Occlusion Detection Using Warping and Cross-Checking Constraints for Stereo Matching

Yo-Sung Ho and Woo-Seok Jang

Abstract In this paper, we propose an occlusion detection algorithm which estimates occluded pixels automatically. In order to detect the occlusion, we obtain an initial disparity map with an optimization algorithm based on the modified constant-space belief propagation (CSBP) which has low complexity. The initial disparity map gives us clues for occlusion detection. These clues are the warping constraint and the cross check constraint. From both constraints, we define a potential energy function for occlusion detection and optimize it using an energy minimization framework. Experimental results show that the result of the occlusion detection from the proposed algorithm is very close to the ground truth.

Keywords Disparity map • Cross-checking • Occlusion handling • Stereo matching

1 Introduction

Stereo matching, widely researched topic in computer vision, is one of the most useful ways for acquiring depth information from two images. Stereo matching acquires 3D data by finding the corresponding points in other images for pixels in one image. The correspondence problem is to compute the disparity map that is a set of the displacement vectors between the corresponding pixels. For this problem, two images of the same scene taken from different viewpoints are given and it is assumed that these images are rectified for simplicity and accuracy of the problem. From this assumption, corresponding points are found in same horizontal line of two images.

Y.-S. Ho (✉) • W.-S. Jang
Gwangju Institute of Science and Technology (GIST), 123 Cheomdan-gwagiro, Buk-guGwangju 500-712, Republic of Korea
e-mail: hoyo@gist.ac.kr

J.S. Jin et al., *The Era of Interactive Media*,
DOI 10.1007/978-1-4614-3501-3_30, © Springer Science+Business Media, LLC 2013

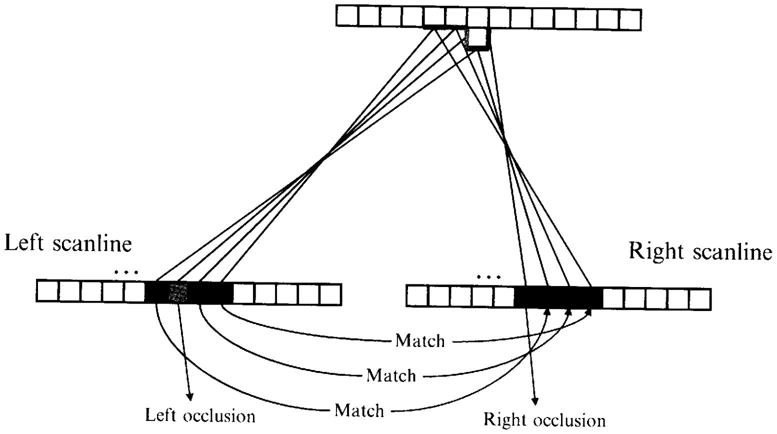

Fig. 1 Matching between two scanlines

Since two images are captured from different positions, occluded pixels exist in stereo images. The occluded pixels are only visible in one image, so accurate estimation of disparity value in these pixels is difficult. Figure 1 illustrates the case of occlusion. Occlusion is an important and challenging in stereo matching. Works that detects occluded pixels and assigns reasonable disparity values to the occlusion pixels is needed for reliable disparity map. The simplest method that detects occluded pixels and estimates disparities of occluded pixels is to use cross-checking [1] and extrapolation. Cross-checking tests that disparity value from the left and right disparity maps are consistent for each pixel. This determines occluded pixels. The disparities of visible pixels are extended into occluded pixels by extrapolation.

Two kinds of constraints have been typically used for occlusion detection in stereo matching. These are the uniqueness constraint and the ordering constraint. The uniqueness constraint [2] uses the fact that the corresponding points between two input images are one-to-one mapping. Several stereo matching methods using uniqueness constraint alternate between occlusion estimation using the estimated disparity map and disparity estimation using the estimated occlusion map [3]. The ordering constraint preserves the order of matching along the scanline in two input images [4]. The ordering constraint has limitation. It is violated in image that contains thin objects or narrow holes.

In this paper, we propose a more accurate occlusion detection method. The proposed method tries to improve conventional methods using above constraints. We do not consider the ordering constraint, because of its ambiguity presented above.

The remainder of this paper is organized as follows. In Sect. 2, we present the proposed occlusion detection method. Section 3 analyzes experimental results. Finally, Sect. 4 concludes this paper.

2 Proposed Method

The proposed method tries to improve conventional methods using the discussed constraints. We do not consider the ordering constraint, because of its ambiguity. We also do not apply occlusion estimation and disparity estimation alternately.

Figure 2 represents the overall framework of the proposed algorithm. First, initial disparity maps are obtained for the left and right images. CSBP is used for optimization of the initial disparity map. Afterward, occlusion is detected using both disparity maps and then, disparity estimation for occluded pixel is performed. Finally, the final disparity map with occlusion handling is generated.

2.1 Initial Disparity Based on Modified CSBP

Many stereo matching algorithms define energy functions and solve it through several optimization techniques such as graph cut and belief propagation. The energy function of Markov Random Field (MRF) is defined as

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

where $D_s(\cdot)$ is the data term of node s. $S_s, t(\cdot)$ is the smoothness term between node 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 in an image and 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. Matching cost as the data term of MRF is defined as

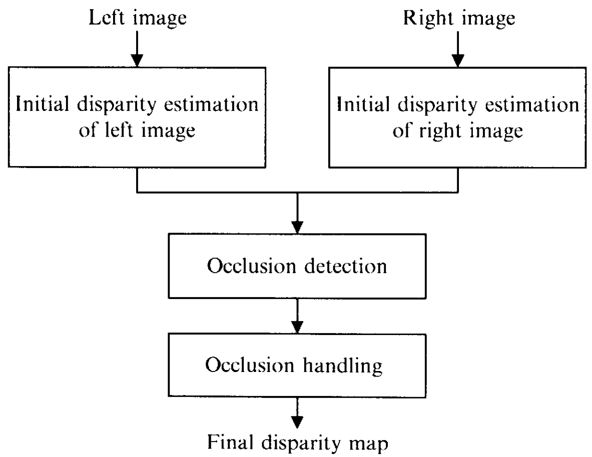

Fig. 2 Overall framework of our proposed method

$$D_s(d_s) = \min(|I_l(x_s, y_s) - I_r(x_s + d_s, y_s)|, T_d) \tag{2}$$

where I_l and I_r are left image and right image, respectively. xs and ys are horizontal and vertical coordinates of the pixel s in the image. d_s is disparity of pixel s. T_d controls the limit of the data cost. The smoothness term in stereo matching is based on the degree of difference between disparities of neighboring pixels. In our method, smoothness term is defined as

$$S_s, t(d_s, d_t) = \min(\lambda|d_s - d_t|, T_s) \tag{3}$$

where T_s is the constant controlling to stop increasing of the cost. λ represents the smoothness strength and is generally represented by a scalar constant. However, this smoothness strength is very sensitive. Thus, we adaptively refine a smooth strength to make our method more practical. First, color differences between pixel s and its neighboring pixels are calculated. High color difference means color edges in the color image. We assume that the color edge is remarkably consistent with the depth edge. 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

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

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

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

where $diff_{mean}$ 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.

$$\lambda' = \lambda \cdot diff_{scale} \tag{6}$$

The energy function is completed. A global optimization method is used to find a disparity value which minimizes the energy at each pixel. We obtain a good result using the belief propagation algorithm. However, general belief propagation is too complex for a reasonable result, 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 $(4TNL^2) = O(TNL^2)$ in standard belief propagation [5]. Thus, it is not good to apply standard belief propagation in real application. Even if quality is a little low, low

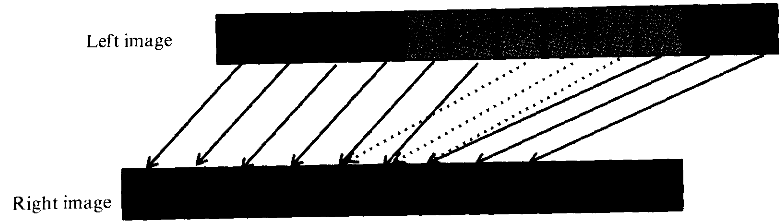

Fig. 3 Warping constraint

complexity algorithm is needed. One of the fast belief propagation reduce the complexity to O(TNL) by using hierarchical coarse-to-fine manner [6]. This algorithm facilitates real-time computation if it uses GPU implementation. However, we want a lower complexity algorithm. CSBP [7] is the fastest and takes the smallest memory among fast belief propagation algorithms. Its complexity depends on only constant space, that is O(1). However, CSBP cannot make enough quality for result. Since we sufficiently refine the result of CSBP, we select the CSBP algorithm as our optimization algorithm.

2.2 *Occlusion Detection*

We present two different kinds of constraints for occlusion detection. These are the warping constraint and the cross check constraint. In the warping constraint, all pixels in the left image are projected to the right images using the left disparity map for the left occlusion map. If the multiple pixels in the left image are projected on only one pixel in the right image, all but one pixel are occluded pixels. In this case, if the disparity map is reliable, the pixel which has the largest disparity value among the multiple matching pixels is the visible pixel and the rest of the pixels are the occluded pixels. However, our initial disparity map based on the modified CSBP is not perfect. Thus, we consider all multiple matching pixels as the candidates of the occluded pixels. Figure 3 illustrates the warping constraint. Red pixels can be regard as the candidates of the occluded pixels.

In order to find the accurate occlusion map with inaccurate disparity map, we define the energy function for warping constraint. The energy function is based on possibility.

$$E_g(D_l) = \sum_s w_b|o_s - G_l(s, D_l)| \tag{7}$$

where $G_l(s, D_l)$ is a binary map by the warping constraint. Multiple matching pixels in the left image are set to "1". 0_s is the occlusion value by assumption. When pixel s is supposed to the occluded pixel, the occlusion value $0s$ is set to "1". w_b is the

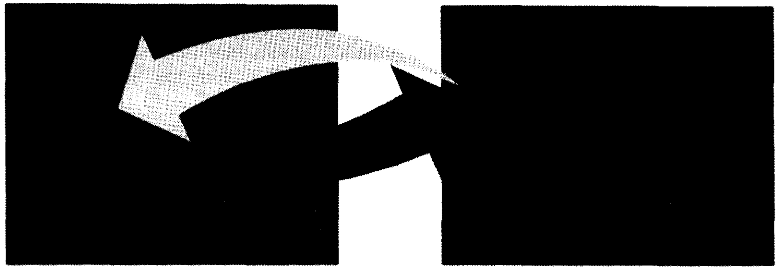

Fig. 4 Cross check constraint

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 left image is not occluded pixel, the disparity values from the left and the right disparity maps should be consistent. Figure 4 illustrates the cross check constraint. Corresponding points have the same disparity value in both images.

The energy function for cross check constraint is defined as

$$E_c(D_l, D_r) = \sum_s |o_s - C_l(s; D_l, D_r)| \tag{8}$$

$$\left\{ \begin{array}{l} C_l = 0,\ if\ D_l(x_s) = D_r(x_s - D_l(x_s)) \\ C_l = 1,\ otherwise \end{array} \right\} \tag{9}$$

where D_l and D_r are the left disparity map and right disparity map respectively. x_s is a pixel in left image. When C_l is "0", it means that the disparity value of the current pixel is reliable.

The final energy function for occlusion detection is defined as

$$E_o = \sum_s (1 - o_s)D_s(d_s) + \lambda_O o_s + \lambda_g E_g(D_l) + \lambda_c E_c(D_l, D_r) + \sum_{s,t \in N(s)} \lambda_s |o_s - o_t| \tag{10}$$

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

3 Experimental Results and Analysis

Table 1 lists the values for the parameters used in the proposed method. These parameters are acquired by experiments to balance energy terms. As I presented before, the smoothness strength (λ) is adaptively set.

In order to evaluate the performance of our proposed method, we follow the methodology which measures the percentages of bad matching pixels [8]. First, we evaluate the occlusion map. The occlusion map in Fig. 5 illustrates the visual comparison of our occlusion maps with ground truth. In order to show the superiority of the proposed method, we also presented the other method using uniqueness constraint [9]. Since its algorithm was adapted to fit our framework, it does not produce exactly the same result as in their method. Table 2 shows the percentage of

Table 1 Parameters for experiment

T_d	T_s	λ_o	λ_G	λ_C	λ_s
30	105	7.5	3	12	4.2

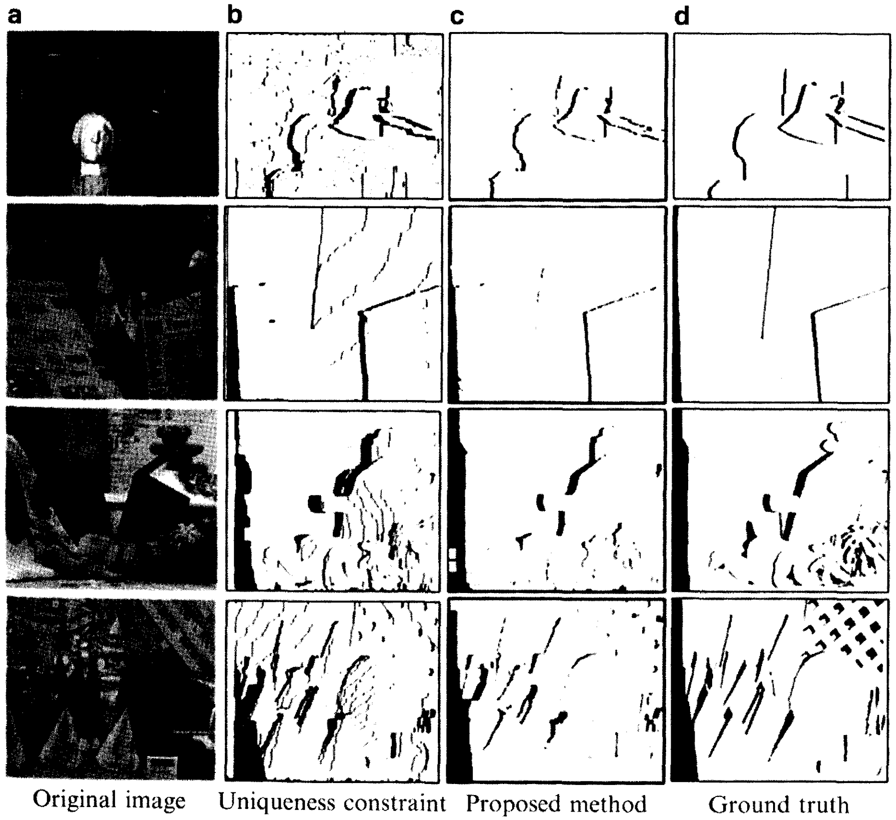

Fig. 5 Result of occlusion detection

Table 2 Evaluation for occlusion map (Error rate, %)

Image	Uniqueness constraint	Proposed method
Tsukuba	5.80	**1.74**
Venus	4.14	**1.16**
Teddy	11.32	**4.75**
Cone	14.09	**6.78**

Table 3 Performance comparison of disparity map (Error rate, %)

Image	CSSP	Uniqueness constraint	Proposed method
Tsukuba	4.17	3.10	2.30
Venus	3.11	2.79	1.54
Teddy	20.20	18.02	13.62
Cone	16.50	15.49	12.70

Table 4 Comparison with the other methods (Error rate, %)

Image	Proposed method	GC + occ [3]	CCH + SegAggr [11]	VarMSOH [12]
Tsukuba	2.30	2.01	2.11	5.23
Venus	1.54	2.19	0.94	0.76
Teddy	13.62	17.40	14.30	14.30
Cone	12.70	12.40	12.90	9.91

the mismatching pixels between the methods and ground truth. These results verify that our occlusion detection method is a high performance method.

In order to check effectiveness of our method, we assigned reasonable disparity values to the occluded pixels. For this, we used the disparity allocation method [10]. It uses the potential energy function which is based on color dissimilarity and spatial distance from the current occluded pixel. This method extends disparity values of visible pixels to the occluded pixels. Objective evaluation is presented in Table 3. We compare our proposed method with the other methods which have good performance with occlusion handling. Table 4 shows comparison results.

Figure 6 shows final disparity map assigned by the disparity allocation method [10]. Figure 6 demonstrates that the proposed method based on CSBP improves the quality considerably.

4 Conclusion

In this paper, we proposed the occlusion detection method for stereo matching. Constant space belief propagation was applied for optimization of an initial disparity map basically. It is efficient since it reduces the complexity. We modified CSBP for better result. Occlusion detection was performed by the warping constraint and the cross check constraint. The experimental results show that the proposed method

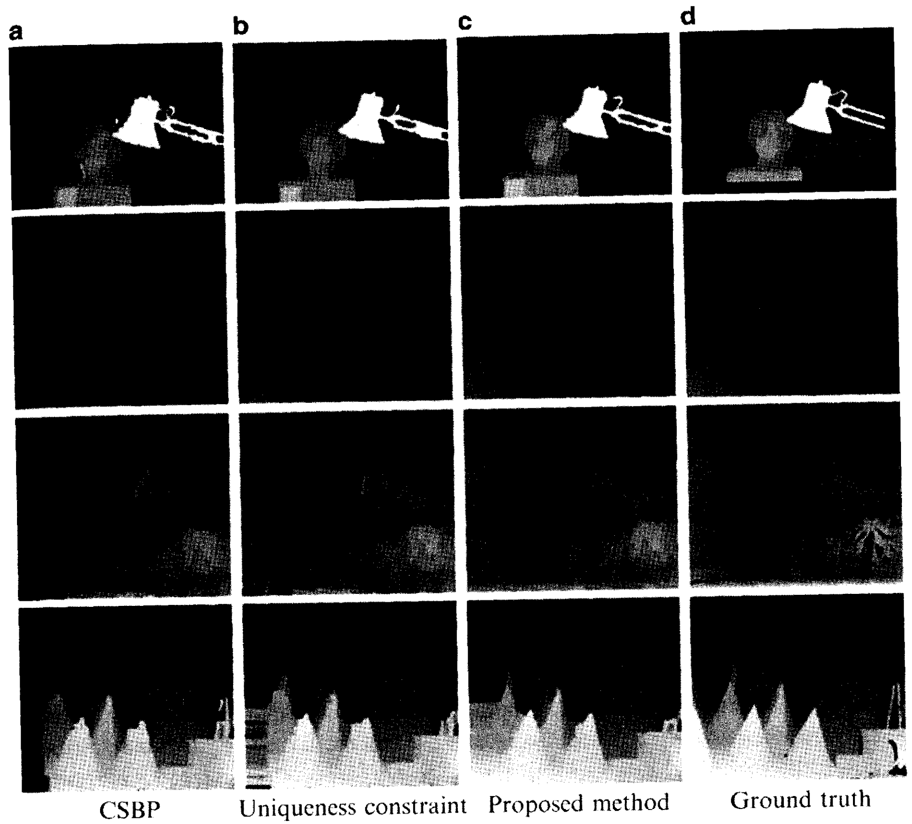

Fig. 6 Result of final disparity map

works very well for stereo images. If we perform stereo matching again using the obtained occlusion from proposed method, a more accurate disparity map can be obtained.

Acknowledgement This work was supported by the National Research Foundation of Korea (NRF) grant funded by the Korea government (MEST) (No. 2011-0030822).

References

1. Hirschmuller, H., Innocent, P.R., Garibaldi, J.M.: Real-time correlation-based stereo vision with reduced border errors. In: International Journal of Computer Vision, vol. 47(1/2/3), pp. 229–246. (2002)
2. Marr, D., Poggio, T.A.: Cooperative computation of stereo disparity. In: Science, vol. 194 (4262), pp. 283–287. (1976)

3. Kolmogorov, V., Zabih, R.: Computing visual correspondence with occlusions using graph cuts. In: IEEE International Conference on Computer Vision, pp. 508–515. (2001)
4. Bobick, A., Intille, S.: Large occlusion stereo. In: International Journal of Computer Vision, vol.33(3), pp. 181–200. (1999)
5. Sun, J., Zheng, N.N., Shum, H.Y.: Stereo matching using belief propagation. In: IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 25(7), pp. 787–800. (2003)
6. Yang, Q., Wang, L., Yang, R., Wang, S., Liao, M., Nister, D.: Real-time global stereo matching using hierarchical belief propagation. In: British Machine Vision Conference, pp. 989–998. (2006)
7. Yang, Q., Wang, L., Ahuja, N.: A constant-space belief propagation algorithm for stereo matching. In: IEEE Computer Society Conference on Computer Vision and Pattern Recognition, pp. 1458–1465. (2010)
8. Scharstein, D., Szeliski, R.: A taxonomy and evaluation of dense two-frame stereo correspondence algorithms. In: International Journal of Computer Vision, vol. 47(1), pp. 7–42. (2002)
9. Min, D., Yea, S., Vetro, A.: Occlusion handling based on support and decision. In: International Conference on Image Processing, pp. 1777–1780. (2010)
10. Jang, W., Ho, Y.: Disparity map refinement using occlusion handling for 3D scene reconstruction. In: International Conference on Embedded System and Intelligent Technology, pp. 213–216. (2011)
11. 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, 2009.
12. R. Ben-Ari and N. Sochen, "Stereo Matching with Mumford-Shah Regularization and Occlusion Handling," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 32, no. 11, pp. 2071–2084, Nov. 2010.