A study of stereo matching algorithm on low texture and depth discontinuity regions

Melvin Gan Yeou Wei[1,2], Rostam Affendi Hamzah[1], Nik Syahrim Nik Anwar[2], Adi Irwan Herman[3]
[1]Department of Electronic Engineering Technology, Fakulti Teknologi Kejuruteraan Elektrik and Elektronik,
Universiti Teknikal Malaysia Melaka, Melaka, Malaysia
[2]Department of Electrical Engineering, Fakulti Kejuruteraan Elektrik, Universiti Teknikal Malaysia Melaka, Melaka, Malaysia
[3]Product and Test Engineering, Texas Instruments, Melaka, Malaysia

ABSTRACT

This article studies the performance of the proposed stereo matching algorithm on complex regions. These regions are areas with very limited information for the matching process which are low texture, and depth discontinuity regions. In this study, each algorithm uses different matching cost computation (MCC) techniques, but for cost aggregation (CA), disparity optimization (DO) and disparity refinement (DR), the technique remains the same. The MCC are absolute difference (AD), the combination of absolute difference and gradient matching (AD+GM) and census transform (CT). Then, for CA, DO and DR, they are minimum spanning tree (MST), winner take all (WTA) and bilateral filter (BF), respectively. The results are presented and discussed in this article. Hence, thru this study the robust method can be estimated at the MCC stage.

This is an open access article under the [CC BY-SA](#) license.

Corresponding Author:

Rostam Affendi Hamzah
Department of Electronic Engineering Technology, Fakulti Teknologi Kejuruteraan Elektrik and Elektronik
Universiti Teknikal Malysia Melaka
Taman Tasik Utama, 75450 Ayer Keroh, Melaka, Malaysia
Email: rostamaffendi@utem.edu.my

1. INTRODUCTION

Generally, stereo matching algorithm consists of 3 structures which are local based algorithm, semi-global based algorithm and global based algorithm. Fundamentally, the basic stereo matching algorithm consists of 4 stages and they are matching cost computation, cost aggregation, disparity optimization and disparity refinement. By referring to the evaluation from [1], there are improved methods such as dynamic programming (DP), scanline optimization (SO), simulated annealing (SA) and graph cut (GC) which were introduced and replacing the winner takes all (WTA) approach. Global based often skip cost aggregation by defining global energy function. Yet, this approach remains unpopular due to complexity of the algorithms at that time. Then, through the comparative study by [2], the study reveals that adding the support weight (weight) for the support region (support window) on cost aggregation technique also produced an accurate disparity map. This technique is well-known among the researchers which this approach normalizes the support region weight with the designated or suitable energy on the pixel of interest. Finally, semi-global based algorithm, which proposed by [3] was introduced through the combination of local based technique and global based technique.

The basic taxonomy of stereo vision disparity map (SVDM) algorithm mainly include 4 stages, yet one crucial stage always remains, matching cost computation (MCC). Here, left image known as reference image will correspondence with the right image; or the targeted image to produce the disparity map [4].

However, by just using matching cost computation, the disparity map produces high error. Therefore, further processing such as cost aggregation, disparity optimization and disparity refinement is required to increase the accuracy of the final disparity map. Matching cost computation is classified into three type: pixel matching, block matching and feature matching. Currently, the pixel-based matching and block matching are well-known matching techniques. Pixel-based matching technique is simple and fast in computation execution, yet the results are mostly unfavorable due to high error [5]. Examples of pixel matching techniques are absolute difference (AD), squared difference (SD) and gradient matching (GM). The corresponding process for pixel matching is one to one pixel matching, which only involves the pixel of interest. However, for block matching, the corresponding process involves multiple pixels which are pixel of interest and the surrounding or neighboring pixels in the elements of the support window [6]. The support window is also reference as "block" or "window". For block matching, multiple pixels (pixel of interest and surrounding pixels) are first aggregated. Only then, the aggregated pixel of interest from each images; left image and right image will correspondence with each other. Examples of block matching techniques are sum of absolute difference (SAD), sum squared difference (SSD) normalized cross correlation (NCC), rank transform (RT) and census transform (CT) according to the literature survey of [7]. Block matching techniques produces better disparity map compared to the pixel matching technique, but the computation time will be longer and closely tied to the sizes of the support window. In addition, the accuracy of disparity map is heavily related to the size of the support window. Improper size selection will also result to the edge fattening, edge blurring and severe depth discontinuity [8].

Feature matching is also another option for MCC. Instead of matching through pixel to pixel corresponding process, feature-based technique correspondences through visual feature, statistical characteristics and transformation structures [9]. When comparing to pixel matching and block matching, feature matching algorithm is more complex and the result is not so favorable, thus the lack of interest among researcher. Besides that, another option is combining multiple matching techniques. This approach is also introduced as joint matching costs. The combination can be include; multiple pixel matching techniques; multiple block matching techniques; multiple feature matching techniques and blending of any matching techniques. An example of blending matching technique was proposed by [10], where the matching cost stage was combined using speeded up robust features (SURF) matching [11] and CT [12]. Combining multiple matching techniques will increase the robustness of the algorithm, but the computational load and processing time will also increase as the number of technique combined increases. This research compares three type of matching cost computation, and they are pixel matching AD, combination of multiple pixel matching-absolute difference and gradient matching (AD+GM), and block matching CT.

A functioning algorithm only required two basic stages, matching cost computation and disparity optimization; more specifically the WTA approach. However, the results are very poor and most objects in the disparity maps are unidentified or unrecognized [13]. In this article, the performance comparison at MCC stage is presented. To provide more reliable comparison, two edge preserving techniques are introduced in the algorithm framework. There are minimum spanning tree (MST) for cost aggregation and bilateral filter (BF) for disparity refinement stage where the main function is to preserve the edges. Therefore, the comparison of three algorithms at MCC stage will have different matching cost computation approaches but the same cost aggregation, disparity optimization and disparity refinement. The main contribution of this article is to provide information on performance comparison and to determine the robust method at MCC on the complex regions for matching process. The complex regions for this process comprises illumination differences, low texture, and depth discontinuity regions. These regions are very difficult to be matched due to lack of pixel information and that's why an algorithm need robust features

2. METHOD

The state-of-the-arts matching techniques are the AD, the combination of AD+GM and CT. The cost aggregation (CA), disparity optimization (DO) and disparity refinement (DR) stages are fixed to a pre-determined method to observe the performance changes at the MCC stage. The selected method uses MST at CA and DO uses WTA approach. Finally, disparity refinement uses BF with weighted median (WM).

2.1. Matching cost computation

Stage 1 of algorithm is matching cost computation. Here three different matching techniques are utilised and they are AD, the combination of AD+GM and CT. The first algorithm uses pixel to pixel matching technique, which is AD, proposed by [14]. AD matching is presented in (1),

$$AD(p,d) = |(I_l(p) - I_r(p-d)|$$ (1)

where p is (x,y); the position of the targeted pixel and d is disparity value. Next, I_l is left image and I_r is right image.

The matching technique is then given a cut-off point, as implemented by [15], which resulting to AD with a threshold and represented as $AD'(p, d)$. The equation for $AD'(p, d)$ is presented in (2),

$$AD'(p,d) = \begin{cases} \tau_{AD}, if\ AD(p,d) > \tau_{AD} \\ AD(p,d), otherwise. \end{cases} \quad (2)$$

where τ_{AD} is the cut-off point. By introducing a threshold, these eliminating outliers. Next, the algorithm uses a combination of multiple pixel difference matching techniques. They are the combination of AD and GM, proposed by [16]. The GM matching extracts the gradient from the pixel of the input image. The G_x, gradient at horizontal direction and G_y, gradient at vertical direction are presented in (3) and (4),

$$G_x = [1 \quad 0 \quad -1] * I \quad (3)$$

$$G_y = \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix} * I \quad (4)$$

where the I is targeted image and * represents convolution operation. After that, gradient magnitude, m presented in (5) is obtained through the component of gradient from G_x and G_y.

$$m = \sqrt{G_x{}^2 + G_y{}^2} \quad (5)$$

From (5), the modulus from gradient magnitude is implemented on the input images where left image is m_l and right image is m_r. Through the gradient displacement of x-direction and the static position of y-direction, the cost for gradient matching, $GM(p, d)$ obtained and presented in (6),

$$GM(p,d) = |m_l(p) - m_r(p - d)| \quad (6)$$

where p is (x, y); the position of the targeted pixel, d is disparity value, m_l is gradient of left image and m_r is gradient of right image. By adding a cut-off point, the gradient matching with threshold, $GM'(p, d)$ is presented in (7),

$$GM'(p,d) = \begin{cases} \tau_{GM}, if\ GM > \tau_{GM} \\ GM, otherwise. \end{cases} \quad (7)$$

where τ_{GM} is the cut-off point. This also contributed in eliminating outliers for GM matching. By combining $AD'(p, d)$ and $GM'(p, d)$, the combined matching cost function, $CMCC(p, d)$ is presented in (8),

$$CMCC(p,d) = AD'(p,d) + \alpha GM'(p,d) \quad (8)$$

where α is a constant for manipulating the sensitivity of illumination differences. The other algorithm uses CT [17]. The pre-matching operates by converting the target pixel and the surrounding pixels within the support window into a bit string. The census bitstring is presented in (9),

$$Census\ (p) = Bitstring_{(i,j)\epsilon\ w}(I(i,j) \geq I\ (p)) \quad (9)$$

where p is (x, y); the position of the targeted pixel and d is disparity value. Next, w is the size for the support window, while i is the position of target pixel in the element of w and j is the coordinate of surrounding pixels in the element of w. Then $I(i, j)$ is the intensity value for the targeted pixel in the element of w and $I\ (p)$ is the intensity value for the targeted pixel. The matching cost is computed using hamming distance between census bit strings of the input images and presented in (10),

$$CT(x,y,d) = \sum_{(x,y)\epsilon\ w} Hamming(Census_{ref}(p), Census_{tar}(p - d)) \quad (10)$$

where p is (x, y); the position of the targeted pixel and d is disparity value. Meanwhile, $Census_{ref}$ is census bit string for the reference's images; left image and $Census_{tar}$ is census bit string for the target image; right image.

2.2. Cost aggregation

Stage 2 of algorithm is CA. The technique applied at this stage is MST, referring from [18]. The proposed segmentation technique at this stage function mainly as edge preserver. The procedures for MST and color image segmentation is shown as:

- The equation for spanning tree is $G = [V, E]$, referring to a non-orientation graph of object. Next, V refers a set of vertices that corresponds to the data set, and E refers the edge connects the vertices, and each edge $e_m = (x_i, x_j)$ connects to a pair of vertices. The MST from G is presented in (11),

$$MST = (A, T) | A = V, T = \{e_1, \dots e_{n-1}\} \text{ with } m(MST) = min\{m(tree) | tree = (V, T')\} \quad (11)$$

- The introduction of d_y, the threshold, edges with larger weights than the threshold are removed from MST, forming the forest F of V, presented in (12),

$$F = \left\{ (V, E') \middle| E' = T - \{e' | m(e') > d_y\} \right\} \quad (12)$$

- The adding of all the trees into F is presented in (13),

$$\{(V_i, T_i) | i = 1, 2, \dots m\}, F = \bigcup_{i=1}^{m} (V_i, T_i) \quad (13)$$

where, $\bigcup_{i=1}^{m} V_i = V, \bigcup_{i=1}^{m} T_i = E'$

- Each (V_i, T_i) can be considered as a cluster $C_i = X_i$, where $T_i = \bigcup_k (e_k' | e' < d_y)$

Then, the function for matching cost computation, $M(p, d)$, from previous stage will represents the set of vertices that corresponds to the data set for spanning three equations, $G = [V, E]$. By substituting $M(p, d)$ into V, (14) is presented.

$$G = [M(p, d), E] \quad (14)$$

In this research, $M(p, d)$ will variance depending on the matching technique. Here, the matching technique can be AD matching; $AD'(p, d)$ or combination of AD and GM; $CMCC(p, d)$ or CT techniques; $CT(x, y, d)$. The cost aggregation equation, $CA(p, d)$ is presented in (15).

$$CA(p, d) = [M(p, d), E] \quad (15)$$

2.3. Disparity optimization

Stage 3 of algorithm is disparity optimization and the technique applied is WTA, implemented by [19], [20]. By using WTA approach, each pixel of the disparity map is optimized with minimum disparity value. WTA is presented in (16):

$$d(p) = arg\, min_{d \in D} CA(p, d) \quad (16)$$

where $d(p)$ is the disparity value at the coordinate of (x, y), D is the range of disparity on an image and $CA(p, d)$ is data obtain from previous stage, cost aggregation for this research.

2.4. Disparity refinement

Stage 4 of algorithm is DR. This stage consists of two main processes which are the post-processing and final disparity map filtering. Post processing, implemented by [21] and [22], normally involves left right (LR) consistency checking and pixels filling-in process. LR checking is applied to detect outliner or the invalid pixels. This process started at left reference disparity map image that coincides with the right reference disparity map. The mismatched values between the two are defined as invalid pixel. LR checking is presented in (17).

$$|d_{LR}(p) - d_{RL}(p - d_{LR}(p))| \leq \tau_{LR} \quad (17)$$

Next, the pixels filling-in process will replace the invalid pixels. Here invalid disparity value is detected and is replaced with the nearest valid disparity value. Since, the left image is pre-set as the reference image, the process started from the left side and then towards the right side. Furthermore, the valid value is set to be position at the same scanning line. The pixels filling-in process is presented in (18),

$$d(p) = \begin{cases} d(p-i), d(p-i) \le d(p+j), \\ d(p+j), otherwise. \end{cases} \tag{18}$$

where $d(p)$ is disparity value of coordinate p. Next, $(p-i)$ is the position of the first valid disparity on the left side and $(p+j)$ is the coordinate of the first valid disparity on the right side. Then, the process continued with disparity refinement step and the technique applied is BF with weighted median. Bilateral filter, $B(p, q)$ proposed by [23] is presented in (19),

$$B(p,q) = exp\left(-\frac{|p-q|^2}{\sigma_s^2}\right) exp\left(-\frac{|d(p)-d(q)|^2}{\sigma_c^2}\right) \tag{19}$$

where (p, q) is the position for target pixel and surrounding pixels. Next, $|p - q|$ is spatial Euclidean and $|d(p) - d(q)|^2$ is Euclidean. Then, σ_s^2 is spatial distance and σ_c^2 is colour similarity. BF function mainly as an edge preserving filter and to further improve the accuracy of disparity map, $B(p, q)$ is then transform into summation of histogram, $h(p, d_r)$ and is presented in (20),

$$h(p, d_r) = \sum_{q \epsilon w_p | d(q)==d_r} B(p, q) \tag{20}$$

where d_r is disparity range and w_p is window size with the radius ($r \times r$) at centred pixel of p. After that, WM filtering was further implemented to achieve higher accuracy, inspired by [24].

3. RESULTS AND DISCUSSION

This article uses the Middlebury standard benchmarking dataset that contains the training images for parameter settings. These images are widely used by the researchers due to its complexity and different characteristic of features [25], [26]. The detailed results for the AD+GM-MST-BF and CT-MST-BF algorithms are presented in Figure 1, Table 1 and Table 2. Figure 1 presents the qualitative and quantitative disparity map results of selected training images to study the performances of the algorithms. Next, for the quantitative results, Table 1 presents the detailed result for nonocc error and Table 2 presents the detailed for all error. Mao and Gong [27], explained nonocc error represents the error of invalid disparity values on the non-occluded regions and all error represents the error of invalid disparity values on all pixels of the disparity map image.
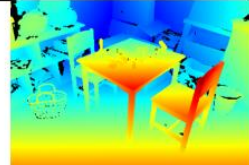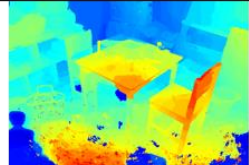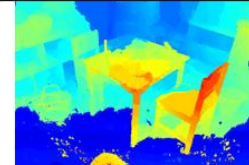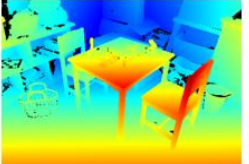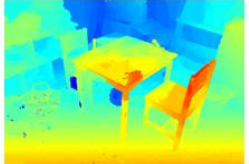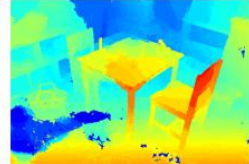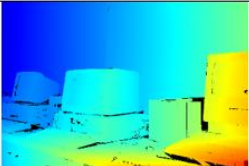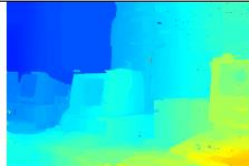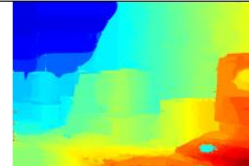

Ground truth	AD+GM - MST - BF	CT - MST - BF
Playtable		
all error (%)	19.5	42.0
PlaytableP		
all error (%)	4.94	13.6
Vintage		
all error (%)	14.3	19.8

Figure 1. The results on the highly low texture region using the Middlebury dataset

Table 1. The results on nonocc error

Algorithms	AD+GM-MST-BF	CT-MST-BF
Weight average	7.58	7.83
Adiron	4.88	3.25
ArtL	6.81	4.54
Jadepl	15.3	13.3
Motor	4.24	3.65
MotorE	7.42	3.75
Piano	5.82	4.30
PianoL	12.7	9.99
Pipes	7.26	6.21
Playrm	8.12	5.30
Playt	17.1	41.9
PlaytP	4.12	11.6
Recyc	4.43	3.12
Shelvs	11.0	6.94
Teddy	3.38	2.64
Vintage	13.4	19.0

Table 2. The results on all error

Algorithms	AD+GM-MST-BF	CT-MST-BF
Weight average	10.6	10.8
Adiron	5.92	3.78
ArtL	12.4	8.43
Jadepl	30.2	28.9
Motor	6.52	6.40
MotorE	9.24	6.51
Piano	6.44	5.14
PianoL	13.0	10.4
Pipes	12.2	11.3
Playrm	13.2	9.26
Playt	19.5	42.0
PlaytP	4.94	13.6
Recyc	4.66	3.48
Shelvs	11.2	7.27
Teddy	4.79	3.82
Vintage	14.3	19.8

3.1. CT method is sensitive to highly low texture region

Low texture regions are regions that are plain and barely have no contrast in texture information [28]. Matching process for low texture regions is difficult due to pixels of such region are almost similar and resulted to more than one corresponding point. Therefore, corresponding to the correct pixels can be more challenging. To demonstrate that CT technique is sensitive to highly low texture region, Figure 1 presents few selective disparity maps, accompanied by their respective all errors, for both AD+GM-MST-BF algorithm and CT-MST-BF algorithm. According to Table 2, most disparity maps from CT-MST-BF algorithm have best accuracy. However, three final disparity maps, Playtable, PlaytableP, and Vintage do not achieve that. In contrast, they have the highest all error when compared with another algorithm. The final disparity maps-Playtable, PlaytableP, and Vintage for AD+GM-MST-BF algorithm achieved better results with all errors of 19.5%, 4.94%, and 14.3% respectively. However, CT-MST-BF algorithm have the higher all errors of 42.0%, 13.0%, and 19.8% respectively. Through comparison, there are huge differences of 22.5%, 8.06%, and 5.5% respectively. The huge differences affected the avg all error by 0.02%, causing AD+GM-MST-BF algorithm to ranked first, with avg all error of 10.6% and CT-MST-BF algorithm ranked lesser with average all error of 10.8%. By observing the final disparity maps for Playtable, PlaytableP, and Vintage, region that are highly low texture are full of invalid pixels. The highly low texture regions are the carpet floor for both Playtable and PlaytableP. Then, for Vintage, the upper left area-the white wall. As mentioned previously, MST and BF mainly act as edge preserver. Therefore, CT technique are sensitive to highly low texture region.

3.2. CT method reduces depth discontinuity

Depth discontinuity represents object that appeared to be "discontinued" or "disappearing" [28]. This are quite common for elongated object such as pipes, paint brush and more. The object is so thin that the pixels of the object are wrongly substitute by surrounding pixels during matching process. For severe cases, the pixels of the object are substitute completely by the surrounding pixels, resulting to object "disappearing". To explain that CT technique reduces depth discontinuity, Figure 2 presents several disparity maps and their respective all errors, for both AD+GM-MST-BF algorithm and CT-MST-BF algorithm.
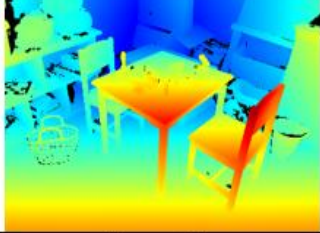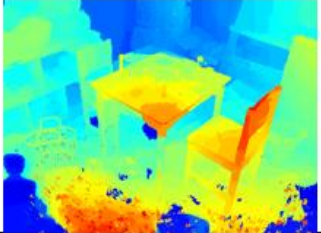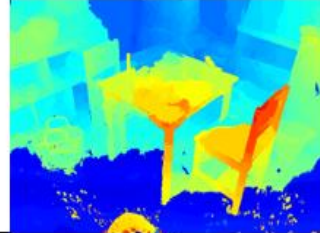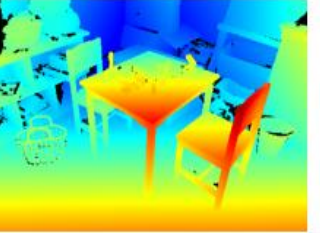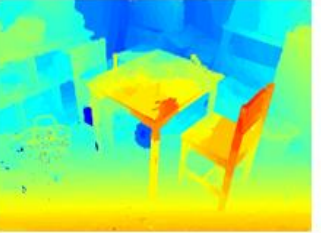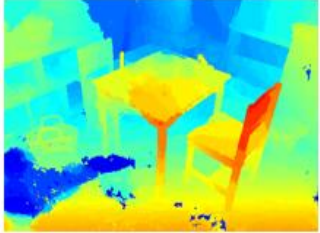
Figure 2. The results on the depth discontinuity region using the Middlebury dataset

In this Figure 2, the selected final disparity maps-ArtL and Pipes for AD+GM-MST-BF and CT-MST-BF algorithms present clear depth discontinuity, especially for AD+GM-MST-BF algorithm. In addition to that, Pipes represents "elongated object" precisely. The elongated objects on final disparity map-ArtL for AD+GM-MST-BF algorithm "disappear" almost completely. These are examples of severe depth discontinuity. Most elongated objects are absent throughout the whole disparity map. However, CT-MST-BF algorithm shows more trances of elongated objects, such as the paint brushes, showing a better respond toward depth discontinuity. The quantitative results of all error also proves that the CT-MST-BF algorithm is better, with 8.43% error while AD+GM-MST-BF algorithm have a higher error of 12.4%. Then, for the final disparity map-Pipes for both algorithms, the elongated objects still suffer depth discontinuity, but they are still recognizable as the elongated objects are thicker. By observing the pipe at the right side of the final disparity map, CT-MST-BF algorithm presents a clearer elongated object compared to AD+GM-MST-BF algorithm. The quantitative results, avg all also validates the claimed as AD+GM-MST-BF algorithm has higher error of 12.2% while CT-MST-BF algorithm is lower, 11.3%. Again, as mentioned before MST and BF mainly function as edge preserver. Therefore, CT technique reduces depth discontinuity.

4.   CONCLUSION

Based on this study, the comparison of algorithms show that the AD+GM-MST-BF algorithm, equipped with the combination of multiple pixel matching technique achieved the best accuracy. Following next is, CT-MST-BF algorithm, equipped with block matching technique. Besides that, there are other findings in this study such as the AD+GM method is sensitive to illumination differences, the CT method is sensitive to highly low texture region and at the same time, the CT method is capable to reduce the error on depth discontinuity region. Hence, it can be concluded that the AD+GM-MST-BF algorithm producss the best result in this experiment. However, at the MCC stage, CT has good results on the depth discontinuity which this region is very complex to be matched.

REFERENCES

[1] D. Scharstein, R. Szeliski, and R. Zabih, "A taxonomy and evaluation of dense two-frame stereo correspondence algorithms," in *Proceedings - IEEE Workshop on Stereo and Multi-Baseline Vision, SMBV 2001*, 2001, pp. 131–140, doi: 10.1109/SMBV.2001.988771.

[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, Dec. 2011, doi: 10.12928/telkomnika.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 (IJECE)*, vol. 6, no. 6, p. 2818, Dec. 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," in *Proceedings - 2016 1st International Conference on Information Technology, Information Systems and Electrical Engineering, ICITISEE 2016*, Aug. 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," *Philosophical Transactions of the Royal Society B: Biological Sciences*, vol. 371, no. 1697, p. 20150264, Jun. 2016, doi: 10.1098/rstb.2015.0264.

[6] H. X. Haifeng Xi and W. Cui, "Wide baseline matching using support vector regression," *TELKOMNIKA (Telecommunication Computing Electronics and Control)*, vol. 11, no. 3, p. 597, Sep. 2013, doi: 10.12928/telkomnika.v11i3.1144.

[7] S. S. N. Bhuiyan and O. O. Khalifa, "Efficient 3D stereo vision stabilization for multi-camera viewpoints," *Bulletin of Electrical Engineering and Informatics (BEEI)*, vol. 8, no. 3, pp. 882–889, Sep. 2019, doi: 10.11591/eei.v8i3.1518.

[8] Q. Yang, "A non-local cost aggregation method for stereo matching," in *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, Jun. 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," in *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, Jun. 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, Apr. 2020, doi: 10.12928/TELKOMNIKA.V18I2.14842.

[11] R. A. Hamzah, H. N. Rosly, and S. Hamid, "An obstacle detection and avoidance of a mobile robot with stereo vision camera," in *International Conference on Electronic Devices, Systems, and Applications*, Apr. 2011, pp. 104–108, doi: 10.1109/ICEDSA.2011.5959032.

[12] S. Ahmed, M. Hansard, and A. Cavallaro, "Constrained optimization for plane-based stereo," *IEEE Transactions on Image Processing*, vol. 27, no. 8, pp. 3870–3882, Aug. 2018, doi: 10.1109/TIP.2018.2823543.

[13] Y. Lee and C. M. Kyung, "A memory- and accuracy-aware gaussian parameter-based stereo matching using confidence measure," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 43, no. 6, pp. 1845–1858, Jun. 2021, doi: 10.1109/TPAMI.2019.2959613.

[14] L. Kong, J. Zhu, and S. Ying, "Local stereo matching using adaptive cross-region-based guided image filtering with orthogonal weights," *Mathematical Problems in Engineering*, vol. 2021, pp. 1–20, May 2021, doi: 10.1155/2021/5556990.

[15] J. Žbontar and Y. Le Cun, "Computing the stereo matching cost with a convolutional neural network," in *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, Jun. 2015, vol. 07-12-June-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, no. 1–3, 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, p. 159, Dec. 2016, doi: 10.3390/sym8120159.

[18] W. Yuan, C. Meng, X. Tong, and Z. Li, "Efficient local stereo matching algorithm based on fast gradient domain guided image filtering," *Signal Processing: Image Communication*, vol. 95, p. 116280, Jul. 2021, doi: 10.1016/j.image.2021.116280.

[19] R. A. Hamzah, M. G. Y. Wei, and N. S. N. Anwar, "Stereo matching based on absolute differences for multiple objects detection," *Telkomnika (Telecommunication Computing Electronics and Control)*, vol. 17, no. 1, pp. 261–267, Feb. 2019, doi: 10.12928/TELKOMNIKA.v17i1.9185.

[20] S. S. Wu, C. H. Tsai, and L. G. Chen, "Efficient hardware architecture for large disparity range stereo matching based on belief propagation," in *IEEE Workshop on Signal Processing Systems, SiPS: Design and Implementation*, Oct. 2016, pp. 236–241, doi: 10.1109/SiPS.2016.49.

[21] A. G. S Fakhar, H. M Saad, K. A Fauzan, H. R Affendi, and A. M Aidil, "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, Jun. 2019, doi: 10.11591/ijece.v9i3.pp1805-1813.

[22] D. Scharstein and H. Hirschmüller, "Middlebury stereo evaluation," 2021. [Online]. Available: https://vision.middlebury.edu/stereo/eval3/.

[23] N. Einecke and J. Eggert, "Anisotropic median filtering for stereo disparity map refinement," in *VISAPP 2013 - Proceedings of the International Conference on Computer Vision Theory and Applications*, 2013, vol. 2, 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," in *ISIEA 2010 - 2010 IEEE Symposium on Industrial Electronics and Applications*, Oct. 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," in *Proceedings - International Conference on Pattern Recognition*, 2012, pp. 356–359, doi: 10.1109/ICPR.2012.9489043.

[26] M. Kitagawa, I. Shimizu, and R. Sara, "High accuracy local stereo matching using DoG scale map," in *2017 Fifteenth IAPR International Conference on Machine Vision Applications (MVA)*, May 2017, pp. 258–261, doi: 10.23919/MVA.2017.7986850.

[27] W. Mao and M. Gong, "Disparity filtering with 3D convolutional neural networks," in *2018 15th Conference on Computer and Robot Vision (CRV)*, May 2018, pp. 246–253, doi: 10.1109/CRV.2018.00042.

[28] Y. Li and S. Fang, "Removal-based multi-view stereo using a window-based matching method," *Optik*, vol. 178, pp. 1318–1336, 2019, doi: 10.1016/j.ijleo.2018.10.126.

BIOGRAPHIES OF AUTHORS

Melvin Gan Yeou Wei ⬤ 🔳 SC ↻ was born in 1992, Melvin Gan Yeou Wei graduated from Universiti Teknikal Malaysia Melaka where he received his B.Eng. majoring in Electrical and Master Degree in Mechatronic Engineering respectively. Currently, he is pursuing Ph.D. in the Universiti Teknikal Malaysia Melaka. He can be contacted at email: melvin.gyw92@gmail.com.

Rostam Affendi Hamzah ⬤ 🔳 SC ↻ 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 Ph.D. 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.

Nik Syahrim Nik Anwar ⬤ 🔳 SC ↻ was born in 1981, Nik Syahrim graduated from University of Applied Science Heilbronn, Germany where he received his Diploma in Mechatronik und Mikrosystemtechnik majoring in Mechatronics in 2006. In 2010 he received his M.Sc. majoring in Mechatronics from the University of Applied Science Aachen, Germany. In 2018, he received Ph.D. majoring in Electrical from the Universiti Sains Malaysia. Currently he is a senior lecturer in Universiti Teknikal Malaysia Melaka teaching Electrical and Mechatronics subjects. He can be contacted at email: syahrim@utem.edu.my.

Adi Irwan Herman ⬤ 🔳 SC ↻ 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.