A Fast Block Matching Algorthim for Stereo Correspondence

Tangfei Tao[1,2], Ja Choon Koo[1],Member,IEEE, Hyouk Ryeol Choi[1] ,Member, IEEE

[1]School of Mechanical Engineering,
Sungkyunkwan University,
300 Cheoncheon-dong, Jangan-gu, Suwon,
Gyeonggi-do 440-476, Korea
jckoo@skku.edu,
Corresponding Author: choi@me.skku.ac.kr

[2] School of Mechanical Engineering,
Xi'an Jiaotong University,
Xi'an, Shaanxi, 710049, P.R.China
taotangfei@mail.xjtu.edu.cn

Abstract—Stereo correspondence is one of the most active research areas in computer vision. An improved block matching approach for fast stereo correspondence is proposed, where the matching criterion is the sum of absolute difference (SAD). The approach tries to progressively confine the matching operation during the searching process, terminates unnecessary computations of matching criteria between the reference block in the left image and the ineligible candidate block in the right image, and eliminates the ineligible block as early as possible while ensuring the optimal disparity of each pixel. The performance of the new algorithm is evaluated by carrying out a theoretical analysis, and by comparing it with the stereo correspondence method based on the standard block matching. Simulation results demonstrate that the disparities obtained by this algorithm are identical to that using standard block matching method, and a reduction of over 55% in computational cost is achieved.

Keywords— **stereo correspondence, block matching, disparity, SAD, computer vision**

I. INTRODUCTION

Stereo correspondence problem has historically been, and continues to be, one of the most investigated topics in computer vision, and a larger number of literatures on it have been published[1]-[6]. The correspondence problem in computer vision concerns the matching of points, or other kinds of primitives, in two or more images such that the matched elements are the projections of the same physical elements in 3D scene, and the resulting displacement of a projected point in one image with respect to the other is termed as disparity. Similarity is the guiding principle for solving the correspondence problem; however, the stereo correspondence problem is an ill-posed task, in order to make it tractable, it is usually necessary to exploit some additional information or constraints. The most popular constraint is the epipolar constraint [7], which can reduce the search to one-dimension rather than two. Other constraints commonly used are the disparity uniqueness constraint and the continuous constraint.

The existing techniques for general two-view stereo correspondence roughly fall into two categories: local method and global method [6]. Local methods use only small areas/neighborhoods surrounding the pixels, while global methods optimize some global (energy) function. Local methods, such as block matching [8], gradient-based optimization [9], and feature matching [10] can be very efficient, but they are sensitive to locally ambiguous regions in images (e.g., occlusion regions or regions with uniform texture). Global methods, such as dynamic programming [11], intrinsic curves [12], graph cuts [13], and belief propagation [14] can be less sensitive to these problems since global constraints provide additional support for regions difficult to match locally. However, these methods are more expensive in their computational cost.

The block matching method is one of the most popular local methods because of its simplicity in implementation. The basic idea of block matching for stereo correspondence is as follows: to estimate the disparity of a point in the left image, firstly, we define a reference block surrounding this point; and then, find the closest matched block, within a search rang in the right image, using a pre-specified matching criterion; thus, the relative displacement between the reference block and the closest matched block constitutes the disparity of the point being evaluated. The commonly used matching criteria are the sum of absolute differences (SAD), the sum of squared differences (SSD) and the normalized SSD [6].

With the given matching criteria, the correspondence problem results in essentially a search problem, and the standard search method for block matching is an exhaustive search, where the matching criterion is calculated for all pixels at all possible search positions. This strategy can guarantee that the best-matched block is found with respect to the chosen criterion. However, the computation loads of such methods are very demanding, even by using the epiploar lines constraint, and therefore many different fast algorithms have been developed [15].

In order to obtain the disparity of a point, the candidate block in the search range that best matches the reference block is of the main interest, not the exact value of the matching criterion between the reference block and all the possible candidates. In other words, it is no necessary to calculate the accurate matching error for the entire region at all search positions, the computation for every single candidate block can be terminated, and this block should be rejected whenever the partial error measure between this candidate block and the reference block exceeds a previously found smaller error measure. Considering this, an improved block matching for fast stereo correspondence is proposed in this paper.

The remainder of the paper is organized as follows: Section 2 presents the improved block matching approach for fast stereo correspondence. Section 3 shows the performance evaluation of the proposed method, and section 4 gives conclusions.

II. THE PROPOSED BLOCK MATCHING APPROACH

A. Main idea of the new algorithm

Here, it is assumed that a binocular stereo pair of images is captured, and the stereo pair has been rectified already, so the disparity value is surveyed along the identical horizontal scan line under the epipolar constraint [7].

For simplicity, RB (reference block) is used to represent the small block surrounding the point (x, y) being considered in the left image, CB (candidate block) to denote a corresponding block surrounding the point $(x, y+d_h)$ in the right image, where d_h is the possible disparity. Let $RB(x+i, y+j)$ and $CB(x+i, y+d_h+j)$ represent the intensities of the pixel with coordinate $(x+i, y+j)$ in the RB in the left image, and the pixel with coordinate $(x+i, y+d_h+j)$ in the CB in the right image, respectively, where $-n \le (i, j) \le n$. Assuming that the block size is $(2n+1)\times(2n+1)$ and, $d_{\max}$ is the maximum allowed disparity, that is, the search range is from $-d_{\max}$ to $d_{\max}$. Here, the sum of absolute difference (SAD), due to its computational efficiency [6], is used to evaluate the similarity between these two blocks RB and CB . The sum of absolute difference is defined as (1), where $|d_h| \le d_{\max}$.

$$SAD(x, y, d_h) = \sum_{i=-n}^{n}\sum_{j=-n}^{n}\left|RB(x+i, y+j) - CB(x+i, y+d_h+j)\right| \tag{1}$$

The candidate block which minimizes the SAD function, is selected as the best matched block, and the corresponding disparity d_h is the disparity of the point (x, y) being evaluated.

Let $SAD_j(x, y, d_h)$ represent the sum of absolute difference of the jth column between the RB and CB with disparity d_h , then $SAD_j(x, y, d_h)$ can be expressed as (2), where $|d_h| \le d_{\max}$.

$$SAD_j(x, y, d_h) = \sum_{i=-n}^{n}\left|RB(x+i, y+j) - CB(x+i, y+d_h+j)\right| \tag{2}$$

So, (1) can be rewritten as (3), where $|d_h| \le d_{\max}$.

$$SAD(x, y, d_h) = \sum_{j=-n}^{n} SAD_j(x, y, d_h) \tag{3}$$

Let $SAD^k(x, y, d_h)$ represent the accumulated sum of the absolute difference of the first k column(s) (assume that the columns of the images are scanned from left to right) between the RB and CB with disparity d_h , then $SAD^k(x, y, d_h)$ can be defined as (4).

$$\begin{gathered} SAD^k(x, y, d_x) = \sum_{j=-n}^{-n+k} SAD_j(x, y, d_h) \\ where \quad 0 \le k \le 2n \quad and \quad |d_h| \le d_{\max} \end{gathered} \tag{4}$$

So, when k is $2n$, $SAD^{2n}(x, y, d_h)$ denotes the exact SAD between these two blocks RB and CB , that is,

$$\begin{aligned} SAD(x, y, d_h) &= SAD^{2n}(x, y, d_h) \\ &= \sum_{j=-n}^{-n+2n} SAD_j(x, y, d_h) = \sum_{j=-n}^{n} SAD_j(x, y, d_h) \end{aligned} \tag{5}$$

Obviously, (6) holds true.

$$\begin{aligned} & SAD^0(x, y, d_h) \le SAD^1(x, y, d_h) \le SAD^2(x, y, d_h) \\ & \le \ldots \le SAD^{2n}(x, y, d_h) = SAD(x, y, d_h) \end{aligned} \tag{6}$$

Let $SAD_{\min}(x, y, d_{so-far})$ denote the minimum SAD obtained so far, for the corresponding best-matched block with disparity d_{so-far} , within the search window during the previous search process. For any other untested candidate block, the algorithm first calculates $SAD^0(x, y, d_h)$, and then compares $SAD^0(x, y, d_h)$ with $SAD_{\min}(x, y, d_{so-far})$. If $SAD^0(x, y, d_h)$ is greater than the $SAD_{\min}(x, y, d_{so-far})$, the exact $SAD(x, y, d_h)$ between these two blocks RB and CB is surely greater than $SAD_{\min}(x, y, d_{so-far})$, it means that this candidate block can be eliminated at once from being considered as the best matched block, and it is unnecessary to calculate the exact $SAD(x, y, d_h)$ between them. Otherwise, $SAD^1(x, y, d_h)$ is calculated and then checked. If $SAD^1(x, y, d_h)$ is greater than the $SAD_{\min}(x, y, d_{so-far})$, for the same reason as above, this candidate block can be eliminated and the calculating of $SAD(x, y, d_h)$ can be terminated. If it is not, $SAD^2(x, y, d_h)$ is calculated, and then checked. The process is repeated until this candidate block is discarded or the last column level is arrived. If a candidate block cannot be eliminated before calculating $SAD^{2n}(x, y, d_h)$, the exact SAD between the reference block and this candidate block with disparity d_h will be calculated. If $SAD^{2n}(x, y, d_h)$ is still less than $SAD_{\min}(x, y, d_{so-far})$, it indicates a new best-matched candidate block appears at that point, and then the current $SAD_{\min}(x, y, d_{so-far})$ is replaced by this $SAD^{2n}(x, y, d_h)$. By such an updating strategy of the so far minimum SAD, $SAD_{\min}(x, y, d_{so-far})$, an increasingly tighter bound for the matching criteria is obtained, and this will progressively confine the search space and, hence, the computation of matching error between the reference block and the remaining ineligible candidate block(s) can be terminated and the ineligible block(s) can be discarded as early as possible.

According to such a strategy, it is clear that a large number of non-best candidates can be eliminated before the arrival of calculating $SAD^{2n}(x, y, d_h)$, at last, there are only a small number of candidate blocks remaining for the calculating of the exact SAD. Since evaluating the exact SAD between two blocks requires considerably more time than when evaluating the $SAD^k(x, y, d_h)$ between the fist k column(s) of the two blocks, the elimination of a number of blocks before their exact $SADs$ are evaluated can save a great deal of computational time. Furthermore, no assumption, which can sacrifice the optimal estimation, is made on the matching criteria and searching strategy, therefore, the algorithm eliminates only blocks with 100% certainty of elimination using standard block matching, and it will not exclude the optimum matched block and the optimum disparity,

Hence, according to this idea, the block matching approach can be improved to greatly reduce the computation burden of the matching process without sacrificing the optimality of the solution.

B. The prediction of the initial disparity

The choice of the initial disparity d_{init} affects the efficiency of this algorithm to some extent, and a good initial choice can avoid unnecessary computation as much as possible and eliminate the non-optimal candidate blocks as early as possible. In this work, considering the continuous assumption of disparity, the initial disparity d_{init} of the pixel (x, y) being evaluated is initialized according to the following strategy.

1) If this pixel (x, y) is the upper left pixel in the left image, that is, if $(x, y) = (0,0)$, let its $d_{init} = 0$;

2) If this pixel (x, y) is the left but not the upper pixel in the left image, i.e., if $(x > 0, y = 0)$, let its d_{init} equal the disparity of its upper adjacent pixel $(x-1,0)$, which has been obtained;

3) In other cases, let the pixel's d_{init} equal the disparity of its left adjacent pixel $(x, y-1)$, which has been obtained already.

This strategy has considered the continuous assumption of disparity to the greatest extent, so it can give the best initial disparity estimation and speed up the proposed method reducing the computational burdens.

III. PERFORMANCE EVALUATION

The improved block matching algorithm can guarantee to get the same optimal result as the standard block matching, because during the search process, about the matching criteria and search strategy, there is no assumption sacrificing the optimal result; the algorithm eliminates only blocks that with 100% certainty will be eliminated using standard block matching. Theoretically, the proposed algorithm can reduce the computation burden obviously, because this algorithm can terminate the unnecessary computation of matching evaluation, which require very expensive computation costs as early as possible and eliminate as many non-optimal search positions in the search range as possible.

Further more, some simulations have been conducted to validate this algorithm by comparing the computation costs of the correspondence method using the proposed algorithms and that based on the standard block matching. A variety of images have been tested, two pairs of them, including Map, Tsukuba, are listed [5] [16]. Due to saving space, the images on the website at http://vision.middlebury.edu/stereo/data/ have not been presented here. Two important factors, the search range (the maximum disparity allowed) and the window size, must be determined properly to improve the overall performance of the algorithm, and are referred to in [5],[6], the window size and search range set for each image are listed in Table 1.

TABLE I. COMPUTATIONAL COSTS COMPARISON OF THE PROPOSED METHOD AND THE STANDARD BLOCK MATCHING METHOD

Image Pairs	**Image Size**	**Search Range**	**Window Size**	**Average Executing Time(Sec.)**	
				Standard	***Proposed***
Map	284x216	29	7x7	4.46	1.93
			9x9	7.55	3.36
Tsukuba	384x288	15	7x7	3.69	1.67
			9x9	6.99	2.99

As shown in Table 1, to complete the disparities estimation for all pixels in the left image, compared with the standard block matching method, the improved block matching method is much faster. The computational costs of the stereo correspondence method based on the improved block matching are reduced by approximately 55% to 58% in comparison with that using standard block matching. The result also shows that the window size and search range are two important factors that affect the computation burden.

IV. CONCLUSION

The block matching algorithm using SAD as the matching criteria has been improved for fast stereo correspondence in this paper. By evaluating the lower bound, which becomes increasingly tighter for the matching criteria, the improved method can successively terminate unnecessary computations of the matching criteria between the reference block in the left image, and the ineligible candidate blocks in the right image, and then progressively eliminate the ineligible blocks as early as possible, while ensuring the optimal disparity estimation. In addition, the proposed method can further speed up the elimination of ineligible candidate blocks by efficiently using the continuous constraint of disparity to predict the initial disparity of each pixel. With far fewer computational requirements, the stereo correspondence algorithm using the improved block matching achieves the same disparity estimation accuracy as the correspondence method based on the standard block matching. The simulation results show that the computations using the proposed method are reduced by over 55%, compared with those based on the standard block matching.

ACKNOWLEDGMENT

This work is supported by the Brain Korea 21(BK21) Program of Ministry of Education and Human Resources Development, and the authors are grateful to the owners of the images used in this paper.

REFERENCES

[1] S.T. Barnard and M.A. Fischler, "Computational stereo," ACM Computing Surveys, vol. 14, pp. 553–572, 1982.

[2] Umesh R. Dhond and J. K. Aggarwal, "Structure from stereo-A review,"IEEE Trans. Systems, Man, and Cybernetics, vol. 19, no. 6, pp. 1489–1510,1989.

[3] Juyang Weng,Narendra Ahuja, and Thomas S. Huang, "Matching two perspective view," IEEE Trans. Pattern Analysis and Machine Intelligence, vol.4,no. 8,pp.806–825,1992.

[4] A. Koschan, "What is new in computational stereo since 1989: A survey of current stereo papers," Technical Report 93-22, Technical Univ. of Berlin, 1993.

[5] Daniel Scharstein, Richard SZeliski, "A taxonomy and evaluation of dense two-frame stereo correspondence algorithms,"International Journal of Computer Vision,vol. 47,no.1,pp.7–42,2002.

[6] Myron Z. Brown, Darius Burschka, and Gregory D. Hager, "Advances in computational stereo," IEEE Trans. Pattern Analysis and Machine Intelligence, vol. 25, no. 8, pp.993–1001,2003.

[7] D. V. Papadimitriou and T. J. Dennis,"Epipolar line estimation and rectification for stereo image pairs," IEEE Trans. Image Processing,vol.5,no.4, pp.672–676, 1996.

[8] O. Faugeras, B. Hotz, H. Matthieu, T. Vieville, Z. Zhang, P. Fua, E.Theron, L. Moll, G. Berry, J. Vuillemin, P. Bertin, and C. Proy, "Real time correlation-based stereo: algorithm, implementations and applications," INRIA Technical Report 2013, 1993.

[9] B.D. Lucas and T. Kanade, "An iterative image registration technique with an application to stereo vision," Proc. Int'l Joint Conf. Artificial Intelligence, pp. 674-679, 1981.

[10] C. Schmid and A. Zisserman, "The geometry and matching of curves in multiple views," Proc. European Conf. Computer Vision, pp. 104-118, 1998.

[11] Clark F. Olson, "Maximum-likelihood image matching," IEEE Trans. Pattern Analysis and Machine intelligence, vol. 24, no. 6,pp.853–857,2002.

[12] C. Tomasi and R. Manduchi, "Stereo matching as a nearest-neighbor problem," IEEE Trans. Pattern Analysis and Machine Intelligence, vol. 20, pp. 333-340, 1998.

[13] V. Kolmogorov and R. Zabih, "Computing visual correspondence with occlusions using graph cuts," Proc. Int'l Conf. Computer Vision, 2001.

[14] Jian Sun, Nan-Ning Zheng, and Heung-Yeung Shum, "Stereo matching using belief propagation," IEEE Trans. Pattern Analysis and Machine intelligence, vol. 25, no. 7, pp.1–14,2003.

[15] Harris Sunyoto, Wannes van der Mark,and Dariu M. G avrila, "A comparative study of fast dense stereo vision algorithms,"IEEE Intelligent Vehicles Symposium, University of Parma, Parma, Italy pp.319-324,June 14-17,2004.

[16] http://vision.middlebury.edu/stereo/data/.