

Development of depth map from stereo images using sum of absolute differences and edge filters

Rostam Affendi Hamzah¹, Muhd Nazmi Zainal Azali¹, Zarina Mohd Noh², Madiha Zahari¹,
Adi Irwan Herman³

¹Fakulti Teknologi Kejuruteraan Elektrik and Elektronik, Universiti Teknikal Malaysia Melaka, Malaysia

²Fakulti Kejuruteraan Elektronik and Kejuruteraan Komputer, Universiti Teknikal Malaysia Melaka, Malaysia

³Product and Test Engineering Department, Texas Instruments, Batu Berendam, Melaka, Malaysia

Article Info

Article history:

Received Jul 28, 2021

Revised Nov 22, 2021

Accepted Dec 2, 2021

Keywords:

Depth map reconstruction

Edge filter

Image processing

Stereo matching algorithm

Sum of absolute differences

ABSTRACT

This article proposes a framework for the depth map reconstruction using stereo images. Fundamentally, this map provides an important information which commonly used in essential applications such as autonomous vehicle navigation, drone's navigation and 3D surface reconstruction. To develop an accurate depth map, the framework must be robust against the challenging regions of low texture, plain color and repetitive pattern on the input stereo image. The development of this map requires several stages which starts with matching cost calculation, cost aggregation, optimization and refinement stage. Hence, this work develops a framework with sum of absolute difference (SAD) and the combination of two edge preserving filters to increase the robustness against the challenging regions. The SAD convolves using block matching technique to increase the efficiency of matching process on the low texture and plain color regions. Moreover, two edge preserving filters will increase the accuracy on the repetitive pattern region. The results show that the proposed method is accurate and capable to work with the challenging regions. The results are provided by the Middlebury standard dataset. The framework is also efficiently and can be applied on the 3D surface reconstruction. Moreover, this work is greatly competitive with previously available methods.

This is an open access article under the [CC BY-SA](https://creativecommons.org/licenses/by-sa/4.0/) license.



Corresponding Author:

Rostam Affendi Hamzah

Fakulti Teknologi Kejuruteraan Elektrik and Elektronik, Universiti Teknikal Malaysia Melaka (UTeM)

Jalan Hang Tuah Jaya, 76100 Durian Tunggal, Melaka, Malaysia

Email: rostamaffendi@utem.edu.my

1. INTRODUCTION

Depth map contains important information for many applications such as range estimation, size measurement and 3D surface reconstruction. This article introduces depth map estimation from stereo images which is part of stereo vision field of study. Stereo vision is the most important field in computer vision and it provides various algorithms for computing different image processing related field of studies. Basically, these stereo images go through a stereo matching process to get the depth map. Based on several literature works in Bhuiyan and Khalifa [1] with previous published articles, the depth map is also called as a disparity map. The process is using two stereo images, the scene depth can be obtained from two separate points with some baseline displaced values. The correlation values of left image compared with right image are the result of stereo matching. The depth map is determined using the different intensity of pixel values on the map or the output from the stereo matching process based on the disparity map [2], [3]. The stereo matching function for computing the exact depth map is very challenging and difficult. Based on the works in Hamzah *et al.* [4]

the challenges are due to repetitive pattern, low texture area and plain color regions. These area or regions are hard to be matched and remains a challenge for the researchers to get an accurate result.

The process of stereo matching is one of the most interesting tasks in computer vision research area. This process begins with the matching of corresponding points from two images (i.e., left and right input images). Essentially, the framework of the matching process was proposed by Vedamurthy [5]. It has four stages which begins with matching cost, aggregation stage, disparity computation and disparity map refinement step at the last stage. Based on the proposed work in [5], there are three major methods to develop the stereo matching framework. The first is local method which uses all four stages in developing the framework. The optimization of local method employs winner-takes-all (WTA) strategy. In difference, global method uses energy minimization approach from the Markov random field (MRF) technique. This method utilizes three stages which excludes the aggregation stage. The third method is semi-global which this method combines both of the local and global methods. Semi-global employs all four stages [6].

Commonly, the depth map of global method is calculated by manipulating the global energy feature. The graph cuts (GC) [7], belief propagation (BP) [8] and dynamic programming (DP) [9] are the algorithms with global method. An algorithm to identify the depth discontinuities from pairs of stereo images as implemented in [9]. Their approach is able to execute the dynamic programming acceleration. A new stereo matching based on segmentation using graph cuts, which is used by assigning disparity planes to each segment to achieve the optimal solution [10]. Suggested integration of cost allocation-filtering approaches and global energy minimization approaches to encourage the increasing of stereo matching calculation using a two-stage energy minimization algorithm based on MRF modeling. This method is successfully solved the problem of stereo matching in occlusion areas. While previous approaches can effectively yield precise disparities in stereo matching, It isn't easy to enforce them, and complex scenes may fail to implement them. In addition, the learning-based methods are not reliable but are dependent on the training data. Eventhough global optimization achieves high accuracy of disparity estimation, huge computation complexity and effecting prolonged time that makes the global method limits its operation in real-time applications. Furthermore, the semi global method is also implemented using MRF approach where the framework is similar like global method with the cost aggregation stage is added in the framework.

Matching cost computation is the first stage which produces preliminary depth map. This stage generates high noise due to corresponding process of left and right images at each coordinate. There are several available methods in current research such as pixel-based matching cost [11], feature-based matching cost [12] and block matching cost [13]. All of these techniques have its own advantages and disadvantages as stated in [14]. Pixel based is fast but produced high noise compared to the other matching cost techniques. Feature based technique creates sparse depth map which only discovers image features such as edges or boundaries. The last matching cost technique is block-based method which this method is able to produce high accuracy if the windows size is suitably selected. The aggregation stage for stereo matching algorithm is very important to remove the preliminary noise after matching cost process. In common, this stage is applied in local and semi-global methods. The aggregation process implements the filtering by summing or averaging the cost volume in a support window. Some of available algorithms utilize the segmentation to improve the accuracy. Žbontar and LeCun [15], the adaptive support windows are applied to increase the accuracy by adjusting the neighboring pixels intensity. Hirschmüller *et al.* [16] has proposed a cost aggregation approach based on the segment-tree for non-local stereo matching, leading to improvements in both the precision of disparities and the processing speed. In placing more emphasis, [17], [18] a cross-scale design was suggested to enhance the cost averaging for effective stereo matching. Recursive edge-aware (REAF) [19] filters provided for precise and effective stereo matching.

The third stage is disparity optimization which this stage normalizes the disparity value and convert it to the intensity of depth pixel on the map. Local based methods embrace this stage with the same approach as implemented in [20]. Global and semi global methods as implemented in [21] skip this stage due to these methods minimize the disparity with the energy minimization approach similar to the Markov random field (MRF) technique. For the final stage of the stereo matching framework, the depth map refinement stage is taking place. This stage removes remaining noise on the depth map as implemented in [22]. It has two sequential processes at this stage which are invalid pixels detection and final depth map improvement or filtering. Einecke and Eggert [23] the segment-based approach was utilized based on fixed plane from the initial depth pixel on each segment. The assumption of depth pixels is varied smoothly and continuously within each homogeneous color segmentation to improve the accuracy. The rest of this paper is organized as follows. Next section explains the methodology of the proposed work in this article and followed with the results and discussion. Last part is the conclusion of the performance work in this article.

2. RESEARCH METHODOLOGY

The proposed framework in this article is shown in Figure 1. Stage 1: the corresponding process begins with the sum of absolute differences (SAD) cost function to get the preliminary disparity map. This function uses 17x17 windows size with block matching technique. Proper selection of windows sizes produces more accurate results and it should improve the efficiency of the corresponding process. Stage 2: at this stage the preliminary disparity map will be filtered to remove the noise or invalid pixels. The guided filter (GF) [24] will be used in the framework. The edge preserving filter of GF is one of the non-linear types that enhances the produced results. Hence, the used of this type of filter eliminates the invalid pixels on the disparity map and capable to increase the accuracy. Stage 3: this step introduces the optimization level of disparity assortment. The proposed work in this article uses winner-takes-all strategy where the normalization of disparity value is selected based on the minimum number. Stage 4: the final stage in the proposed framework is to further improve the accuracy by two continuous processes. The processes are fill-in invalid pixels then followed by the final filtering process which uses the bilateral filter (BF) [25]. The purpose of fill-in process is to replace the invalid pixel with neighboring valid pixel. This process makes the disparity map more accurate and is capable to reduce the error on the disparity map. The BF is used in the framework due to this filter is able to eliminate remaining noise and maintained the object edges. The capability of this filter at this stage is upsurge the accuracy on the disparity map.

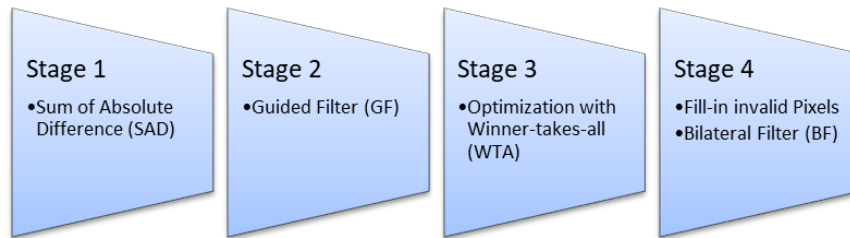


Figure 1. The stages of the proposed algorithm

2.1. Matching cost computation

The SAD is used in the proposed framework as shown in Figure 1 which this stage produces preliminary disparity map. Thus, early stage is very important where the corresponding points between two pixels of left and right images take place. One of the major problems for matching procedure at this step is corresponding points on the textureless regions. To minimize the mismatched problem during the matching process, block matching technique or window-based technique is applied in this framework with the SAD function. The advantage of block matching technique is accomplished to reduce the error on the textureless region. The consistency of weight at every pixel in the SAD windows increases the matching correctness and at the same time reduces the mismatched error. In (1) explains the SAD function which this article uses the red, green, and blue (RGB) images as an input stereo images.

$$SAD(x, y, d) = \frac{1}{M} \sum_{(x,y) \in M} |I_l^i(x, y) - I_r^i(x - d, y)| \tag{1}$$

Where the (x, y, d) is the pixel of interest with the disparity value, the left and right images are I_l and I_r respectively, the RGB channels denoted by i of right and left images and the SAD support window size represented by M at the size of (17×17) .

2.2. Cost aggregation

This stage is the most important part which removes the noise from preliminary depth map. Fundamentally, this stage will filter out the noise from matching process and should be proficient to keep the object edges. Some of illogical and ambiguities pixels are formed during the corresponding process. Therefore, at this step the filter must be strong and robust against any error at this stage. The GF is employed due to deep maintaining object boundaries and at the similar time capably to eliminate noise particularly on the low texture and flat color regions. In (2) is the equation of BF applied in this paper.

$$GF_{(p,q)}(I) = \frac{1}{w^2} \sum_{(p,q) \in w_c} \left(1 + \frac{(I_p - \mu_c)(I_q - \mu_c)}{\sigma_c^2 + \epsilon} \right) \tag{2}$$

Where $\{\epsilon, \sigma, \mu, I, p, q, w, c\}$ denoted by $\{\text{center pixel of } w, \text{variance value, constant parameter, reference image (left input image), mean value, coordinates of } (x,y), \text{window support size, neighboring coordinates}\}$. The GF

is engaged in this paper in line for efficient noise removal with fast execution of time processing. The GF improves the precision at the object boundaries. The final calculation of this step is provided by (3),

$$CA(x, y, d) = SAD(x, y, d)GF_{(p,q)}(I) \quad (3)$$

where the CA is the cost aggregation, $SAD(x, y, d)$ signifies the input of the first step and $GF_{(p,q)}(I)$ signifies the core of the GF.

2.3. Disparity optimization

This stage optimizes the depth map by using the winner-takes-all (WTA) strategy. The strategy involves the selection of the minimum disparity value is normalized by using the floating-point number from the cost aggregation stage. The WTA is usually used in the local methods due to fast execution [17], [25]. The WTA calculation is provided by (4).

$$d_{x,y} = arg \ CA(x, y, d) \quad (4)$$

Where $C(x, y, d)$ denotes the second stage of aggregation step and D represents a set of valid disparity values for an image. Fundamentally, after this stage the disparity map still contains noise or invalid pixels. Thus, this map needs to be enhanced or to be refined to get the best results.

2.4. Refinement stage

This stage is the last part of the proposed algorithm which has two continuous post-processes. It starts with hole filling process or technically the process of invalid pixel replacement with the valid pixel value. This article uses the nearest valid pixel value (i.e., close neighbor) to fill-in the hole or to replace the invalid pixel locations. After this process, some unwanted pixels or artifacts emerged. Therefore, the depth map needs to be smoothed to filter out the artifacts. Thus, the bilateral filter (BF) is utilized. This filter is very robust against the low texture and repetitive pattern regions. It is because the characteristic of this filter is able to remove the noise and at the same time maintained the object edges. The BF function is shown by (5) as follows:

$$BF_{p,q} = \sum_{q \in w_B} \exp\left(-\frac{|p-q|^2}{\sigma_s^2}\right) \exp\left(-\frac{|I_p - I_q|^2}{\sigma_c^2}\right) \quad (5)$$

where w_B and q are the size of window of BF and neighboring pixels and p is the location pixel of interest at (x,y) respectively. The $p-q$ represent spatial Euclidean interval and $I_p - I_q$ signifies the Euclidean distance in color space. The σ_s indicates a factor of spatial adjustment and σ_c corresponds to similarity factor for the color detection. Therefore, the final result of the depth map is represented by (5) which the $BF_{p,q}$ is the depth output d at the location pixel of interest p .

2.5. 3D Reconstruction

This section explains the application of the depth map. It will prove the accuracy of the implied depth map in this article based on qualitative measurement. From the depth map result in (5), a library from the OpenCV will be utilized for 3D surface reconstruction. It based on the (6):

$$D = \frac{bf}{d} \quad (6)$$

where D represents the depth, b and f are the stereo camera baseline and focal values respectively, and d denotes the disparity value. Therefore, the 3D surface reconstruction in this article formulated as (7) as follows:

$$3D = \frac{bf}{BF_{p,q}} \quad (7)$$

3. RESULTS AND DISCUSSION

This segment describes the investigational results and analysis on the implementation of the proposed framework. The experimental has been conducted based on the C++, 10 gen intel core i5 processor and 8GBHz random access memory (RAM). This article uses a standard benchmarking evaluation image which was provided by the Middlebury stereo assessment system. This database contains 15 images of training images which these images have different characteristics such as various illumination, low texture, discontinuity, and plain color regions. The all and nonocc errors are the attributes to the performance

measurement of the proposed algorithm. The measurement is based on the weight average error of testing and training images which is provided by the Middlebury online webpage. Figure 2 demonstrates the depth map results in grayscale color of the challenging regions. The results show 3 images with clear matching results on the depth map. Low texture region of the chair on Adirondack image is clearly observed in the depth map result. The plain color region of recycle bin image is also visible on the depth map. For the repetitive region, the Jadeplant image is selected due to its repetitive leaves on the image. The depth map result shows the leaves are well detected with its object sizes. Based on this figure, the proposed work in this article is strong against the challenging regions. Overall, 15 training images have been used for performance measurement. There are Recycle, Adirondack, Teddy, Playroom, Jadeplant, MotorcycleE, Pipes, Piano, PianoL, Playtable, ArtL, PlaytableP, Shelves, Motorcycle and Vintage as displayed in Figure 3. This figure shows almost accurate depth map results that were constructed based on RGB color channels. Fundamentally, the red color implies the objects is closer to the stereo camera. Meanwhile blue color region is far away from the stereo camera. It is also applied on the depth map with the grayscale color scheme.







Challenging Regions	Left Reference Image	Depth Map Results (Grayscale)
Low texture region for the Adirondack image		
Plain color region for the Recycle image		
Repetitive region for the Jadeplant image		

Figure 2. The depth map results for the challenging regions using training images

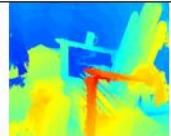
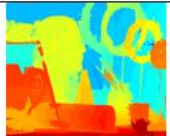
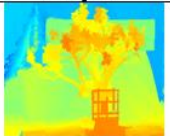
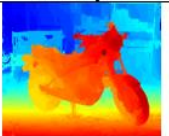
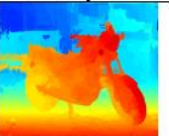
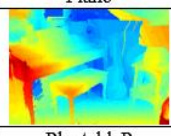
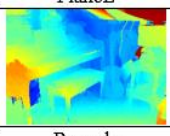
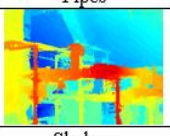
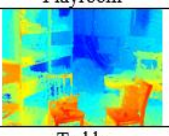

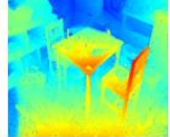

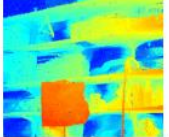
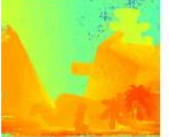

Adirondack	ArtL	Jadeplant	Motorcycle	MotorcycleE
				
Piano	PianoL	Pipes	Playroom	Playtable
				
PlaytableP	Recycle	Shelves	Teddy	Vintage
				

Figure 3. The disparity map results from the proposed work using the training images

Tables 1 and 2 are the quantitative measurement results based on the proposed work in this article using the Middlebury standard stereo evaluation system. There are 15 training images with the quantitative results provided by the Middlebury through online system. Table 1 tabulates the results of the nonocc error which also displays the proposed algorithm produces 8.83% of average error. It ranks at second behind the morphological processing stereoscopic visualization (MPSV) method with 8.81% and more accurate than the accurate dense stereo matching (ADSM) and binary stereo matching (BSM) methods with 8.95% and 13.40% respectively. For Table 2, the all error is presented where the evaluation is made with the same methods in Table 1 to make the comparison process more consistent. The proposed work produces the lowest average error compared to the ADSM, BSM and more importantly the proposed work is more accurate than the MPSV. It shows that the proposed work in this article is competitive with other available methods. There are several complex images such as Jadeplant, Piano, PianoL and Playtable produce the lowest average error compared with other methods in Table 2.

Figure 4 demonstrates the results of 3D surface reconstruction from a depth map result. Fundamentally, the accuracy of the depth map produced by the proposed framework is very crucial. It determines the quality of 3D surface reconstruction and overall algorithm features. Based on this figure, the motorcycle depth map (grayscale) is obviously discovered where the contour of depth detection is accurately presented for both conditions (i.e., front view and top view). The motorcycle is closer to the stereo camera and is clearly separated from the background grayscale tones. It shows that the background objects are far away from the motorcycle position. It can be proved by the 3D surface reconstruction from top view image where the depth distance is well-positioned. The depth layers are also showed in this image which indicates the matching process from the first to the final stage of the proposed work is efficiently established. Fundamentally, the pixel intensities are accurately positioned on the depth map of each object. Hence, this helps to accurately reconstruct the 3D surface with accurate depth estimation.

Table 1. The comparison of current available methods with the nonocc error from the Middlebury

Algorithms	MPSV [17]	Proposed algorithm	ADSM [18]	BSM [25]
Adiron	3.83	7.07	13.3	7.27
ArtL	6	7.01	6.1	11.4
Jadepl	19.7	14.6	15	30.5
Motor	5.85	4.39	3.67	6.67
MotorE	5.53	5.08	5.67	6.52
Piano	5.68	6.37	7.08	10.8
PianoL	34.3	12.9	20.6	32.1
Pipes	9.59	9.41	6.57	10.5
Playrm	5.86	14.8	13.2	12.5
Playt	15.3	9.63	23.1	24.4
PlayP	4.2	7.59	3.55	12.8
Recyc	4.59	7.76	5.76	7.42
Shelvs	13	18	17.2	16.4
Teddy	3.7	5.01	3.05	4.88
Vintge	14.3	17	10.1	32.8
Ave	8.81	8.83	8.95	13.4

Table 2. The comparison of current available methods with the all error from the Middlebury

Algorithms	Proposed algorithm	ADSM [18]	MPSV [17]	BSM [25]
Adiron	9.06	14.3	5.87	12.7
ArtL	10.1	10.6	9.43	28.7
Jadepl	30.6	34.1	40.2	58.7
Motor	7.02	6	9.11	14.8
MotorE	8.23	8	8.8	14.7
Piano	6.98	7.37	7.03	16
PianoL	13.4	20.4	34.2	35.8
Pipes	15.5	12.1	15.8	24.5
Playrm	19.3	16.9	8.58	29.4
Playt	11.9	25.5	16.9	31
PlayP	9.71	5.84	5.89	20.2
Recyc	8.29	5.83	6.78	12.1
Shelvs	17.9	17.2	13.7	19.2
Teddy	5.83	4.11	4.82	14.3
Vintge	17.3	11.1	16.8	39.3
Ave	12.1	12.3	12.7	23.5

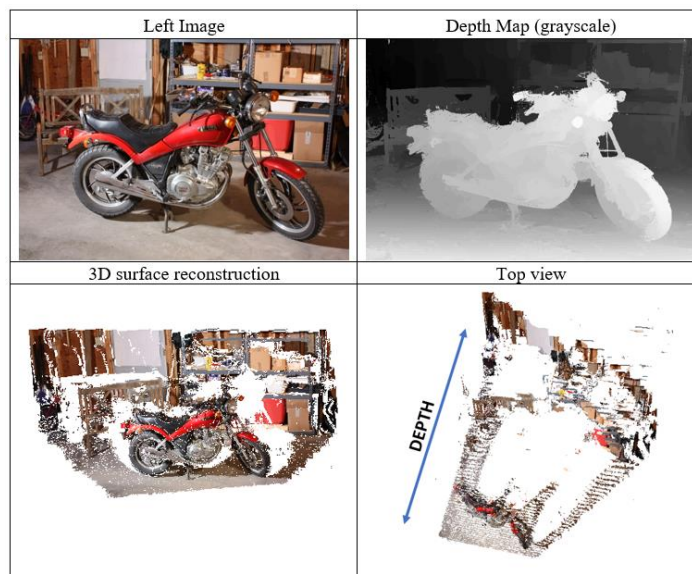


Figure 4. The depth map result with the 3D surface reconstruction of the motorcycle image

4. CONCLUSION

This paper proposes a framework for depth map reconstruction utilizing stereo images. This work established four stages of algorithm framework producing depth map or disparity map. The framework starts with matching cost calculation, cost aggregation, optimization and refinement stage. The proposed work in this paper used sum of absolute differences (SAD), guided filter (GF), winner-takes-all (WTA) and bilateral filter (BF) respectively. Based on the investigational study from a standard benchmark, the proposed work in this paper is capable to work with the regions that have difficulty to be matched such as low texture, plain color and repetitive pattern regions. It can be proved in from the provided results; the depth maps are completely reconstructed using sample standard images. The performance of the proposed framework is also measured based on the quantitative measurement. From this measurement result, it can be seen the proposed work is competitive with current published works with 8.83% and 12.10% of nonocc and all errors respectively. These quantitative measurements are provided from the standard benchmarking evaluation system from the Middlebury. Finally, this proposed work is also accomplished to be used for 3D surface reconstruction which is presented in this article, the motorcycle image has been used. The depth estimation from the 3D reconstruction result shown precise objects location and depth contour. Hence, the projected work in this paper can be used as a complete algorithm for the depth map algorithm and viable with other available methods.

ACKNOWLEDGEMENT

This work was sponsored by the Universiti Teknikal Malaysia Melaka and Ministry of Higher Education Malaysia with grant number FRGS/1/2020/TK0/UTEM/02/12.




REFERENCES

- [1] S. S. N. Bhuiyan and O. O. Khalifa, "Efficient 3D stereo vision stabilization for multi-camera viewpoints," *Bulletin of Electrical Engineering and Informatics*, vol. 8, no. 3, pp. 882-889, 2019, doi: 10.11591/eei.v8i3.1518.
- [2] W. Budiharto, A. Santoso, D. Purwanto, and A. Jazidie, "Multiple moving obstacles avoidance of service robot using stereo vision," *TELKOMNIKA Telecommunication Computing Electronics and Control*, vol. 9, no. 3, pp. 433-444, 2011, doi: 10.12928/telkonnika.v9i3.733.
- [3] E. Winarno, A. Harjoko, A. M. Arymurthy, and E. Winarko, "Face recognition based on symmetrical half-join method using stereo vision camera," *International Journal of Electrical and Computer Engineering*, vol. 6, no. 6, pp. 2818-2827, 2016, doi: 10.11591/ijece.v6i6.pp2818-2827.
- [4] R. A. Hamzah, H. Ibrahim, and A. H. A. Hassan, "Stereo matching algorithm for 3D surface reconstruction based on triangulation principle," *2016 1st International Conference on Information Technology, Information Systems and Electrical Engineering (ICITISEE)*, 2016, pp. 119-124, doi: 10.1109/ICITISEE.2016.7803059.
- [5] I. Vedamurthy *et al.*, "Recovering stereo vision by squashing virtual bugs in a virtual reality environment," *Phil. Trans. R. Soc. B*, 2016, doi: 10.1098/rstb.2015.0264.
- [6] H. X. H. Xi and W. Cui, "Wide baseline matching using support vector regression," *TELKOMNIKA Telecommunication Computing Electronics and Control*, vol. 11, no. 3, pp. 597-602, 2013, doi: 10.12928/telkonnika.v11i3.1144.




- [7] D. Scharstein and R. Szeliski, "A taxonomy and evaluation of dense two-frame stereo correspondence algorithms," *International Journal of Computer Vision*, vol. 47, pp. 7-42, 2002, doi: 10.1023/A:1014573219977.
- [8] Q. Yang, "A non-local cost aggregation method for stereo matching," *2012 IEEE Conference on Computer Vision and Pattern Recognition*, 2012, pp. 1402-1409, doi: 10.1109/CVPR.2012.6247827.
- [9] C. Rhemann, A. Hosni, M. Bleyer, C. Rother, and M. Gelautz, "Fast cost-volume filtering for visual correspondence and beyond," *CVPR 2011*, 2011, pp. 3017-3024, doi: 10.1109/CVPR.2011.5995372.
- [10] R. A. Setyawan, R. Soenoko, M. A. Choiron, and P. Mudjirahardjo, "Matching algorithm performance analysis for autocalibration method of stereo vision," *TELKOMNIKA Telecommunication Computing Electronics and Control*, vol. 18, no. 2, pp.1105-1112, 2020, doi: 10.12928/telkomnika.v18i2.14842.
- [11] C. Richardt, H. Kim, L. Valgaerts, and C. Theobalt, "Dense Wide-Baseline Scene Flow from Two Handheld Video Cameras," *2016 Fourth International Conference on 3D Vision (3DV)*, 2016, pp. 276-285, doi: 10.1109/3DV.2016.36.
- [12] R. A. Hamzah, H. N. Rosly, and S. Hamid, "An obstacle detection and avoidance of a mobile robot with stereo vision camera," *2011 International Conference on Electronic Devices, Systems and Applications (ICEDSA)*, 2011, pp. 104-108, doi: 10.1109/ICEDSA.2011.5959032.
- [13] Q. Yang, P. Ji, D. Li, S. Yao, and M. Zhang, "Fast stereo matching using adaptive guided filtering," *Image and Vision Computing*, vol. 32, no. 3, pp. 202-211, 2014, doi: 10.1016/j.imavis.2014.01.001.
- [14] J. Kowalczyk, E. T. Psota, and L. C. Perez, "Real-Time Stereo Matching on CUDA Using an Iterative Refinement Method for Adaptive Support-Weight Correspondences," in *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 23, no. 1, pp. 94-104, Jan. 2013, doi: 10.1109/TCSVT.2012.2203200.
- [15] J. Žbontar and Y. LeCun, "Computing the stereo matching cost with a convolutional neural network," *2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2015, pp. 1592-1599, doi: 10.1109/CVPR.2015.7298767.
- [16] H. Hirschmüller, P. R. Innocent, and J. Garibaldi, "Real-time correlation-based stereo vision with reduced border errors," *International Journal of Computer Vision*, vol. 47, pp. 229-246, 2002, doi: 10.1023/A:1014554110407.
- [17] N. Ma, Y. Men, C. Men, and X. Li, "Accurate dense stereo matching based on image segmentation using an adaptive multi-cost approach," *Symmetry*, vol. 8, no. 12, pp. 159-168, 2016, doi: 10.3390/sym8120159.
- [18] M. Kitagawa, I. Shimizu, and R. Sara, "High accuracy local stereo matching using DoG scale map," *2017 Fifteenth IAPR International Conference on Machine Vision Applications (MVA)*, 2017, pp. 258-261, doi: 10.23919/MVA.2017.7986850.
- [19] R. A. Hamzah, M. G. Wei, and N. S. Anwar, "Stereo matching based on absolute differences for multiple objects detection," *TELKOMNIKA Telecommunication Computing Electronics and Control*, vol. 17, no. 1, pp. 261-267, 2019, doi: 10.12928/telkomnika.v17i1.9185.
- [20] S. Wu, C. Tsai, and L. Chen, "Efficient Hardware Architecture for Large Disparity Range Stereo Matching Based on Belief Propagation," *2016 IEEE International Workshop on Signal Processing Systems (SiPS)*, 2016, pp. 236-241, doi: 10.1109/SiPS.2016.49.
- [21] A. G. S. Fakhar, "Development of portable automatic number plate recognition (ANPR) system on Raspberry Pi," *International Journal of Electrical and Computer Engineering*, vol. 9, no. 3, pp. 1805-1813, 2019, doi: 10.11591/ijece.v9i3.pp1805-1813.
- [22] S. Daniel and S. Richard, Middlebury Stereo Evaluation-Version 3, 2021, [Online]. Available: <https://vision.middlebury.edu/stereo/eval3/> (accessed: Mar 2021).
- [23] N. Einecke and J. Eggert, "Anisotropic median filtering for stereo disparity map refinement," *Proceedings of the International Conference on Computer Vision Theory and Applications*, 2013, pp. 189-198, doi: 10.5220/0004200401890198.
- [24] R. A. Hamzah, R. A. Rahim, and H. N. Rosly, "Depth evaluation in selected region of disparity mapping for navigation of stereo vision mobile robot," *2010 IEEE Symposium on Industrial Electronics and Applications (ISIEA)*, 2010, pp. 551-555, doi: 10.1109/ISIEA.2010.5679404.
- [25] K. Zhang, J. Li, Y. Li, W. Hu, L. Sun, and S. Yang, "Binary stereo matching," *Proceedings of the 21st International Conference on Pattern Recognition (ICPR2012)*, 2012, pp. 356-359.

BIOGRAPHIES OF AUTHORS






Rostam Affendi Hamzah    graduated from Universiti Teknologi Malaysia where he received his B.Eng majoring in Electronic Engineering. Then he received his M. Sc. majoring in Electronic System Design Engineering and PhD majoring in Electronic Imaging from the Universiti Sains Malaysia. Currently he is a lecturer in the Universiti Teknikal Malaysia Melaka teaching digital electronics, digital image processing and embedded system. He can be contacted at email: rostamaffendi@utem.edu.my.






Muhammad Nazmi Zainal Azali    currently pursuing the M.Sc. degree in Electronic Engineering from Universiti Teknikal Malaysia Melaka. His current research interests focusing on stereo vision and digital image processing. He is also interested in electronic soldering and circuit. He can be contacted at email: nazmi_z@icloud.com.






Zarina Mohd Noh    received the Ph.D. degree from Universiti Putra Malaysia. She is currently a senior lecturer also M.Sc. Co-Supervisor at Universiti Teknikal Malaysia Melaka. Her research interest includes image processing and computer embedded system engineering. She can be contacted at email: zarina.noh@utem.edu.my.



Madiha Zahari    was born in Melaka, Malaysia. She received the Bachelor degree in Electrical & Electronic Engineering from University Technology of Petronas in 2006 and Master of Engineering in Industrial Electronic and Control from University of Malaya. She is currently work as a Lecturer at Universiti Teknikal Malaysia Melaka. She can be contacted at email: madiha@utem.edu.my.



Adi Irwan Herman    graduated in 2015 with a Bachelor Degree in Computer Engineering Technology (Computer System) from the Universiti Teknikal Malaysia Melaka. Currently, he works with Texas Instrument more than 5 years with his current research interests are computer engineering related field of studies. He can be contacted at email: adiirwanherman@gmail.com.