An Efficient Algorithm for Stereo Correspondence Matching

Mozammel Chowdhury, Junbin Gao, and Rafiqul Islam

***Abstract*—Stereo matching has gained the popularity in computer vision and image processing. The objective of stereo correspondence matching is to obtain dense depth information of objects for 3D reconstruction and modeling. Several stereo correspondence algorithms have been developed in last couple of years. However, they are not suitable for real time applications due to their limitations of high computational cost. This paper proposes an efficient algorithm for stereo correspondence matching, which is fast and capable of tackling additive noise. The algorithm estimates average disparity within a search range and attains the benefit of mean filtering. Experimental evaluations demonstrates the effectiveness of our proposed algorithm comparable to the existing stereo methods.**

***Index Terms*—Stereo matching, window costs, disparity, 3D reconstruction, modeling, virtual reality.**

I. INTRODUCTION

Stereo correspondence matching is a popular and important tool in computer or robot vision which is essential for determining three-dimensional depth information of objects by using a pair of left and right images captured by a stereo camera system. Dense depth information are required in applications like robot navigation and control, multimedia, virtual reality, teleconferencing, 3D modeling and sensing, traffic scene analysis and so on [1].

Stereo vision technique is similar in concept to human binocular vision. In stereo, two images of the same scene are taken from slightly different viewpoints using two cameras placed in the same lateral plane [2]. Consequently, for most pixels in the left image there is a corresponding pixel in the right image in the same horizontal line. The difference in the coordinates of the corresponding pixels is known as disparity d, as shown in Fig. 1, can be expressed by the following equation.

$$d = x_L - x_R \tag{1}$$

The disparity is inversely proportional to the distance of the objects from the camera position [3]. Using the disparity, the depth z can be defined by the following equation:

$$z = \frac{bf}{d} \tag{2}$$

where d is the disparity, z is the distance of the object point from the camera (the depth), b is the base distance between the left and right cameras, and f is the focal length of the camera lens. Fig. 2 shows that the two images of an object are obtained by the left and right cameras observing a common scene. This pair of stereo images allows us to obtain the 3D information or the distance of the object.

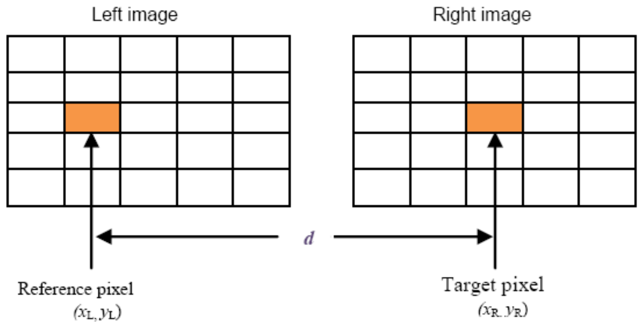


Fig. 1. The pixel x_R in the right image is the corresponding pixel of x_L in the left image.

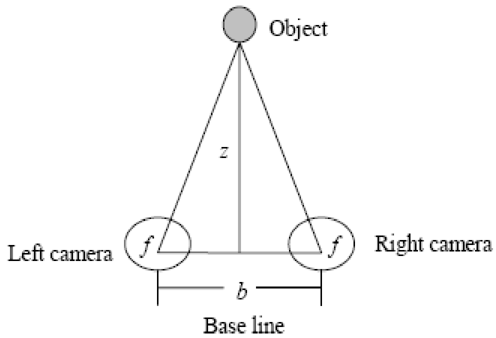


Fig. 2. The stereo imaging technique.

Stereo matching algorithms can be classified into two categories: global and local algorithms. Global algorithms [4]-[7] rely on iterative schemes that carry out disparity assignments on the basis of the minimization of a global cost function. These algorithms yield accurate and dense disparity measurements but exhibit a very high computational cost that renders them unsuited to real-time applications. Local algorithms [8]-[10], also called area-based algorithms, calculate the disparity at each pixel on the basis of the photometric properties of the neighboring pixels. Compared to global algorithms, local algorithms yield significantly less accurate stereo disparity but can run fast enough to be deployed in many real-time applications. In this paper, we have proposed a fast and efficient local algorithm for stereo correspondence matching.

The pipeline of the proposed stereo system is shown in Fig. 3. Our stereo estimation pipeline consists of the following stages: (i) Input of the stereo image pair, (ii) RGB to Gray scale conversion, (iii) Window cost computation, (iv)

Manuscript received August 11, 2015; revised November 11, 2015.

Mozammel Chowdhury, Junbin Gao, and Rafiqul Islam are with the School of Computing and Mathematics, Charles Sturt University, Australia (e-mail: mochowdhury@csu.edu.au).

DOI: 10.7763/IJCTE.2017.V9.1114

Disparity estimation, (v) Final disparity map.

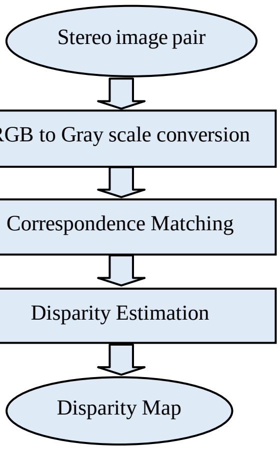


Fig. 3. The processing steps of the proposed stereo system.

II. Correspondence Matching and Disparity Estimation

The local or window-based stereo algorithms conventionally estimate disparity or depth information based on matching windows of pixels in the left and right image sequences by using sum of square differences (SSD), sum of absolute differences (SAD), or normalized correlation techniques [3]. To determine the correspondence of a pixel in the left image using SSD, SAD or normalized correlation techniques the window costs are computed for all candidate pixels in the right image within the search range. The pixel in the right image that gives the best window cost i.e., the minimum SSD or SAD value or the maximum correlation value is the corresponding pixel of the left image.

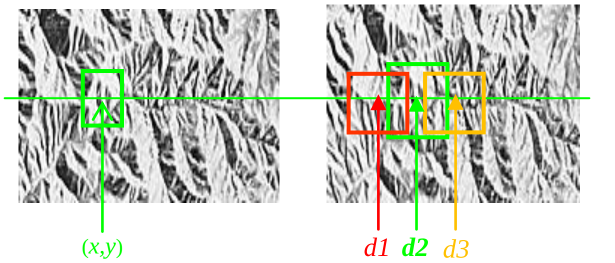
The window cost $W_C(x,y,d)$ of a reference pixel at position (x,y) in the left image with disparity d, can be estimated by any one of the following measures. In this work, we just calculate the SAD score between a window centered at position (x, y) in the left image and the corresponding window centered at position $(x+d, y)$ in the right image.

$$W_c^{SAD}(x,y,d)=\sum_{i=-m}^{m}\sum_{j=-n}^{n}\left|f_L(x+i,y+j)-f_R(x+i+d,y+j)\right| \quad (3)$$

where $f_L(x,y)$ and $f_R(x,y)$ are the intensities of the pixels at position (x,y) in the left and right images, respectively. $(2m+1)$ and $(2n+1)$ are the width and height of the rectangular window, respectively.

III. Proposed Algorithm

This section presents our proposed stereo matching algorithm. The proposed method is an improvement of our previous method [11]. In this work, for a reference pixel (x,y) in the left image we search the correspondence pixel $(x+d, y)$ in the right image within a searching range $-d_{max}$ to $+d_{max}$. To achieve a substantial gain in accuracy with less expense of computation time, our algorithm perform searching operation for correspondence matching only on the diagonally neighbor pixels within a square window in the right image rather than employing conventional direct search upon all pixels within the window. Thus, several disparity values are obtained within the search range and the average disparity of these values is computed. Fig. 4 shows that within the range $-d_{max}$ to $+d_{max}$ (d_{max} = 1) in the right image, five windows are selected for estimation of window costs for a reference pixel (x, y) in the left image. After finding several disparities (d_1, d_2, d_3), average disparity d is obtained. Through this procedure, disparities are calculated for all reference pixels in the left image.



(a) Reference pixel (x, y) in the Left image. (b) Disparity values within the selected windows in the Right image within the

Fig. 4. Disparity calculation for the candidate pixel in the Left image by searching the corresponding pixels in the right image within the searching range from $-d_{max}$ to $+d_{max}$ (d_{max} = 1).

The proposed Average Disparity algorithm is pointed out below:

Algorithm Proposed stereo correspondence matching

1. For each candidate pixel (x, y) in the left image, search the corresponding pixel on same epipolar line in the right image within a search range employing a square diagonal window:

   for $d' = -d_{max}$ to $+d_{max}$ **do**

   if

   $|f_L(x,y)-f_R(x+d',y)| <$ threshold **then**

   Calculate $W_C(x,y,d')$

2. Find d such that, $d = \arg\min W_C(x,y,d')$
3. Calculate the average disparity d_{avg}.
4. Repeat steps 1 to 3 to calculate disparities of all pixels in the left image.

TABLE I: Disparity Estimation Accuracy

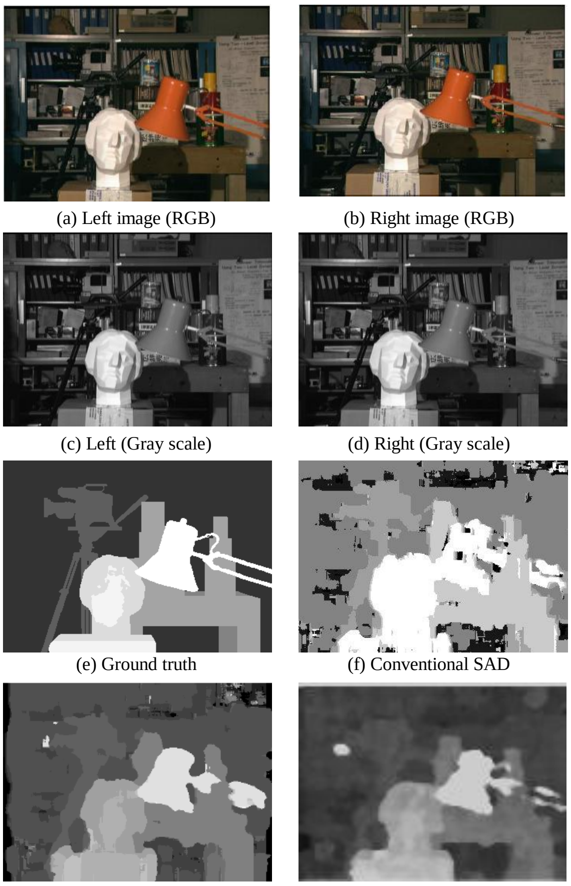
"D:\Debug\image_LR2.exe"

Threshold value (pixel)	Window size (pixel)	Accuracy (%) correct matching
Without threshold	3 x 3	51.4 %
	7 x 7	67.0 %
	11 x 11	72.3 %
δ=5	3 x 3	49.6 %
	7 x 7	62.1 %
	11 x 11	64.6 %
δ=10	3 x 3	50.5 %
	7 x 7	64.3 %
	11 x 11	67.7 %
δ=15	3 x 3	50.9 %
	7 x 7	65.4 %
	11 x 11	69.2 %
δ=20	3 x 3	51.1 %
	7 x 7	65.9 %
	11 x 11	70.1 %
δ=25	3 x 3	51.1 %
	7 x 7	66.1 %
	11 x 11	70.5 %
δ=30	3 x 3	51.2 %
	7 x 7	66.3 %
	11 x 11	70.9 %

Press any key to continue_

IV. Experimental Results

In order to demonstrate the effectiveness of the algorithm, we present the processing results with a standard image pair, including with the ground-truth values for quantitative comparison with other methods. The images are taken from Middlebury stereo datasets [12]. Fig. 5 shows the Tsukuba Head image set with ground truth and corresponding disparity map. In this figure, the ground truth image is histogram equalized for visualization purpose. Experiments are carried out on an Intel Core i3, 1.6 GHz PC with 4 GB RAM. The algorithm has been implemented using Visual C++. Disparities are estimated by SAD method. Table I summarizes the obtained disparity estimation results. The accuracies shown in this table represent the percentage of correct disparities (i.e. same value as that of ground truth).



(a) Left image (RGB) (b) Right image (RGB)

(c) Left (Gray scale) (d) Right (Gray scale)

(e) Ground truth (f) Conventional SAD

(g) Average method [11] (h) Proposed method

Fig. 5. Stereo images (Tsukuba Head) with ground truth.

TABLE II: COMPARISON OF COMPUTATION TIME

Disparity Estimation Method	Window size	Computation time (second)
Conventional SAD	3×3	1.00
Average Disparity [11]	3×3	0.51
Proposed Method	3×3	0.32

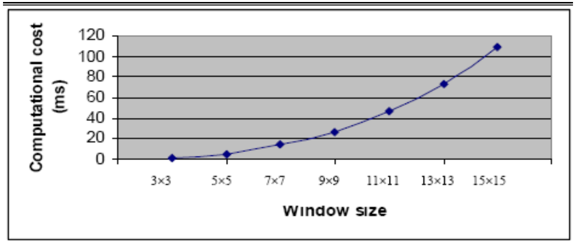


Fig. 6. Computational cost versus window size.

The computational time versus window size graph is shown in Fig. 6, which reveals that the computational cost increases with the size of the window. Therefore, a window of size 3×3 performs better results than any other window size. Table II shows the summary of comparison between traditional window-based method and our proposed method. Experimental results predict that our proposed method provides a reduction of about 70% of computational time.

V. CONCLUSION

A fast stereo correspondence algorithm has been developed that analyses color images to estimate the disparity map. This is achieved by finding the corresponding pixel in a pair of image sequences and computing the disparity values. Numerous stereo vision applications for obtaining 3D scene of objects require fast computation of dense stereo disparity. To cope with this problem, a very simple and fast technique is proposed in this research which is applicable in real time systems. Experimental results demonstrate the effectiveness of the proposed algorithm and confirm that our method is capable of reducing computation time of about 50% with no appreciable degradation of accuracy. Our next plan is to improve the algorithm than can be able to cope with occlusion problem and mismatch of epipolar geometry.

REFERENCES

[1] N. Lazaros, G. C. Sirakoulis, and A. Gasteratos, "Review of stereo vision algorithms: From software to hardware," *International Journal of Optomechatronics*, vol. 2, pp. 435–462, 2008.

[2] M. H. Chowdhury, Md. A. M. Mondal, and Md. A.-A. Bhuiyan, "3D imaging using stereo vision," in *Proc. the 7th International Conference on Computer and Information Technology (ICCIT'04)*, Dhaka, Bangladesh, 2004, pp. 307-312.

[3] M. Chowdhury and M. Bhuiyan, "Fast window based stereo matching for 3D scene reconstruction," *The International Arab Journal of Information Technology*, vol. 10, no. 4, 2013.

[4] M. Bleyer and M. Gelautz, "Simple but effective tree structures for dynamic programming-based stereo matching," in *Proc. Int. Conf. Computer Vision Theory and Applications*, pp. 415–422, 2008.

[5] O. Veksler, "Stereo correspondence by dynamic programming on a tree," in *Proc. IEEE Conf. Comput. Vis. Pattern Recog.*, 2005, pp. 384–390.

[6] Q. Yang, "Stereo matching using tree filtering," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 37, no. 4, pp. 834-846, 2015.

[7] J. Sun, N. Zheng, and H. Y. Shum, "Stereo matching using belief propagation," *IEEE Trans. Pattern Anal. Mach. Intell*., vol. 25, no. 7, pp. 787–800, 2003.

[8] L. De-Maeztu, S. Mattoccia, A. Villanueva, and R.l Cabeza, "Linear stereo matching," in *Proc. IEEE International Conference on Computer Vision*, 2011, pp. 1708-1715.

[9] D. Stefano, Luigi, M. Marchionni, and S. Mattoccia, "A fast area-based stereo matching algorithm," *Image and Vision Computing*, vol. 22, no. 12, pp. 983–1005, 2004.

[10] A. Geiger, M. Roser, and R. Urtasun, "Efficient large-scale stereo matching," in *Proc. Asian Conference on Computer Vision*, 2010.

[11] M. Mozammel Hoque Chowdhury and Md. Al-Amin Bhuiyan; "A new approach for disparity map determination," *Daffodil International University Journal of Science and Technology*, Dhaka, Bangladesh, vol. 4, no. 1, pp. 16-19, 2009.

[12] Middlebury stereo datasets. [Online]. Available: http://vision.middlebury.edu/stereo/data/

Mozammel Chowdhury is a doctoral researcher at the School of Computing and Mathematics, Charles Sturt University, Australia. Prior to this, Mr. Chowdhury was a faculty member at the Department of Computer Science and Engineering, Jahangirnagar University, Bangladesh. He has expertise in computer vision, machine intelligence, image analysis, cyber-security, tele-health, e-learning and e-governance. Mr. Chowdhury received his B.Sc. (honors) and MS (research) degrees in computer science and engineering from

Jahangirnagar University. He has published around 40 research articles in reputed journals and conference proceedings. He is an editorial board member of some international and national journals.

Junbin Gao received the B.Sc. degree in computational mathematics from the Huazhong University of Science and Technology (HUST), China in 1982, and the Ph.D. degree from the Dalian University of Technology, China, in 1991. He was a lecturer and senior lecturer of Computer Science with the University of New England, Australia, from 2001 to 2005. From 1982 to 2001, he was an associate lecturer, a lecturer, an associate professor, and a professor with the Department of Mathematics at HUST. He is currently a professor of computing science with the School of Computing and Mathematics, Charles Sturt University, Australia. His main research interests include machine learning, data mining, Bayesian learning and inference, and image analysis.

Rafiqul Islam is a senior lecturer in the School of Computing and Mathematics and the leader of Cyber Security Research Area, Faculty of Business, at Charles Sturt University, Australia. Dr. Islam has expertise in network & information security, cyber-security, sensor network and machine learning. He received his PhD and postdoctoral experience from Deakin University, Australia. Dr. Islam was a lead researcher on various research projects including several ARC funded projects. Dr. Islam published more than 90 refereed research papers. Dr. Islam is also an editorial board member for an international journal and the general chair and a technical program committee (TPC) member of various international conferences.