Development of Double Stage Filter (DSF) for Stereo Matching Algorithms and 3D Vision Applications

Teo Chee Huat, Nurulfajar bin Abd Manap, Masrullizam bin Mat Ibrahim
Department of Electronic Engineering, Faculty of Electronics and Computer Engineering
University Teknikal Malaysia Melaka, Melaka, Malaysia
nicteo@student.utem.edu.my

Abstract—**A part of the stereo matching algorithms development is mainly focused on overcoming unwanted aspects such as noises, unwanted regions and occlusions. In this paper, a new technique which is called Double Stage Filter (DSF) is introduced. This technique is a hybrid algorithm which consists of dynamic programming and block matching. The main feature of DSF is mainly its function at the post-processing stage that is to remove the noises and horizontal stripes, obtained from the raw disparity depth map of dynamic programming. In order to remove the unwanted aspects, a two-stage filtering process is applied. In this DSF algorithm, segmentation process is also required to segment the optimized raw disparity depth map into several parts according to the pixel colours. The first filter block is applied to remove the noises of the segmented parts before merging. Meanwhile, the second filter is used to remove the unwanted region of the outliers on segmented parts after merging processes. The new disparity depth map of DSF is evaluated in Middlebury Stereo Vision page with a few evaluation functions, such as similarity structural (SSIM), peak to signal noise ratio (PSNR) and mean square errors (MSE). At the end of this paper, the performance of DSF is compared with other techniques.**

Index Terms—**stereo matching, disparity, depth map, double stage filter, dynamic programming, median filter, segmentation**

I. INTRODUCTION

Stereo matching is one of the popular topics in the field of computer vision and has gained interest among researchers. There are many stereo matching algorithms that have been developed, which can be found in the Middlebury Stereo Vision Website by Scharstein and Szelinski [1]. This particular page provides various datasets of stereo images and standard evaluation for researchers to compare the results obtained from their proposed stereo matching algorithm with the others. According to Scharstein [1], most stereo matching algorithms consist of four common steps such as matching cost computation, cost aggregation, disparity computation optimization and disparity refinement.. However, the four steps are not necessary to be included in a stereo matching algorithm as it depends on the design organized by the researchers.

In the development of stereo matching algorithm, the most preferable step which attracts the interest from the researchers is on the post-processing step. Post-processing is the step of disparity refinement where the raw disparity map is obtained from the computation of correspondence and optimization. The raw disparity map contains some noises, mismatches and holes from occlusion that need to be smoothed out and removed. There are various types of post-processing that can be performed, such as the sub-pixel estimation, median filter, cross-checking and surface fitting [2, 3, 4, 5]. Some algorithms with high processing speed may obtain a relatively low precision of disparity depth map [6, 7, 8].

The main aim of this research is to develop a stereo matching algorithm with less complexity in computation, but simultaneously able to reduce the unwanted aspects of a raw disparity depth map. Basic block matching (BBM) will be used to determine the corresponding match using the sum of absolute difference (SAD) [9]. Dynamic programming acts as an optimization to compute the minimum cost of matching pixels between the corresponding scanlines. However, there is an issue with this approach, particularly on the consistency of inter-scanline of the raw disparity depth map obtained by dynamic programming [2, 7, 10, 11, 12]. There are horizontal lines that appear on the raw disparity depth map obtained due to the scanline process from dynamic programming. In order to reduce the lines on the raw disparity depth map, segmentation is needed to segment the raw disparity depth according to pixel colors of the contents in the map. Several segmented parts filtered by the basic median filter merged to produce a new disparity map. After merging, the second stage of median filter will be applied to the new disparity map to remove the existing outliers. As there are two stages of median filter involved, this algorithm is known as Double Stage Filter (DSF).

This paper consists of six sections. Section 2 includes the overview of the block diagram and algorithm structures. Section 3 covers the proposed stereo matching algorithm of Double Stage Filter (DSF). Section 4 discusses the improvement of DSF in obtaining a disparity depth map in comparison to the basic block matching, dynamic programming algorithm and Middlebury Ranking Stereo Page are presented. In Section 5, various performance evaluation techniques on the algorithms in obtaining the disparity depth map are presented. Lastly, Section 6 concluded with suggestions for future work.

II. OVERVIEW OF ALGORITHM DESIGN STRUCTURE

The overview of Double Stage Filter (DSF) algorithm is shown in Figure 1. DSF algorithm consists of two major parts: basic stereo matching and double stage filter. For the first part, it consists of three basic components, which are needed in a stereo matching algorithm according to Scharstein [1] such as matching cost computation, cost aggregation and disparity computation/optimization. In the step of matching cost computation, there are two types that are commonly used namely, the pixel-based matching cost [13, 14, 15, 16] and the area-based matching cost [17]. For matching cost computation, there are various similarity measurement techniques that are usually applied, such as the Sum of Absolute Differences (SAD), Sum of Squared Differences (SSD), Normalized Cross Correlation (NCC) and Sum of Hamming Distance (SHD). Cost aggregation is the absolute differences within the types of windows like square windows, multiple windows, shiftable windows [7, 18], adaptive windows [19, 20, 21, 22] and windows with constant disparity [23]. The function of cost aggregation depends on the content in the image, such as the position of the objects, shape, and weights [24].
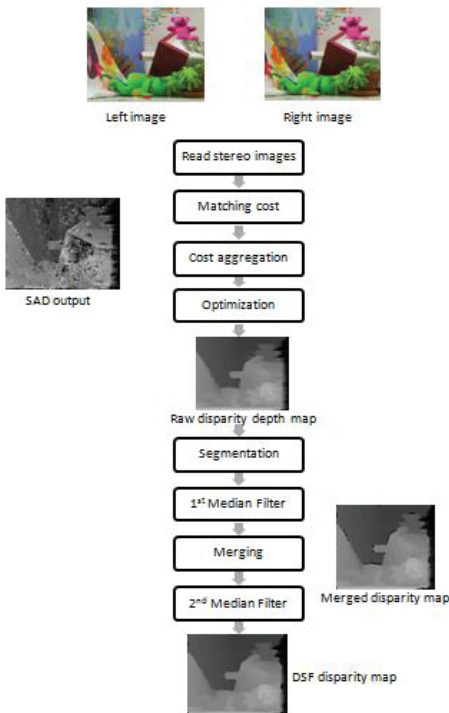

Figure 1: Block diagram design of DSF algorithm

In this algorithm, SAD is used as the matching cost to compare the image regions between each block with the region of interest because of its simplicity in calculation. However, the matching cost may cause noises since each pixel selects its disparity from all other pixels independently.

Cost aggregation depends on the window-based to sum up or take the average over the cost values from disparity space image (DSI) [25]. Fixed window of cost aggregation is applied in this algorithm as it is commonly used for applications which concern with real-time output due to fast and easy implementation. The size of the window needs to be chosen appropriately since a small window does not provide good results on less-textured regions, while a bigger size window does not have the ability to capture tiny and thin objects.

The dynamic programming method is used in DSF algorithm as an optimization in order to reduce the noises on the disparity image obtained from SAD with fixed window [26]. This method has its own constraint to smooth or remove the edges of the objects boundaries found in the noisy disparity image [27, 28]. It will refine the disparity image between rows in horizontal direction. The direction of dynamic programming in refining the noises of a disparity image between rows may cause lines to appear on the raw disparity depth map.

In the second part of DSF algorithm, the major concern is to remove the lines and reduce the remaining noises of the disparity depth map resulting from dynamic programming. Segmentation is used in this algorithm to segment the raw disparity depth map into several parts according to the pixel colors in the map. The selection on the particular range of pixel color only will be segmented and filtered by a median filtering process. Median filtering is a non-linear operation, which is most frequently used in image processing to reduce noises while preserving the edges of an image [29]. Several segmented parts will go through the first stage of median filtering with suitable window size in part-by-part to remove noises on every part of the disparity map. According to [30], it is important to choose a suitable size window as the use of large windows could blur the depth boundaries and could not capture tiny contents of image while small windows are less accurate in providing good results in less-textured regions due to its disability in capturing sufficient variation intensity. The step of merging will add up the filtered parts into a new disparity map. However, there will be remaining outliers between the filtered parts after the merging step. In the last step, the second stage of median filtering is applied to remove the outliers and the remaining noises of the new disparity map.

III. DOUBLE STAGE FILTER ALGORITHM

This section will describe the whole algorithm of Double Stage Filter (DSF) in detail. In the matching cost, a basic block matching (BBM) is used to find the corresponding pixels between the stereo pairs. The block matching will approximate the motion vector and the displacement of the correspondence stereo where the pixel value of the target image will be taken as the corresponding pixel for reference image [9]. It will also reduce the matching errors between the positions of the block, position (x, y) of the target image I_t, while $(x+u, y+v)$ is the position of reference image as I_{t-1} where u and v as the motion vector and these variables can be summarized as the sum of absolute difference (SAD) [9].

$$(u,v) \equiv \sum_{j=0}^{S-1} \sum_{i=0}^{S-1} |I_t(x+i,y+j) - I_{t-1}(x+u+i, y+v+j)| \qquad (1)$$

where S represents the block size while i and j representing the pixels. The $SAD_{(X,Y)}(u,v)$, (a, b) acts as the motion vector estimation to compare each position in $(x + u, y + v)$ for minimization [9].

$$(a,b) \equiv \arg\min_{(u,v)\in Z} SAD_{(x,y)}(u,v) \qquad (2)$$

where Z is equal to $\{(u, v)|$ -B $\le$ u, v $\le$ B and $(x + u, y + v)$ representing the correct position of pixel in the reference image, $I_{t-1}\}$ while B acts as an integer to look for suitable range. The basic block matching is most frequently used in image processing due to its less computation in obtaining a disparity depth map. However, the prediction on the right order position of images is not accurate and this shows that block matching is not able to guarantee the global matching errors.

On the cost aggregation, SAD with fixed window is used to aggregate the matching cost in this algorithm due to its simplest strategy. The fixed window has its own characteristic which assumes the frontal-parallel surfaces indirectly, ignores the depth discontinuities, and is unable to detect on uniform areas and repetitive patterns [31]. These characteristics of fixed window can be solved in the step of optimization and filtering process.

In the step of disparity computation or optimization, the dynamic programming (DP) method is chosen, which is used to optimize energy function of non-deterministic polynomial-time hard (NP-hard) for the purpose of smoothing [1, 27]. Global optimization can be categorized into one-dimension and two-dimension optimization methods. One dimension optimization is a traditional method, in which its smoothing strategy focuses on a horizontal direction. Its evaluation is on the disparity which comes with a pixel that is dependent on all other pixels on the same scanline while independent on the disparity that focuses on the other scanlines. This shows that it is not a purely global optimization; however, but it is still frequently used by researchers due to its simple and fast implementation in obtaining a good disparity depth map.

Another optimization method is the two-dimension optimization where it smoothes the stereo images in both directions, vertical and horizontal by using simulated annealing, continuation techniques and mean-field annealin;, however, unfortunately these methods are less capable to optimize the equation (1) [32, 33, 34].

$$E(d) = E_{data}(d) + E_{smooth} \qquad (3)$$

where $E_{data}(d)$ is stated as the disparity function through the correspondence pixels of disparity map, d which is minimized from its input data when there is similarity on intensities while it will be maximized when the correspondence pixels have slightly different intensities. E_{smooth} acts as the assumptive of smoothness of algorithm, which measures the disparity between the pixels on pixel grid [1]. However, there are two other techniques which are able to optimize the equation (1), graph-cuts and belief propagation as it can obtain better results

by referring to the data on ground truth [35, 36]. This research focuses on a simple and low complexity of the stereo matching algorithm. One dimension of the optimization dynamic programming is chosen to be used in the DSF algorithm as the energy minimization issue can be solved efficiently. Disparity space image (DSI) can be defined as a data structure to explain how the dynamic programming algorithm finds matches and occlusions simultaneously [21]. Although DP is a highly efficient method in obtaining good disparity depth map, it has difficulty in inter-scanline consistency that appears as horizontal "streaks" in the obtained disparity map. Researchers have proposed several approaches in improving the inter-scanline consistency, unfortunately these approaches produce only a slight reduction of the horizontal "streaks" [11, 12, 37]. In DSF, the step after the optimization can be used to remove the horizontal "streaks" in the raw disparity depth map obtained from the dynamic programming.

The step after optimization is segmentation, where in this step the raw disparity depth map that is obtained from the dynamic programming will be segmented into several parts according to the pixel colors based on the disparity range set for a particular stereo pair of images. Every segmented part depends on the range of pixel colors in the disparity map as shown in Figure 2. Each segmented part from the raw disparity depth map will be filtered up using median filtering. Median filtering works by sliding pixel by pixel through the image and replacing each pixel value with a median value of the neighboring pixels, which act as a "window". The median value is formed by sorting the pixel values from the window in the order of numerical and replacing the pixel with the center pixel value [29]. Filtered parts are then merged into a new disparity map by adding up all filtered images of each segmented part as shown in Figure 3(a). The new merged disparity map may contain cracks and holes due to the merging process, therefore the second stage of median filtering is applied to remove the outliers and the existing noises. The final disparity depth map is obtained with the horizontal "streaks" completely removed as shown in Figure 3(b).

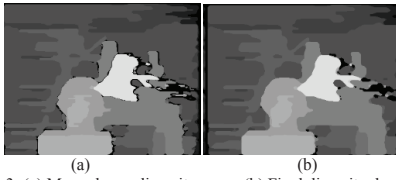Figure 2: Sample of segmented part of raw disparity depth map for Tsukuba

Figure 3: (a) Merged new disparity map ; (b) Final disparity depth map

IV. COMPARISON AMONG STEREO MATCHING ALGORITHMS

This section will discuss the comparison of the disparity depth map, which is obtained from the three stereo matching algorithms: basic block matching (BBM), dynamic programming (DP), and double stage filter (DSF). The developers of online performance evaluation in Middlebury Stereo Vision Page provide a variety of datasets including the stereo images with the ground truth images [1]. This Middlebury evaluation system is frequently used by researchers to evaluate their own algorithm. In this paper, four stereo pairs of images from the Middlebury Stereo Vision Page were used as the input data for the three stereo matching algorithms for evaluations: Tsukuba, Venus, Teddy and Cones.

In order to evaluate the stereo matching algorithms, the disparity depths obtained from the DSF algorithms were uploaded onto the Middlebury page with a specified scale of integer factors. The disparity maps will be generated by the evaluation system and the root-mean-squared (RMS) error and the percentage of bad matching pixels will be calculated by its calculation engine. The equations for the two quality measures are as follows [1]:

i) Root-mean-squared (RMS)

$$R = \left(\frac{1}{N} \sum_{(x,y)} |d_c(x,y) - d_T(x,y)|^2 \right)^{\frac{1}{2}}$$
(4)

where N as the total number of pixels, $d_c(x,y)$ represents the computed disparity map while $d_T(x,y)$ represents the ground truth map.

ii) Percentage of bad matching pixels

$$B = \left(\frac{1}{N} \sum_{(x,y)} |d_c(x,y) - d_T(x,y) > \delta_d \right)$$
(5)

where δ_d represents the tolerance for disparity error. The value of 1.0 is used for this equation as this value is fulfilled by few studies from other researchers [38, 39, 40]. There are three types of regions which are computed by the pre-processing on the ground truth image and the reference image such as textureless regions, occluded regions and depth discontinuity regions. Many researchers have used this platform to evaluate their stereo matching algorithms as it provides an overview performance on their algorithms in comparison to other algorithms.
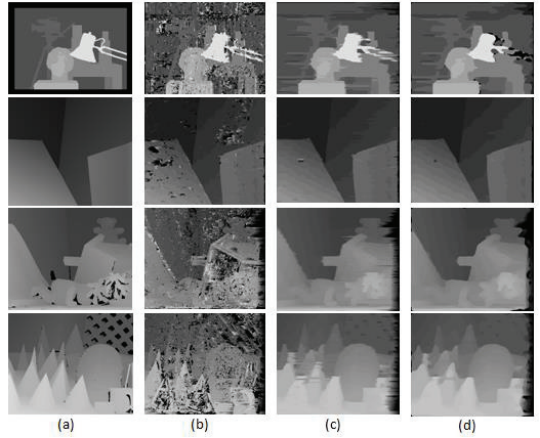

Figure 4: Results of three stereo matching algorithms by the Middlebury benchmark datasets : Tsukuba, Venus, Teddy and Cones. The first column images are the ground truths. The second column images show the results from BBM. The third column images show the results from DP. The fourth column images show the results from DSF algorithm.

Figure 4 shows the result of the disparity maps derived from BBM, DP and DSF with the ground truth images of four datasets from Middlebury. The result in the second column in Figure 4 used the fixed window SAD of 7x7 for ('Venus'), and 3x3 for ('Tsukuba', 'Teddy' and 'Cones'). Meanwhile, there are two window sizes for each stereo pair in DSF algorithm due to the two stages of median filtering as post-processing. For the first stage of median filtering, the window size of 7x7 for ('Tsukuba' and 'Venus') and 11x11 for ('Teddy' and 'Cones') are used. Meanwhile, for the second stage of median filtering, the window size of 7x7 for ('Tsukuba') and 11x11 for ('Teddy', 'Cones' and 'Venus') are applied. These parameters are selected to obtain the best result for disparity maps. In the second column, it shows that the disparity maps, which are obtained from BBM contain noises all around the map. DP algorithm successfully removed the noises of BBM, unfortunately there are still horizontal 'steaks' appeared in the disparity maps as shown in the third column. In the fourth column in Figure 4 it indicates that the proposed algorithm, DSF has completely removed the horizontal 'streaks' of DP algorithm.

Table 1

Middlebury ranking for DSF algorithm and the values shows the percentage of bad pixels which errors are more than 1 pixels. For Venus image, the proposed

Algorithm	Tsukuba			Venus			Teddy			Cones			Avg (%)
	N-o	All	Disc	N-o	All	Disc	N-o	All	Disc	N-o	All	Disc	
PhaseBased	4.26	6.53	15.4	6.71	8.16	26.4	14.5	23.1	25.5	10.8	20.5	21.2	15.3
IMCT	4.54	5.90	19.8	3.16	3.83	23.2	18.0	23.1	35.3	12.7	18.5	27.9	16.3
FW-DLR	4.87	5.89	22.9	2.50	3.22	18.3	18.2	18.7	37.2	24.2	27.9	42.1	18.8
BioDEM	6.57	8.43	28.1	3.61	4.80	33.7	13.2	21.3	34.5	6.84	16.0	19.8	16.4
DSF	4.76	6.19	21.9	1.85	2.72	24.2	16.0	23.2	34.6	22.3	29.4	44.0	19.3
SSD+MF	5.23	7.07	24.1	3.74	5.16	11.9	16.5	24.8	32.9	10.6	19.8	26.3	15.7
SO	5.08	7.22	12.2	9.44	10.9	21.9	19.9	28.2	26.3	13.0	22.8	22.3	16.6
DP	5.36	7.03	22.1	2.41	3.25	25.0	14.5	22.1	30.2	26.5	33.4	47.9	20.0
PhaseDiff	4.89	7.11	16.3	8.34	9.76	26.0	20.0	28.0	29.0	19.8	28.5	27.5	18.8
STICA	7.70	9.63	27.8	8.19	9.58	40.3	15.8	23.2	37.7	9.8	17.8	28.7	19.7
LCDM+AdaptWgt	5.98	7.84	22.2	14.5	15.4	35.9	20.8	27.3	38.3	8.9	17.2	20.0	19.5
Infection	7.95	9.54	28.9	4.41	5.53	31.7	17.7	25.1	44.4	14.3	21.3	38.0	20.7
BBM	24.9	26.0	30.5	8.90	9.77	26.1	40.7	45.4	51.9	34.3	40.2	56.7	33.0

algorithm shows good result in non-occluded (N-o) region.

* N-o = Non-occluded regions
* Dis = Depth discontinuities

Table 1 shows the performance and ranking of the stereo matching algorithms for the disparity maps in the Middlebury Stereo Vision page. The ranking depends on the average rank for bad pixels percentage. The BBM algorithm in the last ranking was due to its basic method of SAD where the results obtained were inaccurate and contained errors such as non-occluded and depth discontinuities. However, after the BBM was optimized by DP and refined by the proposed algorithm, DSF, the ranking has gone up to 7 places while 2 places above DP algorithm. The results shown in Table 1 also indicate our algorithm was competitive with the other stereo matching algorithms as our proposed algorithm focused on less computational complexity. Therefore, the algorithm was designed with basic methods of pre-processing and post-processing for easy applicable on any stereo matching systems. Meanwhile, for other algorithms, which have better ranking and have higher computation, its algorithm is suitable for complex content of stereo pair like the Teddy and Cones. Therefore, the algorithm is designed with basic methods of pre-processing and post-processing for easy applicable on any stereo matching systems.

Figure 5 shows the evaluation results from Middlebury for non-occluded regions, which is based on bad pixels with absolute disparity error more than 1. The first column shows the sample images for evaluation provided by Middlebury page where the white areas show non-occluded region, while the black areas show the occluded and border region. The second column shows the non-occluded regions of BBM algorithm that contain many errors. From observation, DP and DSF algorithm (from the third and fourth column respectively) indicate that Tsukuba and Venus datasets have less errors. There are many incorrect disparities on the datasets of Teddy and Cones due to the complexity of the content and texture regions in the dataset. The results show DP algorithm was able to smooth and remove noises, however it still produced horizontal 'streaks' on the disparity map obtained. The results in fourth column proved that DSF algorithm is able to improve the disparity depth map while removing the horizontal 'streaks' of DP algorithm. There might contain some cracks and holes on some regions of the disparity map obtained by DSF

algorithm due to segmentation and merging of disparity. This can be solved using matching
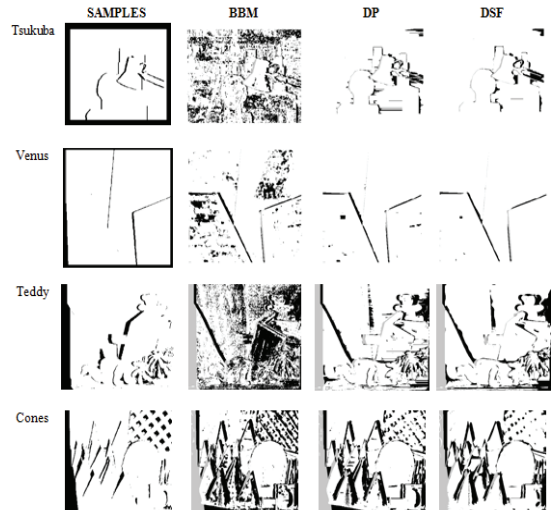
methods such as graph cut and segmented matching.

Figure 5: Non-occluded regions with occluded and border regions

V. EVALUATION ON STEREO MATCHING ALGORITHMS

There are a few other evaluation techniques used in this paper to evaluate the three stereo matching algorithms. The first technique is the Mean Squared Error (MSE), which is used to measure the average of squared errors between the computed disparity map and the ground truth.

$$MSE = \frac{1}{MN} \sum_{y=1}^{M} \sum_{x=1}^{N} [I_1(x,y) - I_2(x,y)]^2 \qquad (6)$$

where M and N represent the rows and columns in the input images of I_1 and I_2 respectively. As the value of MSE is low, it shows that the cumulative squared error is decreased [40].

Another evaluation technique is Peak Signal to Noise Ratio (PSNR), it comes together with MSE since its calculation equations involve the values of MSE. PSNR is used to determine whether an algorithm is able to produce a good result by comparing the effects of image quality using the image enhancement algorithms [40] .

$$PSNR = 10 \; log_{10}(\tfrac{R^2}{MSE}) \qquad (7)$$

In PSNR, R is the maximum fluctuation of data type in the input data, which is equal to 1 if the input image contains double-precision floating-point while it is equal to 255 for 8-bit unsigned integer data type. The higher the PSNR, the better the quality of the disparity map obtained, and it indicates that more noises are removed. The last technique in this research used to evaluate the performance of proposed algorithm and other algorithms is the Structural Similarity Index Metric (SSIM). It is a technique for measuring the similarity between the computed disparity map and the ground truth. SSIM is developed to improve the problem on peak signal-to-noise ratio (PSNR) and mean squared error (MSE) which are inconsistent to the human eye perception [41]. MSE and PSNR are techniques which assume the perceived errors, while SSIM approximates on image degradation as perceived change in structural information. Structural information is the spatially close pixels in an image which have strong inter-dependencies to carry important information about the structural content in the visual data [42].

$$SSIM(x,y) = \frac{\left(2\mu_x\mu_y + C_1\right)(2\sigma_{xy} + C_2)}{(\mu_x^2 + \mu_y^2 + C_1)(\sigma_x^2 + \sigma_y^2 + C_2)}$$

(8)

In SSIM equation, μ_x and μ_y are the average of x and y, σ_x^2 and σ_y^2 are the variance of x and y, σ_{xy} represent the covariance of x and y. $C_1 = (k_1 L)^2$ and $C_2 = (k_2 L)^2$ are the variables to optimize the division with denominator where L represents the dynamic range of pixel values while $k_1 = 0.01$ and $k_2 = 0.03$ as default. The results calculated from SSIM index should be in a decimal value between -1 to 1.

Table 2
MSE values for disparity maps of Middlebury datasets

	BBM	DP	DSF
Tsukuba	3537.69	2134.93	2233.33
Venus	496.92	242.16	276.55
Teddy	3157.40	1308.55	1538.73
Cones	4750.06	1580.13	1860.02

Table 3
PSNR values for disparity maps of Middlebury datasets

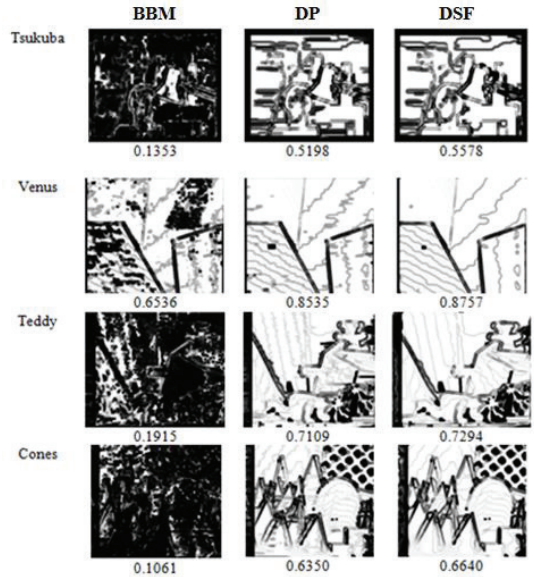|	BBM	DP	DSF
Tsukuba	12.64	14.84	14.64
Venus	21.17	24.30	23.71
Teddy	13.14	16.96	16.26
Cones	11.36	16.14	15.44

Figure 6: SSIM index map and values for disparity maps of Middlebury datasets

Based on Table 2, it shows that the values of MSE for BBM algorithm are much higher than DP and DSF algorithm. This indicates that BBM algorithm obtain disparity maps that contain more errors compared to DP and DSF algorithm. Meanwhile, the MSE values for DP and DSF algorithm for all the datasets are slightly closer. In Table 3, the values of PSNR for BBM algorithm are much lower than the other two algorithms. This indicates that the disparity maps obtained by BBM algorithm have lower quality than the disparity maps that obtained by other two algorithms. BBM algorithm has difficulty in removing the noises compared to DP and DSF algorithm. Figure 6 shows the SSIM index maps with SSIM index values for all the datasets. It clearly shows that the results obtained by DSF algorithm are better than DP and BBM algorithm. The SSIM index values of DSF algorithm is nearly to 1 which indicate that the disparity maps obtained have a higher similarity to the ground truth images of the datasets compared to the other two algorithms. Besides, the SSIM index maps of DSF algorithm proved that the horizontal 'streaks' and the remaining noises in the SSIM index maps of DP algorithm are reduced and removed. BBM algorithm has the lowest values of the SSIM index for all the datasets among the stereo matching algorithms and the overview of SSIM index maps is almost black, which contains high noises. From the three evaluation techniques, it shows that DSF algorithm has higher potential in removing noises and horizontal 'streaks' compared to BBM and DP algorithm.

VI. CONCLUSION

The Double Stage Filter (DSF) algorithm is designed with the aim to improve the raw disparity maps in the post-processing stage. This algorithm is applied together with

additional basic techniques in segmentation, median filtering and merging and design of a less complex computation stereo matching algorithm. The disparity maps obtained are evaluated on Middlebury Stereo Vision page and evaluated with MSE, PSNR and SSIM. DSF algorithm performed well as compared to DP and BBM algorithms. In the ranking of Middlebury, DSF has improved up to 7 places from BBM algorithm while 2 places above DP algorithm. DSF is also competitive with other existing algorithms besides DP and BBM algorithms. From the observation on the results of DSF algorithm, it is able to remove noises, horizontal 'streaks' of DP algorithm and improve the depth discontinuities of disparity depth maps. For future work, the DSF algorithm may be implemented for most of 3D applications due to its simple implementation and less computational complexity of algorithm design. For recommendation, median filtering may be replaced with other advanced type of filtering to obtain different results of the disparity depth map.

REFERENCES

[1] D. Scharstein, R. Szeliski, and R. Zabih, "A taxonomy and evaluation of dense two-frame stereo correspondence algorithms," International Journal of Computer Vision, vol. 47, pp.7-42, 2002.

[2] Y. S. Heo, K. M. Lee, and S. U. Lee, "Robust Stereo matching using adaptive normalized cross-correlation," IEEE Trans. Pattern Anal. Mach. Intell., vol. 33, no. 4, pp. 807–822, 2011.

[3] Z. Ma and K. He, "Constant Time Weighted Median Filtering for Stereo Matching and Beyond," pp. 2–9, 2013.

[4] P. Fua., "A parallel stereo algorithm that produces dense depth maps and preserves image features," Machine Vision and Applications, vol.6, pp. 35–49, 1993.

[5] W. Li and B. Li, "Virtual view synthesis with heuristic spatial motion," vol. 1, pp. 1508–1511, 2008.

[6] C. Zhou, A. Troccoli, and K. Pulli, "Robust stereo with flash and no-flash image pairs," Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit., pp. 342–349, 2012.

[7] A. F. Bobick and S. S. Intille, "Large occlusion stereo," IJCV, vol.33. no. 3, pp. 181–200, 1999.

[8] F. Calakli, A. O. Ulusoy, M. I. Restrepo, G. Taubin, and J. L. Mundy, "High resolution surface reconstruction from multi-view aerial imagery," Proc. 2nd Jt. 3DIM/3DPVT Conf. 3D Imaging, Model. Process. Vis. Transm. 3DIMPVT 2012, pp. 25–32, 2012.

[9] H. Jaspers and H. Wuensche, "Fast and Robust B-Spline Terrain Estimation for Off-Road Navigation with Stereo Vision," pp. 140–145, 2014.

[10] Y. C. Wang, C. P. Tung, and P. C. Chung, "Efficient disparity estimation using hierarchical bilateral disparity structure based graph cut algorithm with a foreground boundary refinement mechanism," IEEE Trans. Circuits Syst. Video Technol., vol. 23, no. 5, pp. 784–801, 2013.

[11] X. Tan, C. Sun, X. Sirault, R. Furbank, and T. D. Pham, "Stereo matching using cost volume watershed and region merging," Signal Process. Image Commun., vol. 29, no. 10, pp. 1–13, 2014.

[12] N. Anantrasirichai, C. N. Canagarajah, D. W. Redmill, and D. R. Bull, "Dynamic Programming for Multi-View Disparity / Depth Estimation Abstract," no. 1, pp. 269–272, 2006.

[13] M. J. Hannah, "Computer Matching of Areas in Stereo Images," PhD thesis, Stanford University, pp. 152-171, 1974.

[14] A. Donate, X. Liu, and E. G. Collins, "Efficient path-based stereo matching with subpixel accuracy," IEEE Trans. Syst. Man, Cybern. Part B Cybern., vol. 41, no. 1, pp. 183–195, 2011.

[15] M. Sabatini, R. Monti, P. Gasbarri, and G. B. Palmerini, "Adaptive and robust algorithms and tests for visual-based navigation of a space robotic manipulator," Acta Astronaut., vol. 83, pp. 65–84, 2013.

[16] P. M. Jodoin, M. Mignotte, and C. Rosenberger, "Segmentation framework based on label field fusion," IEEE Trans. Image Process., vol. 16, no. 10, pp. 2535–2550, 2007.

[17] L. Di Stefano, M. Marchionni, and S. Mattoccia, "A fast area-based stereo matching algorithm," Image Vis. Comput., vol. 22, no. 12, pp. 983–1005, Oct. 2004.

[18] R. D. Arnold, "Automated stereo perception," Technical Report AIM-351, Artificial Intelligence Laboratory, Stanford University, pp, 67-78, 1983.

[19] M. Debella-gilo and A. Kääb, "Locally adaptive template sizes for matching repeat images of mass," Institute of Geosciences , University of Oslo , Oslo , Norway," no. 1, pp. 4281–4284, 2011.

[20] S.-Y. Lee, J.-Y. Sim, C.-S. Kim, and S.-U. Lee, "Correspondence Matching of Multi-View Video Sequences Using Mutual Information Based Similarity Measure," IEEE Trans. Multimed., vol. 15, no. 8, pp. 1719–1731, 2013.

[21] X. Mi, "Stereo Matching based on Global Edge Constraint and Variable Window Propagation," pp. 936–940, 2012.

[22] G. P. Fickel, C. R. Jung, T. Malzbender, R. Samadani, and B. Culbertson, "Stereo matching and view interpolation based on image domain triangulation," IEEE Trans. Image Process., vol. 22, no. 9, pp. 3353–3365, 2013.

[23] C. C. Pham and J. W. Jeon, "Domain transformation-based efficient cost aggregation for local stereo matching," IEEE Trans. Circuits Syst. Video Technol., vol. 23, no. 7, pp. 1119–1130, 2013.

[24] F. Tombari, S. Mattoccia, L. Di Stefano, and E. Addimanda, "Classification and evaluation of cost aggregation methods for stereo correspondence," 2008 IEEE Conf. Comput. Vis. Pattern Recognit., pp. 1–8, Jun. 2008.

[25] F. Tombari, L. Di Stefano, S. Mattoccia, A. Mainetti, and D. Arces, "A 3D Reconstruction System Based on Improved Spacetime Stereo," no. December, pp. 7–10, 2010.

[26] R. A. Hamzah, K. A. A. Aziz, and A. S. M. Shokri, "A Pixel to Pixel Correspondence and Region of Interest in Stereo Vision Application," pp. 193–197, 2012.

[27] O. Veksler, "Stereo Correspondence by Dynamic Programming on a Tree," 2005 IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit., vol. 2, no. 1, pp. 384–390.

[28] C. S. Park and H. W. Park, "A robust stereo disparity estimation using adaptive window search and dynamic programming search," Pattern Recognit., vol. 34, no. 12, pp. 2573–2576, Dec. 2001.

[29] Lim, Jae S., "Two-Dimensional Signal and Image Processing," Englewood Cliffs, NJ, Prentice Hall, pp. 469-476, 1990.

[30] N. A. Manap and J. J. Soraghan, "Disparity Refinement Based on Depth Image Layers Separation for Stereo Matching Algorithms," Centre for Excellence in Signal and Image Processing, vol. 4, no. 1, 2012.

[31] J. Zhao and J. Katupitiya, "A multi-window stereo vision algorithm with improved performance at object borders," Proc. 2007 IEEE Symp. Comput. Intell. Image Signal Process., pp. 66–71, 2007.

[32] L. Zhu, Y. Chen, Y. Lin, C. Lin, and A. Yuille, "Recursive Segmentation and Recognition Templates for 2D Parsing," Adv. Neural Inf. Process. Syst. 21, vol. 34, no. 2, pp. 1985–1992, 2009.

[33] I. Gerace and R. Pandolfi, "A color image restoration with adjacent parallel lines inhibition," Proc. - 12th Int. Conf. Image Anal. Process. ICIAP 2003, pp. 391–396, 2003.

[34] A. Gharib and A. Harati, "Toward Application of Extremal Optimization Algorithm in Image Segmentation," pp. 167–172, 2012.

[35] G. Vogiatzis, C. Hernández, P. H. S. Torr, and R. Cipolla, "Multiview stereo via volumetric graph-cuts and occlusion robust photo-consistency," IEEE Trans. Pattern Anal. Mach. Intell., vol. 29, pp. 2241–2246, 2007.

[36] D. Min and K. Sohn, "An asymmetric post-processing for correspondence problem," Signal Process. Image Commun., vol. 25, no. 2, pp. 130–142, 2010.

[37] C. Cassisa, "Local vs global energy minimization methods: Application to stereo matching," 2010 IEEE Int. Conf. Prog. Informatics Comput., vol. 2, pp. 678–683, 2010.

[38] Y. Yin, M. Jin, and S. Yi Xie, "A stereo pairs disparity matching algorithm by mean-shift segmentation," 3rd Int. Work. Adv. Comput. Intell. IWACI 2010, pp. 639–642, 2010.

[39] S. Zhang, S. Liu, Y. Mao, and X. Wang, "Global optimization for bidirectional stereo matching with occlusion handling," Proc. 2012 Int. Conf. Meas. Inf. Control. MIC 2012, vol. 2, pp. 553–557, 2012.

[40] C. S. Panchal and A. B. Upadhyay, "Depth Estimation Analysis Using Sum of Absolute Difference Algorithm," pp. 6761–6767, 2014.

[41] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image quality assessment: from error visibility to structural similarity," IEEE Trans. Image Process., vol. 13, no. 4, pp. 600–12, Apr. 2004.

[42] Y. Tao, H. Lin, F. Dong, C. Wang, G. Clapworthy, and H. Bao, "Structure-aware lighting design for volume visualization," IEEE Trans. Vis. Comput. Graph., vol. 18, no. 12, pp. 2372–2381, 2012.