2010 10th IEEE International Conference on Computer and Information Technology (CIT 2010)

Disparity Adjustment for Local Stereo Matching

Xia Hu*, Caiming Zhang*†, Wei Wang*,and Xifeng Gao*
*School of Computer Science and Technology
Shandong University, Jinan, China, 250101
Email: huxia@mail.sdu.edu.cn
†School of Computer Science and Technology
Shandong Economics University, Jinan, China, 250014
Email: czhang@sdu.edu.cn

***Abstract*—Local stereo matching methods still play an important part as they are simple and fast. Some local methods perform well and even better than most global methods. But they usually achieve accuracy at the expense of speed. Simple local methods are fast, but exhibit systematic errors. In this paper, we focus on the invalid regions of traditional window-based matching and present a new solution to improve the initial matching result according to the identification of invalid regions and weight-based disparity adjustment. Our solution provides an easy way to increase the accuracy with little additional computation. The experimental results are evaluated using the Middlebury stereo test bed, showing that our method achieves competitively performance. Moreover our algorithm modules could be easily expanded into other matching algorithms as post-processing.**

***Index Terms*—Bincular vision; stereo matching; local method; disparity adjustment;**

I. Introduction

Stereo matching is the problem of finding matching points in two or more images of the same scene, usually assuming known camera geometries. Nowadays it is still one of the most heavily investigated topics in computer vision. Stereo matching has recently experienced significant progress as many new algorithms are presented. Thanks to the Middlebury testing benchmark [1], researchers could also compare their algorithms against all the state-of-the-art algorithms.

Stereo matching algorithms are currently classified into two groups [1]: local and global methods. Typically, global ones can achieve a higher degree of accuracy in retrieving disparity information [2], [3], [4], [5], [6]. However, most global correspondence methods are computationally expensive and sometimes need many parameters that are hard to determine. As a result, area-based local methods still play an important part in real-time applications as they are often simple and fast. Several real time systems have been developed using area-based local methods [7], [8], [9]. Generally, local algorithms yield significantly less accurate disparity maps, but many local algorithms can achieve high accuracy too [10], [11], [12], [13]. Most of them are based on segmentation [14] or adaptive weights [12]. Unfortunately, the general speedy methods [7], [8], [9] are hardly applied to speed these local algorithms.

In this paper, we focus on the area-based approaches, which yield a dense disparity map by matching small image patches as a whole, relying on the assumption that nearby points usually have similar displacements. According to the research upon the problems with traditional window-based methods, we present a method to find the invalid regions and also give a new weight-based disparity adjustment method for the traditional window-based local methods. Our solution provides an easy way to increase the accuracy with little additional computation.

Our method assumes that surfaces vary usually smoothly within real images, except at object borders. We use the color similarity and geometric proximity [12]. In Section II, we will discuss these basic theories of our algorithm in detail. The proposed method is composed of three parts: initial local matching, invalid regions definition, weight-based disparity adjustment. We give a detailed explanation for each part in Section III, and then show some experimental results in Section IV. In Section V, we discuss the proposed method and conclude the paper.

II. Our method

We aim at increasing the accuracy for the traditional window-based methods with little additional computation. Generally, traditional window-based methods exhibit systematic errors especially in some special regions such as weakly-texture areas and occluded/half-occluded areas. In order to improve the matching result, we present a method to find these invalid regions which mainly contain occluded/half-occluded areas and discontinuity areas. Then we define a new weight function to adjust the disparities in these regions.

A. Basic Theory

In this section, we firstly discuss the main problems with traditional window-based methods and then give the matching constraints and assumptions used in our method.

Our method is based on the research upon the problems with traditional window-based methods. H.Hirschm$\ddot{u}$ller analyzes the problems with correlation-based stereo [7]. Actually all the window-based methods share the similar problems. The traditional window-based methods yield a dense disparity map by matching fixed square windows as a whole, relying on the assumption that nearby points within the support window usually have similar displacements. Therefore, depth variations will introduce errors in the calculation. This usually happens at depth discontinuities and weakly-texture regions. Whether the introduced error can be neglected or leads finally to the

978-0-7695-4108-2/10 $26.00 © 2010 IEEE
DOI 10.1109/CIT.2010.248

wrong decision depends on the similarity between the object, the occluded and visible part of the background, which is covered by the support window. The wrong decisions usually lead to the extending of objects horizontally and the fuzzy disparities in the weakly-texture areas(Fig.1).

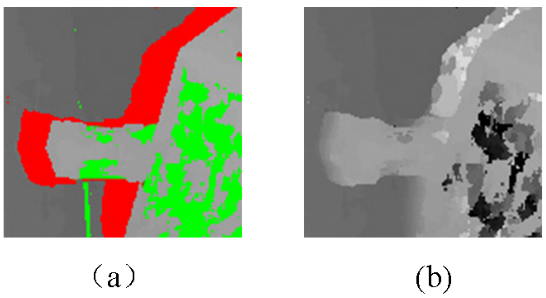


Fig. 1. Wrong decisions for Teddy (a) Some wrong decisions at depth discontinuities(Marked in RED) and weakly-texture regions(Marked in GREEN) in the disparity map with traditional window-based matching; (b) Disparity map with traditional window-based matching

So we could identify these invalid regions according to the occluded and boundary information. Meantime, we could adjust the disparities of invalid regions using the nearby information as they often have the similar disparities in the small neighborhood.

To sum up, our algorithm is based on the following matching constraints:

Epipolar Constraint: every point on a given epipolar line in one image must correspond to a point on the corresponding epipolar line in the other image. In our paper, we assume that all the image pairs are rectified, that is, the epipolar lines are in the horizontal directions.

Smooth Constraint: disparities vary usually smoothly within real images, except at object borders. As we identify the invalid regions especially at object borders in advance, we could effectively adjust the disparities of invalid regions using the nearby information. We use this constraint to identify the invalid regions with initial matching in the second step (see Section III.B.).

Boundary Constraint: every point at an object border in one image should correspond to a point at the corresponding border in the other image. We use this constraint as a necessary complement to the smooth constraint which does not work at object borders. We use both smooth constraint and smooth constraint to define our weight function in the third step (see Section III.A.).

Besides the above matching constraints, we also use the left-right check as its effectiveness of finding the invalidating occlusions and mismatches.

B. *Overview of the approach*

Our method can be partitioned into three blocks (Fig.2): initial matching, identification of invalid regions, and the weight-based disparity adjustment. Firstly, we use the traditional window-based matching as our initialization. Meanwhile, we do the left-right check in this step. Then, we identify the invalid regions whose disparities should be altered. Our identification is based on the result of the left-right check. Finally, we will assign a new disparity value for every invalid pixel defined in the former step. In this work, we define a new weight function to assign disparity values of the invalid regions defined above.

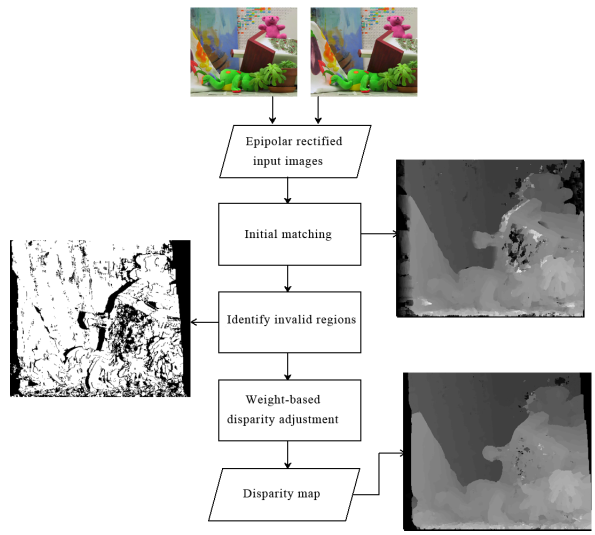


Fig. 2. Overview of our method

III. Detailed Description

In this section, we give a more detailed description of the three steps outlined above. Our order of description follows the algorithm procedure.

A. *Initialization*

We choose the traditional window-based methods as our initial matching for their simplicity and rapidity.

Our initial matching is a simple window-based matching method. Firstly, we use the truncated absolute differences (truncated AD) as our cost. After aggregating the matching cost with pixels in a square window whose size is win_size, the disparity with the lowest aggregated cost is selected (winner-takes-all) (Fig.3).

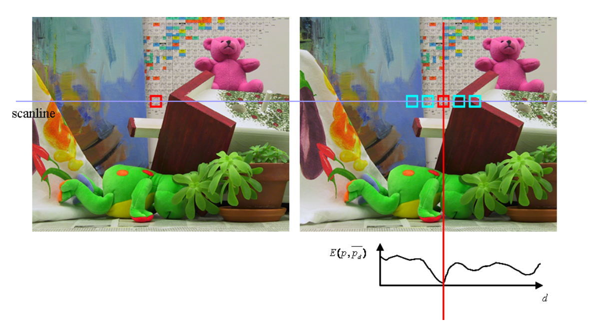


Fig. 3. Initial matching

The dissimilarity between pixel p in the left image and $\overline{p_d}$ in the right image, $E(p, \overline{p_d})$, can be expressed as

$$E(p, \overline{p_d}) = \sum_{q \in N_p, \overline{q_d} \in N_{\overline{p_d}}} \min\{ \sum_{c \in \{r,g,b\}} |I_c(q) - I_c(\overline{q_d})| \, , T\} \quad (1)$$

where I_c is the intensity of the color band c and T is the truncated value that controls the limit of the matching cost.

After the dissimilarity computation, the disparity d_p of pixel p is selected by the WTA(Winner-Takes-All) method as

$$d_p = \arg \min_{d \in S_d} E(p, \overline{p_d}) \quad (2)$$

where $S_d = \{d_{\min}, ..., d_{\max}\}$ is the set of all possible disparities.

Meantime, we do a left-right consistency check, whose result we will use at the second step. By reserving the aggregated costs, we could easily get the result of Left-Right check after a reverse matching computation.

B. Identification of invalid regions

Traditional window-based methods exhibit systematic errors, i.e., blurring blurring of object borders. Based on the discussion in Section II.A, we know that the matching errors usually happen at depth discontinuities and weakly-texture regions. To correct these errors efficiently, we identify these invalid regions in this step.

Our identification of invalid regions is based on the result of Left-Right check for its effectiveness of finding the invalidating occlusions and mismatches. Then, we extend each region to the nearest object borders. We take the nearest left pixel whose intensity of the color band L (lab color space) deflects to the left as its left border while taking the nearest right pixel whose intensity of the color band L deflects to the right as its right border.

So our BG (pixel set of invalid regions identified in this step) is defined as follows:

1) Initialize the BG with the pixels which do not pass the left-right check;
2) For every pixel
   - A left-extending
     - i. Find the nearest left pixel $LP(i, j-d)$ for $P(i, j)$ which satisfies the following condition: $abs(L(i, j-d) - L(i, j-d-1)) < abs(L(i, j-d) - L(i, j-d+1))$
     - ii. Stop if meeting with the pixel which belongs to BG while searching the LP;
     - iii. All the pixels between P and LP are assigned to BG ;
   - B right-extending
     - i. Find the nearest left pixel $RP(i, j+d)$ for $P(i, j)$ which satisfies the following condition: $abs(L(i, j+d) - L(i, j+d+1)) < abs(L(i, j+d) - L(i, j+d-1))$
     - ii. Stop if meeting with the pixel which belongs to BG while searching the RP;
     - iii. All the pixels between P and RP are assigned to BG ;

C. Weight-based disparity adjustment

Surfaces vary usually smoothly within real images, except at object borders, that is to say, the matching usually meets with Smooth Constraint and Boundary Constraint. According to this, we could adjust the disparities using neighborhood information. In the second step, we have already identify the invalid regions. Then we could increase the accuracy of matching by correcting the disparity values of pixels in these invalid regions.

Our disparity adjustment is based on the following two assumptions. In a finite region points which have similar color and short spatial distance should have the similar depths[2], and we also assume that disparity values vary usually smoothly, except at object borders.

We adjust the disparity for each pixel belonging to BG defined in the second step. Firstly, we calculate the weight of each pixel q in a local window N_p whose center is p (the pixel belonging to BG defined in the second step). Then we choose the disparity of the pixel whose weight is largest.

Our weight-based disparity assignment is based on a weight function which can be written as

$$w(p, q) = f_s(\Delta C_{pq}) \cdot f_p(\Delta G_{pq}) \cdot f_d(\Delta D_q) \quad (3)$$

where $f_s(\Delta C_{pq})$ and $f_p(\Delta G_{pq})$ represent the strength of grouping by similarity and proximity, respectively[2]. We particularly introduce the term $f_d(\Delta D_q)$ to represent the strength of ΔD_q which describes the local similarity of pixel q.

The color similarity is defined as

$$f_s(\Delta C_{pq}) = \exp(-\frac{\Delta C_{pq}}{r_c}) \quad (4)$$

where ΔC_{pq} represents the Euclidean distance between two colors in the Lab color space, r_c is 7.

The spatial proximity is defined as

$$f_s(\Delta G_{pq}) = \exp(-\frac{\Delta G_{pq}}{r_p}) \quad (5)$$

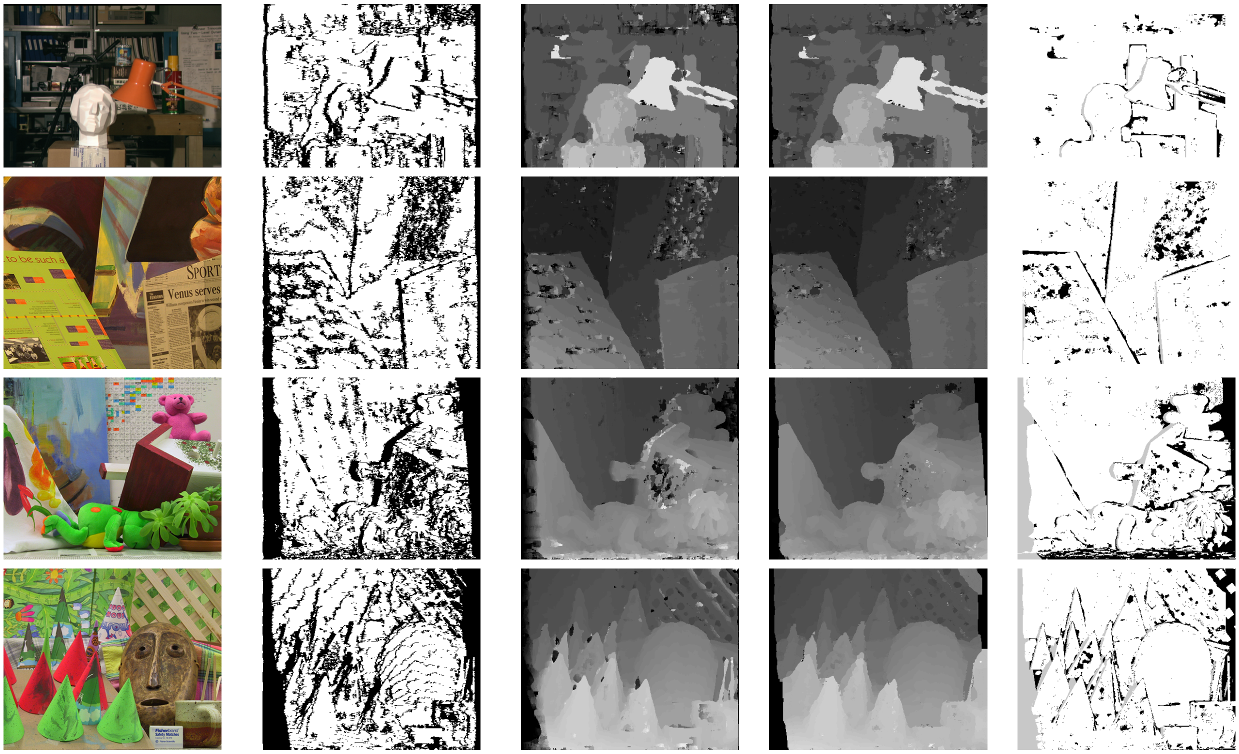
where ΔG_{pq} represents the Euclidean distance between p and q in the left image. The parameter r_p is determined according to the size of the adjustment window as $r_p \propto$ (window size).

The local depth similarity is defined as

$$f_s(\Delta D_q) = \exp(-\frac{\Delta D_q}{r_l}) \quad (6)$$

where ΔD_q describes the local depth similarity of q, and can be expressed as

$$\Delta D_q = \frac{sqrt(\sum_{r \in N_q, r \notin BG} (d_r - d_q)^2)}{N_r} \quad (7)$$



left image | BG (Marked in black) | Initial matching | Our result | 'bad pixel' maps

Performance comparison between our method and traditional local matching

Algorithm	Tsukuba			Venus			Tedday			Cones		
	nonoccl	all	on disc.	nonoccl	all	on disc.	nonoccl	all	on disc.	nonoccl	all	on disc.
Traditional matching	7.44	9.44	18.5	13.4	14.8	32.3	18.7	26.9	31.6	11.9	21.4	20.3
Our result	5.40	7.21	17.5	7.79	8.88	27.3	16.1	22.3	30.2	11.1	18.9	21.9

Fig. 4. Experimental results

where N_q is the pixel set within the square window whose center is q and size is d_win, BG is the pixel set defined in the second step. And d is the disparity for left image with initial matching. N_r is the number of pixel set r.

According to equation (4), (5) and (6), equation (3) can be rewritten as

$$w(p,q) = \exp(-\frac{\Delta C_{pq}}{r_c}) \cdot \exp(-\frac{\Delta D_q}{r_P}) \cdot \exp(-\frac{\Delta D_q}{r_l}) \quad (8)$$

After weight computation, the disparity of each pixel p in BG is assigned as the initial disparity of q whose weight is largest as

$$d_p = \{d_q | \arg \max_{q \in N_p, q \notin BG} w(p,q)\} \quad (9)$$

where d is the initial disparity. N_p respects the pixel set within the square window whose center is p and size is win_size.

IV. Experiments

In this section, we report the experimental result of our solution whose initial matching is traditional window-based matching. Firstly, we provide all of the parameter settings used in our algorithm which are shown in TABLE I. We use the same parameter settings for all the image pairs. Then we give our experimental results in Fig.4.

Initial matching	win_size	T			
	9	40			
Disparity adjustment	win_size	d_win	r_c	r_p	r_l
	35	15	7	17.5	16

TABLE I
Experimental parameter

V. Discussion and conclusion

In this paper, we focus on the area-based approaches and provide a solution with the weight-based modification for the invalid regions of local methods. And we also proposed a method to detect the invalid regions. The experiments show that our method could improve the calculation precision with little additional computation. Moreover our algorithm modules could be easily expanded into other matching algorithms as post-processing.

Although our algorithm is not the best on accuracy, there is still space left for improvement. For instance, in our algorithm, we make the traditional window-based algorithm as the initial matching whose accuracy is not so good. If the initial matching is improved, the result can be much better too. In particular, we plan to introduce the solution into other local methods and we will also try other adjustment methods for the invalid regions.

Acknowledgment

This work is supposed by National 863 High-Tech programe of China(2009AA01Z304), the National Natural Science Foundation of China(60933008), Shandong Province National Nature ScienceFoundation(No.Z2006G05), National Research Foundation for the Doctoral Program of Higher Education of China (20070422098).

References

[1] D.Scharstein and R.Szeliski, "A taxonomy and evaluation of dense two-frame stereo correspondence algorithms," *IJCV*, vol. 47, no. 1/2/3, pp. 7–42, 2002.

[2] A.Klaus, M.Sormann, and K.Karner, "Segment-based stereo matching using belief propagation and a self-adapting dissimilarity measure," *ICPR*, 2006.

[3] M.Bleyer, C. M.Gelautz, and C.Rhemann, "A stereo approach that handles the matting problem via image warping," *CVPR*, 2009.

[4] Q.Yang, L.Wang, R.Yang, H.Stewnius, and D.Nistr, "Stereo matching with color-weighted correlation, hierarchical belief propagation and occlusion handling," *PAMI*, 2008.

[5] Q.Yang, R.Yang, J.Davis, and D.Nistr, "Spatial-depth super resolution for range images," *CVPR*, 2007.

[6] Z.Wang and Z.Zheng, "A region based stereo matching algorithm using cooperative optimization," *CVPR*, 2008.

[7] H. Hirschmüller, "Improvements in real-time correlationbased stereo vision," *In Proc. IEEE Pacific-Rim Symposium on Image and Video Technology*, 2001.

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

[9] L. Stefano, M.Marchionni, S.Mattoccia, and G.Neri, "A fast area-based stereo matching algorithm," *15th IAPR/CIPRS International Conference on Vision Interface,Calgary,Canada*, no. 27-29, May 2002.

[10] A.Hosni, M.Bleyer, M.Gelautz, and C.Rhemann, "Local stereo matching using geodesic support weights," *ICIP*, 2009.

[11] F.Tombari, S.Mattoccia, and L. Stefano, "Segmentation-based adaptive support for accurate stereo correspondence," *PSIVT*, 2007.

[12] K.-J.Yoon and I.-S.Kweon, "Adaptive support-weight approach for correspondence search," *PAMI*, vol. 28, no. 4, pp. 650–656, 2006.

[13] M.Gerrits and P.Bekaert, "Local stereo matching with segmentation-based outlier rejection," *In CRV*, 2006.

[14] F.Tombari, S.Mattoccia, and L.DiStefano, "Segmentation-based adaptive support for accurate stereo correspondence," *In Proc. IEEE Pacific-Rim Symposium on Image and Video Technology*, 2007.