A Multistage Hybrid Median Filter Design of Stereo Matching Algorithms on Image Processing

Maged Aboali, Nurulfajar Abd Manap, Abd Majid Darsono, Zulkalnain Mohd Yusof
Machine Learning and Signal Processing (MLSP) Research Group,
Centre for Telecommunication Research & Innovation (CeTRI),
Faculty of Electronic and Computer Engineering (FKEKK), Universiti Teknikal Malaysia Melaka (UTeM)
nurulfajar@utem.edu.my

***Abstract*—This paper presents a new method of stereo matching algorithms known as a Multistage Hybrid Median Filter (MHMF). The main challenging problem within computer vision field is stereo matching, considering the many drawbacks and issues resulting from stereo matching, which makes it a difficult and unresolved approach. Large studies and methods have been introduced and attempted to deal with stereo matching problems from a different perspective. However, most of these existing methods still suffer from low accuracy and algorithms complexity. Thus, to provide valuable and efficient solutions to stereovision community and to support researchers in the same line of research, we created a highly efficient and robust algorithms based on Hybrid Median Filter (MHF). In this paper, the (MHMF) is introduced as a newly implemented method of stereo matching algorithms. This method consists of three main stages, in which each stage involves multiple processes. In stage 1, the Basic Block Matching (BBM), Dynamic Programming (DP) algorithms are accomplished. While stages 2 and 3 rely on newly developed post-processing algorithms, which involve Hybrid Median Filtering (MHF), segmentation and merging processes as the main core of our research approach. The significant contribution of this method is on its capability of solving major drawbacks and problems of stereo marching including the high noises, horizontal stripes, multiple unwanted regions and aspects, and occlusions problems. The paper provides an evaluation of our method with some existing algorithms using the most common stereo functions including the MSE, PSNR, and SSIM. Finally, it is found that our developed algorithms achieved high accuracy of disparity depth map with convenient execution time among other compared methods**.

***Index Terms*—Image Processing; Stereo Matching Algorithms; Disparity Depth Map; Multistage Hybrid Median Filter (MHMF).**

I. Introduction

Stereo matching is a significant approach in computer vision field and has been actively investigated by researchers. It is a challenging area of research considering the core part of the numerous 3D stereo applications has highly accurate disparity depth image [1], [2]. A critical and well-addressed problem within the stereo vision area is the ways to find the corresponding pixels to multiple or a pair of stereo images. As one of the passive three-dimensional methods, stereo matching is applied to find the corresponding points between a pair of stereo view images. The principle aims and basis of stereo matching stand on the ability of this method to extract the three-dimensional method and provide three-dimensional 3D reconstruction from two stereo image pairs [3], [4]. To extract the depth of information from a pair of stereo images, a depth map of stereo matching algorithms is contacted by determining the correspondence or “pixel-wise” between those stereo multi-view images for each pixel. In this case, the pixel-wise is defined and represented by the disparity of “a displacement vector of the specific pixel between the two stereo pair images”. All disparity values of each stereo image are stored in the disparity map, in which the disparity map provides the disparity value for every pixel.

Over the past decades, the interest of stereo matching methods has grown motivated by the high demand for extracting three-dimensional 3D information from stereo image pairs captured by two stereo cameras. Numerous emerging applications rely on the 3D information represented as a depth map. Stereo matching approach reconstructs a depth map from the stereo multi-view images. The generated or obtained depth map of stereo matching algorithms is the core part and plays a significant role in numerous applications such as the 3D reconstruction, robot vision, automatic navigation, (3D) scanning, machine vision, 3D tracking, virtual reality, automotive vehicles, mobile applications, 3D cinema, monitoring systems, photogrammetry, free-view video, and many more [5], [6].

The massive advances in 3D applications and technologies increased the interest of developing efficient methods that are able to extract three-dimensional 3D information with high accuracy and less execution time. However, stereo matching algorithms are commonly computationally intensive and developing an efficient algorithm of stereo matching on embedded systems is a challenge, especially for real-time applications. Thus, a huge number of studies have been published and numerous methods based on stereo matching approach have been proposed and developed, which can be searched in Middlebury Website Page of stereo vision by Scharstein and Szelinsk [7].

Middlebury Website Page is a global online sharing web-page for performing a standard evaluation of all developed algorithms by providing datasets benchmarking of stereo image for researchers in the same line of research area. It allows researchers to compare their obtained results and performance of their developed stereo matching algorithms among each other.

According to the taxonomy and the categorization scheme by Scharstein and Szelinski for stereo algorithms [7], the existing stereo matching approach for computing disparity depth maps can be categorized into two main types: local methods [8]–[10], which generally use supporting window with a particular size and form to find and calculate the matching cost and to obtain the disparity in cost aggregation. The local methods can be realized with certain low

complexity and ease of implementing to the real-world systems.
In contrast, the global approach [11]–[13] assumes an energy function, particularly the smoothness "enhancement" and data energy terms to obtain the matching disparity such as algorithm to optimize the energy function algorithms [14], the graph cut (GC) [15], the dynamic programming (DP) [16], and the belief propagation (BP) [12], [13]; whereas, global approaches can provide disparity map with high accuracy. However, they are usually computationally expensive [17]. Another approach of finding pixel-wise "correspondence" between a pair of stereo images is the semi-global. This approach follows the main principle of global algorithm, but over multiple 1D scanlines of a pair stereo images rather than the whole one. It utilizes a local approximation to form the matching cost and aggregates using a global cost function. The semi-global algorithms approach adds the constraint of smoothness for the output disparity depth map [18].

Disparity depth map implementation is among the most focused approach of stereo matching algorithms due to its continued and extensive relevance. It is a well-studied topic that highlights the part of stereo matching algorithm and considers obtaining high accurate disparity map is a great challenge. However, a disparity map implementation with high accuracy has been a major demand since its embedded feature is considered unique in many 3D applications and modern technologies. This opens the opportunity for developing a huge number of new algorithms with various structures and different performance. Although a large number of studies and a lot of methods have adopted the advance of stereo matching algorithms toward improving the quality of obtained disparity dense map, the major drawbacks and problems of stereo marching include the high level of noises, the existence of horizontal stripes, and the presence of multiple unwanted regions and aspects. Additionally, the occlusions problems remain as challenges and unresolved. [19], [20].

The structure of this paper has five configurations parts, as follows: In Section I, stereo matching algorithms and its interest are introduced. In Section II, some previous works and progression research are discussed. In Section III, the outlines of our developed method based on Hybrid Median Filter (HMF) for stereo matching algorithms are discussed and described. In Section V, all the obtained results of disparity depth map using our implemented method are discussed in detail. In Section VI, the conclusion of this paper is presented.

II. RELATED WORKS

In this section, we investigate the most previous and significant works related to the scope of our research. The term three-dimensional (3D) reconstruction of stereo matching algorithms is among the approaches that have extensive investigation and great experimented search [18]–[20]. Thus, for the stereo matching domain and specifically for all the four standard steps introduced by Scharstein and Szelinski in their taxonomy [21], various methods and researches have been proposed and published. Nevertheless, this domain is still a great target of interest by numerous scholars and researchers' due to its high demand as a method with high-quality output. Additionally, it is also considered as algorithm with less complexity and computational efficiency. The fundamental of stereo vision research can be categorized into three common approaches: Preprocessing approach, stereo matching approach, and the post-processing approach. Each of these approaches has a series of research studies and focuses [22]. However, among all studies and aspects of stereo matching algorithm domain, scholar and developers paid more attention to disparity refinement step "post-processing" algorithms as a core step for further improving the quality of depth map toward obtaining a highly accurate depth map [23], [24]. Thus, generally, the post-processing step is a critical part of stereo matching algorithms since it aims to smooth the dense map obtained from the optimization step.

A significant detailed and focused discussion for post-processing step of stereo matching algorithms will be presented as follows:

A. Post-processing Algorithms

As mentioned previously, a wide variety of stereo matching algorithms have been deployed by many researchers abroad in order to improve the quality of disparity dense map and advanced algorithms performance. The implementation of disparity depth map from two or more stereo view images depends on two stages: (1) Initial disparity map obtaining, and (2) Disparity map refinement. In stage 1, the initial disparities maps are obtained. Meanwhile in stage 2, the main focus is on improving the quality of disparities maps of stage 1. The raw disparity map, the "initial disparity map" from the optimization step is obtained and often comes with several drawbacks and unwanted aspects including the high noise, occlusion problems, presence of textured regions, existence of horizontal stripes, and non-edge preserving, and many more. Largely, numerous related studies, researches, and methods have been widely investigated, developed and applied to solve the problems and drawbacks of the raw disparity map. Conventionally, some of the major practice methods depend on the left-right consistency technique as in [25], where the disparity map for the left and right of stereo images is calculated, and through the disparity map values, the inconsistencies are detected and eliminated. Birchfield et al. in [26] and Massimo Camplani in [27] introduced and proposed a hole filling to overcome some drawbacks of the raw disparity map.

A different approach has proposed the sub-pixel precision to maximize the resolution of the disparities as in [25]. These approaches allow for maximizing the depth resolution, although more computational cost is needed. More advanced and modern studies attempt to choose the smallest valid disparity for the neighborhood as discussed in [9]. In the same research area, some authors attempt to look for the minimum cost within the range restricted by the closest left and right valid disparity, as described in [28]. Using an efficient and interactive post-processing method Caizhang et al. tried to improve the raw depth map of the limited user input, as illustrated in [29]. Min et al. [30] presented an asymmetric post-processing method using an iterative filtering scheme during the refinement step with a specific asymmetric consistency check and adaptive filtering. There are also further advanced approaches of post-processing algorithms implemented with high performance and have the ability to limit the influence of raw disparity map refinement. Karsten et al. in [31] has introduced and developed a method that involves "unweighted" median filter to remove the number of the false matches with some additional computation processes. In addition, Sun et al. in [32] attempted to improve the quality of disparity map by imposing further consistency constraints to the depth maps. Meanwhile, in [33], Yang et al. suggested and developed a new non-local cost aggregation for representing the refinement of non-local aggregation results on the tree framework. In contrast, the most popular and simple applied approaches use the sub-pixel median filter as an efficient method that has the ability to overcome abrupt disparities [25], [34], [35].

B. Filtering Algorithms

Over the past years, significant and advanced stereo filtering algorithms have been achieved, particularly through the image processing domain. Those algorithms have been developed with different structures and designs, as found in the Middlebury Website Page by Scharstein and Szelinsk [7]. However, and in spite of the huge studies in developing various algorithms to achieve the goal of high accurate and less computation in gaining the disparity map, a unique algorithm with the ability of obtaining a high quality depth map remains so far a unique opportunity and a real challenging task, since the obtaining disparity depth map is mostly affected by a high level of noise. [9], [36], [37].

Thus, during the past decades, a resurgence interest has been seen in developing an efficient and high robust version of filtering algorithms, especially the versions with high abilities in handling occlusion, overcoming the depth discontinuities, removing noise, and overcoming some the unwanted regions and aspects. Meanwhile, many of these algorithms have been developed and conducted in image processing domain, especially for disparity map of stereo matching algorithms as illustrated in [21], [38]. Some other and recent studies that achieved better results are presented in [10], [38], [39], [40].

In this paper, the development of Multistage Hybrid Median Filter (MHMF) of stereo matching algorithms as a new method of stereo hybrid filtering is presented. In this study, a development of MHMF with high accuracy and more computational efficiency as well as a less complex structure on image processing is mainly discussed. Precisely, the Multistage Hybrid Median Filter (MHMF) is a new stereo hybrid filtering method with robust algorithms that perform accurately and efficiently in obtaining and smoothing the disparity depth map with such convenient computation time as a newly implemented method of stereo matching algorithm. The structure of Multistage Hybrid Median Filter (MHMF) consists of three main stages, wherein stage 1, the Basic Block Matching (BBM), Dynamic Programming (DP) algorithms are achieved. Whereas, Basic Block Matching (BBM) is applied in order to perform the matching process and finding corresponding pixels specifically with Absolute Difference (SAD). Besides, the Dynamic Programming (DP) is applied to find the minimum matching cost and operates as an optimization part as well. In contrast, the raw disparity map from optimization process came with multiple drawbacks such as horizontal streaks, particularly from the inter-scanline of Dynamic Programming (DP) Programming [16], [41], [42], that affect the quality of the obtained disparity map. Thus, to minimize the drawbacks of the raw disparity map, new developed post-processing algorithms are presented as stages (2, 3) of our algorithms. These algorithms involve Hybrid Median Filtering (MHF), segmentation and merging processes as the main core of our research approach. Through the post-processing approach, the segmentation part is applied to the segment to obtain raw disparity depth from stage 1 to many segments based on pixel colors contents for each object. Then, all extracted segments proceeded to the hybrid median filter to remove noises and streaks. Next, the merged part is followed to create a new efficient disparity map. In this case, there are three specific steps of Hybrid Median Filtering (MHF), thus our method is named as Multistage Hybrid Median Filter.

III. METHODOLOGY

In this section, we present the description of the implemented method in more details. The developed structure and fundamental algorithms of our approach "Multistage Hybrid Median Filter" are discussed. A clear and complete description of each part of developed algorithms is provided. All the outlines and fundamental aspects of our approach are further explained as follows:

The structure of Multistage Hybrid Median Filter (MHMF) is shown in Figure 1.

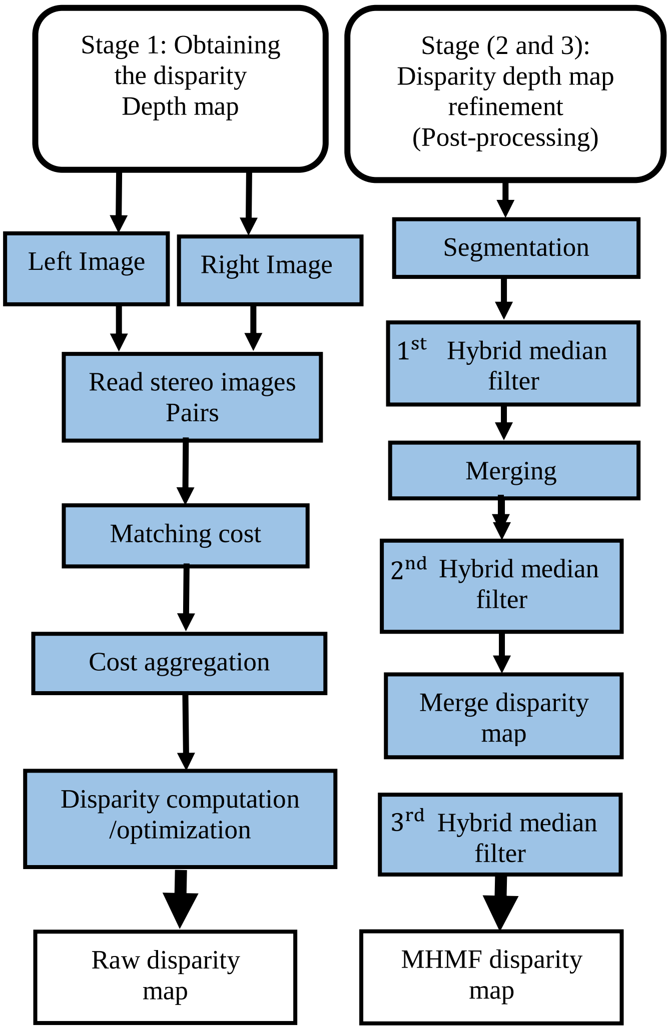


Figure 1: Block diagram and structure of MHMF algorithm

The design consists of three stages: In stage 1, the raw disparity map "initial disparity map" is obtained through three basic stereo matching algorithms considered as a necessity need based on Scharstein and Szelinski [21]. These steps are the matching cost, the cost aggregation process, and the disparity map computation /optimization, in which these steps are the crucial parts of stereo matching algorithms in order to obtain the raw disparity map. Stages 2 and 3 represent the refinement step (post-processing) that focus on refining the output of stage 1 “raw disparity map” and improving its accuracy. Besides, the post-processing of our algorithms is a newly developed approach that involves segmentation part, merging process, and organized hybrid median filter algorithms as the main core of our developed method.

In general, according to [21], most of the stereo matching algorithms are performed in four common steps as follows:

1) Matching cost computation
2) Cost (support) aggregation
3) The disparity computation/optimization
4) The disparity refinement

where each step of these algorithms has a specific influence, and effect on the stereo matching performance, although not all of these algorithms and the basic procedures are required to be included in creating a stereo matching complete framework. However, each step can be executed and structured in different form based on the focus of the study.

A. Matching Cost

In this significant and initial step, the correspondence pixels between the left and right images are determined. The similarity or dissimilarity of each pixel is under the consideration to the candidate’s matching pixel. However, the matching cost function is required to calculate all positions of pixels in the reference image and candidate pixels of the target image. Thus, our developed algorithm functions to find the correspondences pixels between both reference image and the target image. The correspondences are calculated using the Sum of Absolute Differences (SAD) function in which the (SAD) function minimizes the matching errors between all points of the block. The coordinate (x, y) of the target image I_t, and the coordinate of the reference image I_{t-1} are addressed as $(x+u,\ y+v)$, where the u and v represent the motion vector. These parameters are used to calculate the Sum of Absolute Difference (SAD) [43].

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

And for the (a, b) the equation is represented as

$$(a,b) = arg\ min_{(u,v)\in z}\ SAD_{(x,y)}\ (u,v) \tag{2}$$

While, the $Z = \{(u,\ v) \mid -B \leq u,\ v \leq B$ and $(x+u,\ y+v)$ indicate to the valid position of pixel in the reference image I_{t-1}, where B is an integer to find for the specific range. Referring to the SAD equation in (1), the global minimum of matching error can be found. The Basic Block Matching has several features such as simple algorithms structure and less computation time in obtaining the disparity depth map. Thus, it has been used by many researchers as initial matching algorithms toward performing matching process and gaining the disparity depth map [44].

B. Cost Aggregation

In the cost aggregation step, the main target is to minimize the matching uncertainties. The information obtained for a single pixel upon computing the matching cost is not sufficient, especially for precise matching, thus the cost aggregation step is required. The step is based on window-based that chooses the average above the cost values of Disparity Space Image (DSI) [45]. Numerous methods have been proposed and developed in order to perform the cost aggregation by applying the fixed window size. Therefore, the fixed square window has been selected through our approach as it has less complexity and uncomplicated strategy [46]–[48]. The window size has to be chosen appropriately since the small window size does not generate suitable or desired results, specifically in the regions with less-texture. In contrast, the large window does not deal with small objects [49]. However, the fixed square window has its own disadvantages such as ignoring the depth discontinuities, low ability to undesirable and inappropriate for on uniform areas, but these limitations can be resolved through disparity computation and refinement parts [50].

C. Disparity Computation/Optimization

In general, stereo matching algorithms rely on two of the most optimizations approaches, which are either the “local or global” approach. In the local approach, the local Winner Takes All (WTA) strategy is applied to select the disparity for each pixel when the final disparities are computed. The disparities are generally correlated to minimize the cost value which will increase the Signal to Noise Ratio (SNR) and reduce the ambiguity [51], [52]. In contrast, for the global optimization approach, particular assumptions are performed for the depth of field of the scene, which are often expressed in the energy minimization framework. Numerous strategies have been developed for the global optimization, and one of the most well-known global methods that have been extensively used in stereo matching algorithms for the energy minimization is the Dynamic Programming (DP) approach. DP has been widely applied as the term of optimization of disparity depth map.

Dynamic Programming (DP) is effectively and independently used for each scan line “row”. The assumption adopted for the dynamic programming is usually of an ordering constraint between the neighboring pixels of the same row. In our method, the Dynamic Programming (DP) is working as an optimizer and noise minimizing approach for the disparity map from the disparity computation with generally fixed square window. This approach is commonly used to optimize energy function for the nondeterministic polynomial-time hard (NP-hard) for smoothing.

There are two well-known approaches of global optimization: the one-dimension (1D) approach, and the two-dimension (2D) approach. The two-dimension (2D) approach is smoothing the stereo images in images in vertical and horizontal directions in order to estimate the disparity map utilizing continuation approach, the simulated annealing, and the mean-field annealing [53]. However, these approaches often are less capable of optimizing the equation, as shown in (1) [54], [55].

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

The term E_{data} *(d)* of data refers to the disparity function through correspondence pixels of disparity map. It determines the effectiveness of the disparity function to ensure that it is appropriate to fit the stereo image pairs in the part of the overall matching cost. E_{smooth} *(d)* data term refers to the conjecture for smoothness implemented from the method, which usually settles the disparity between the pixels on pixel grid [56]–[58]. The one-dimension optimization approach of Dynamic Programming (DP) is chosen to solve and overcome the minimization problem. Besides, Disparity Space Image (DSI) is known as a data structure to illustrate on Dynamic Programming algorithm to solve occlusions and matches correspondingly.

D. Segmentation Approach

The principle of segmentation refers to the process of partitioning an image into several regions and portions. Therefore, the process of creating parts of an image into segments that are simple and conceptually meaningful for further analysis is known as segmentation term. This segmentation is classified into several and specific regions based on the characteristics of the pixels in the image. In the image segmentation, two simple and basic properties are the discontinuity and similarity. The concept of discontinuity refers to the partition of an image depending on the abrupt changes in intensity, including the isolated points, edges, and lines in an image. The concept and approach in similarity are based on partitioning of similar image into regions according to a set of predefined criteria. Based on the stages (2 and 3) of MHMF algorithms as a core of our method, there are six steps involved, which are the segmentation step, first hybrid median filter, first merging step, second hybrid median filtering, second merging step, and third hybrid median filter. The segmentation part of our algorithms is specifically applied to the raw disparity map from the optimization step. The contents of the raw disparity map are extracted according to the pixel colors. All the pixels' colors are based on their values that rely on the disparity range preferred for computation, in which every extracted segment relies on the range of pixel colors of the disparity map.

*E. Hybrid Median Filter (HMF***)**

The remarkable achievement of median filtering from the past until recent years has brought a huge interest and encouraging research avenues. The deterministic properties and detailed information of median has been studied and investigated by numerous developers and scholars abroad [59]–[62]. The median filter remains as one of the most highlighting and popular applied filters in image processing domain due to several features, including robustness and efficient impulsive noise and its suitability to suppress the impulse noise repeatedly. However, there are some limitations of median filters, especially in giving the same level of smoothness in-homogeneous regions and round off corners. Besides, median filters usually erase lines thinner than half of the width for the neighborhood [62]. Subsequently, large progressions have been made in developing versions of median filter such as weighted median, constant time-weighted median and hybrid median filters in the recent years [63], [64].

Hybrid Median Filter (HMF) is a newly derived version of the median filter. It has a low pass filter and an efficient tool with high ability to remove noises of an image. The Hybrid Median Filter (HMF) has been studied, investigated, and widely applied by many researchers specifically in smoothing in the signal and image processing domain. The filter has shown huge capability in smoothing and removing noises, preserving edges "corners" better than the fundamental version median filter [65], [66]. In addition, the Hybrid Median Filter (HMF) include three specific steps and operation ranking, as follows:

1. The (MR), which represents the median of horizontal and vertical R pixels.
2. The (MD) is the median of diagonal D Pixels, where the filtered value is the median of the two median values and the central pixel C.
3. The median [MR, MD, C]

The data from diverse spatial directions are often ranked separately [67]. Figure 2 (a, b) represents an example of HMF, where n= 7 pixels, as shown in Figure 2 (a), while the n = 5 pixels wide neighborhood is represented in Figure 2 (b) matrix. It includes 25 pixels (as shown in the example neighborhood below) and has to be ranked in the traditional method. In contrast, in the hybrid method, each of the two groups contains only 9 pixels, and the final comparison involves only three values. Even with the additional logic and manipulation of values, the hybrid method is faster than the conventional median [67], [68]. Therefore, the Hybrid Median Filter (HMF) has been selected as a significant part of our method (MHMF) algorithms to smooth each segment portion of the raw disparity map from the segmentation algorithms process, where each segment will be filtered up using hybrid median filtering. Then, each filtered segment will proceed to the merging process to combine into a new disparity map.

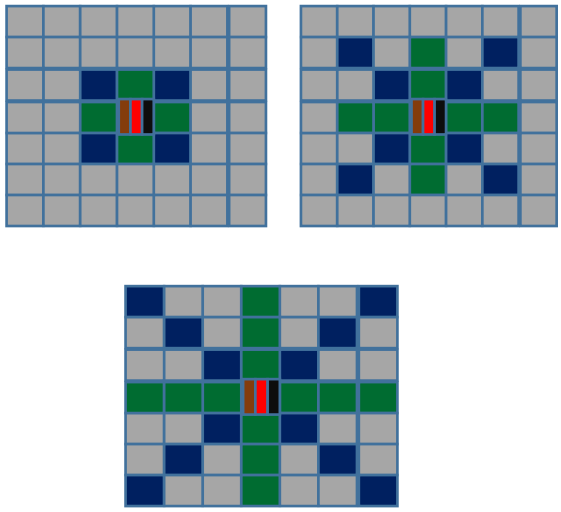


Figure 2: (a) General diagram of neighborhood pixels in hybrid median filter

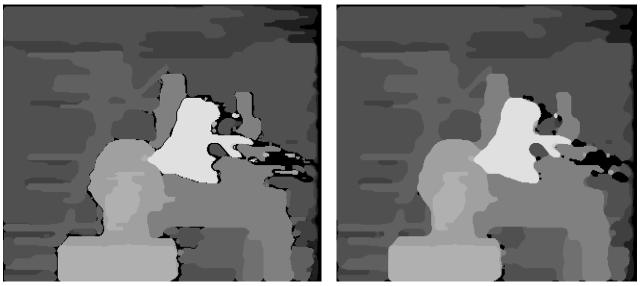
D	*	R	*	D
*	D	R	D	*
R	R	C	R	R
*	D	R	D	*
D	*	R	*	D

Figure 2: (b) HHF with a 5 x 5 box (default value).

F. Merging Algorithm Approach

The merging approach is a specific step applied particularly within the first and second time of hybrid median filtering of our method. All the extracted segments will proceed to the hybrid median filter in order to remove all the streaks and noises found in each segment. Then, after the hybrid median filtering is performed for each extracted segment, all the filtered segment will be merged up to become a disparity map. Based on our developed method MHMF, the first merging is to combine the extracted segments that go through a hybrid median filtering process to become a new disparity map.

Some noises of the extracted segments remained after the first merging step and the second hybrid median filtering. Thus, second merging is done to construct a third hybrid median filtering. Consequently, the final disparity map is obtained with the streaks completely removed and the unwanted aspects such as occlusion, depth discontinuities, texture-less regions are reduced. Figures 3 (a) and (b) represent the result of the merge process and the final disparity depth map of our developed method.



(a) (b)

Figure 3: (a) Tsukuba merged disparity map; (b) Final Tsukuba disparity depth map from MHMF

IV. RESULTS AND DISCUSSION

This section discusses the results of disparity dense map obtained from our developed method, specifically the result from the Basic Block Matching (BBM), Dynamic Programming (DP), and our developed method “Multistage Hybrid Median Filter” (MHMF). All the input images datasets of the stereo matching algorithms are a standard set of images from the Middlebury page “Tsukuba, Venus, Teddy, and Cones”. Three main evaluation functions of stereo matching algorithms have been applied to analyze the performance of the three stereo matching algorithms: The Basic Block Matching (BBM), Dynamic Programming (DP), and our developed method “Multistage Hybrid Median Filter” (MHMF). The standard image stereo dataset used for the evaluation process is Tsukuba dataset since the stereo pair is easy to be evaluated, especially through the unwanted aspects which generated. Besides, the results from Tsukuba dataset is gathered in a small short period due to its simple contents. The first function applied is the Mean Squared Error (MSE), used significantly to obtain the average of squared errors between the generated disparity map and the ground truth.

$$\mathrm{MSE} = \frac{1}{\mathrm{MN}} \sum_{\mathrm{y}=1}^{\mathrm{M}} \sum_{\mathrm{x}=1}^{\mathrm{N}} [\mathrm{I}_1(\mathrm{x},\mathrm{y}) - \mathrm{I}_2(\mathrm{x},\mathrm{y})]^2 \qquad (4)$$

The parameters M and N from Equation (4) represent the rows and columns of the applied images of I_1 and I_2 respectively. Specifically, the decreasing in the values the MSE indicates that the progressive or cumulative squared error is lower [69].

The second applied function is the Peak Signal to Noise Ratio (PSNR), where the function is linked to MSE function by including its function parameters as part of its PSNR’s structure. The PSNR function is the function to construct the performance of the developed algorithm in gaining the better result using a comparison term for the quality of an image utilizing the image smoothing algorithms

$$PSNR = 10\ Log_{10}\left(\frac{R^2}{MSE}\right) \qquad (5)$$

For the second function “Peak Signal to Noise Ratio” (PSNR), the R parameter represents the maximum fluctuation of the data type for the image data input. The increase in the value of the PSNR indicates better quality of the implemented disparity map with less noise [70].

In addition, the Structural Similarity Index Metric (SSIM) is the third evaluation function applied to analyze the performance of our developed algorithm to some other existing algorithms. This function is commonly applied to calculate the similarity between the obtained disparity map and the ground truth. The SSIM is developed to solve the issue on peak signal-to-noise ratio (PSNR) and mean squared error (MSE), which is less consistent with the human eye perception [42]. MSE and PSNR are evaluation methods that approximate the perceived errors, while SSIM estimate on image degradation recognized as the change in structural data Structural data is the spatially near pixels in an image, which belongs to strong inter-dependencies to show up crucial data regarding the structural content in the visual information.

$$SSIM(X,Y) = \frac{(2\mu_x\mu_y + c_1)(2\sigma_{xy} + c_2)}{(\mu_x^2 + \mu_y^2 + c_1)(\sigma_x^2 + \sigma_y^2 + c_2)} \qquad (6)$$

In the Structural Similarity Index Metric SSIM equation, μ_x and μ_y are the average of x of *x* and *y*, σ_x^2 and σ_y^2 are the variance of *x* and *y*, σ_{xy} represents the covariance of and *y* represents 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 the denominator, where *L* represents the dynamic range of pixel values while *k1* = 0.01 represents the dynamic range of pixel values while *k1* = 0.01 and *k2* = 0.03 as default. The results calculated from SSIM index should be in decimal value between -1 to 1.

A. Performance Analysis Results

In this section, the significant results from the Basic Block Matching (BBM), the Dynamic Programming (DP), and our developed method Multistage Hybrid Median Filter (MHMF) are presented and discussed.

In Figure 4, the result values of the MSE are represented. The results of the computed disparity map are generally proportional to the progressively increase with window sizes among three stereo matching algorithms. The Basic Block Matching (BBM) approach with a diamond shape on the line shows that the MSE values for Tsukuba stereo pairs are progressively decreasing. This indicates that as window size increases, more errors are reduced for BBM algorithm. However, there is an inverse relationship between the Dynamic programming (DP) approach and the BBM algorithm as the MSE values for Tsukuba stereo pair that produced by DP algorithm increased with the extension of the window size respectively. Results from DP algorithm indicates that it performs well when using a smaller size of the window rather than the bigger window size due to its scanline optimization on each row of the pixel. The smaller size of the window may scan on the content of image precisely and more errors can be reduced compared to the bigger window size. For our developed method "MHMF", the MSE values are specified in a standard range and almost closer to dynamic programming but has better performance. However, based on our study, a window size of 7 X 7 is chosen for Tsukuba stereo image as the convenient window sizes are able to capture well on tiny and thin objects of the content in an image. However, the selection of window size relies on the complexity of the content of the image.

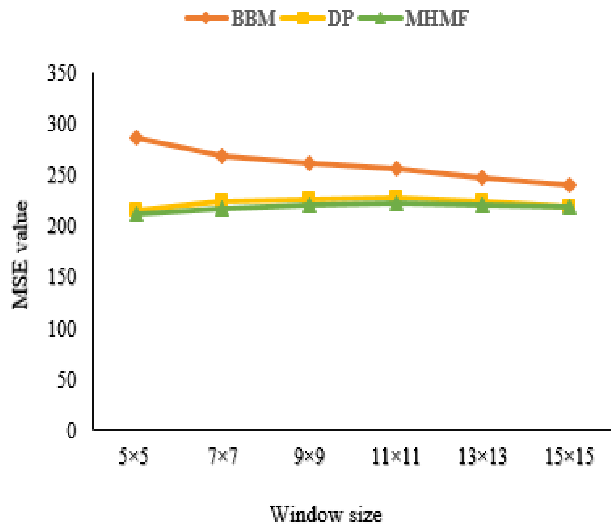


Figure 4: MSE value for Tsukuba stereo image

Based on the obtained results shown in Figure 5, it is clear that the PSNR values computed from BBM algorithm gradually increased with the increase of window sizes. This indicates that noises can be minimized with the increase of window size for BBM algorithm. Besides, for the DP algorithm, the generated PSNR values are gradually decreased and then gradually increased, which means the values obtained from the DP algorithm are not stable. Less noise is removed, which indicates that DP algorithm can only operate efficiently on reducing errors on the image with particularly suitable (small) window sizes compared to BBM algorithm. For our method "MHMF", the window sizes did not much exert effect on the performance of the whole algorithm. The graph shows that the PSNR values obtained from MHMF algorithm remain stable.

The results for the SSIM are based on the obtained results represented in Figure 6. The values from SSIM indicate that the values are gradually increased. This shows that BBM algorithm requires a large window size to obtain a better result of disparity depth map. However, the SSIM values of disparity map for each window size of BBM algorithm are lower than the SSIM values of disparity map for DP algorithm and MHMF algorithm. Besides, for the DP algorithm, the SSIM values obtained remain stable, while the SSIM values for the MHMF algorithm are not much different among the window sizes.

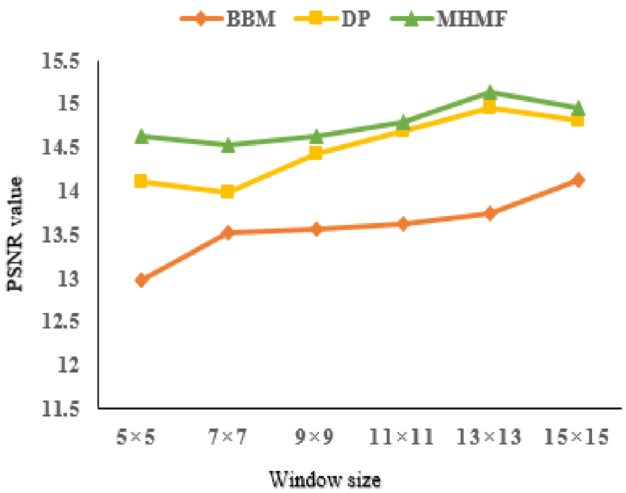


Figure 5: PSNR value for Tsukuba stereo images

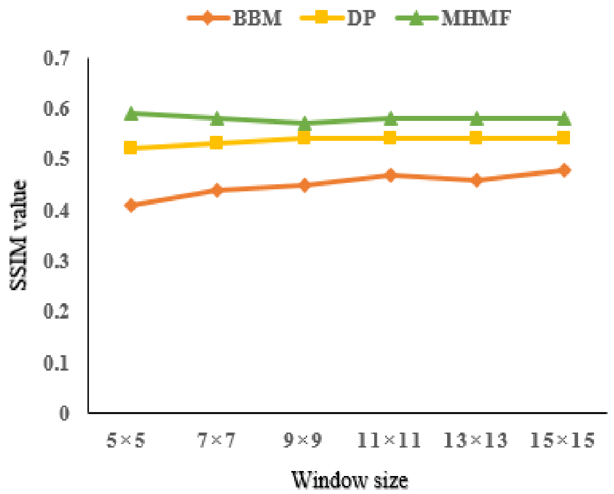


Figure 6: SSIM value for Tsukuba stereo image

B. Computation Analysis Results

The method used to compute the process execution time for all the three evaluated algorithms, the Basic Block Matching (BBM), the Dynamic Programming (DP), and our proposed approach Multistage Hybrid Median Filter (MHMF) is based on the tic toc on the MATLAB platform, which is a method applied to calculate the running time of algorithms. In this research works, the tic toc is mainly applied to compute only the main functions in the algorithm without including the sub-functions such as reading data and showing out figures. The results presented in Figure 7, shows that the computation process of the BBM algorithm is the fastest compared to the other two algorithms due to its simple algorithm. However, the results obtained from this algorithm have too much noise. Besides, while the average computation time taken for DP algorithm is approximate about one minute, whereas the average computation time of the MHMF algorithm is approximately one minute plus as the MHMF algorithm is a combination of BBM and DP algorithms, implying its computation time is a bit longer compared to the rest two algorithms. The difference of time taken for the computation between DP and MHMF algorithms is only a few seconds, although the few seconds are worth for MHMF algorithm to obtain better results of the disparity map compared to BBM and DP algorithms.

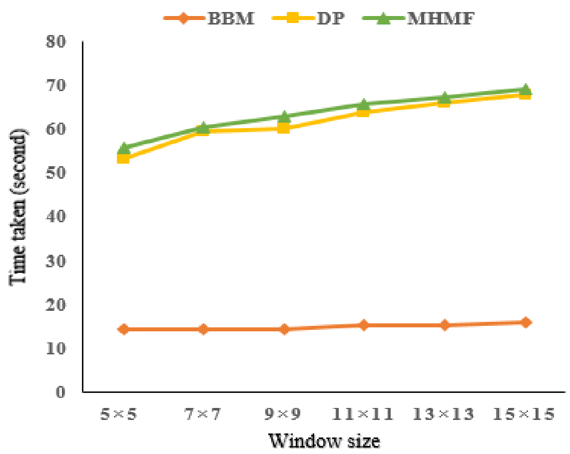


Figure 7: Computation time taken (in seconds)

V. CONCLUSION

In this paper, a new method of stereo matching algorithms upon stereo hybrid filtering is presented. A Multistage Hybrid Median Filter (MHMF) of stereo matching algorithms with high accuracy and less complexity of obtaining disparity depth map has been implemented to be embedded in numerous stereo matching applications. As one of the significant aims of providing valuable and efficient solutions to stereo vision community and support researchers in the same line of research, we contributed and developed a highly efficient and robust algorithm method based on Hybrid Median Filter with high capability in improving the quality of disparity depth map. The structure of the Multistage Hybrid Median Filter (MHMF) involved the Basic Block Matching (BBM), Dynamic Programming (DP) algorithms, Hybrid Median Filtering (MHF), segmentation and merging processes. The obtained results of the disparity depth map used several functions including MSE, PSNR, and SSIM. From the evaluation and observation of the results as well as the previous discussion, our proposed approach MHMF is significant as it can achieve high accuracy of disparity depth map with less algorithm complexity and computational efficiency among other algorithms. According to the graph results of the evaluation functions, it is clear that disparity map obtained by MHMF is closer to the ground truth image rather than the other algorithms. The MHMF algorithms have the ability of remove noises, horizontal 'streaks' of DP algorithm and minimize the depth discontinuities of disparity depth maps. For future research work, the MHMF can be embedded in numerous stereo matching applications.

ACKNOWLEDGEMENT

Authors would like to thank Universiti Teknikal Malaysia Melaka (UTeM) and Ministry of Science, Technology and Innovation (MOSTI) for supporting this research under Science Fund 06-01-14-SF0125 L00026.

REFERENCES

[1] S. Zhu and L. Yan, “Local Stereo Matching Algorithm with Efficient Matching Cost and Adaptive Guided Image Filter,” Vis. Comput., vol. 33.9, pp. 1087–1102, 2017.
[2] G. Hong and B. Kim, “A Local Stereo Matching Algorithm Based on Weighted Guided Image Filtering for Improving the Generation of Depth Range Images,” Displays, vol. 49, pp. 80–87, 2017.
[3] M. Aboali, N. A. Manap, A. M. Darsono, and Z. M. Yusof, “Review on Three-Dimensional (3-D) Acquisition and Range Imaging Techniques,” Int. J. Appl. Eng. Res., vol. 12.10, pp. 2409–2421, 2017.
[4] S. Hussain and R. Modi, “Advancement in Depth Estimation for Stereo Image Pair,” Int. J. Innov. Res. Comput. Commun. Eng. 1.4, 2013.
[5] I. Lee and B. Moon, “An Improved Stereo Matching Algorithm with Robustness to Noise Based on Adaptive Support Weight,” J. Inf. Process. Syst., vol. 13.2, pp. 256–267, 2017.
[6] R. Dieny, J. Thevenon, J. Martinez-del-rincon, and J. Nebel, “Bioinformatics Inspired Algorithm for Stereo Correspondence,” 2011.
[7] D. Scharstein and R. Szeliski, “A Taxonomy and Evaluation of Dense Two-Frame Stereo Correspondence Algorithms,” nternational J. Comput. Vis., vol. 47.1-3, pp. 7–42, 2001.
[8] A. Hosni, M. Bleyer, and M. Gelautz, “Secrets of adaptive support weight techniques for local stereo matching q,” Comput. Vis. Image Underst., vol. 117, no. 6, pp. 620–632, 2013.
[9] C. Rhemann, A. Hosni, M. Bleyer, C. Rother, and M. Gelautz, “Fast Cost-Volume Filtering for Visual Correspondence and Beyond,” IEEE Trans. Pattern Anal. Mach. Intell. 35.2 504-511, 2013.
[10] L. De-maeztu, S. Mattoccia, and R. Cabeza, “Linear stereo matching,” Comput. Vis. (ICCV), 2011 IEEE Int. Conf. on. IEEE, 2011.
[11] N. Komodakis, G. Tziritas, and N. Paragios, “Fast , Approximately Optimal Solutions for Single and Dynamic MRFs,” Comput. Vis. Pattern Recognition, 2007. CVPR’07. IEEE Conf. on. IEEE, 2007.
[12] J. Sun, N. Zheng, and S. Member, “Stereo Matching Using Belief Propagation,” IEEE Trans. pattern Anal. Mach. Intell. 25.7 787-800. 2003, 2003.
[13] B. Potetz, “Efficient Belief Propagation for Vision Using Linear Constraint Nodes,” Comput. Vis. Pattern Recognition, 2007. CVPR’07. IEEE Conf. on. IEEE, 2007.
[14] L. Torresani, V. Kolmogorov, and C. Rother, “A dual decomposition approach to feature correspondence,” IEEE Trans. pattern Anal. Mach. Intell. 35.2 259-271, 2013.
[15] L. Wang, H. Q., Wu, M., Zhang, Y. B., & Zhang, “Effective stereo matching using reliable points based graph cut,” Vis. Commun. Image Process. (VCIP), 2013, IEEE, 2013.
[16] A. F. Bobick and S. S. Intille, “Large Occlusion Stereo,” Int. J. Comput. Vis., vol. 33.3, pp. 181–200, 1999.
[17] D. Chen, M. Ardabilian, L. Chen, and S. Member, “A Fast Trilateral Filter-Based Adaptive Support Weight Method for Stereo Matching,” IEEE Trans. Circuits Syst. Video Technol., vol. 25.5, pp. 730–743, 2015.
[18] B. Khomutenko and P. Martinet, “Direct Fisheye Stereo Correspondence Using Enhanced Unified Camera Model and Semi-Global Matching Algorithm,” Control. Autom. Robot. Vis. (ICARCV), 2016 14th Int. Conf. on. IEEE, 2016.
[19] J. V. C. I. R, R. Affendi, H. Ibrahim, A. Hasni, and A. Hassan, “Stereo matching algorithm based on per pixel difference adjustment , iterative guided filter and graph segmentation q,” J. Vis. Commun. Image Represent., vol. 42, pp. 145–160, 2017.
[20] J. Menant, G. Gautier, M. Pressigout, L. Morin, and J. Nezan, “An Automatized Method to Parameterize Embedded Stereo Matching Algorithms,” J. Syst. Archit., vol. 80, pp. 92–103, 2017.
[21] D. Scharstein and R. Szeliski, “A taxonomy and evaluation of dense two-frame stereo correspondence algorithms,” Int. J. Comput. Vis. 47.1-3 7-42, 2002.
[22] S. Choi, J. Jeong, and D. H. Hwang, “Post-Processing Algorithms for Real-time Active Stereo Vision,” Consum. Electron. (ISCE 2014), 18th IEEE Int. Symp. on. IEEE, 2014.
[23] X. Wang, Y. Tian, H. Wang, and Y. Zhang, “Stereo Matching by Filtering-Based Disparity Propagation,” PLoS One, vol. 11.9, p. e0162939, 2016.
[24] S. Choi, J. Jeong, J. Chang, and H. Shin, “Implementation of Real-Time Post-Processing for High-Quality Stereo Vision,” ETRI J., vol. 37.4, pp. 752–765, 2015.
[25] P. Fua, I. Sophia-antipolis, R. Lucioles, and F.-V. Cedex, “A parallel stereo algorithm that produces dense depth maps and preserves image features,” Mach. Vis. Appl., vol. 6.1, pp. 35–49, 1993.
[26] T. C. Birchfield S, “A pixel dissimilarity measure that is insensitive to image sampling,” IEEE Trans. Pattern Anal. Mach. Intell., vol. 20.4, pp. 401–406, 1998.
[27] M. Camplani, L. Salgado, M. Camplani, and L. Salgado, “Efficient Spatio-Temporal Hole Filling Strategy for Kinect Depth Maps,” Three-Dimensional Image Process. Appl. 8290, 2012.
[28] B. P. Zhou X, “Radiometric invariant stereo matching based on relative gradients,” Image Process. (ICIP), 2012 19th IEEE Int. Conf. (pp. 2989-2992). IEEE, 2012.
[29] C. Lin, C. Varekamp, K. Hinnen, and G. De Haan, “Interactive Disparity Map Post-processing,” Image Process. (ICIP), 2012 19th IEEE Int. Conf. on. IEEE, 2012.
[30] D. Min, J. Oh, and K. Sohn, “Asymmetric post-processing for stereo correspondence,” Pattern Recognition, 2008. ICPR 2008. 19th Int.

Conf. on. IEEE, 2008.

[31] D. Maier and M. Reinhard, "Calculating Dense Disparity Maps from Color Stereo Images , an Efficient Implementation," Int. J. Comput. Vis. 47.1-3 79-88, 2002.

[32] X. Sun, X. Mei, S. Jiao, M. Zhou, and H. Wang, "Stereo Matching with Reliable Disparity Propagation," 3D Imaging, Model. Process. Vis. Transm. (3DIMPVT), 2011 Int. Conf. on. IEEE, 2011.

[33] Q. Yang, "A Non-Local Cost Aggregation Method for Stereo Matching," Comput. Vis. Pattern Recognit. (CVPR), 2012 IEEE Conf. (pp. 1402-1409). IEEE, 2012.

[34] H. Komatsu, "Technical Bulletin Interferometry : Principles and Applications of Two-Beam and Multiple-," p. 26, 1991.

[35] T. C. Huat, A. Manap, and M. Ibrahim, "Development of Double Stage Filter (DSF) for Stereo Matching Algorithms and 3D Vision Applications," J. Telecommun. Electron. Comput. Eng., vol. 7.2, pp. 39–46, 2015.

[36] S. Park, M. Park, and K. Yoon, "Confidence-based Weighted Median Filter for Effective Disparity Map Refinement," Ubiquitous Robot. Ambient Intell. (URAI), 2015 12th Int. Conf. on. IEEE, 2015.

[37] Q. Yang, "Stereo Matching Using Tree Filtering," IEEE Trans. Pattern Anal. Mach. Intell., vol. 37.4, pp. 834–846, 2015.

[38] N. Lazaros, G. C. Sirakoulis, and A. Gasteratos, "Review of Stereo Vision Algorithms : From Software to Hardware," J. Real-Time Image Process. 11.1 5-25, 2008.

[39] K. Yoon and S. Member, "Adaptive Support-Weight Approach for Correspondence Search æ," IEEE Trans. Pattern Anal. Mach. Intell., vol. 28.4, pp. 650–656, 2006.

[40] W. H. . Huat, T.C., bin Abd Manap, N. and Saad, "Analysis on Segment-Based Double Stage Filter Algorithm for Stereo Matching," J. Eng. Appl. Sci., vol. 11, no. 10, pp. 6240–6245, 2016.

[41] N. Cross-correlation, "Robust Stereo Matching Using Adaptive," IEEE Trans. Pattern Anal. Mach. Intell., vol. 33, pp. 807–822, 2011.

[42] Y. Wang, C. Tung, and P. 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, pp. 784–801, 2013.

[43] Y. S. Chen, Y. P. Hung, and C. S. Fuh, "Fast block matching algorithm based on the winner-update strategy," IEEE Trans. Image Process., vol. 10.8, pp. 1212–1222, 2001.

[44] L. Di Stefano, M. Marchionni, S. Mattoccia, G. Neri, and B. Sede, "A Fast Area-Based Stereo Matching Algorithm," Image Vis. Comput., vol. 22.12, pp. 983–1005, 2004.

[45] F. Tombari, L. Di Stefano, S. Mattoccia, A. Mainetti, and D. Arces, "A 3D Reconstruction System Based on Improved Spacetime Stereo," Control Autom. Robot. Vis. (ICARCV), 2010 11th Int. Conf. on. IEEE, 2010.

[46] M. Debella-gilo and A. Kääb, "Locally adaptive template sizes for matching repeat images of Earth surface mass movements," ISPRS J. Photogramm. Remote Sens., vol. 69, pp. 10–28, 2012.

[47] S. Lee, J. Sim, C. Kim, S. Member, and S. Lee, "Correspondence Matching of Multi-View Video Sequences Using Mutual Information Based Similarity Measure," IEEE Trans. Multimed., vol. 15.8, pp. 1719–1731, 2013.

[48] X. Mi, "Stereo Matching based on Global Edge Constraint and Variable Window Propagation," Image Signal Process. (CISP), 2012 5th Int. Congr. on. IEEE, 2012.

[49] S. J. Manap NA, "Disparity refinement based on depth image layers separation for stereo matching algorithms," J. Telecommun. Electron. Comput. Eng., vol. 4(1), pp. 51–64, 2012.

[50] J. Zhao and J. Katupitiya, "A multi-window stereo vision algorithm with improved performance at object borders," Comput. Intell. Image Signal Process. 2007. CIISP 2007. IEEE Symp. on. IEEE, 2007.

[51] A. A. Cigla, C., & Alatan, "Information permeability for stereo matching," Signal Process. Image Commun., vol. 28.9, pp. 1072–1088, 2013.

[52] K. Zhang, S. Member, J. Lu, and Q. Yang, "Real-Time and Accurate Stereo : A Scalable Approach with Bitwise Fast Voting on CUDA," IEEE Trans. Circuits Syst. Video Technol., vol. 21.7, pp. 867–878, 2011.

[53] D. G. and F. Girosi., "Parallel and deterministic al- gorithms from MRF's: Surface reconstruction.," IEEE Trans. Pattern Anal. Mach. Intell., vol. 13.5, pp. 401–412, 1991.

[54] L. L. Zhu, Y. Chen, Y. Lin, C. Lin, and A. Yuille, "Recursive Segmentation and Recognition Templates for 2D Parsing," Adv. Neural Inf. Process. Syst., 2009.

[55] I. Gerace, V. Vanvitelli, R. Pandolfi, and V. Vanvitelli, "A Color Image Restoration with Adjacent Parallel Lines Inhibition," Image Anal. Process. 2003. Proceedings. 12th Int. Conf. on. IEEE, 2003.

[56] R. Gouiaa and J. Meunier, "3D reconstruction by fusioning shadow and silhouette information," Comput. Robot Vis. (CRV), 2014 Can. Conf. on. IEEE, 2014.

[57] D. Neilson and Y. H. Yang, "A component-wise analysis of constructible match cost functions for global stereopsis," IEEE Trans. Pattern Anal. Mach. Intell., vol. 33.11, pp. 2147–2159, 2011.

[58] F. Tombari, S. Mattoccia, L. Di Stefano, and E. Addimanda, "Classification and evaluation of cost aggregation methods for stereo correspondence," Comput. Vis. Pattern Recognition, 2008. CVPR 2008. IEEE Conf. on. IEEE, 2008.

[59] L. A. Aranda, P. Reviriego, and J. A. Maestro, "Error Detection Technique for a Median Filter," IEEE Trans. Nucl. Sci., 2017.

[60] R. Kunsoth, "Modified Decision Based Median Filter for Impulse Noise Removal," Wirel. Commun. Signal Process. Netw. (WiSPNET), Int. Conf. on. IEEE, 2016.

[61] B. Xiong and Z. Yin, "A Universal Denoising Framework With a New Impulse Detector and Nonlocal Means," EEE Trans. Image Process., vol. 21, pp. 1663–1675, 2012.

[62] C. Lin, J. Tsai, and C. Chiu, "Switching Bilateral Filter With a Texture / Noise Detector for Universal Noise Removal," IEEE Trans. Image Process., vol. 19.9, pp. 2307–2320, 2010.

[63] N. C. Gabbouj, M., Coyle, E. J., & Gallagher Jr, "An overview of median and stack filtering," Circuits, Syst. Signal Process., vol. 11.1, pp. 7–45, 1992.

[64] S. Arastehfar, A. A. Pouyan, and A. Jalalian, "An enhanced median filter for removing noise from MR images," J. AI Data Min., vol. 1.1, pp. 13–17, 2013.

[65] M. R. Rakesh, B. Ajeya, and A. R. Mohan, "Hybrid Median Filter for Impulse Noise Removal of an Image in Image Restoration," Int. J. Adv. Res. Electr. Electron. Instrum. Eng., vol. 2.10, 2013.

[66] Q. Yang, "A Hybrid Median Filter for Enhancing Dim Small Point Targets and Its Fast Implementation," Multimed. Signal Process. (CMSP), 2011 Int. Conf. IEEE, vol. 1, pp. 239–242, 2011.

[67] D. M. Bappy and I. Jeon, "Combination of hybrid median filter and total variation minimisation for medical X-ray image restoration," IET Image Process., vol. 10.4, pp. 261–271, 2015.

[68] Z. A. Mustafa, B. A. Abrahim, and Y. M. Kadah, "Modified Hybrid Median Filter for Image Denoising," Radio Sci. Conf. (NRSC), 2012 29th Natl. IEEE, 2012.

[69] A. B. Panchal, C. S., & Upadhyay, "Depth Estimation Analysis Using Sum of Absolute Difference Algorithm," Int. J. Adv. Res. Electr. Electron. Instrum. Eng., vol. 3.1, 2015.

[70] M. Khaparde, A., Naik, A., Deshpande, M., Khar, S., Pandhari, K., & Shewale, "Performance Analysis of Stereo Matching Using Segmentation Based Disparity Map," ICDT 2013 Eighth Int. Conf. Digit. Telecommun., 2013.