Performance Analysis between Basic Block Matching and Dynamic Programming of Stereo Matching Algorithm

Maged Aboali, Nurulfajar Abd Manap, Abd Majid Darsono, Zulkalnain Mohd Yusof

Machine Learning and Signal Processing (MLSP) Research Group, Center of Telecommunication Research & Innovation (CETRI), Fakulti Kejuruteraan Elektronik Dan Kejuruteraan Komputer (FKEKK), Universiti Teknikal Malaysia Melaka (UTeM), Malaysia.

nurulfajar@utem.edu.my

***Abstract*— One of the most important key steps of stereo vision algorithms is the disparity map implementation, where it generally utilized to decorrelate data and recover 3D scene framework of stereo image pairs. However, less accuracy of attaining the disparity map is one of the challenging problems on stereo vision approach. Thus, various methods of stereo matching algorithms have been developed and widely investigated for implementing the disparity map of stereo image pairs including the Dynamic Programming (DP) and the Basic Block Matching (BBM) methods. This paper mainly presents an evaluation between the Dynamic Programming (DP) and the Basic Block Matching (BBM) methods of stereo matching algorithms in term of disparity map accuracy, noise enhancement, and smoothness. Where the Basic Block Matching (BBM) is using the Sum of Absolute Difference (SAD) method in this research as a basic algorithm to determine the correspondence points between the target and reference images. In contrast, Dynamic Programming (DP) has been used as a global optimization approach. Besides, there will be a performance analysis including graphs results from both methods presented in this paper, which can show that both methods can be used on many stereo vision applications.**

***Index Terms*— Basic Block Matching (BBM) algorithms; disparity map accuracy Dynamic Programming (DP); Performance analysis.**

I. INTRODUCTION

One of the challenging problems of computer vision community is the stereo matching algorithm of stereo vision. It's a long-standing and attraction issue through numerous researchers and groups in stereo vision field. In general, a stereo matching principle is to create the disparity depth map of two multiple images of the same scene which captures from slightly different viewpoints. The importance of the disparity map came from the ability of presents and provides geometrical information and details of objects in the captured scene and estimated through the process of stereo computation utilizing a pair of images. However, and in spite of a huge amount of stereo matching algorithms for implementing the disparity map in last decades, the computation processes of accurate disparity remain a challenging task.

Besides, the key point behind the interest of stereo matching is certainly apparent in stereo vision, since it involves in general with wide range of applications of image processing and particularly in (e.g. 3D scene reconstruction, robotic vision, image-based rendering, 3D virtual reality, mapping and simultaneous localization, and more) [1]–[3]. Thus, the stereo matching algorithm continues to be one of the most active and heavily investigated areas of research in stereo vision. It concerns with computing the disparity by searches of correspondence pixel pairs of stereo vision images, where both pixels are originated from the camera of the same object view in the three- dimensional 3D world [4]. Furthermore, the pixel correspondence problems were one real challenge and drawback of stereo matching. Thus, a vast amount of stereo systems, techniques, and algorithms with diverse principles have been extensively researched, and proposed by many scholars from different countries abroad [5], [6].

Moreover, decreasing the complexity and costs of the matching process were other attracted terms, since they allow to gain better research results and efficiently improve the practicability of the processes for the stereo matching algorithm [7], [8]. In addition, the recent advances and massive progress in developing and enhancing the disparity depth map have brought different methods of stereo matching algorithms, which differ in their performance in term of speed and accuracy as important conditions in gaining precise output result of stereo images in computer vision. These methods can be searched in stereo datasets such as Middlebury page, which is a global sharing website that provides a standard evaluation for many developed methods and benchmark datasets of stereo images for researchers worldwide [9].

Deriving the depth information and gaining the objects details from the capturing images with low cost and less complexity are main objectives of stereo matching approach. Where these details and information are integral for numerous of two-dimensional 2D and three-dimensional 3D applications. However, the low accuracy of stereo matching algorithms is particularly affected in the quality of disparity depth map that influences the performance of desired applications [10]. Thus, many stereo matching algorithms are developed to overcome the previous issue, where most of these algorithms have been surveyed properly by Scharstein and Szeliski [9].

This research paper particularly deals with Dynamic Programming (DP) and Basic Block Matching (BBM) algorithms of stereo image pairs in term of accuracy and speed of implementing the disparity map. The paper is a performance analysis between both techniques, and presenting the accuracy of both methods over one another. Besides, the paper provides more explanation and discussion in details for both methods presented in this paper supported with graphs results. The clear discussion and comparison can show the proposing of dynamic programming method to be used on many applications on stereo vision as part of image processing area. Both methods have been proceeding using MATLAB software platform, while the images utilized through this paper divide into two groups. The first group is a standard set of images from the Middlebury website page, and the second group is set of images that captured using the (mv Bluefox) camera.

II. Matching Basic Concept and Steps Algorithms

The basic concept of stereo matching algorithms originated particularly from the stereo vision systems with the partition of the left and right side, that act as source input data for the aim of stereo matching.

In general, stereo vision algorithms have four standard steps including (1) matching cost computation, where the matching costs for assigning diverse disparity hypotheses to the different pixels are calculated (2) cost aggregation, by aggregating initial matching costs spatially over support regions, (3) disparity computation and optimization, where the best or unique disparity hypothesis for each pixel is computed, thus global or local cost function is minimized , and lastly (4) disparity refinement, in which the created disparity map is post-processed to eliminate the mismatches or to perform sub-pixel disparity estimates [9].

On another hand, these steps are not necessary to be applied in total on developing stereo matching algorithms, where each step can be implemented in different ways and it depends on the focused task and the design of algorithm by the researchers which can bring different effects on the output result. Besides, the stereo matching process includes several steps after reading the stereo pairs as an input data, where the process starts with matching cost as an initial step to compute the correspondence pixels between the reference and a target image of stereo pairs. Besides, there are two different types of matching cost: the Area-based cost of matching [11], and the Pixel-based cost of matching [12].

Then, matching cost will be summed on the cost aggregation step using multiple types of windows with specific and constant disparity map. Furthermore, the step of cost aggregation works on particular requirements including, user-specified orientation window, automatically detected window, and the pixels inside windows [13]. Subsequently, the optimization step will look for desirable disparity assignment like the preferable area with in the disparity space image that can decrease the cost function on stereo image pairs.

Eventually, the disparity refinement addressed as the last step, which necessary as a key step to remove the mismatches or increase the resolution due to the occlusion [14]. All these particular steps in the flow chart of Figure 1 are illustrating the stereo matching algorithm processes as key steps. However, for the developing of stereo matching algorithms, not all these steps are necessary to be included, it depends on the implementation of the desired system or the specific task requirements.

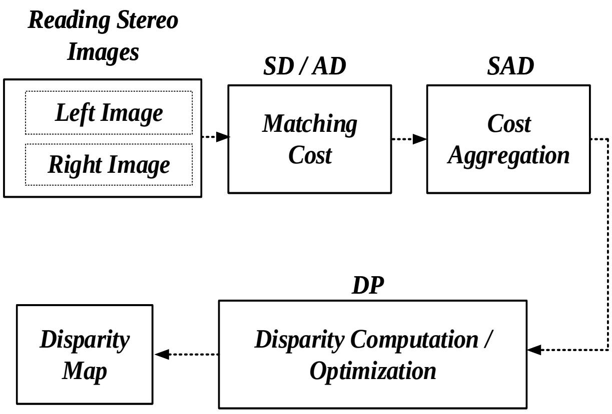


Figure 1: Basic concept and steps of stereo matching algorithm

The steps of stereo matching algorithms from Figure1 can be further detailed as follows:

A. The Matching Cost

For the aim of defining the similarity of the reference pixel and candidate pixel matching of stereo multiple images, the matching cost function is required in order to compute for all the locations or positions of the right pixel and left pixel. Figure 2 gives an overview of pixel-based matching, where for the pixel of the left image is represented as P^l and the right image pixel represented as P^r, and both are referring to the matching pixel intensities in the left and right image planes respectively, for the same scene at point P. Besides i^l and j^l are the coordinate positions of the pixel P^l. While i^r and j^r are the coordinate positions of the pixel P^r. Moreover, a great progress has paid in this approach, where various methods have been applied frequently for the pixel-based matching process including Square Difference (SD), the Absolute Difference (AD) [15], and the Truncated Absolute Differences (TAD) [16].

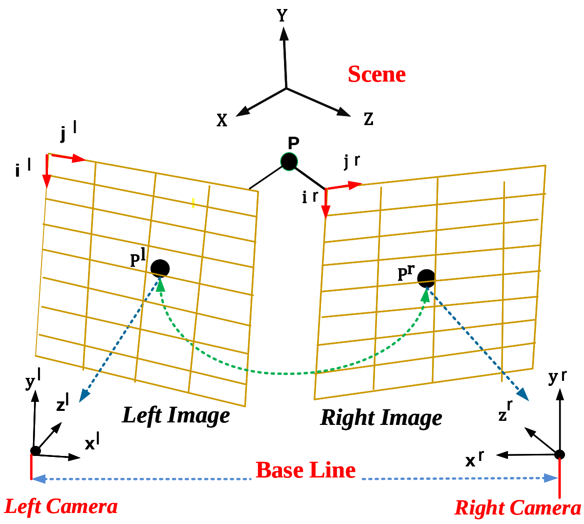


Figure 2: Overview of pixel-based matching process and Epipolar geometry

Many research theories and equations with various mathematical and computational properties have been applied for computing the pixel-based matching cost and finding the best match with using matching criteria. However, the most typical include the Sum of Absolute Difference (SAD), the Sum of Squared Difference (SSD), the Normalized Cross Correlation (NCC), the Rank Transforms (RT), and Census Transforms (CT) [16].

In addition, based on the previous methods, several specific systems and structures algorithms for the matching cost of stereo matching algorithms are investigated and proposed, some of them are new concepts, whereas others are inspired from the previous researches and works. These techniques along with their features, issues, and limitations have been surveyed properly by many researchers within stereo algorithms such as in [17], [18].

However, most of these techniques are using the Absolute Difference (SAD), the Sum of Squared Difference (SSD), the Normalized Cross Correlation (NCC) as typical similarity measures. Where the (SAD) method is computationally fast and the algorithms are easy to be developed, which makes template process even faster. Besides, the correspondence is accomplished by selecting the windows of the required dimension in cost matrix and adding the difference between all elements over the entire windows [9], [19].

The (SAD) method has less time consuming, thus many applications are using the (SAD) method to obtain the best match. However, this technique has its own restrictions, where the critical matches used only for the reference image while other points of stereo pairs possibly are matched with multiple points. The technique also does not perform well for images with high texture [9]. In contrast, in Sum of Squared Differences (SSD), the differences are squared and aggregated within the square window.

Hence, the measure includes a higher computational complexity in comparing to the SAD algorithm method as it has numerous multiplication. Furthermore, for the Normalized Cross Correlation (NCC) method of the cost aggregation step, a window of desirable size is obtained and proceed over the cost matrix or the entire image. Thus, correspondence is determined by dividing the normalized summation of the product of intensities over the entire window. Besides, the NCC measure includes a higher computational complexity in comparing to SAD and SSD algorithms since it has various operations such as division, square root, and multiplication. According to [20], [21] the equations of the area-based matching cost functions are given as follows: For the Sum Absolute Difference (SAD)

$$SAD = \sum_{(i,j)\in W} |I_1(i,j) - I_2(x+i, y+j)| \tag{1}$$

For Sum Square Difference (SSD)

$$SAD = \sum_{(i,j)\in W} |I_1(i,j) - I_2(x+i, y+j)| \tag{2}$$

While for the Normalized Cross Correlation (NCC)

$$NCC = \frac{\sum_{(i,j)\in W} I_1(i,j)\ x\ I_2(x+i, y+j)}{\sqrt[2]{\sum_{(i,j)\in W} I_1^2(i,j)\ x\ \sum_{(i,j)\in W} I_2^2(x+i, y+j)}} \tag{3}$$

Where I_1 refers to the reference image, while I_2 indicates to the target image, W indicates the square window for aggregation.

Table 1 provided a comparison between most popular and typical similarity measures SAD, SSD, and NCC. However, and according to the discussion previously and by referring to Table 1, the SAD provides multiple features with less limitations among all other presented methods. While it considers as an exemplary similarity measure over another methods to provides a convenient match.

Table 1
Comparison between most popular and typical similarity measures [18], [22], [23]

Method	Advantages	Disadvantages
Sum of Absolute Differences (SAD)	- High speed - Achieve reasonable quality - Less complex algorithms	- Sensitive to outliers - Does not work well for images with high texture
Sum Square Difference (SSD)	- Fast & Algorithms are easy. - Can be used for gray-level image applications	- Sensitive to outliers - Not accurate - Does not provide good results under adverse environment
Normalized Cross Correlation (NCC)	- More robust under illumination changes - Widely used in object recognition	- Computationally slow - Tends to blur depth discontinuities - Complex algorithms

B. Disparity Computation and Optimization

Further improvement for the disparity estimation quality has found more interest by the researchers and developers, where they paid more attention in computation/optimization part and presented some hopeful and different methods. These methods are classified into two main classes local method and global. While in the local method [24]–[27], the main concern is on the matching cost computation and the step of aggregation cost.

It utilizes the Winner Takes All (WTA) by picking or selecting each pixel, where the disparity is correlated to the minimum cost value, so as to increase the Signal to Noise Ratio (SNR) in order to reduce the ambiguity, such as those implemented by Cigla and Alatan [28], Zhang et al [29]. Based on their findings and outcomes, the disparity maps gained through this stage still contain errors especially in the form of undesired pixels and occluded regions.

However, there are addressed limitations upon this method as its impose only on matches of the reference image, but for the rest of pixels of the target image for stereo pairs, it may match to multiple pixels. Besides, in local methods and due to due to aggregation is performed through summation or averaging over support regions, their accuracy is sensitive to noise and unclear regions. In contrast, the global method is a framework to look for the disparity *d* that minimize the global energy or energy function over the disparity computation phase such as pixel-based matching cost be selecting the desired surface within the Disparity Space Image (DSI). Within global methods, certain assumptions are made regarding the depth of field for the scene that is often expressed or presented in an energy minimization framework. The huge effort in the global approach is often expended through the disparity computation stage, thus the aggregation part is usually skipped [9].

In addition, for the global methods, a few strategies have been proposed and generated such as implemented global approach using a Graph Cut (GC) algorithm to optimize the energy function [30]. The Belief Propagation (BP) [31]. Another well-known global technique that is extensively applied with a stereo matching algorithm for energy minimization is a Dynamic Programming (DP) approach. The DP is executed for each scan line (row) independently and effectively, where assumption adopted (DP) is that of an ordering constraint between neighboring pixels of the same row. Thus, the Dynamic Programming (DP) technique has been selected as an optimization part of for the stereo matching algorithm through this research paper.

While by referring to the equation (4) the data term E_{data} *(d)* refers to the disparity function, which finding out how effectively the disparity function is appropriate in fitting the stereo image pairs in the part of the overall matching cost. In addition, E_{smooth} *(d)* is representing the conjecture for smoothness implemented from the method [24], [32]–[34].

The equation (4) is representing for both E_{data} *and* E_{smooth} as following:

$$E\,(d) = E_{data}\,(d) + E_{smooth}\,(d) \tag{4}$$

C. Refinement of Disparity

The disparity map estimates of stereo correspondence algorithms are implemented during the step of disparity computation or optimization in some discretized space. However, usually these disparities present with unwanted occlusions and regions and undesirable aspects such as noises, which need to be corrected and identified. Thus, many of stereo algorithms have been created and improved to gain better disparity maps [35]. One of these stereo correspondence algorithms is by utilizing the sub-pixel interpolation, where it developed to interpose the cost of matching with the parabola function. Sub-pixel computation consists of many steps to perform including adding a curve to the costs of matching in the discrete disparity stages to smooth the resolution of the output gained from the stereo matching algorithms and by using the iterative gradient descent [9], [36].

III. MAIN ALGORITHM STRUCTURE AND OUTLINES

Among the existing methods of stereo matching algorithms and based on the details and discussion presented in the previous section, Dynamic Programming (DP) and Basic Block Matching (BBM) are the methods that have been experimented through this research to gain the disparity depth map of stereo image pairs. And these methods can be further discussed in the next two parts

A. Basic Block Matching

The Basic Block Matching (BBM) method is utilized to determine the correspondence pixels' points between the reference image and the target image of stereo image pairs. While through this experimental research the Basic Block Matching (BBM) method is using the Sum of Absolute Differences (SAD) as a basic algorithm to perform the corresponding process for both image groups standard set, where the first group are taken from Middlebury, while the second group is captured by (mv Bluefox) camera.

It is a common method for determining the correspondence on stereo matching algorithms and along the experimental, the pixel point value of the target image is mainly predicted as the corresponding pixel in the reference image of stereo pairs, while the displacement of the corresponding pixels or as motion vector to be computed or estimated utilizing the block matching. The block matching is mainly utilized to minimize the matching errors of the block at the point or position of $(x,\, y)$ of the target image, I_t while for the point or position of the reference image, I_{t-1} which will have addressed as $(x{+}u,\, y{+}v)$ where u and v are the motion vectors. Hence, these variable defined can be summed up as the Sum of Absolute Difference (SAD) [37]. However, the p refers to the block size, as (p x p) and in order to minimize the $SAD_{(x,y)}$ $(u,\, v)$, the (a, b) is known or defined as the motion vector estimation for comparing and determining the SAD for each position, $(x{+}u,\, y{+}v)$ of the experimented datasets

$$SAD_{(x,y)}(u,v) = \sum_{j=0}^{p-1}\sum_{i=0}^{p-1}\left|I_t(x+i, y+j) - I_{t-1}(x+u + i, y+v+j)\right| \tag{5}$$

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

$$(a,b) = arg\ min_{(u,v)\in z}\, SAD_{(x,y)}\,(u,v) \tag{6}$$

Where Z = {$(u,\ v)$ | -B ≤ $u,\ v$ ≤ B and $(x{+}u,\ y{+}v)$ are representing the valid or preferable position of pixel in the reference image, I_{t-1} while B is an integer to find or search for the range. And by referring to the SAD equation (1), the global minimum of matching error can be determined.

B. Dynamic programming

For the disparity optimization step, the dynamic programming algorithms have been selected as global optimization algorithm during this paper research as this algorithm optimize energy function to be NP-hard for the aim of smoothness and enhancement. The global optimization can be classified into two types including one-dimension and two-dimensions optimization categories. Where for the optimization of one-dimension it focusses on the pixel that based on other pixels on the same scanlines, but independent on the disparity that focuses on other scanlines.

However, the one-dimension considers as a traditional technique of optimization and it is not truly global optimization, where the smoothness of this method focuses only on horizontal direction. In contrast, the optimization of two-dimension is a more effective method, since it smoothing the stereo images in the vertical and horizontal directions to estimate the disparity map using continuation method, simulated annealing, and mean-field annealing [38]. But these methods not quite enough for optimizing the equation shown on (4). Moreover, there are two techniques or methods that appropriate or compatible in the optimizing of the equation in (4). The first method is the belief propagation and the second is the graph-cuts as both methods have the ability to gain better results appropriately to ground truth data of stereo matching algorithms [39], [40].

During this research, during experimental results, the dynamic programming on tree is utilized since its more effective and efficient as a one-dimension optimization. Where, the tree graph for the DP can be indicated and represented as $T\,(V, E)$ where E refers to the edges, while V refers to vertices. Where the desired efficiency of the dynamic programming on tree derives or starts with its optimization on the energy function

$$E(d) = \sum_{a \in V} m(d_a) + \lambda \sum_{(a,b)\in E} S(d_a, d_b) \qquad (7)$$

Where a refers to the pixel in the left image of stereo and the Where a refers to the pixel in the left image of stereo and the d_a represents the value of disparity map, d at the pixel of a. By assuming the part $m\ (d_a)$ is the penalty of matching for relating the d_a to the pixel of a, which consider the absolute difference between the pixel. For the a in the left stereo image and a pixel that shift on the right stereo image can represented as $m\ (d_a)\ a \in V$. While, by assuming the part $s\ (d_a, d_b)$ as the penalty of smoothing for the disparity of the d_a and d_b to the pixel p and q, where the variables can be represented as S $(d_a, d_b\ \ (a, b) \in E)$. Thus, in order to gain minimum energy from the equation presented on (7), put the h as the root vertex of tree, thus it present as $h \in V$ and by assuming z$\in V$ as the number of edges for the root of distance between h and z. However, for all the node of z which belongs to a origin as p(z), while the depth is in equal to the depth of z-1. But in term if it is not a root, the energy minimum value of the equation represents in (7) have a sub-tree rooted at the edge in the edge and, z, and the p(z) which represented as $dp(z)$ [41], [42]

$$E_z\left(d_{z(z)}\right) = \min_{d_z \in D} \left(m(d_v) + s\left(d_z, d_{a(z)}\right) + \sum_{W \in Cz} E_w(d_z) \right) \qquad (8)$$

Furthermore, the Cz represent the children set of z and for the optimal disparity of the root node, h can be gained as shown by equation (9)

$$L_h = \arg \min_{d_h \in D} (m(d_h) + \sum_{W \in C_h} E_w(d_z)) \qquad (9)$$

But if the z term is a node without children, thus the Cz is empty and both functions l_z and E_z can be evaluated instantly. Let put J as the maximum depth in the tree, thus the energy function of equation (7) is then optimized by evaluating both functions E_z and l_z for every node z at the specific depth J. However, after evaluation on the functions, begin with the evaluating on the same functions for all the nodes with the specific depth of J-1 because of any child w has the depth of J, this is generally the evaluation on E_w and L_w. Then, the next step is to continue evaluating the both functions L_z and E_z in minimizing the order for the depth until it come or reach to the root for the disparity computation purpose.

IV. Cameras Configuration and Disparity Implementations

A. *Cameras Configuration and Image Capturing*

In this particular section, a stereo vision system was successfully built based on stereo vision principle as shown in Figure 3. The system is generally consisting of two cameras 'left and right' related to each other by a horizontal distance is known as (baseline). While, the camera used in this research is (mv Bluefox), which applied to capture our own stereo datasets. Table 2 illustrates the camera particular specifications.

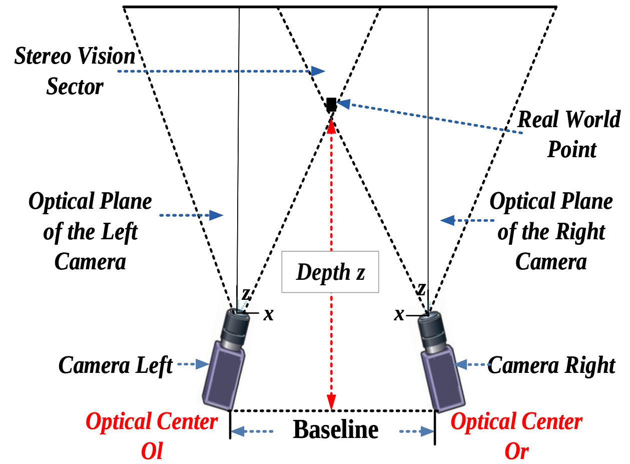


Figure 3: Stereo vision system camera configuration

Table 2
(mv Bluefox) camera Specifications

Camera Specifications	
Interface	Compact industrial camera series with USB 2.0
Type	The mv Bluefox is a compact industrial CCD & CMOS camera
Resolution	0.3M (640 x 480)
Sensor	CMOS sensor -200w with 110 dB high dynamic range (HDR)
Driver	The driver in combination with FPGA to reduces the PC load to a minimum
Memory	8 M pixels' memory
FPS	60

B. *Disparity Implementation*

In this part, the concept and steps of stereo matching algorithms are illustrated. The depth map generated from the stereo matching algorithms, by using the Basic Block Matching (BBM) and the Dynamic Programming (DP) is relying on input stereo pair images taken with (mv Bluefox) camera. Where after reading the stereo image pairs as input datasets, the process starts with matching cost as an initial step to compute the correspondence pixels between reference and target image pair. Then, matching cost will be summed on the cost aggregation step using multiple types of windows with specific and constant disparity map.

Subsequently, the optimization step will look for desirable disparity assignment like the preferable area with in the disparity space image that can decrease the cost function on stereo image pairs. Eventually, the disparity refinement addressed as the last step, which necessary as a key step to remove the mismatches or increase the resolution due to the occlusion. All these steps are represented in Figure 1 from the initial step of reading the stereo pairs to the implementation of the disparity depth map.

C. Disparity evaluation methods

Through this paper research, two types of evaluation approaches have been used:

a. Objective Evaluation

The objective evaluation approach has been performed with applying two of the evaluation functions to evaluate the performance of all results obtained from both stereo matching algorithms Dynamic Programming (DP) and Basic Block Matching (BBM). However, the dataset utilized for the evaluation process is Tsukuba, since this stereo dataset is easy to be analyzed because the simplicity of its contents and the data results are gathered in a short period of time. Where the first evaluation function is defined as Mean Squared Error (MSE) which used to determine the average of squared errors between the obtained disparity map and the original ground truth.

$$MSE = \frac{1}{MN}\sum_{y=1}^{M}\sum_{x=1}^{N}[I_1(x,y) - I_2(x,y)]^2 \qquad (10)$$

Where the M and N parameters are referring to the rows and columns of the applied images of $I1$ and $I2$ respectively. Where the decreasing in the of MSE value indicates that the progressive or cumulative squared error is lower [23]. In other hand, the second evaluation function is the Peak Signal to Noise Ratio PSNR, the function is related MSE function by containing the function parameters as part of its structure. The function constructs the performance of developed algorithm in gaining the better result using a comparison term for the quality of an image utilizing the image smoothing algorithms.

$$PSNR = 10\, Log_{10}\left(\frac{R^2}{MSE}\right) \qquad (11)$$

For the PSNR function, R parameter is referring to the maximum fluctuation of the data type of the image input data. While the increase or the higher value of the PSNR indicate the better quality of the implemented disparity map with less noise [43].

b. Subjective Evaluation

For the subjective evaluation approach, the evaluation is performed for our own stereo pair the datasets that have been captured using the (mv Blue FOX) camera. The results obtained can only be evaluated subjectively by human's eyes observation on the disparity depth maps.

V. RESULT IMPLEMENTATION AND EVALUATION

This section presents and explains the final result of the disparity map, where during this experimental research two stereo matching algorithms have been experimented which are the Basic Block Matching (BBM) and Dynamic Programming (DP). The paper evaluates the disparity output result implemented by both methods as well as compares the smoothness of disparity map obtained by the BBM and DP over one another. However, the stereo image pairs used through this paper are divided into two parts. The first part is a standard set of images from the Middlebury page, and the second part is set of images that captured using the (mv Bluefox) camera. Thus, for clear evaluation in more details of the disparity map obtained by using both methods and by applying the two parts of image datasets, the result can be divided into the following sections:

A. Results of Disparity Depth Based on Middlebury Benchmark Datasets Objective Evaluation

This particular section presents and shows the results of disparity depth maps gaining from the stereo matching algorithms, using the Basic Block Matching (BBM) and the Dynamic Programming (DP). However, the input stereo pair images applied for the implementing on the stereo matching algorithms are taken from the Middlebury benchmark datasets including Teddy, Venus, Tsukuba, Cones. Where Figure 4 shows the output result of the disparity map for each image dataset using both methods.

From Figure 4, and based on the observation on the results from stereo matching algorithms Basic Block Matching (BBM) as well as Dynamic Programming (DP) the accuracy that relies on smoothness is differ between both methods, where the result from the DP algorithm is more smoothness in compare to the basic block matching method which includes some noise in the implemented disparity map.

However, the effectiveness out of DP algorithm in enhancing or smoothing the disparity depth map is also relying on the appropriate disparity range DR of the stereo image pairs applied.

In addition, Table 3 illustrates the portable disparity range for each dataset of the stereo image applied during this research for the image taken from the Middlebury datasets. Where the disparity ranges are gained from the experimental by using stereo matching algorithms and perceiving on the output accuracy for all disparity ranges applied on for each dataset coding.

Besides, from the observation of the results obtained in Table 4 it clear that the higher of disparity range for stereo images datasets, the longer time is taken for running on the stereo matching algorithms. Furthermore, Table 4 presented the result of the time taken for each stereo matching algorithm per second for the various sets of stereo images. Where for the dynamic programming, the computation efficiency is the low, while the computation efficiency for the basic block matching is high.

Thus, we conclude that Dynamic Programming (DP) provides more robust and accurate result in compare the Basic Block Matching (BBM), while the computational process of DP is certainly slow.

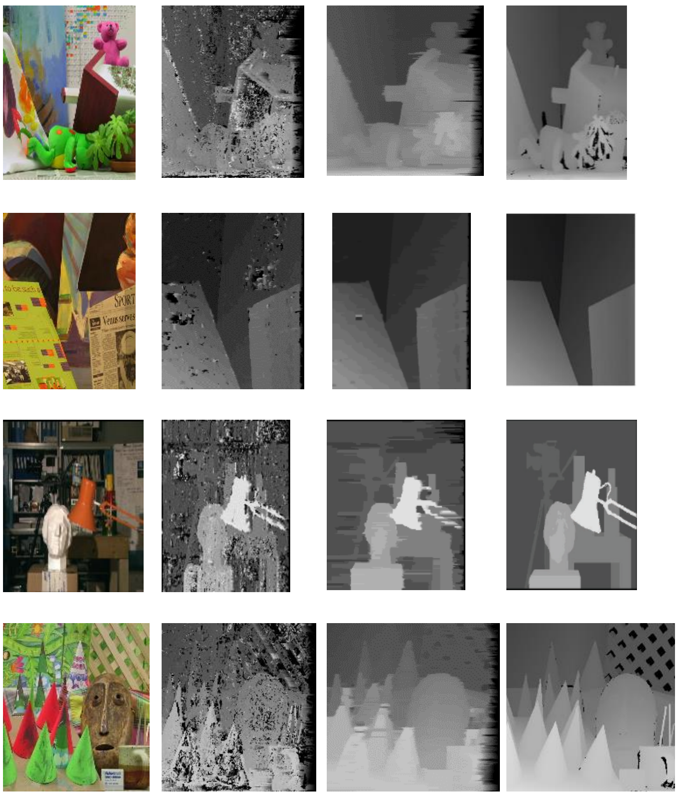
However, Basic Block Matching (BBM) is achieving low accurate result comparing Dynamic Programming (DP), but it has a high speed in proceeding the computational process. Moreover, the hardware part used to run the simulation of this experimental result for both stereo matching algorithms is the portable computer with integrated of processor Intel (R) Core(TM) i5-7200U CPU @2.5GHz 2.7GHz.

Table 3
Specific Disparity Ranges of Stereo Image Pairs

Stereo Images	Specific Disparity Range
Tsukuba	16
Teddy	59
Venus	19
Cones	50

Table 4
Time Taken Per Second for Different Set of Stereo Images of Stereo Matching Algorithms

Stereo Image	Basic Block Matching	Dynamic Programming
	Time Taken Per Second	
Tsukuba	17	54
Teddy	33	176
Venus	25	107
Cones	29	167



(a) (b) (c) (d)

Figure 4: Results of stereo matching algorithms by using the Middlebury benchmark datasets. Teddy, Venus, Tsukuba, Cones. Where the first column images are the original source images. The second column shows the disparity map result from the basic block matching. The third column shows the disparity map from dynamic programming. The fourth column shows the ground truths of images

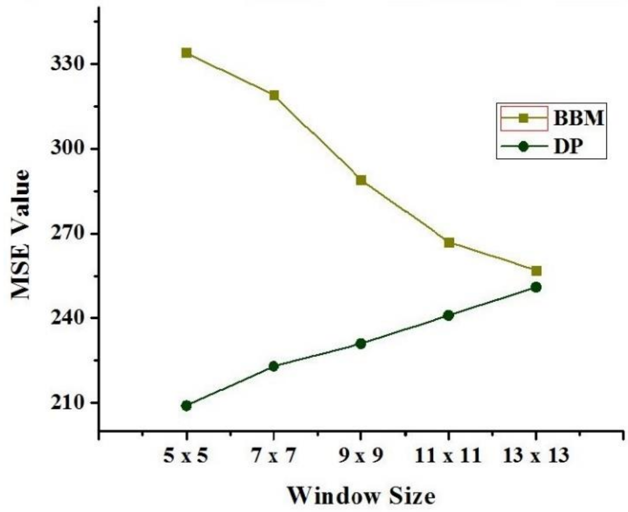


Figure 5: Value of MSE for Tsukuba datasets

Figure 5 above represents the value from MSE function, the value indicates that results out of MSE function for gaining and determining the disparity map are directly proportional to the gradually increasing of the window sizes for the two stereo matching algorithms: Basic Block Matching (BBM), Dynamic Programming (DP). However, the graph values of the BBM algorithm of the MSE values for Tsukuba stereo pair are gradually decreasing and this indicates that as window size increases, more errors are decreased for Basic Block Matching (BBM) algorithm.

In contrast, DP algorithm of the MSE values is raising proportional to the window size extending for Tsukuba stereo pairs. Thus, the results out of DP algorithm refers that the smaller size of window the better results are obtained due to its scanline optimization on each row of pixels, while the big size window might have missed some scanning on small objects in the disparity map structure or contents while smaller window may scan on content of image precisely and more errors can be reduced. besides, window size chosen is based on the complexity of the content of an image.

In addition, Figure 6 represents the values out of the PSNR faction obtained for the two stereo matching algorithms for Tsukuba datasets. Based on the results from Figure 6, the graph is clearly illustrating that the values computed from the PSNR function for the BBM algorithm is gradually raised with increasing of window sizes, and this indicates that more noises are reduced proportionally with the increase of window size for BBM algorithm.

While in the term of DP algorithm the values generated out of PSNR are gradually decreasing which indicates to the less removed noise, and this obviously represents that an efficiently operating of DP algorithm in reducing the errors of an image can only be with smaller window sizes.

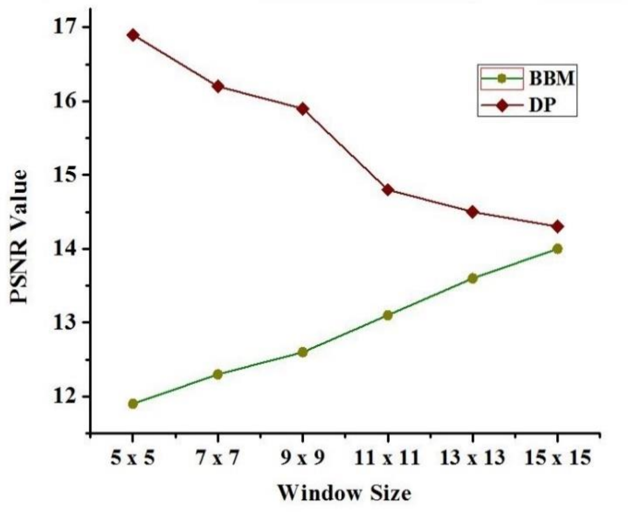


Figure 6: The value PSNR for Tsukuba datasets

For calculating the time through the two stereo matching algorithms, the tic toc computation method has been used. Where the tic toc method is applied to measure and calculate the complete execution time for an algorithm among the two stereo matching algorithms respectively. However, during this particular research, the method for the calculating the time (tic toc) is only used for computing the main functions in the algorithm, without not including some sub functions such as reading data part and showing out Figures.

Figure 7 below represents the time taken per second for the computation of the BBM and DP algorithms, whereas the graph shows the BBM is faster in comparison to DP algorithm due to its simplicity algorithm, but the results out of the BBM algorithm has much noises. Meaning while the time taken for DP algorithm computation is near to one minute.

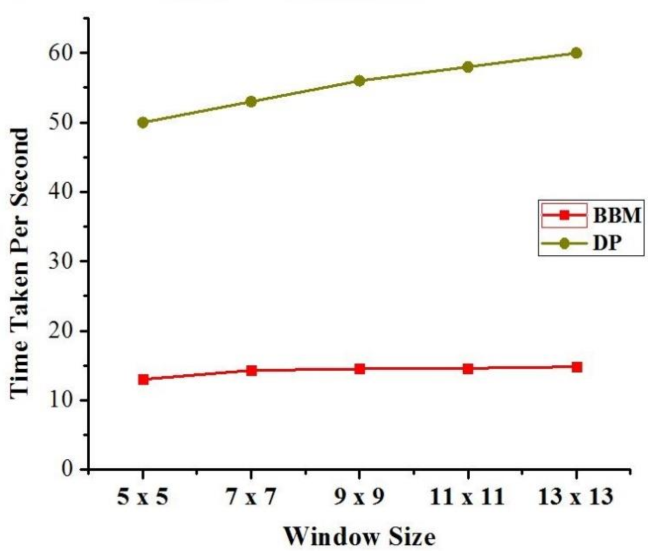


Figure 7: Time taken per second for computation of the BBM and DP algorithms

B. Results of Disparity Depth Map Based on Image Dataset Captured by (mv Bluefox) Camera Subjective Evaluation

In this section, depth maps gaining from the stereo matching algorithms, by using the Basic Block Matching (BBM) and the Dynamic Programming (DP) are based on our own stereo datasets. The stereo image datasets are taken using (mv Bluefox) camera, which applied for implementing on the stereo matching algorithms. However, the (mv Bluefox) camera is only capable of capturing out left and right images without creating the images ground truth. Besides, Figure 8 shows the output result of the disparity maps for each image dataset by applying both methods. Besides, for subjective evaluation of the result in Figure 8, the output result obtained with (mv Bluefox) camera can only be evaluated and analysis by interested researcher or human’s eyes observation for the depth disparity maps.

Where by observing the results from stereo matching algorithms of Basic Block Matching (BBM) and Dynamic Programming (DP), the result from the DP algorithm is more smoothness in compare to the BBM method which is noisier. In addition, the efficiency of dynamic programming algorithm in smoothing the disparity depth map is depending also on the appropriate camera baseline as can be observed for the stereo image pairs applied with specific baseline.

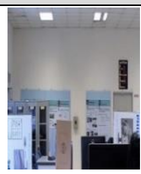
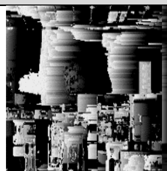
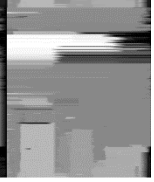
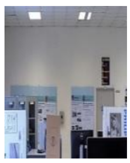
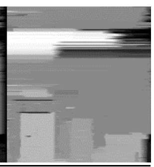
(a) Baseline in (cm)	(b) Original Image	(c) Basic Block Matching	(d) Dynamic Programming
Disparity range = (20)			
Disparity range = (25)			
Disparity range = (30)			

Figure 8: Results of stereo matching algorithms for dataset captured using (mv Bluefox) camera: Where (a) is the disparity ranges for each dataset coding. While (b) shows the original datasets images. In (c) the disparity map result from the basic block matching. In (d) shows the disparity map from dynamic programming.

VI. DISCUSSION

From the observation and evaluation of the output result of the disparity depth map, along with this comparative analysis of Basic Block Matching (BBM) and Dynamic Programming (DP) of stereo matching algorithms, it can be clearly seen that dynamic programming is a more efficient method of smoothing the depth map. Where DP is a more robust method and more capable of minimizing matching errors. Besides the method is time-consuming but often achieve an accurate result of the disparity map. In contrast, Basic Block Matching (BBM) method considered computationally faster and less time-consuming, while the method has a poor performance in smoothing disparity map and reducing matching error Table 5 presents a characteristic comparison between Basic Block Matching (BBM) and Dynamic Programming (DP) of stereo matching algorithms.

Table 5
Characteristic comparison between Basic Block Matching (BBM) and Dynamic Programming (DP) of stereo matching algorithms.

Characteristic	Basic Block Matching (BBM)	Dynamic Programming (DP)
Computational running process	Fast	Slow
Disparity map obtaining	Less accurate	More accurate
Algorithms structure	Not complex	Not complex
Ability to reduce matching errors	Less satisfactory	More satisfactory
Quality depth map for 3D applications	Less effective	More effective
Removing defective stripes.	No	Yes
Working under illumination changes and texture	Does not work well	Better

VII. CONCLUSION

In this paper, a comparative analysis of Basic Block Matching (BBM) and Dynamic Programming (DP) of stereo matching algorithms have been presented. Both methods have been investigated and experimented in term of disparity map accuracy implementation, noise enhancement, and smoothness. Besides, the Basic Block Matching (BBM) is applied in this research to perform a basic matching algorithm to determine the correspondence points between the target and reference image sets. While the Dynamic Programming (DP) has been used as a global optimization approach. The performance of the algorithm of both methods was tested based on the output results of the disparity depth map using two type of evaluation: objective and subjective evaluations. However, the objective evaluation is performed for the standard dataset taken from Middlebury database, while subjective evaluation is done for our own images captured by (mv Bluefox) camera. Where from results obtained for subjective evaluation part, and objective evaluation with applying evaluation functions including MSE and PSNR, it is clear that dynamic programming algorithm is more capable of minimizing the matching errors and gaining better output results in comparison to the basic block matching. Thus, by relying on the comparison of both stereo matching algorithms, dynamic programming is more portable for getting the more satisfying effect of disparity depth map and particularly in removing noise of the visible defective stripe. lastly, depth map from the dynamic programming algorithm is quite efficient to be applied for a wide range of applications comparing to basic block matching algorithms.

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] D. Murray and C. Jennings, “Stereo vision based mapping and navigation for mobile robots,” Robot. Autom. 1997. Proceedings., 1997 IEEE Int. Conf. on. Vol. 2. IEEE, 1997.

[2] M. D. Yang, D. W., Chu, L. C., Chen, C. W., Wang, J., & Shieh, “Depth-Reliability-Based Stereo-Matching Algorithm and Its VLSI Architecture Design,” IEEE Trans. Circuits Syst. Video Technol. 25.6 1038-1050, 2015.

[3] M. E. N. Yubo, Z. Guoyin, M. E. N. Chaoguang, L. I. Xiang, and M. A. Ning, “A Stereo Matching Algorithm Based on Four-Moded Census and Relative Confidence Plane Fitting,” Chinese J. Electron. 24.4 807-812, 2015.

[4] L. Zhang, “Fast Stereo Matching Algorithm for Intermediate View Reconstruction of Stereoscopic Television Images,” IEEE Trans. circuits Syst. video Technol. 16.10 1259-1270, 2006.

[5] 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. 12.10 2409-2421, 2017.

[6]S. Hussain and R. Modi, “Advancement in Depth Estimation for Stereo Image Pair,” Int. J. Innov. Res. Comput. Commun. Eng. 1.4, 2013.

[7] Q. Zhang, P. An, Y. Zhang, L. Shen, and Z. Zhang, “An improved depth map estimation for coding and view synthesis,” Image Process. (ICIP), 2011 18th IEEE Int. Conf. on. IEEE, 2011.

[8] M. Rziza, A. Tamtaoui, L. Morin, and D. Aboutajdine, “Estimation and Segmentation of a Dense Disparity Map for 3D Reconstruction,” Acoust. Speech, Signal Process. 2000. ICASSP’00. Proceedings. 2000 IEEE Int. Conf. on. Vol. 4. IEEE, 2000.

[9] 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.

[10] M. Gong and Y. Yang, “Fast Stereo Matching Using Reliability-Based Dynamic Programming and Consistency Constraints,” Comput. Vision, 2003. Proceedings. Ninth IEEE Int. Conf. on. IEEE, 2003.

[11] L. Di, M. Marchionni, and S. Mattoccia, “A fast area-based stereo matching algorithm,” Image Vis. Comput. 22.12 983-1005, 2004.

[12] A. Donate, X. Liu, and E. G. Collins, “Efficient path-based stereo matching with subpixel accuracy,” IEEE Trans. Syst. Man, Cybern. Part B 41.1 183-195, 2011.

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

[14] G. Li, “Stereo matching using normalized cross-correlation in LogRGB space,” Comput. Vis. Remote Sens. (CVRS), 2012 Int. Conf. on. IEEE, 2012.

[15] R. Long, Y. Lei, and Z. Z. L. G. Jiaqi-fei, “An Improved Stereo Match Algorithm Based on Support-Weight Approach,” Instrum. Meas. Comput. Commun. Control (IMCCC), 2014 Fourth Int. Conf. on. IEEE, 2014.

[16] 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.

[17] B. Tippetts, D. Jye, L. Kirt, and J. Archibald, “Review of stereo vision algorithms and their suitability for resource-limited systems,” J. Real-Time Image Process. 11.1 5-25, 2016.

[18] D. Kumari and K. Kaur, “A Survey on Stereo Matching Techniques for 3D Vision in Image Processing,” Int. J. Eng. Manuf 4 40-49, 2016.

[19] 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.

[20] J. Cai, “Fast Stereo Matching : Coarser to Finer with Selective Updating,” Image Vis. Comput. New Zealand, 2007.

[21] S. N. Sinha, D. Scharstein, and R. Szeliski, “Efficient high-resolution stereo matching using local plane sweeps,” Proc. IEEE Conf. Comput. Vis. Pattern Recognition, 2014.

[22] H. Hirschm and D. Scharstein, “Evaluation of Cost Functions for Stereo Matching,” Comput. Vis. Pattern Recognition, 2007.

CVPR'07. IEEE Conf. on. IEEE, 2007.
[23] L. Gomes Filho, J. C., Peiter, A. S., Pimentel, W. R. O., Soletti, J. I., Carvalho, S. H. V., & Meili, "Biodiesel production from Sterculia striata oil by ethyl transesterification method.," Ind. Crop. Prod. 74 767-772, 2015.
[24] 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.
[25] 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.
[26] 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.
[27] L. De-maeztu, S. Mattoccia, and R. Cabeza, "Linear stereo matching," Comput. Vis. (ICCV), 2011 IEEE Int. Conf. on. IEEE, 2011.
[28] A. A. Cigla, C., & Alatan, "Information permeability for stereo matching," Signal Process. Image Commun. 28.9 1072-1088, 2013.
[29] 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. 21.7 867-878, 2011.
[30] 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.
[31] B. Potetz, "Efficient Belief Propagation for Vision Using Linear Constraint Nodes," Comput. Vis. Pattern Recognition, 2007. CVPR'07. IEEE Conf. on. IEEE, 2007.
[32] M. Bleyer and M. Gelautz, "A layered stereo matching algorithm using image segmentation and global visibility constraints," ISPRS J. Photogramm. Remote Sens. 59.3 128-150, 2005.
[33] R. Gouiaa and J. Meunier, "3D reconstruction by fusioning shadow and silhouette information," Comput. Robot Vis. (CRV), 2014 Can. Conf. on. IEEE, 2014.
[34] D. Neilson and Y. H. Yang, "A component-wise analysis of constructible match cost functions for global stereopsis," IEEE Trans. pattern Anal. Mach. Intell. 33.11 2147-2159, 2011.
[35] 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.
[36] P. Ochs, J. Malik, and T. Brox, "Segmentation of moving objects by long termvideo analysis," IEEE Trans. pattern Anal. Mach. Intell. 36.6 1187-1200, 2014.
[37] Y. S. Chen, Y. P. Hung, and C. S. Fuh, "Fast block matching algorithm based on the winner-update strategy," IEEE Trans. Image Process. 10.8 1212-1222, 2001.
[38] D. G. and F. Girosi., "Parallel and deterministic al- gorithms from MRF's: Surface reconstruction.," IEEE Trans. Pattern Anal. Mach. Intell. 13.5 401-412, 1991.
[39] V. Kolmogorov and R. Zabih, "Multi-camera scene reconstruction via graph-cuts," Comput. Vision—ECCV 2002 8-40, 2002.
[40] P. Felzenszwalb and D. Huttenlocher, "Belief Propagation for Early Vision," Int. J. Comput. Vision, Vol. 70, No. 1, 2006.
[41] O. Veksler, "Stereo Correspondence by Dynamic Programming on a Tree," Comput. Vis. Pattern Recognition, 2005. CVPR 2005. IEEE Comput. Soc. Conf. on. Vol. 2. IEEE, 2005.
[42] T. Hu, B. Qi, T. Wu, X. Xu, and H. He, "Stereo matching using weighted dynamic programming on a single-direction four-connected tree q," Comput. Vis. Image Underst. 116.8 908-921, 2012.
[43] 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.