International Journal of Engineering and Advanced Technology (IJEAT)
ISSN: 2249 – 8958 (Online), Volume-9, Issue-5, June 2020

Stereo Correspondence Estimation by Two Dimensional Real Time Spiral Search Algorithm

Md. Abdul Mannan Mondal, Mohammad Haider Ali

Abstract: ***This paper presents a new searching algorithm titled "Two Dimensional Real Time Spiral Search Algorithm (2DRTSSA)" to compute the stereo correspondence or dense disparity map of two rectified images. The proposed algorithm can estimate the minimum stereo correspondence or disparity among all the window costs of a fixed axis from minimum to maximum range of that axis. It can also simultaneously calculate the dense disparity of another axis with the same range of axis. So the proposed method calculates stereo correspondence two dimensionally at a time and thus it increases the speed and accuracy over the existing state-of-the-arts methods of one dimensional and left-right searching strategy. The 2DRTSSA method calculates firstly the two window costs; one is along with the +x direction and another is along with –y direction .The minimum disparity of estimated two window costs and their distance parameters are remaining contribute in final selection. The rest of two window costs of –x direction and +y direction are also calculated using the same procedure. The minimum disparity of newly estimated two window costs and their distance are remaining contribute in final selection. The process is then repeated for the successive pixels of reference image along with the 2D scan lines from left to right of the whole image. The 2DRTSSA method is able to optimize the speed and accuracy of estimated dense disparity. Experimental results are compared in Section-IV (A), Section-IV (B) and Section-IV(C) with the current state-of-the-arts methods those are tested on Middlebury Standard stereo data set. The proposed 2DRTSSA method establishes the highest speed and accuracy with properly reconstructed 3D of dense disparity image.***

Index Terms: Stereo correspondence, window cost, spiral search, disparity, sum of square differences, normalized correlatio technique.

I. INTRODUCTION

In a binocular vision, stereo correspondences or disparity is the most important factor to track the viewing objects exactly .In most cases it is defined by the parameter *d* and refers to the differences of *x*-axis between left and right images captured by the stereo visional respective eyes. Now a days it challenges to compute the disparity in real time based applications. This parameter can be calculated by the following-

$$d = X_L - X_R \tag{1}$$

Where, X_L is the reference pixel position in left image.
X_R is the corresponding pixel position in right image.

The parameter disparity *d* of stereo images is most essential to compute the depth of information of an object. If B_{dt} is baseline distance of two horizontally placed cameras, *F* is the focal length of camera, the depth *z* of information of an object can be computed by the following-

$$z = \frac{FB_{dt}}{d} \tag{2}$$

The main objectives of stereo matching are to outline the disparity map based on all stereo correspondences stereo images. The dense disparity map is essential for robotic vision, pedestrian navigation, 3D tracking and reconstruction. Disparity is normally estimated by using Sum of Square Differences (SSD), Sum of Absolute Differences (SAD), or Normalized Correlation Techniques (NCT). Window cost aggression technique is mostly used because of its efficiency and very simplicity of implementation. In case of window system there are major problems for selecting the accurate size and shape of operating windows [1- 2]. In order to manage these problems some authors implemented adaptive windows of variable shapes and sizes[3-7] that causes more computational time for different mask or kernel sizes. To explore the best window cost the authors of [6] and [7] have done the direct search on different window sizes and silhouettes. In [8], color stereo image processing has shown better accuracy but it requires huge computational time. It needs high speed estimation of dense disparity [9-10] for the application of autonomous driving, 3D object recognition, pedestrian detection etc. In Virtual Masking System [9] the computational performance has been upgraded by excluding its false matching correspondences. +45° and -45° searching ideas are employed in [11] that extracts the best correspondences to overwhelmed the window-based problems. The similar very basic idea is presented in [12] but after huge modifications the proposed method differs mainly by the simultaneous searching that able to reconstruction 3D scenes, improves speed and accuracy.

Revised Manuscript Received on April 20, 2020.
* Correspondence Author
Md. Abdul Mannan Mondal*, Department of Computer Science and Engineering, University of Dhaka, Dhaka, Bangladesh. Email: mannan_mondal@yahoo.com
Mohammad Haider Ali, Department of Computer Science and Engineering, University of Dhaka, Dhaka, Bangladesh. Email: haider@du.ac.bd

The better classifications have opened by Scharstein and Szeliski [13] and many novel approaches have been presented here. Primarily matching algorithms are classified with respect to spare output and dense output. Feature based methods that based on segments or edges between stereo images result sparse output.

Such type of output has the limitations both speed and accuracy due to their disadvantages causes it dreadful for many applications. Dense disparity estimation algorithms are divided into local and global ones. Local methods are also known as area based stereo matching that can perform better speed compare to global methods. According to this, disparity is being calculated at a point in a fixed window. Global methods are also known as intensity or energy based stereo matching that can perform better accuracy compare to local methods. According to this method, the global cost function is reduced as minimum as possible. This cost function synthesizes image data and smoothness terms. Besides these some algorithms are not fallen into above mentioned two categories. Recently, neural adaptive stereo matching [15] are done by trained neural networks based on window size and shape. One dimensional cellular automation filter [18] makes the algorithm more adaptive to each window. Yoon et al. [17] used SAD method and a left-right consistency check that performed like real time system. Yoon's method can process 7 *fps* for image resolution of 320×240 with levels of disparity 32. The experiment has been implemented by an Intel Pentium 4 processor at 2.66GHz. The uses of Cellular Automata (CA) are presented in [18].Real-time extraction of disparity maps was demonstrated by this method. It can process the input image with a speed of 40 *fps*. This method also based on SAD matching cost. A window-based method uses different support-weights in [19]. To decrease the image obscurity [20], the support-weights of the pixels are accustomed based on geometric proximity and color similarities. The research work presented in [20] based on unified framework that supports the fusion of any partial knowledge such as matching features and surfaces about disparities. In order to use of standard dynamic programming it combined edge, corner and dense stereo corresponding algorithm to act as a controller points. According to Bayesian estimation theory [22], the continuity, coherence, occlusion constraints and the adjacency principal are taken into considerations to assign disparities. Dense disparity map is estimated by an adapting window based method [23]. The recent works of related problem to the matching costs are stated in [24] and [25]. The authors used bilateral filter to determine the cost aggregation and in order to reduce the computational cost they also limit the label space. The work in [26] can be considered as a cost aggregation method by guided image filter. The average runtime [26] of the four standard Middlebury data sets is 960 milliseconds reported in [29]. So the run time of single image pair like Tsukuba or Venus is about (960÷4) 240 milliseconds. Disparity space image (DSI) structure and gradient information has been combined as a new technique is first time introduced by Nadia Baha and Slimane Larabi [27]. They used DSI technique with adaptive window – support. Another approach is introduced by themselves as DSI and refinement. The experimental results take time 0.2 *second* and 0.39*second* respectively for processing Tsukuba head image pair. A new geodesic *o* (1) filter is employed in [28] for the reliable disparity propagation. Such type of filter is very operative for the cost matching. As it is state-of –the-art method and the speed of this method has been justified on the Middlebury standard data set, so we can compare this paper to our proposed 2DRTSSA method.

Xun Sun et al. [28] perform the experiment on PC furnished with a 3.0 GHz Intel i5 CPU, 8 GB of memory and a Geforce GTX 580 graphics card. The processing time on Middlebury standard data set is only 9 milliseconds.

A cost aggression has been adaptively estimated on a tree structure derived from the stereo image pair to preserve depth edges [29]. This latest idea is launched by Q.Yang [29] in which shortest distances measure the similarity between two pixels on the tree. The average run time of the four standard Middlebury data sets are 90 milliseconds using the tree filtering method. But he et al. [29] mentioned in same section that the runtime is *7* milliseconds on average on the Middlebury data sets. For identically comparison to our proposed method we consider his second result, it takes *7* milliseconds on average on the Middlebury data sets. Q. Yang tested his experiment on a MacBook Air laptop computer with a 1.8 GHz Intel Core i7 CPU and 4 GB memory.

Another recent method achieves the state-of-the–arts result on Middlebury stereo data sets that performs stereo matching as a two steps energy-minimization algorithm [30]. The running time of this method is 3 seconds only for Tsukuba data set and 20 second for Teddy data set on a PC having an Intel Core i5-4300U 1.9-GHz CPU and a 6-GB RAM. Semi-global matching and cost is refined by cross-based aggression [31] has been introduced by J. Zbontar and Y. LeCun . Y. LeCun el at. [31] also uses left-right consistency check to remove the inaccuracies. The experiment performs on KITTI stereo data set. With the above review we found that the researchers employed window-based techniques, tree structure, energy minimization and geodesic *o* (1) filter to calculate the matching costs one dimensionally. But in our proposed method we calculate *2D matching costs simultaneously*. This is the main difference between the proposed state-of-the –art method and other latest methods. The mentioned recent methods are very similar to our proposed method but differing mostly in optimal searching technique. Beside these analysis, the work in [28] requires preprocess and the works in [27], [29], [31] needed post processing steps like refinement, filtering and histogram equalization. Finally the only work in [30] is tested without post processing. The proposed 2DRTSSA method also runs without preprocessing and post processing. The experimental disparity maps are directly eligible to compare with ground truth dense disparity. In our research we have done that the experimental proposed cost contributes to achieve the state-of-the –art results. So considering the direction of search similarity, identical stereo data set (Middlebury Standard data set) and hardware platform we can consider the papers of [14], [27], [28], [29] and [30] to compare the state-of-the-arts of proposed 2DRTSSA method.

Retrieval Number: D8592049420/2020©BEIESP
DOI: 10.35940/ijeat.D8592.069520
Journal Website: www.ijeat.org

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication

The main contribution of this paper is fully two dimensional search based window cost speedy method with following advantages-

- Experimental results demonstrate that the 3D reconstruction of output disparity map is very similar to ground truth dense disparity.
- We observe **1052** ***fps*** for input images with 384×288 (Tsukuba head pair) pixel resolution.
- The proposed method's experimental results have been compared with some state-of-the- arts methods and it claims that the proposed 2DRTSSA method is currently the state-of-the-arts both in computational time (only 0.95*ms*) and speed for Tsukuba stereo pair with upgrading the accuracy of **93.8%** and only **6.2%** bad pixels in percentage with threshold 1.

II. PROPOSED 2DRTSSA SEARCH METHOD

The innovative 2DRTSSA search method can be explained as a co-ordinate geometric concept .The search ranges are outlined in Figure 1 that shows the search coordinates range ($-C_{xmin}$, $-C_{ymin}$) to ($+C_{xmax}$, $+C_{ymax}$) instead of using $-d_{max}$ to $+d_{max}$ in one dimensional existing system. Accordingly, the proposed method is operated in two dimensionally at a time. In first phase, first search is done concurrently in the 1st and 3rd quadrants of right image as indicated in Fig. 1(a). In second phase, second search is performed in the second and fourth quadrants of right image .In both cases the searching commences from the starting point (say $-C_{xmin}$,0) to the ending point ($+C_{xmax}$,0) as shown in Fig. 1(b). Every iteration program sequence tends to reach to the origin point. Each reference pixel of reference image (left image) is hunted in the two axial coordinate's points according to the stated search method. According to the proposed 2DRTSSA method, each pixel of reference image is firstly compared with negative *x*-direction of right image as well as positive *y*-direction of right image tailed by one pixel gap. Suppose in two cases, two separated window costs are determined as d_1 and d_2, respectively. Secondly, the same pixel is compared with positive *x*-direction of right image as well as negative *y*-direction of right image tailed by one pixel gap. So, another two distinct costs are determined as d_3 and d_4, respectively. All of the experimentally estimated disparities $\{d_1, d_2, d_3, d_4 \ldots d_{+Cxmax}\}$ are passed to the minimum function of array d_i. As a final point, the final disparity d is selected from the set of elements $W_C(x,y,d_i)$,*i.e.* $W_c(x,y,d) \in W_c(x,y,d_i)$.Therefore the stereo correspondence or disparity of a reference pixel of left image is $P(x,y)=d$.

The process is then repeated for the successive pixels of reference image along with the 2D scan lines from left to right of the whole image. With the above mentioned strategies the proposed method avoids the repetition of redundant comparisons and false matching's causes to increase the computational speed and accuracy.

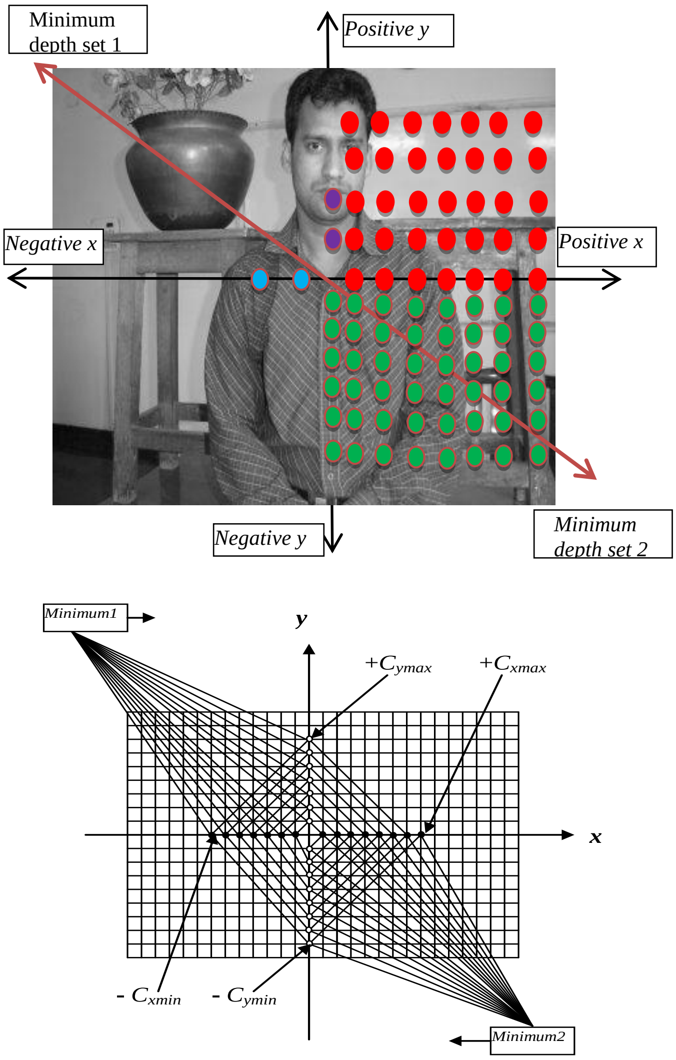


Fig. 1. Illustration of 2DRTSSA search method with co-ordinate prefecture.

III. ALGORITHM AND FLOWCHART OF 2DRTSSA

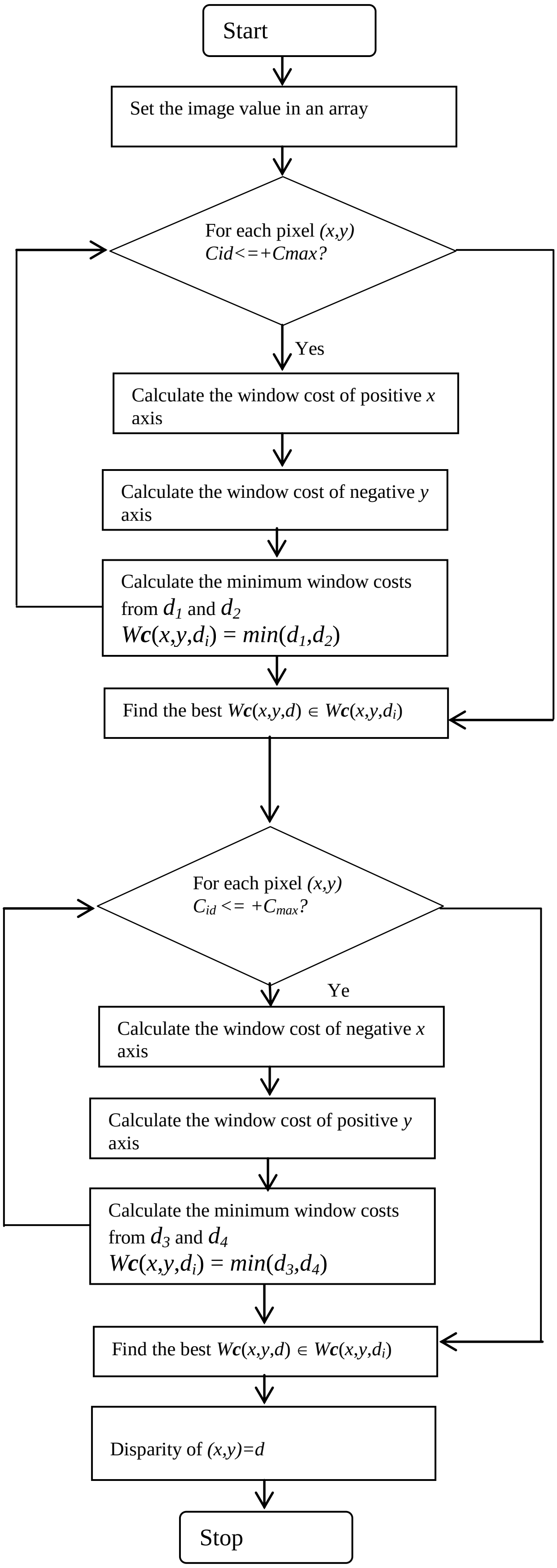
1. Initialization: $P(x, y) =0$.
2. Repeat step 3 to 5 for each pixel $P(x,y)$
3. Repeat step for $C_{id} = -C_{min}$ to $+C_{max}$ *do*
   Compute window cost $W_c((x+C_{id}*2),y,d_1)$
   Compute window cost $W_c(x, (y+c_{id}*(-2)+1), d_2)$
   $Wc(x,y,d_i) = minimum(d_1,d_2)$
4. [Exit step 3]
5. Find the *minimum cost* $W_c(x,y,d) \in W_c(x,y,d_i)$
6. [close step 2]
7. *Final Disparity of* $P(x,y) = d$
8. Exit

The key idea of this algorithm states that the search is divided into two regions which are well defined in step 3. One cost aggression is estimated along the *x* axis on photometric point $((x+C_{id}*2), y)$ while the other cost aggression is estimated along the *y* axis on photometric point $(x, (y+c_{id}*(-2)+1)$.Two axis are selected simultaneously by the expression in the first bracket in step 3.

Retrieval Number: D8592049420/2020©BEIESP
DOI: 10.35940/ijeat.D8592.069520
Journal Website: www.ijeat.org

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication

So the proposed method searches a reference pixel on two probable space at a time within a finite range C_{id} .On the contrary, the existing state-of-the-arts algorithms search a reference pixel only one space at a time. So this idea makes the proposed method faster than existing methods.



IV. COMPUTATIONAL COMPLEXITY ANALYSIS

The computational complexity of 2DRTSSA algorithm is $O(n\times w/2)$, where n is total number of candidates pixels and w window size of mask. Two matching costs are estimated two different co-ordinates (x and y) at the same time. The main idea of this method is that 50% window costs can be estimated by instruction pipelining. So two pipelining calculations (in *C++* code) executed in one instruction per cycle. That is the main reason to reduce the computational time of proposed method. The required memory depends only the size of n i.e. it directly proportional to image size. It apparently seems to require more space for two window costs at a time. But actually the proposed algorithm compares instantly two window costs, selects the minimum one window cost, and discard the other. The total run time for the Tsukuba head image pair is 0.95 milliseconds on the hardware of Intel Core i3, 2.3 GHz processor with 4 GB DDR3 RAM.

V. EXPERIMENTAL RESULTS

The experiments are performed on Middlebury standard stereo images of Tsukuba stereo pair. The computational time, speed and accuracy of the proposed algorithm have been justified over the said stereo images of Tsukuba head. The experimental dense disparity maps are estimated from left and right image applying 2DRTSSA is shown in Fig. “4-9”. Fig.10. shows standard dense disparity of ground truth image.

Experiments are performed on Intel Core i3, 2.3 GHz processor with 4 GB DDR3 RAM. The algorithm is performed by Visual C++ programming language. To determine the correspondence of a pixel of reference image, the window costs are estimated for the candidates’ pixels of right image within the search range -10 to +10. The experimental results state that the proposed algorithm is currently the best cost aggression method among the existing state-of-the-art methods. The top performer algorithms are reported in [27], [28], [29] and [30]. All are ranked by Middlebury benchmark [33]. So we have to prove the claim by comparing the time and speed with the top performer algorithms which demonstrated in Table I. The disparity maps of the Middlebury data set for Tsukuba head are estimated by proposed 2DRTSSA method are illustrated in Fig. 2(b). Table-I shows that the proposed 2DRTSSA algorithm outperforms the current and earlier top performer algorithms. Moreover, the proposed method is faster than all others top performer algorithms. The accuracy of the proposed algorithm for Tsukuba head is **93.8%** i.e. the bad pixel in percentage with the error threshold is only **6.2%** which is almost the same of the top algorithms. Little variation of accuracy occurs due to orientations of pixel redundancy. The experimental results are analyzed in three phases are stated below-

A. *Experiment 1: Observation of 3D Reconstruction and Objects Recognition of Experimental Output.*

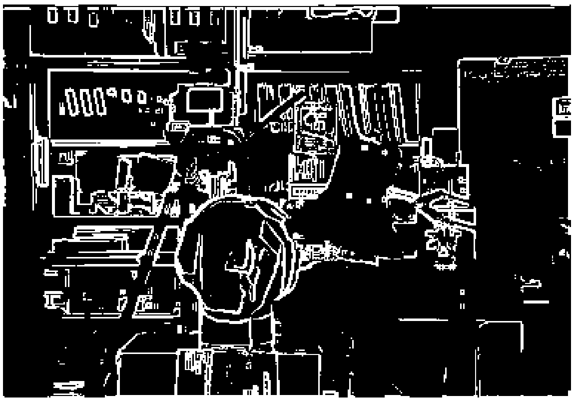
The Tsukuba stereo pair of input images contains different objects at different depth of positions.

Retrieval Number: D8592049420/2020©BEIESP
DOI: 10.35940/ijeat.D8592.069520
Journal Website: www.ijeat.org

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication

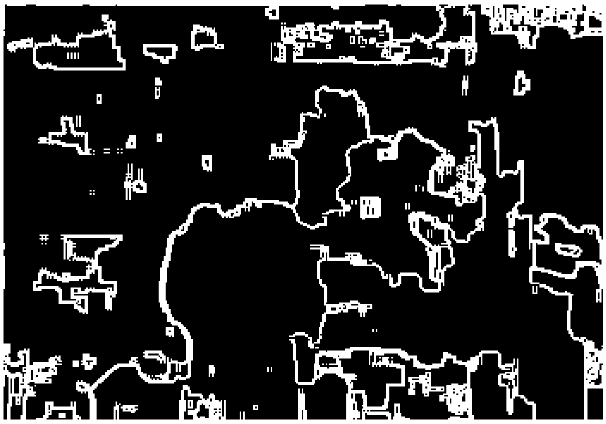
Background and foreground objects are situated at different depth. Almost overlapping objects are found in background of Tsukuba stereo pair those are occlusions and poor objects. However these stereo pair also contains some special regions like head of the statue, table lamp and video camera. These types of regions are really quite difficult to separate from other objects by stereo matching process. So the first challenge is to distinguish the different depth by marking the different gray level value of output image. Nearest object is shown by more white color and farthest object is shown by dark grey level value or black. It is worth observing that the 3D structure of output image has been reconstructed clearly in Fig. 2(b) where the face of the statue, table lamp, video camera as well as interesting objects are recognized (seen) easily. Moreover the objects' depths are observing on nearest objects are seen by more white color and deeper objects are seen by dark grey level value or black as shown in Fig. 2(b). Consequently, the camera and its trestle nearest objects such as face (head) of Tsukuba, table lamp are visualized by all most white color. On the contrary, the camera and its trestle's farthest objects such as video camera, book shelf, background wall of Tsukuba stereo pair are reconstructed with all most black color.

Object borders are clearly recognized in estimated dense disparity image, i.e. border localization problem of article [14] are solved in the proposed method. The output images are further fed into the object's detection algorithm and the output image's objects border are identified which are illustrated in Fig. 2(a) and Fig. 2(c).



(a)

(b)



(c)

Fig. 2. Localized objects' borders.

The details of estimated dense disparity image are compared to ground truth image of Tsukuba head. The experimental results obtained by the 2DRTSSA method on Tsukuba head are very similar to their ground truth image. The estimated dense disparity's 3D structure is recovered and its objects border are correctly identified which are outlined in Fig. 2(a) and Fig. 2(c). So the result ensures that the similar depths are found in estimated dense disparity which is outlined in Fig. 2(b).

B. *Experiment2: Computational Cost Calculation and Comparison with Other State-of-the-Arts Methods:*

Disparities of reference image are estimated by Sum of Square Difference (SSD) technique using 2DRTSSA search algorithm without any pruning for different window sizes. The disparities are estimated with the search range from -10 to +10.The effects of said search are investigated with respect to computational costs and speed (in *fps*). The computational costs and speed (in *fps*) performances of proposed 2DRTSSA method has been compared with other state-of-the-art-methods [27-30] and proposed method's performances have been tested in the machine of Intel Core i3, 2.3 GHz processor laptop with 4 GB DDR3 RAM.

The 2DRTSSA's experimental results have been compared with the result of methods those are tested on Middlebury standard data set. The ranking results in Table-I indicate that the proposed 2DRTSSA method is ranked 1st out of existing four states-of-the- arts methods [27-30]. It shows the highest speed **1052 *fps*** and lowest computational time **0.95 *ms*** among the four latest methods with lower configuration of machine. So it is claimed that the proposed method is currently the state-of-the-arts method for Tsukuba head image pair with 3X, 7X, 9X, and 3157X faster than the methods of [14], [29], [28] and [30] respectively.

Table-I: Numerical Comparison of proposed 2DRTSSA and existing state-of-the-art methods.

Methods	Machine	Computational time (Millisecond)	Speed (fps)	Rank
2DRTSSA [proposed]	2.3 GHz , Intel Core i3	0.95	1052	1
Fast Area Based method[14]	2.3 GHz , Intel Core i3	3.05	328	2
Tree filtering [29]	1.8 Ghz, Intel Core-i7	7	143	3
Edge-aware & Geodesic filter[28]	3.0Ghz,Intel Core-i5+Geforce GTX card	9	111	4
DSI & Adaptive Support[27]	2.2Ghz,Core Duo	200 (0.2 sec.)	5	5
Energy Minimization[30]	1.9 Ghz, Intel Core-i5	3000 (3 sec.)	0.33	6

C. *Experiment 3: Accuracy of Proposed 2DRTSSA Method*

The accuracy of this algorithm has been justified over standard stereo images of Tsukuba head. Table-II illustrates the accuracy of proposed 2DRTSSA method applied on standard stereo images of Tsukuba head and also represents the accuracy and bad pixels in percentage with error threshold 1.

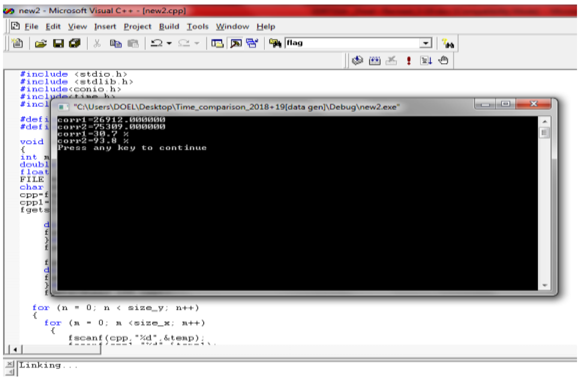


Fig. 3. Run time snapshot of 2DRTSSA method for accuracy.

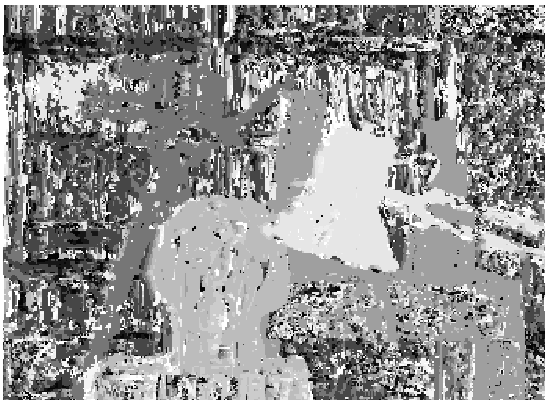


Fig. 4. Dense disparity map for window size 3×3

Fig. 5. Dense disparity map for window size 5×5.

Fig. 6. Dense disparity map for window size 7×7.

Fig. 7. Dense disparity map for window size 9×9.

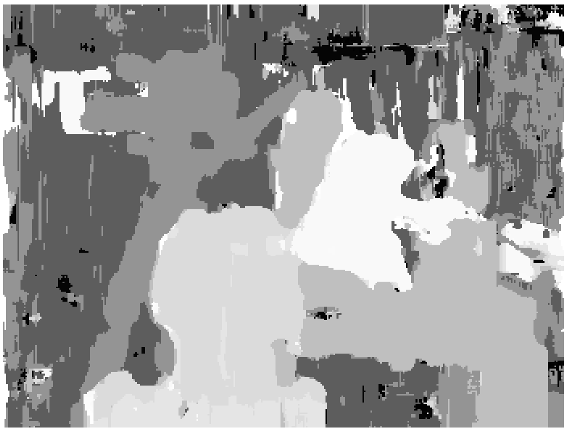


Fig. 8. Dense disparity map for window size 11×11.

Retrieval Number: D8592049420/2020©BEIESP
DOI: 10.35940/ijeat.D8592.069520
Journal Website: www.ijeat.org

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication

Fig. 9. Dense disparity map for window size 15×15.

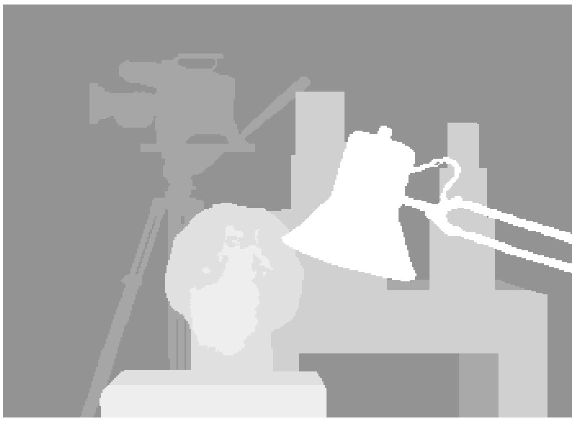


Fig. 10. Dense disparity map of ground truth image.

In order to estimate the accuracy of the said method we have tested the experimented using different mask sizes which results are outlined from Fig. "4-9". From Table-II the numerical evaluations confirm that the bad pixel in percentage is only **6.2%** with error threshold 1. But using the same condition bad pixels in percentage were 6.33%, 7.88%, and 7.18 % reported in [28], [29] and [30] methods respectively for Tsukuba head with the experiments of Middlebury stereo data sets.

Table- II: Accuracy of 2DRTSSA for Tsukuba stereo pair.

Window size (pixel)	Accuracy (%)	Bad Pixels (in %) with the Error Threshold 1
3×3	73.3	26.7
5×5	79.8	20.2
7×7	92.1	7.9
9×9	93.2	6.8
11×11	**93.8**	**6.2**
15×15	83.6	16.4
17×17	83.3	16.7
19×19	72.9	27.1
21×21	**72.4**	**27.6**

Table-II suggests that accuracy is gradually increased if widow size increases .But this is valid from window size 3×3 to window size 11×11 only. The lowest accuracy started from **73.3%** for 3×3 window size and the highest accuracy occurs at **93.8%** for widow size 11×11 with bad pixel in percentage only **6.2%.** After that if we increase widow size (like 15×15, 17×17 etc.) the accuracy is decreased and it reaches at **72.4%** for widow size 21×21. The graphical interpretation is also illustrates in Fig. 11. that also prompted the ***best operating window*** size which is 11×11 for best accuracy of robotic vision.

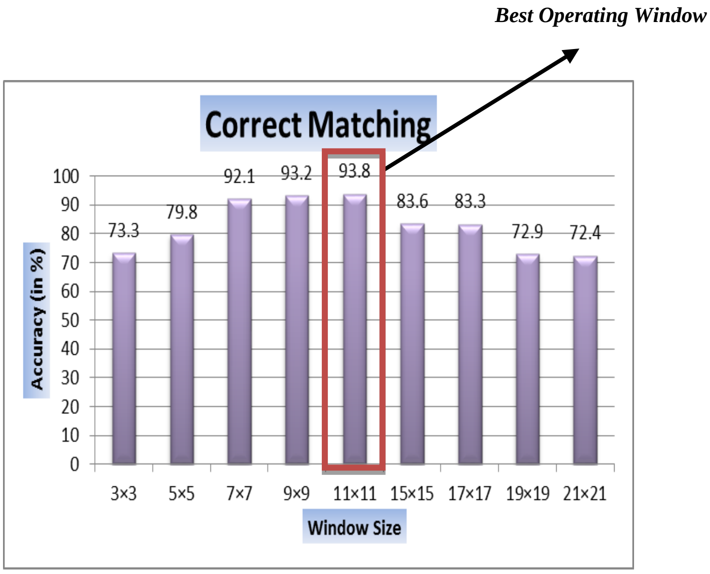


Fig. 11. Illustrates the correct matching of estimated dense disparity with ground truth image of Tsukuba head.

VI. CONCLUSION

The main contribution of the proposed method is to speed up the computational time. Its aim is also to improve the strength of the window based cost aggression method in order to use in real time application. The speed of our algorithm is **1052** ***fps*** for input images of Tsukuba head image pair. So it can scan, calculate, process and display output **1052** frame/second for the case of standard Tsukuba head image pair. We implement it by 2D parallel costs estimation causes to reduce the computational costs.

Moreover, the 2DRTSSA algorithm does not require additional programmable 3D hardware like 3D Graphics Processing Unit (GPU). The proposed 2DRTSSA method demonstrates the state-of-the-arts results and outdoes the existing top methods. In future we will consider fuzzy based technique to extend the research works for better realization of its behaviors.

REFERENCES

1. S. T. Barnard and M. A. Fischler, “Stereo vision” in *Encyclopedia of Artificial Intelligence*, John Wiley, New York, 1987, pp. 1083-1090.
2. W. Hoff and N. Ahuja, “Surfaces from stereo: Integrating feature matching, stereo correspondence estimation and contour detection,” *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 11, no. 2, pp.121-136, 1989.
3. 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.
4. Olga Veksler, “Stereo matching by compact windows via minimum ratio cycle,” in *Proceedings of the IEEE International Conference on Computer Vision*, 2001,pp. 540-547 .
5. Md.Abdul Mannan Mondal, Md.Al-Amin Bhuiyan “Disparity Estimation By A Two-Stage Approximation Real Time Algorithm” *The International Management and Technology Conference (IMT)*, December 2004,pp: 12-17.
6. S. S. Intille and A. F. Bobick, “Stereo correspondence-space images and large occlusion stereo,” in *Proceedings of the European Conference on Computer Vision (ECCV)*, 1994,pp. 179-186.
7. A. Fusiello and V. Roberto, “Efficient stereo with multiple windowing,” in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (ICVPR)*, 1997, pp. 858-863.
8. K. Muhlmann, D. Maier, J. Hesser, R. Manner, “Calculating dense stereo correspondence maps from color stereo images, an efficient implementation,” in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (ICVPR)*, 2001. pp. 30-36.
9. Md.Abdul Mannan Mondal, Md. Mustafa Kamal, Md. Al- Amin Bhuiyan,”Stereo correspondence Estimation By Real Time Approximation Algorithm Using Virtual Masking System.”, in *Proceedings of 8th International Conference on Pattern Recognition and Information Processing (PRIP’05)*, May 2005,pp.20-25.
10. Md. Mozammel Hoque Chowdhury, Md.Abdul Mannan Mondal, Md.Al-Amin Bhuiyan, “3D Imaging with stereo vision”,in *Proceeding of 7th International Conference on Computer & Information Technology*, December 2004, pp: 307-312.
11. Md.Abdul Mannan Mondal, Md. Mustafa Kamal, Md.Ariful Hyder, Md.Al-Amin Bhuiyan, “Stereo Correspondence Estimation by Diagonal Mask Searching ”, in *proceeding of 8th International Conference on Computer & Information Technology*, December 2005. pp: 379-383.
12. Md.Abdul Mannan Mondal, Md. Haider Ali “On Stereo Correspondence Estimation: A Spiral Search Algorithm”, in *proceeding International Conference on Signal and Information Processing (ICSIP 2010)*, IEEE Catalog Number: CFP1095L-ART, 2010. pp 204-207.
13. D.Schartein ,R.Szeliski, “ A taxonomy and evaluation of dense two frame stereo correspondence algorithms”, *Int. J. Comput. Vis.*, vol.47,no.1-3,pp.7-42,2002.
14. Lugi Di Stefano,Massimiliano Marchionni, Stefano Mattoccia, “A fast Area Based Stereo Matching Algorithm” , *Image and vision Computing 22*,pp.983-1005,2004.
15. Elisabetta Binaghi, Ignazio Gallo, Giuseppe Marino, Mario Raspanti, “Neural adaptive stereo matching”, *Pattern Recognition Letters* vol. 25,pp. 1743-1758,2004.
16. Abijit S.Ogale and Yiannis Aloimonos , “Shape and the Stereo Correspondence Problem”,*Int. Journal of Computer Vission.* Vol.65,no.3,pp.147-167,2005.
17. Sukjune Yoon,Sung-Kee Park, Sungehul Kang, Yoon Keun Kwak, “Fast correlation –based stereo matching with the reduction of systematic errors”, *Pattern Recognition Letters* vol.26, pp. 2221 2231,2005.
18. L. Kotoulas, A. Gasteratos,G.Ch.Sirakoulis , C. Georoulas, and I. Andreadis, “Enhancement of Fast Acquired Disparity Maps using a 1-D Cellular Automation Filter”, in *proceedings of the fifth IASTED International Conference on Visualization, Imaging and Image Processing*, September,2005.
19. Kuk-Jin Yoon and I.S. Kweon, “Adaptive Support –Weight Approach for Correspondence Search”, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol.28,no. 4, pp.650-656,April 2006.
20. P.H.S. Torra, A. Criminisi, “Dense stereo using pivoted dynamic programming”, *Image and Vision Computing* vol.22, pp. 795-806, 2004.
21. Lazaros Nalpantidis,Georgios Ch. Sirakoulis and Antonios Gasteratos, “Rieview of stereo matching algorithms for 3D vision”, in *16th International Symposium on Measurement and Control in Robotics ISMCR,* 2007, pp.116-124.
22. Salvador Gutierrez,Jose Luis Marroquin, “Robust approach for disparity estimation in stereo vision”, *Image and Vision Computing* vol.22, pp. 183-195, 2004.
23. Michael Bleyer, Margrit Gelautz, “ A layered stereo matching algorithm using image segmentation and global visibility constraints”, *ISPRS Journal of Photogrammetry and Remote Sensing* vol. 59,pp.128-150,2005.
24. D. Min, J. Lu, and M. N. Do, “A revisit to cost aggregation in stereo
25. matching: How far can we reduce its computational redundancy?” in
26. *Proc. IEEE ICCV*, Nov. 2011, pp. 1567-1574.
27. D.Min, J. Lu, and M. N. Do, “Joint histogram-based cost aggregation for stereo matching,” *IEEE Trans. Pattern Anal. Mach. Intell.*, vol.35, no. 10, pp. 2539–2545, Oct. 2013.
28. A. Hosni, C.Rhemann, M. Bleyer, C. Rother, and M. Gelautz.“ Fast cost volume filtering for visual correspondence and beyond”, *PAMI*,vol. 35,pp.504-511, 2013.
29. Nadia Baha & Slimane Larabi “Accurate real-time disparity map computation based on variable support window ” *International Journal of Artificial Intelligence & Applications (IJAIA),* Vol.2, No.3, pp. 22-33, July 2011.
30. Xun Sun , Xing Mei , Shaohui Jiao , Mingcai Zhou , Zhihua Liu , Haitao Wang “Real-time local stereo via edge-aware disparity propagation” *Pattern Recognition Letters* , vol.49,pp.201-206, 2014.
31. Q. Yang, “Stereo matching using tree filtering” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol.37, no.4,pp. 834-846, 2015.
32. Mikhail G. Mozerov & van de Weijer, “Accurate stereo matching by two-step energy minimization”, *IEEE Transactions on Image Processing*,vol.24,no.3,pp. 1153-1163, March 2015.
33. Zbontar, J., & LeCun, Y. “ Computing the stereo matching cost with a convolutional neural network”, in *Proceedings of the IEEE conference on computer vision and pattern recognition*, , October 2015, pp.1592-1599.
34. Zbontar, J., & LeCun, Y. “Stereo matching by training a convolutional neural network to compare image patches”. *Journal of Machine Learning Research*,vol.17,no.2,pp. 1-32 , April 2016.
35. D. Scharstein, R. Szeliski, Middlebury stereo evaluation – version
36. 2, 2010, Available: http://vision.middlebury.edu/stereo/eval/

AUTHORS PROFILE

Md. Abdul Mannan Mondal has completed his B.Sc. (Hons.) and M.S. from the dept. of *Computer Science and Engineering, Jahangirnagar* University, Savar, Dhaka, Bangladesh. His research field was in image processing in M.S. program and also research interest includes stereo image processing and computer vision.

He is currently working as an Associate Professor of Computer Science and Engineering, Dhaka City College affiliated by National University, Bangladesh and also the Ph.D. student of University of Dhaka under the supervision of Mohammad Haider Ali. Mr. Mannan is a Senior Member of *International Association of Computer Science and Information Technology*(IACSIT), Singapore. He has completed Short-Term Training Programme on *Knowledge Discovery in Data Bases: Data, Information & Knowledge* (DInk’10) from the Centre for Soft Computing Research (CSCR), Indian Statical Institute, India.

Mohammad Haider Ali has completed his *Doctor of Engineering* from the dept. of *Electronics and Information Engineering*, Toyohashi University of Technology, Japan in the field of *Computer Graphics*. His research interest includes Computer Graphics and Image Processing. He is currently working as a professor in the dept. of Computer Science and Engineering, University of Dhaka and also the Ph.D. supervisor of Md. Abdul Mannan Mondal. Dr. Mohammad Haider Ali is the fellow of *Bangladesh Computer Society,* and former Student Member of *Institute of Electronics, Information and Communication Engineering* (IEICE), Japan.