Barbara, in 1990. He joined the Electronics and Telecommunications Research Institute (ETRI), Daejeon, Korea, in 1983. From 1990 to 1993, he was with Philips Laboratories, Briarcliff Manor, NY, where he was involved in development of the advanced digital high-definition television system. In 1993, he rejoined the Technical Staff of ETRI and was involved in development of the Korea direct broadcast satellite digital television and high-definition television systems. Since 1995, he has been with the Gwangju Institute of Science and Technology, Gwangju, Korea, where he is currently a Professor in the Department of Information and Communications. His research interests include digital image and video coding, image analysis and image restoration, advanced coding techniques, digital video and audio broadcasting, 3-D television, and realistic broadcasting.