Stereo for Recovering Sharp Object Boundaries

Jeonghee Jeon[1], Choongwon Kim[2] and Yo-Sung Ho[1]

[1] Kwangju Institute of Science and Technology (K-JIST),
1 Oryong-dong Puk-gu, Kwangju, 500-712, KOREA
{jhjeon, hoyo}@kjist.ac.kr
[2] Chosun University,
375 Seosuk-dong Dong-gu, Kwangju, 501-758, KOREA
{cwkim}@chosun.ac.kr

Abstract. In this paper, we propose a stereo matching algorithm using multiple windows to recover sharp object boundaries. In order to achieve this goal, we employ commonly used techniques, such as left-right consistency, uniqueness constraint, and expansible multiple windows. By adding a new idea of multiple windows over these techniques, we can minimize the boundary overreach, which is usually caused by the unsuitable window size or shape. We demonstrate that the algorithm using a new multiple windows can display the dense and sharp disparity map and the windows adaptively used according to features of boundary edges. A problem generated by the disposition order of multiple windows is also presented.

1 Introduction

Stereo matching based on correlation or sum of squared differences (SSD) is a basic technique to obtain a dense map from images [1] [3] [4] [5] [7] [8] [10] [11]. Although this technique yields dense depth maps, it fails within occluded areas and/or poorly textured regions. Any window-based stereo matching in these regions has problems of low reliability. In general, the fattening or thinning operation along the object boundary has a problem that a window can contain both foreground and background surfaces with different disparities [2]. Kanade and Okutomi addressed the problem of choosing the right support region for high reliability with adaptive window [3]. At each point, a rectangular window is grown to an optimal size based on an estimate of disparity uncertainty in the current window. A greedy algorithm is used to select the best of the four possible directions to grow the window at each step. They also presented a multiple-baseline stereo to determine a single point in a region with repetitive patterns [4]. Fuseiello, et al. proposed a method for choosing the right support region by the multiple window approach [5]. For each pixel, they perform the correlation operation with nine different windows, and obtain the disparity from the window of the smallest SSD value. The basic idea of this scheme is that a window yielding a smaller SSD is more likely to cover a constant depth region.

In this paper, we propose a stereo matching algorithm using multiple windows to recover sharp object boundaries. Relevant techniques and a new algorithm are de-

M.H. Shafazand and A M. Tjoa (Eds.): EurAsia-ICT 2002, LNCS 2510, pp. 248–255, 2002.

scribed in Section 2. In Section 3, we are shown experimental results using synthetic and real stereo pairs, and we summarize the proposed algorithm in Section 4.

2 The Algorithm

2.1 New Multiple Windows

In recent years, stereo techniques using multiple windows have been proposed to recover precise object boundaries and display disparity maps efficiently [5] [8]. However, they do not use windows of diagonal edges that are often encountered in the image. They use only symmetric windows. The stereo algorithm using multiple windows can compute more accurate similarity than one using a single window. However, it is difficult to calculate similarity exactly in the boundaries of diagonal edges because the shapes of the windows are suitable for detecting horizontal and vertical edges. We introduce new windows of diagonal shapes, as shown in Fig. 1.

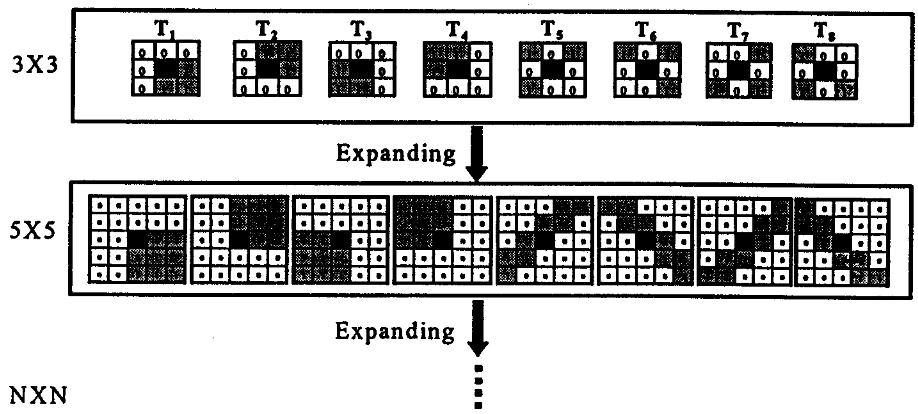


Fig. 1. Multiple windows

In Fig. 1, we show four windows T_1~ T_4 to detect horizontal, vertical, and corner edges that are borrowed from the literature [5] [8]; however, windows T_5~ T_8 for diagonal edges are newly introduced. All windows shown in Fig. 1 can be regularly expanded for stereo matching. We notice that the multiple windows have the same number of gray cells if their sizes are equal. All gray cells are considered to calculate similarity of each window and the darker gray cells are the pixels that we want to find in other images. In this paper, we assume that all stereo pairs are fully rectified by camera parameters [14].

2.2 Similarity Measure

In the area-based stereo algorithm, the similarity measure of each pixel usually uses the well-known SSD.

$$S_T(x, y, d) = \sum_{i,j \in w} [I_L(x+i, y+j) - I_R(x+i+d, y+j)]^2 \tag{1}$$

where I_L, I_R, d, w, and T are left, right images, disparity, gray cells within window, and one of multiple windows, respectively. The best match for each point in the image using SSD can be found by comparing the square window centered at this point against the window of equal size centered at points that lie on the corresponding scanline in the other image. The SSD across the window is used as the similarity measure. The location that minimizes this measure is selected as the best match, and the disparity is stored. From the similarity of each window, we determine the disparity from the window generating the smallest SSD.

$$S(x,y,d) = \arg\min_{T=1}^{8} S_T(x,y,d) \tag{2}$$

We can determine the disparity by Eq. (1) and Eq. (2); however, there are a few problems, as shown in Fig. 2(b).

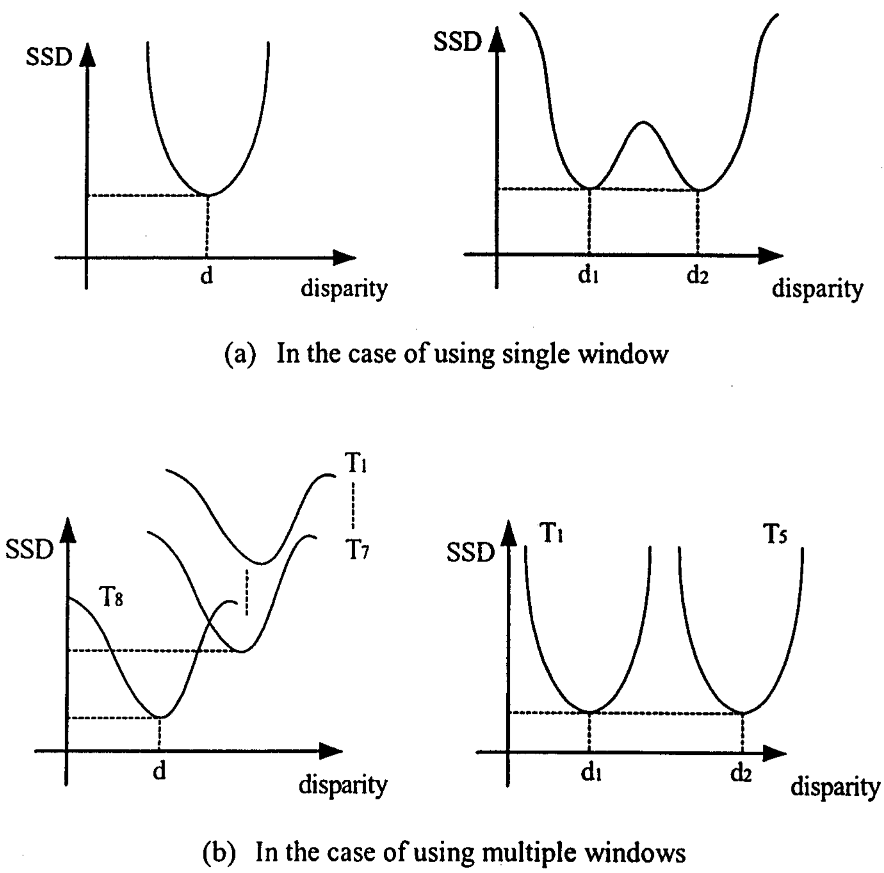


Fig. 2. Problems in determining a disparity

In the left graph of Fig. 2(a), the disparity can be determined uniquely at d. Marr and Poggio described uniqueness constraint to minimize false matches [9]. The constraint is that a given pixel or feature from one image can match no more than one pixel or feature from the other image. However, if the number of pixels having the same minimum value is two or more, as shown the right graph of Fig. 2(a), we cannot determine disparity exactly. It was mainly happened in image regions of regular patterns or uniform intensity values. Okutomi and Kanade presented a multiple-baseline stereo to select a single point in the region [4]. We have described a method to estimate single disparity at a region of uniform intensity by extending the window size [10]. In this paper, our method simply expands the window size to four directions to refer more pixels if there are multiple local minima within the search range. With the operation, the bulk of multiple local minima are disappeared as the SSSD function of multiple-baseline stereo [4].

In case of multiple windows, a decision of disparity is faced with a new problem. In the left graph of Fig. 2(b), we know that the disparity can be easily estimated as d if the minimum of SSD in each window is different. However, as shown in right graph of Fig 2(b), we cannot decide the disparity of window with the smallest SSD accurately. The disparities are usually determined by priority of the windows. We call the problem disposition order. As shown in right graph of Fig. 2(b), disposition order of the windows can be an important factor if the SSD of each window is same but disparities are different each other. Disposition order is priority of windows to estimate a disparity from the SSDs. If priority of T_1 is higher than T_5, disparity is estimated as d_1. However, if priority of T_1 is lower, disparity is determined as d_2. For the reasons, any algorithm using multiple windows cannot avoid the problem, which is originated by disposition order of multiple windows. Thus, it should be carefully determined to estimate right disparity. We have selected the disposition order as Fig. 1. This selection is made from the assumption that object's boundaries in the scene of real world have mostly vertical and horizontal lines.

In order to detect occlusion, we use left-right consistency. The principle is that a valid match point should be equally matched in left-right and right-left direction [7]. Each point on one image can match at most one point on the other image, and the matched points have the same disparity in both directions, respectively. Therefore, we can easily predict an occluded pixel or region by checking consistency.

2.3 Proposed Algorithm

We described relevant techniques and characteristics of the multiple windows. In order to estimate single disparity, we firstly use a smaller window, which can be expanded to four directions according to a situation of uniqueness constraint. The windows can be also expanded according to condition of left-right consistency that a valid disparity should be equally existed in both directions of left-right and right-left direction. Disposition order of multiple windows is considered.

All points of image are basically examined whether uniqueness constraint or left-right consistency are satisfied or not. If even one of two conditions is not satisfied, the size of window is enlarged to find a unique match point by referencing more pixels. Through stereo matching by using uniqueness constraint, left-right consistency, and expanding multiple windows, we know that the algorithm has a possibility to minimize boundary overreach. The proposed stereo matching procedure to minimize boundary overreach is as following.

Step 1: Set all the points to FALSE.
Step 2: If a point has FALSE, start to stereo matching using the Eq. (1) and Eq. (2).
Step 3: Check uniqueness constraint and left-right consistency.
Step 4: If two conditions are simultaneously satisfied, save the disparity.
Step 5: Mark the point as TRUE.
Step 6: Process Step 2 ~ Step 5 until the last point of image with FALSE.
Step 7: Expand window size for the point with FALSE and iterate Step 2 ~ Step 6.
Step 8: Display disparity map or windows map, that indicates which window is used for searching disparity from the proposed eight windows.

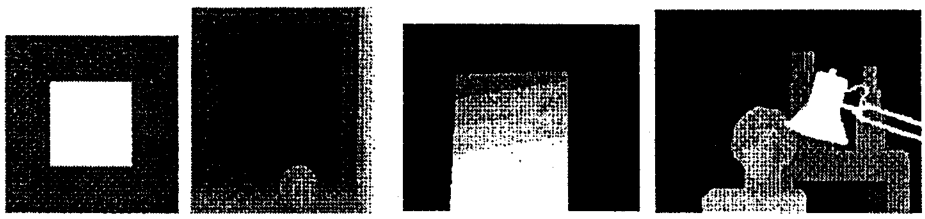
3 Experimental Results

We described stereo matching algorithm using multiple windows. The central problems for recovering precise object boundaries are to select appropriate window size and shape as described in Section 1 and Section 2. These problems are important issues in the field of stereo matching irrespective of using single or multiple windows. The issues are found in many literatures [3] [6] [7] [8] [11]. We do not test the problems originated by window size or shape of square or rectangular because the problems have excellently described in many papers [3] [6] [11].

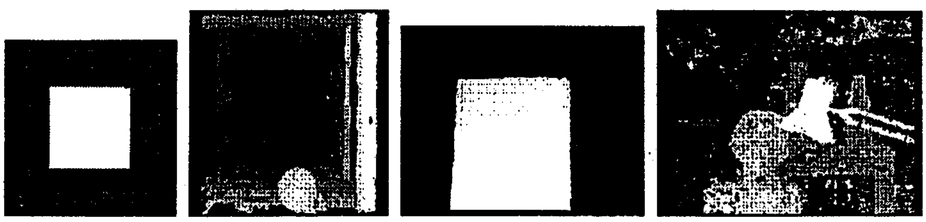
In this section, we perform simulation on various kinds of gray-level image to evaluate performance of the proposed algorithm. The images used are classified into synthetic and real image as shown in Fig. 3(a): synthetic images are Random dot, Corridor, and Microsoft and real image is Tsukuba downloaded from web site [15][16]. The Random dot stereo pairs have features that square object of foreground and background are shifted to right direction with six and two, respectively. The Corridor stereo is clean synthetic image without noise and has various objects including such as straight lines, curves, and circle. The stereo pairs of Microsoft are image that object is slanted, and the Tsukuba stereo pairs have many features such as narrow objects, regions of uniform intensity, and so on.

The algorithm is firstly evaluated by disparity map and the influence of disposition order is secondly assessed by disparity map and window maps, that represents which window is selected out of the proposed eight windows. Fig. 3 presents various disparity maps by using our algorithm and Symmetric Multi-Window (SMW) stereo algorithm, respectively. The SMW algorithm by Fusiello, et. al uses nine windows [5]. The SMW algorithm is an adaptive, multiple windows scheme using left-right consistency to compute disparity and its associated uncertainty. From Fig. 3, it is found that the proposed algorithm produces better results in the aspect of sharp boundaries but the SMW algorithm gives smoother disparity map. It is noticed that disparity maps by our algorithm presents gray-level to only the points with "TRUE" and black to other points with "FALSE".

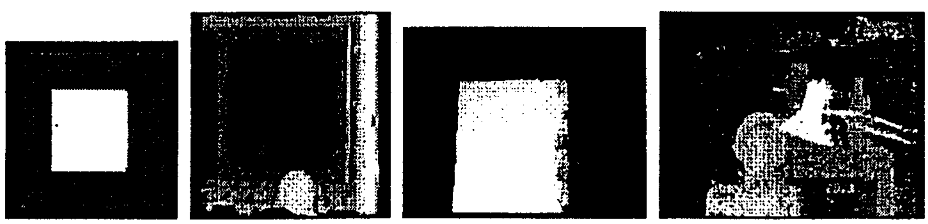
(a) Random dot, Corridor, Microsoft and Tsukuba stereo images



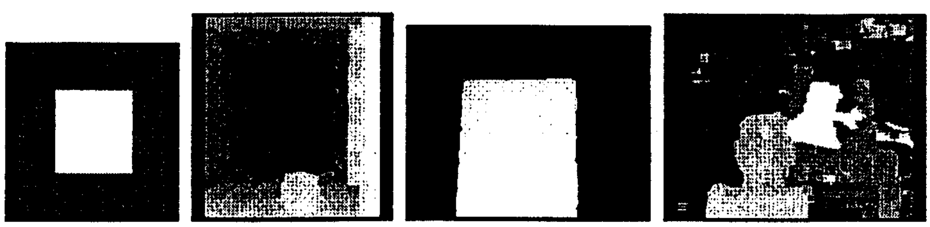
(b) Ground truth



(c) Disparity maps by our algorithm when the priority has order of $\mathbf{T}_1 \rangle \ldots \rangle T_8$

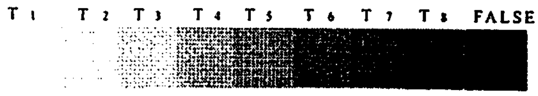


(d) Disparity maps when the priority has order of $\mathbf{T}_1 \langle \ldots \langle T_8$

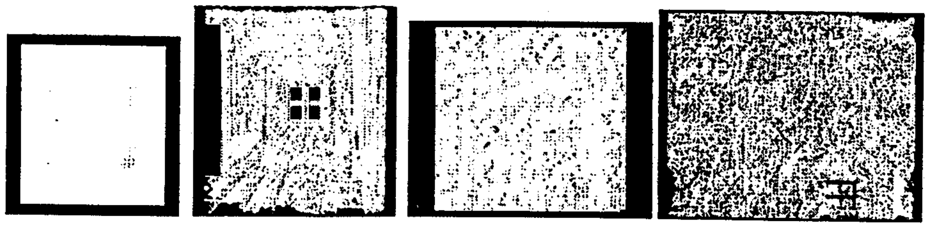


(e) Disparity maps by SMW

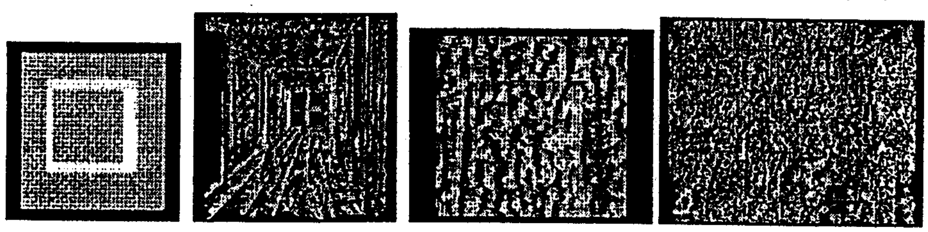
Fig. 3. Stereo image, ground truth, and disparity maps



(a) Gray level bar



(b) Window maps by our algorithm when the priority has order of $T_1 \rangle \ldots \rangle T_8$



(c) Window maps when the priority has order of $T_1 \langle \ldots \langle T_8$

Fig. 4. Gray level bar and window maps

From the disparity maps of Fig. 3, we can find that our algorithm minimizes the boundary overreach. In case of object's boundary of Fig. 3(c) and (e), our results show steeper boundaries than those of SMW. However, performance of two algorithms cannot be simply compared because the window size of SMW is not known. We can show an influence of disposition order using window maps that indicate which window is used among the multiple windows. A careful comparison between Fig. 3(c) and Fig. 3(d) shows that disposition order has a great influence on the disparity map and out assumption about disposition order is valid. In Fig. 4(a), gray level bar represents which each window is used. The window maps are shown in Fig. 4(b) and Fig. 4(c). From the window maps, we know that the proposed windows are correctly used according to edge features. Especially, the window map of the Corridor containing objects of diagonal edges demonstrates the proposed algorithm uses appropriately the windows of T_5~T_8.

4 Conclusions

In this paper, we have proposed stereo matching algorithm using multiple windows to recover sharp object's boundary. The algorithm assumes that qualities of depth are affected by an appropriate window size, shape, and disposition order of multiple windows. We have made new multiple windows of horizontal, vertical, and diagonal edges. The windows are to be expanded according to uniqueness constraint and left-right consistency. We carefully determined disposition order of the windows because of an effect on estimation of disparity. The experimental results have shown that our algorithm presents not only dense disparity maps but also sharp at the boundaries of object. We also demonstrated that the proper window is used depending on the properties of boundaries like straight line, curve, circle and diagonal edge. These windows can be used in image retrieval, segmentation, motion estimation, image processing, etc.

Acknowledgement.
This work was supported in part by the Korea Science and Engineering Foundation (KOSEF) through the Ultra-Fast Fiber-Optics Networks (UFON) Research Center at Kwangju Institute of Science and Technology (K-JIST), and in part by the Ministry of Education (MOE) through the Brain Korea 21 (BK21) project.

References

1. Dhond, U. R., Aggarwal, J.K.: Structure from Stereo-A Review, IEEE Transactions on System, Man, and Cybernetics, Vol. 19, No. 6, Nov./Dec. (1989) 1489-1510.
2. Otha, Y., Tamura, H.: Mixed Reality, Springer-Verlag (1999).
3. Kanade, T., Okutomi, M.: A Stereo Matching Algorithm with An Adaptive Window: Theory and Experiment, IEEE Transaction on Pattern Analysis and Machine Intelligence, Vol. 16, No. 9, Sept. (1994) 920-932
4. Okutomi, M., Kanade, T.: A Multiple-Baseline Stereo, IEEE Transaction on Pattern Analysis and Machine Intelligence, Vol. 15, No. 4, (1993) 353-363.
5. Fusiello, A., Roberto, V., Trucco, E.: Efficient Stereo with Multiple Windowing, Proceeding of CVPR, (1997) 858-863
6. Scharstein, R., Szeliski, R.: Stereo Matching with Nonlinear Diffusion, International Journal of Computer Vision, Vol. 28, No. 2, July (1998) 155-174.
7. Fua, P.: A Parallel Stereo Algorithm that Produces Dense Depth Maps and Preserves Image Features, Machine Vision and Application, Vol. 6, No.1, (1993) 35-49.
8. Okutomi, M., Katayama, Y., Oka, S.: A Simple Algorithm to Recover Precise Object Boundaries and Smooth Surface, IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'2001), Vol. 2, (2001) 138-144.
9. Marr, D., Poggio, T.: Cooperative Computation of Stereo Disparity, Science, Vol. 194, Oct. (1976) 283-287.
10. Jeon, J., Kim, K., Kim, C., Ho, Y. S.: A Robust Stereo-Matching Algorithm Using Multiple-Baseline Cameras, IEEE Pacific Rim Conference on Communications, Computers and Signal Processing (PACRIM 2001), Vol. I, Aug. (2001) 263-266.
11. Kang, S.B., Szeliski, R., Chai, J,: Handling Occlusions in Dense Multi-View Stereo, Technical Report MSR-TR-2001-80, Microsoft Research, Sept. (2001).
12. Gonzalez, G. C., Woods, R. E.: Digital Image Processing, Addison Wesley (1992).
13. Frei, W., Chen, C. C.: Fast Boundary Detection: A Generalization and a New Algorithm, IEEE Transaction on Computers, Vol. C-26, No. 10, Oct. (1977) 988-998.
14. Tsai, R.Y.: A versatile Camera Calibration Technique for High-Accuracy 3D Machine Vision Metrology Using Off-the-Self TV Cameras and Lenses, IEEE Journal of Robotics and Automation, Vol. RA-3, No. 4, Aug. (1987) 323-344.
15. http://www.middleburry.edu/stereo.
16. http://www-dbv.cs.uni-bone.de/stereo_data.