

A NEW APPROACH FOR DISPARITY MAP DETERMINATION

M. Mozammel Hoque Chowdhury and Md. Al-Amin Bhuiyan

Department of Computer Science and Engineering

Jahangirnagar University, Savar, Dhaka-1342, Bangladesh.

E-mail: mozammel_ju@yahoo.com, alamin_bhuiyan@yahoo.com

Abstract: This article presents a new method to determine disparity map useful for three-dimensional (3D) scene reconstruction. The main task behind the computation of disparity map is stereo correspondence matching. In recent years, several stereo matching algorithms have been developed to find corresponding pairs in two images: left and right images captured by a stereo camera. But these algorithms exhibit a very high computational cost. With a view to reduce the computation time and produce a smooth and detailed disparity map, a fast and new approach based on average disparity estimation is proposed in this research, which can tackle additive noise. Experimental results confirm that the method achieves a substantial gain in accuracy with less expense of computation time.

Key Words: Disparity map, Stereo correspondence, Stereo Vision, 3D Scene Reconstruction.

1. Introduction

An important research topic in image processing and computer vision is stereo correspondence matching. The stereo correspondence problem is one of the main tasks for obtaining disparity map. Determinations of disparity maps are required in applications like robot navigation and control, machine vision, medical imaging and 3D reconstruction.

Stereo correspondence algorithms can be classified into two categories: *global* and *local* algorithms. *Global* algorithms [1] 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 [2-5], 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 disparity maps but can run fast enough to be deployed in many real-time applications.

In stereo correspondence matching, since two images of the same scene are taken from slightly different viewpoints using two cameras, placed in the same lateral plane, so, 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, which is inversely proportional to the distance of the objects from the camera. The disparity can be defined by the following equation:

$$d = \frac{bf}{z}. \quad (1)$$

where 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. Figure 1 shows that the two images of an object are obtained by the left and right cameras observing a common scene. Finding the same points in two images such that the matched points are the same projections of a point in the scene is called stereo correspondence matching and is the fundamental computational task underlying disparity map determination.

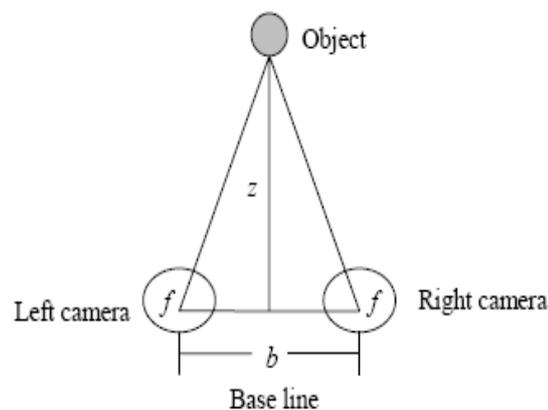


Figure 1. Disparity estimation technique

2. Stereo Disparity Estimation

Stereo correspondence or disparity is conventionally determined based on matching windows of pixels by using sum of square

differences (SSD), sum of absolute differences (SAD), or normalized correlation techniques [6-14]. 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 is the corresponding pixel of the left image.

The basics of stereo correspondence matching are as follows:

For each epipolar line

For each pixel in the left image

- *compare with every pixel on same epipolar line in right image*
- *pick pixel with minimum match cost*

The computation of window cost is given by Eq. (2).

$$W_C(x, y, d) = \begin{cases} \sum_{i=1}^{w_x} \sum_{j=1}^{w_y} [f_L(x+i, y+j) - f_R(x+i+d, y+j)]^2, & \text{for SSD} \\ \sum_{i=1}^{w_x} \sum_{j=1}^{w_y} |f_L(x+i, y+j) - f_R(x+i+d, y+j)|, & \text{for SAD} \\ \frac{\sum_{i=1}^{w_x} \sum_{j=1}^{w_y} f_L(x+i, y+j) f_R(x+i+d, y+j)}{\sqrt{\sum_{i=1}^{w_x} \sum_{j=1}^{w_y} f_L^2(x+i, y+j) \sum_{i=1}^{w_x} \sum_{j=1}^{w_y} f_R^2(x+i+d, y+j)}}, & \text{for Correlat} \end{cases} \quad (2)$$

Where $f_L(x, y)$ and $f_R(x, y)$ are the intensities of the pixel at a position (x, y) the left and right images, respectively, $W_C(x, y, d)$ is the window cost of a pixel at position (x, y) in the left image with disparity d , w_x and w_y are the window width and height, respectively.

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 indicates the corresponding pixel of the pixel in the left image. In this research, window cost calculation is performed based on SSD algorithm. In direct search, it requires to compute the window costs i.e., the SSD values

for all candidate pixels within the search range, $-d_{\max}$ to $+d_{\max}$.

For each pixel (x, y) in the left image the disparity estimation algorithm using direct search are described in Figure 2.

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

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

End For

Find best $W_C(x, y, d) \in W_C(x, y, d')$.

Disparity of $(x, y) = d$.

Figure 2. Disparity estimation algorithm for pixel (x, y) in the left image

Window-based stereo matching technique is widely used due to its efficiency and ease of implementation. However, Barnard and Fisher [11] point out a problem in the selection of a window with fixed size and shape. Many researchers proposed adaptive window methods using windows of different shapes and size depending on local variations of intensity and disparity [10]. But in adaptive window algorithms, the computation time is relatively higher than that of fixed window algorithms. To overcome this problem and to achieve a substantial gain in accuracy with less expense of computation time we have proposed a fast and very simple algorithm in this research.

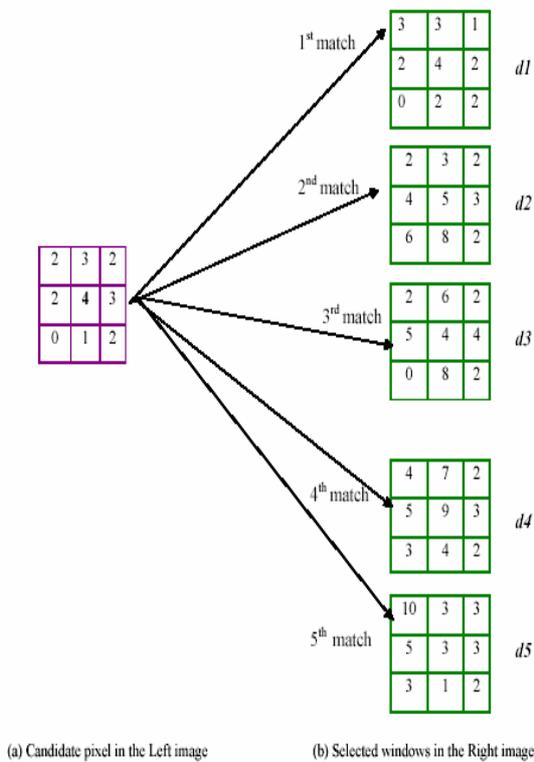
3. Average Disparity Estimation Method (Proposed Algorithm)

To obtain a smooth disparity map with less window cost, we have proposed and implemented a new method called 'average disparity estimation'. In many research works, to estimate the correspondence of a pixel in the left image, the window costs are computed for all candidate pixels in the right image within the search range. But in our proposed method, to determine the correspondence of a pixel in the left image, we just compute the window costs for pixels in the right image whose intensities are different within a certain threshold value to the intensity of the pixel in the left image. Having the window costs for a certain threshold value for a number of selected pixels in the right image within the search range, we then compute the average disparity.

In our proposed method, within a searching range $-d_{\max}$ to $+d_{\max}$ in the right image, several

window costs are calculated for a number of selected pixels to match with the reference pixel in the left image. Then the average window cost is considered for disparity estimation. **Figure 3** shows that within the range $-d_{max}$ to $+d_{max}$ ($d_{max} = 2$) in the right image, five windows are selected for estimation of window costs for a reference pixel (4) in the left image. After finding several disparities ($d1, d2, d3, d4, d5$), average disparity d_{avg} is calculated and placed in to construct the disparity map. This procedure is repeated to estimate disparities for all reference pixels in the left image.

Figure 3. 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 $+d_{max}$ ($d_{max} = 2$).



1. For each pixel (x,y) in the left image, search the corresponding pixel in the right image:
 - I. **for** $d' = -d_{max}$ to $+d_{max}$ **do**
if $|f_L(x,y) - f_R(x+d',y)| < threshold$
 - II. Calculate window costs $W_c(x,y,d')$.
 - III. Determine average disparity d_{avg}

end
2. Repeat 1 to calculate disparities of all pixels in the left image.

Figure 4. Average Disparity estimation algorithm with pruning for pixel (x,y) in the left image.

The average disparity estimation algorithm is shown in Figure 4.

4. Experimental Results and Discussions

In order to demonstrate the effectiveness of the algorithm, we present the processing results from synthetic and real image pairs, including ones with ground-truth values for quantitative comparison with other methods. The images are provided by the Computer Vision and Image Media Laboratory, University of Tsukuba, Japan. Figure 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.

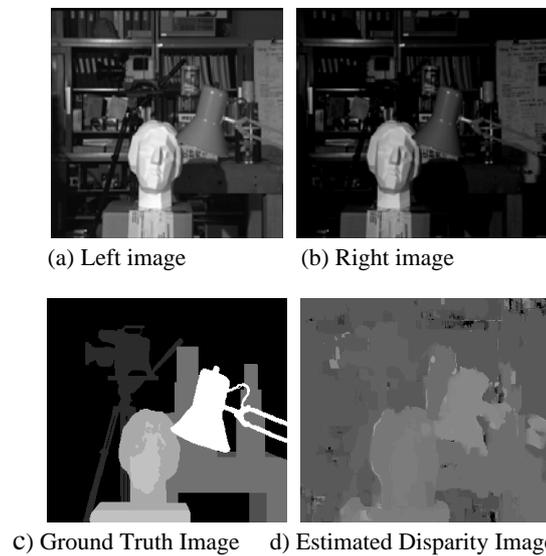


Figure 5. Gray scale stereo images (Tsukuba Head) with ground truth.

Experiments are carried out on a Pentium IV 2.1 GHz PC with 512 MB RAM. The algorithm has been implemented using Visual C++. Disparities of the left image are estimated by SAD method without and with considering a threshold value for different window sizes. Table 1 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). Table 2 shows the summary of comparison between traditional window based method and our proposed method (average disparity estimation method). Experimental results predict that our proposed method provides a reduction of 60% of computational time.

Table 1: Disparity estimation accuracy (in %) for different threshold values(run time snapshot).

"E:\Image Processing\Stereo Matching\Debug\opt1.exe"				
Threshold Value (pixel)	Window Size (pixel)	Accuracy (%)		Computation Time (Second)
		(Correct match)	(Diff. pixel)	
Without threshold	3 x 3	49.1 %	72.7 %	1
	5 x 5	59.6 %	82.9 %	5
	7 x 7	64.9 %	87.1 %	14
	9 x 9	68.2 %	89.3 %	28
	11 x 11	70.1 %	90.1 %	47
	13 x 13	71.4 %	90.4 %	73
· 5	3 x 3	47.4 %	71.1 %	1
	5 x 5	56.5 %	80.3 %	4
	7 x 7	60.2 %	83.3 %	10
	9 x 9	62.0 %	84.5 %	19
	11 x 11	62.8 %	84.4 %	32
	13 x 13	63.2 %	84.1 %	49
· 10	3 x 3	48.4 %	71.9 %	1
	5 x 5	58.1 %	81.4 %	5
	7 x 7	62.5 %	84.8 %	12
	9 x 9	64.7 %	86.3 %	22
	11 x 11	66.1 %	86.7 %	36
	13 x 13	66.8 %	86.6 %	55
· 15	3 x 3	48.7 %	72.2 %	1
	5 x 5	58.6 %	81.8 %	5
	7 x 7	63.4 %	85.5 %	12
	9 x 9	66.1 %	87.1 %	23
	11 x 11	67.5 %	87.5 %	38
	13 x 13	68.2 %	87.4 %	60
· 20	3 x 3	48.9 %	72.4 %	1
	5 x 5	58.9 %	82.1 %	5
	7 x 7	64.0 %	86.0 %	12
	9 x 9	66.9 %	87.7 %	23
	11 x 11	68.3 %	88.1 %	40
	13 x 13	69.1 %	88.0 %	66
· 25	3 x 3	48.9 %	72.4 %	2
	5 x 5	59.1 %	82.3 %	7
	7 x 7	64.1 %	86.2 %	15
	9 x 9	67.0 %	88.0 %	27
	11 x 11	68.5 %	88.5 %	44
	13 x 13	69.3 %	88.3 %	68
15 x 15	70.0 %	88.2 %	105	

Table 2: Computational time reduction (%) with window based method.

Disparity Estimation Method	Window size	Computation time (second)	Computation time reduction (%)
Window Based (Direct Search)	7 x 7	0.95	0
	5x5	0.97	0
	3x3	1.00	0
Average Disparity Method (Proposed)	7 x 7	0.38	60.0
	5x5	0.39	59.7
	3x3	0.41	59.0

5. Conclusions

A fast and very simple stereo matching algorithm has been developed that analyses grayscale or color images to estimate the disparity map for 3D scene reconstruction. 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 easy technique is proposed in this research, which is applicable in fast estimation of stereo disparity. The proposed system has the feature of average filtering and can tackle additive noise in the images.

A total of about 20 images are used to investigate the capacity of the proposed algorithm. Among them only a few images are found false. Most of the images are taken using a digital camera, but some are from scanner, and some from video tapes recorded from different sources. Experimental results demonstrate that the success rate of more than 95% is achieved. Experimental results confirm that the proposed method is capable of reducing computation time of about 60% with achieving a substantial gain in accuracy.

References

- [1] D. Stefano and S. Mattocchia. "Fast stereo matching for the videt system using a general purpose processor with multimedia extensions" In *Fifth IEEE International workshop on Computer Architecture for Machine Perception*, Padova, Italy, Sept. 2000.
- [2] D. Stefano, M. Marchionni, S. Mattocchia, G. Neri, "A Fast Area-Based Stereo Matching Algorithm" 15th IAPR/CIPRS International Conference on Vision Interface Calgary, Canada, May 2002.
- [3] M. Mozammel Hoque Chowdhury, Md. Abdul Mannan Mondal, and Md. Al- Amin Bhuiyan, "3D Imaging Using Stereo Vision", Proc. of the 7th International Conference on Computer and Information Technology (ICCIT'2004), Dhaka, Bangladesh, pp. 307-312.
- [4] M. S. Uddin, Mozammel Hoque Chowdhury, T. Shioyama and Abdul Mannan Mondal, "Fast window-based approach for stereo matching", Journal of Science, Jahangirnagar University, Vol. 27, NO. 1, 2004.
- [5] D. Geiger, B. Ladendorf, A. Yuille, "Occlusions and binocular stereo," International Journal of Computer Vision, vol. 14, pp. 211-226,1995.
- [6] M. J. Hannah, "A system for digital stereo image matching," Photogrammetric Engineering and Remote Sensing, vol. 55, no.12, pp. 1765-1770, 1989.
- [7] M. D. Levine, D. A. O'Handley, G. M. Yagi, "Computer determination of depth maps," Computer Graphics and Image Processing, vol. 2, no. 4, pp. 131-150,1973.

- [8] D. J. Panton, "A flexible approach to digital stereo mapping," *Photogrammetric Engineering and Remote Sensing*, vol. 44, no.12, pp. 1499-1512, 1978.
- [9] G. A. Wood, "Realities of automatic correlation problem," *Photogrammetric Engineering and Remote Sensing*, vol. 49, no. 4, pp. 537-538, 1983.
- [10] T. Kanade and M. Okutomi, "A stereo matching algorithm with an adaptive window: Theory and experiment," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 16, no. 9, 1994.
- [11] S. T. Barnard and M. A. Fischler, "Stereo vision," *Encyclopedia of Artificial Intelligence*, New York: John Wiley, 1987), pp. 1083-1090.
- [12] Olga Veksler, "Stereo matching by compact windows via minimum ratio cycle," *Proceedings of the IEEE International Conference on Computer Vision (ICCV 2001)*, pp. 540-547, 2001.
- [13] S. S. Intille and A. F. Bobick, "Disparity-space images and large occlusion stereo," *Proceedings of the European Conference on computer Vision (ECCV 1994)*, pp. 179-186, 1994.
- [14] A. Fusiello and V. Roberto, "Efficient stereo with multiple windowing," *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (ICVPR 1997)*, pp. 858-863,1997.

M. Mozammel Hoque Chowdhury is working as an Assistant Professor in the Department of Computer Science and Engineering, Jahangirnagar University, Savar, Dhaka, Bangladesh. His research interest includes: Image Processing, Computer Vision, Robotics, 3D Scene Reconstruction, Computer Graphics and Machine Intelligence.

Dr. Md. Al Amin Bhuiyan is a Professor in the Department of Computer Science and Engineering, Jahangirnagar University, Savar, Dhaka, Bangladesh. He got the Dr. Eng. degree in Electrical Engineering from Osaka City University, Japan, He did the post doctoral research from National Institute of Informatics, Japan and University of Hull, UK. His research interest includes: Robotics, Image Processing, Computer Vision, Computer Graphics Pattern Recognition, 3D Scene Reconstruction, Face Recognition, Ukiyoe Photo Identification, Human-robot Symbiosis, Knowledge based Image Understanding, and so on.