Indonesian Journal of Electrical Engineering and Computer Science
Vol. 22, No. 3, June 2021, pp. 1312~1319
ISSN: 2502-4752, DOI: 10.11591/ijeecs.v22.i3.pp1312-1319

Local stereo matching algorithm using modified dynamic cost computation

A. F. Kadmin[1], R. A. Hamzah[2], M. N. Abd Manap[3], M. S. Hamid[4], T. F. Tg. Wook[5]
[1,3,4]Fakulti Kejuruteraan Elektronik dan Komputer, Universiti Teknikal Malaysia Melaka, Melaka, Malaysia
[2,5]Fakulti Teknologi Kejuruteraan Elektrik dan Elektronik, Universiti Teknikal Malaysia Melaka, Melaka, Malaysia

Article Info

Article history:

Received Mar 7, 2021
Revised May 3, 2021
Accepted May 5, 2021

Keywords:

Disparity map
Dynamic cost
Local method
Stereo correspondence
Stereo matching

ABSTRACT

Stereo matching is an essential subject in stereo vision architecture. Traditional framework composition consists of several constraints in stereo correspondences such as illumination variations in images and inadequate or non-uniform light due to uncontrollable environments. This work improves the local method stereo matching algorithm based on the dynamic cost computation method for depth measurement. This approach utilised modified dynamic cost computation in the matching cost. A modified census transform with dynamic histogram is used to provide the cost in the cost computation. The algorithm applied the fixed-window strategy with bilateral filtering to retain image depth information and edge in the cost aggregation stage. A winner takes all (WTA) optimisation and left-right check with adaptive bilateral median filtering are employed for disparity refinement. Based on the Middlebury benchmark dataset, the algorithm developed in this work has better accuracy and outperformed several other state-of-the-art algorithms.

Corresponding Author:

A. F. Kadmin
Faculty of Electrical and Electronic Enginering
Universiti Teknikal Malaysia Melaka
Hang Tuah Jaya, Durian Tunggal 76100, Melaka Malaysia
Email: fauzan@utem.edu.my

1. INTRODUCTION

The stereo-matching algorithm's development framework is a popular topic in the stereo vision architecture to obtain the disparity map. Several important applications apply the disparity map to acquire the object depth measurement, such as in medical image processing, 3D reconstruction, virtual reality, autonomous navigation, and many more [1]. The disparity estimation accuracy evaluation is quite crucial since small inaccuracies lead to no little impact on the result of the 3D application. Thus, it has been a targeted subject by many researchers [2]. The disparity accuracy assessment of the stereo matching algorithm can be compared with other state-of-the-art algorithms using the online evaluation platform such as Middlebury dataset [3]. Most of the algorithms established are based on the traditional framework of Schartein and Szeliski, comprised of four stages; 1) Matching cost, 2) Aggregation of cost, 3) Disparity optimisation, and 4) Final disparity refinement [4]. The framework used an essential cost function of matching to measure the stereo images' corresponding points from two or multiple perspectives [5].

The approach of each stereo matching algorithm can be categorised into three groups: local, global, and semi-global [6]. The local approach delivers a fast computation of cost using the pixel information comparison process between the nearby pixel, contributing to shorter processing time. The local method is mostly used in real-time applications, but it produces low disparity accuracy from mismatches, and high noise sensitivity. While global approach typically provides high accuracy of disparity map by producing the

disparities within the stereo vision system using energy function minimization process. The trade-off disadvantage of this approach is contributing to cost exhaustive, too complex and longer execution time. The third approach is the semi-global, which combines local and global approaches to a more balance disparity accuracy and differs based on the framework.

It is difficult to acquire accurate correspondences due to several factors, especially from low texture regions, radiometric differences from environment illumination variations, and blurry boundaries displayed due to poor segmentation methods. Several factors contributed to these problems, such as specular reflection and non-Lambertian surfaces [7]. The stereo camera light sensor also produced illumination differences, resulting in dissimilar intensity levels corresponding to the same point in 3D space. Besides that, the irregularity of the image captured from the stereo camera also caused radiometric differences. The properties from inconsistent multiple stereo cameras like Gaussian noise, salt, and pepper noise, gain setting, and vignetting are the main factors for radiometric differences.

An improved framework of the local stereo matching algorithm approach is developed to increase the disparity map accuracy based on modified dynamic cost computation. This work used a modified census transform (CT) with a dynamic histogram to provide the cost computation in the matching cost stage, while in the aggregated cost stage, it utilised a fixed-window strategy with bilateral filtering that efficiently retains image depth information and edge in the image. In the disparity selection and refinement stage, a winner-takes-all (WTA) optimisation and a consistency check of left and right with adaptive bilateral median filtering are employed. The rest of this article is organised is being as. Section 2 explained the illumination constraints. The methodology of the framework is shown in section 3. Section 4 will explain the experiment result and the discussion, while the conclusion is summarised in section 5.

2. ILLUMINATION VARIATIONS IN STEREO CORRESPONDENCE

The most challenging stereo matching constraint is to get accurate correspondence estimation information recovered from the 2D plane image due to several factors [8]. The basic stereo vision system used an essential cost function of matching to measure the corresponding stereo images' corresponding points from two or multiple perspectives. It is challenging to establish accurate correspondences between pixels that lie inside low-textured image areas, as shown in Figure 1 for Teddy images left and right [9] due to different illumination environments, amongst other factors, which needs further complicated cost functions to solve the radiometric differences.

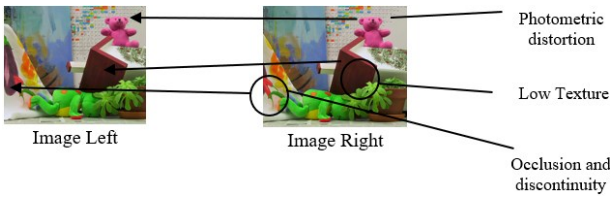

Figure 1. Noise factors in stereo images from Middlebury Teddy dataset

One of the issues is the inconsistency in two points of stereo images that need to be matched. Many attributes contribute to the inconsistency problem because the intensity and colour are different according to the perspective of images that may cause by the other camera sensor characteristics because of camera sensor characteristics or image noise, vignetting, and slightly different settings. Radiometric pre-calibration can only compensate for some of these differences and is not possible in all situations. Among other differences that contributed are from the non-Lambertian surfaces of the reflected light from the camera's different viewing angle. It is quite challenging to determine an accurate stereo correspondence with stereo image pair's viewpoint that consists of exposure to poor conditions and illumination variations [10]. The differences of viewpoint can be minimized by implementing a smaller stereo baseline; however, it will decrease the accuracy of geometric reconstruction. To obtain an accurate depth estimation, this constraint needs to be solved.

It is not easy to estimate accurate correspondences if each viewpoint image is captured under varying illumination and exposure conditions. An implementation of cross-scale pixel matching based on distance transform can improve these issues [11]. The researcher stated the disparity map output from this algorithm is robust at the boundary and low texture regions. Census transform is a noise-sensitive algorithm due to the algorithm executes brightness comparison between the center pixel with neighbouring pixels within a matching window [12]. The researcher proposed the upgraded Star-census transform executed with a symmetrical pattern of brightness comparison between the center pixel with neighbouring pixels divided by a

certain distance along with the matching window. Lim and Lee [13] proposed a fast stereo and robust Patchmatch-based matching under radiometric differences to acquire the disparity measurement from stereo images. That work showed the proposed framework outperforms better than conventional framework by 3.35 per cent of bad pixel errors and 4.71-27.24 times faster, separately. An algorithm utilises an adjustment to the pixel differences'; absolute differences (AD) with iterative guided filter also proposed by [14] focused on radiometric distortions improvement. An algorithm that exploits AD differences is linked together with CT also been developed to minimise the illumination variations effects. The increase of illumination differences contributes to the error matching rate increasing quickly, and thus affecting the disparity map accuracy.

3. PROPOSED FRAMEWORK

Fundamentally, the proposed stereo matching algorithm has been developed is illustrated in Figure 2. This framework utilises the disparity estimation of modified matching cost with dynamic histogram and census signature in the matching cost computation to produce a 2D depth map compared with a traditional framework by [15]. To accomplish the objective of this work, some processes will be conducted. The proposed work methodology consists of five main necessary stages, pre-processing, matching cost, aggregation of cost, disparity optimisation, and final disparity refinement. This framework of stereo images is pre-processed with the 1D Gaussian filter size 1 x 3 to smooth the original images as in (1):

$$G_f(x) = \exp \frac{-x^2}{2\sigma^2} \quad (1)$$

where x is the relative coordinate to the center of kernel and σ is the filter standard deviation.

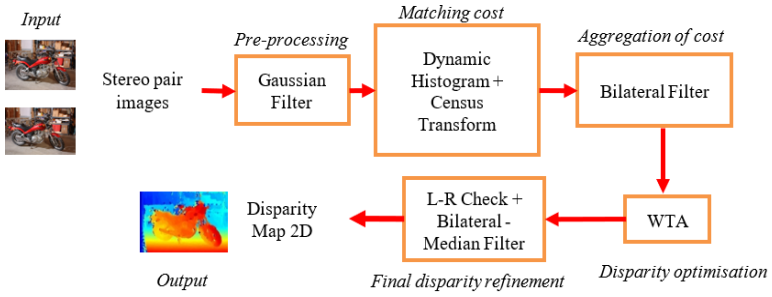

Figure 2. The framework blocks of the developed algorithm

The algorithm generates sub-histogram partition, the gray level allocation for dividing each sub-histogram and employing equalisation [16] in the matching cost computation. The partitions of the sub-histogram are based on local minima parameters. Again, a 1D Gaussian filter size 1x3 is utilised on the main histogram to eliminate unnecessary minima. Afterwards, it creates partitions of sub-histogram based on the points that fall between two local minima, which by default the non-zero histogram of first and last are measured as minima [17]. The partitions will shun domination by some part of the histogram to others. Each sub-histogram is allocated with a specific range of gray level values that spanned an output image's histogram. The gray level allocation is determined based on the gray level span ratio in each occupied sub-histogram. This will avoid over or under enhancement to the stereo image because equalisation works separately on each sub-histogram, as in (2).

$$Y(x,y) = \sum_{k=0}^{i-1} range_{k+1} + (\sum_{k=i}^{i} range_k - \sum_{k=i}^{i-1} range_{k+1}) \; x \; \sum_{k=1}^{i} \frac{n_k}{M} \quad (2)$$

Where the Y (x, y) the output dynamic histogram.

The *range* presents the dynamic range for the histogram while n_k is the quantity of pixels with the intensities of k, and M is the total of pixels. The approach avoids both gray levels of dissimilar sub-histograms mapped to the identical value of gray level in the image output mainly due to the sequential, non-overlapping and specific grey level ranges. So, there is no loss of important information in image details [18]. CT is a non-parametric local transform formula with a local intensity relationship between center pixel and near pixels inside a matching window [19]. The matching window is defined as (2u+1) x (2v+1), so the equation can be presented as (3) and (4) from the Y (x, y) for both reference and target image:

$$T(x,y) = \bigotimes_{i=-u}^{u} \bigotimes_{i=-v}^{v} \xi(I(Y(x,y)), I(Y(x+i,y+i))) \quad (3)$$

$$\xi(q_1,q_2) = \begin{cases} 1, if\ q_1 > q_2 \\ 0, if\ q_1 \leq q_2 \end{cases} \quad (4)$$

where the intensity from the gray level of the histogram for the pixel in Y (x, y) represents by I(Y(x,y)) while the $\otimes$ designates a bitwise catenation, the ξ is the relationship function and can be well-defined where q_1 is the centre pixel while q_2 is the closest neighbour pixel. Then, the Hamming distance is calculated using the Hamming function to determine the differences between both transform vectors that expressed as (5).

$$C_d(x,y) = Hamm(T_l(x,y), T_r(x-d,y)) \quad (5)$$

Where d is the disparity and T_l, T_r represents the left and right transform image vectors.

A pixel's correspondence can be computed by measuring the cost for entire candidates in a window between the target image and reference image. The target image provides the best cost of the window compared with the corresponding reference image. A fixed window applied using the box filter method [20]. The size of the window is established to (2z+1) x (2z+1) so that the value of cost aggregation can be determined using:

$$T^r(x,y,d) = \frac{1}{2z+1} \sum_{u=-z}^{z} C_d(x+u,y,d) \quad (6)$$

$$A^r_{square}(x,y,d) = \frac{1}{2z+1} \sum_{v=-z}^{z} T^r(x,y+v,d) \quad (7)$$

Where the cost matrix is T^r works to store the intermediate outcomes. It assumes the stereo pair images are well rectified, showing the corresponding epipolar lines are on the same height and horizontal. The cost value will be aggregated with the bilateral smoothing filter as in (8).

$$CA(p,d) = \sum_{q \in W_p | d(q)=d_z} A^T_{square}(x,y,d) \quad (8)$$

Where d_z indicates the disparity range and wp represents the size of the matching window, including z x z radius at a centred of p pixel.

This work determines the disparity selection after cost aggregation to minimise each pixel of aggregated corresponding value utilising the WTA approach to obtain the precise disparity map. The WTA approach for local methods can improve the computational cost as applied by [21]. Based on their works, the disparity maps output during this stage still producing errors in the textureless and occluded regions. The expression of WTA is delivered by (9).

$$d = \arg min_{d \in d_z} CA\ (p,d) \quad (9)$$

Where the minimum cost of d from cost aggregation is selected, the cost aggregation signifies the range of allowable disparity values.

The final stage of this work comprises the post-processing step defined as final disparity refinement. This approach is worked from the reference disparity map of the image, which corresponds to the right disparity map's target image. Thus, utilising the similar strategy employed by [22], the location of point p for disparity validation map is expressed by (10).

$$d_0(p) = \begin{cases} 1, if\ \left| d_{LR}(p) - d_{RL}(p - d_{LR}(p)) \right| \leq \tau_{LR} \\ 0, otherwise \end{cases} \quad (10)$$

Where d_{RL} and d_{LR} signify reference and target of the disparity maps. The map comprises the disparity location of valid (i.e., 0=inlier) and invalid (i.e., 1=outlier). In this work, the τ_{LR} is fixed to zero that have a similar setting with [23]. The aim is to produce the minimum error in the final map. Unnecessary noises can be removed using the adaptive bilateral median filter (ABMF) of B(p;q). The ABMF was employed as in [24] as in (8). Their work attained a great accuracy level on the removal of noise. Corresponding to (8) and (10), the value of weighted B(p, q) is adjusted to a summation of the histogram (p, d_z) that contributes to (11).

$$h(p,d_z) = \sum_{q \in W_p | d(q)=d_z} B(p,q) \quad (11)$$

Where d_z indicates disparity range and w_p represents the size of matching window, including the z x z radius at the centred of p pixel. The median of h(p, d_z) valued finalised the disparity value d′ given by (12):

$$d' = med\{d|h(p, d_z)\} \quad (12)$$

The algorithm's quantitative and qualitative performance analysis will be employed using the Middlebury vision benchmark dataset by computing the bad pixel error as in (13). Middlebury dataset was established by [4], that comprises 15 training set images. The level assessment of accuracy for each image depends on two parameters (i.e., bad pixels among all pixels in non-occluded areas (nonocc) and all pixels detected as valid pixels (all). The lower bad pixel percentage for the stereo matching algorithm will better disparity map accuracy, allowing the algorithm's accuracy to be assessed objectively [25].

$$B = \frac{1}{N}\sum_{(x,y)}(|d_c(x,y) - d_T(x,y)| > \delta_d) \quad (13)$$

Where δ_d (eval bad thresh) is a disparity error tolerance, for this work, we used $\delta_d = 1.0$. N is the total of pixels.

4. EXPERIMENT RESULTS AND DISCUSSION

All experiments for this work are performed using the Window 10 platform on desktop personal computer (PC) with Intel Xeon 2.6 GHz processor with 64 GB memory. The Middlebury dataset was documented from real surroundings of indoor and outdoor utilising a stereo vision architecture. Therefore, the architecture and system include very complicated and challenging images for framework evaluation. Figure 3 shows the comparison disparity map matching result on the Middlebury images 'Motorcycle', 'Playtable' and 'Vintage' images, Figures 3 (a) Reference image, (b) Ground-truth, (c) binary stereo matching (BSM), (d) intrinsic curve stereo matching (ICSG), (e) adaptive support weight with guided filter (ISM), (f) adaptive support weight and guided filter (ADSR_GIF), (g) removed normal cross correlation (R-NCC), (h) Proposed framework between several established frameworks such as binary stereo matching (BSM) [26], intrinsic curve stereo matching (ICSG) [27], gradient matching with iterative guided filter (ISM) [28], adaptive support weight with guided filter (ADSR_GIF) [29] and modified network coordination centre (R-NCC) [30]. The object scenes located at the depth are designated step by step, increasing the disparity values based on the final value from closer to farther according to colours assignment.

Although the disparity map for the developed framework is not entirely smoothed than ISM and ADSR_GIF, the uniform and border surface area produces better matching detection. The low texture that contributes mismatching process caused by the plain colour and textureless surface regions and sometimes also produced a constant luminance in large areas is marked in the red box compared with other frameworks shown in Figure 3. The disparity map also shows no streak artefacts, which usually occur in the local method stereo matching framework. The framework performed poorly at low texture regions and occlusion areas, as shown around the yellow box marking. Consequently, it is more difficult and very tricky to develop the algorithm in larger low texture regions due to pixel intensities' likeness to each other. By visually evaluating the developed algorithm's disparity map, the framework able to improve the edge-preserving properties in the disparity map and robust against the illumination variation and discontinuities areas. The disparity maps result has proved the algorithm framework improved the disparity map accuracy compared with other state-of-the-art methods.

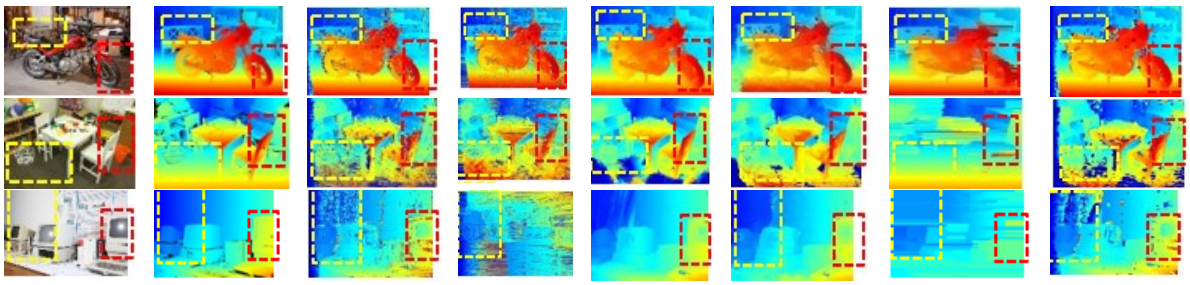

Figure 3. Disparity map comparison for Middlebury dataset 'Motorcycle',' Playtable' and 'Vintage' images, (a) reference image, (b) ground-truth, (c) BSM, (d) ICSG, (e) ISM, (f) ADSR_GIF, (g) R-NCC, and (h) proposed framework

Table 1 presents the Middlebury quantitative evaluation of the developed algorithm compared with several other frameworks. The accomplishment of the proposed framework from the experiment has been measured and compared with other local frameworks in the Middlebury Dataset. The outcomes show the framework is among the lowest average errors that signify the proposed accuracy accomplishment competitiveness. All the compared algorithm except the proposed algorithm developed without the pre-processing stage and dynamic histogram matching cost. It indicates that the developed framework is placed at the top of the comparison table, producing 41.0% and 33.0% of all and nonocc errors-the second method contributed by ADSR_GIF and trailed by ISM, ICSG, BSM and R-NCC. The developed framework decreases in all errors with 5%-10% and nonocc errors with 3%-5%, respectively, compared to the other frameworks. The framework also shows a promising result of bad pixel error for images containing illumination variations and radiometric differences such as 'Recycle', 'Teddy' and 'Vintage'. The pre-processing stage with the matching costs in the proposed framework is robust against the illumination variations, recognizing the regions with different intensity and brightness. The modified dynamic cost of matching cost computation presents a fine contour of disparity levels with the lowest noise displayed on the image. Additionally, the bilateral and median filter adoption also increased the efficiency and preserved the edges of an object.

Table 1. Bad pixel error percentage for nonocc and all under Middlebury dataset

Method	BSM		ICSG		ISM		ADSR_GIF		R-NCC		Proposed	
	Nonocc %	All %	Nonocc %	All %	Nonocc %	All %	Nonocc %	All %	Nonocc %	All %	Nonocc %	All %
Adiron	59.8	39.0	48.7	50.9	37.3	40.0	43.6	40.0	26.2	54.3	29.7	35.2
ArtL	25.8	43.9	15.1	27.2	28.8	37.1	18.6	31.1	14.8	35.3	22.8	38.5
Jadepl	27.9	54.9	44.9	56.4	45.5	54.4	36.7	49.1	30.2	61.9	38.2	52.0
Motor	38.9	38.3	29.1	34.4	32.1	36.3	24.6	39.3	30.9	25.7	27.4	34.9
MotorE	60.6	37.1	25.2	30.2	34.3	38.5	58.6	45.1	72.9	30.2	26.1	33.8
Piano	33.3	43.9	37.3	41.2	44.7	48.1	22.8	42.3	41.6	46.2	35.6	40.5
PianoL	46.9	64.2	53.6	56.4	53.1	56.1	56.3	62.2	77.7	61.2	47.6	51.5
Pipes	37.3	42.2	31.9	41.8	34.2	43.1	49.7	35.1	64.1	46.7	26.7	38.7
Playrm	26.3	54.0	49.6	55.8	44.6	50.9	18.7	48.1	27.4	67.4	43.3	50.9
Playt	64.8	61.5	62.3	65.3	59.8	62.3	56.0	55.6	59.1	45.0	52.0	56.9
PlaytP	51.5	46.9	30.5	33.3	44.5	48.5	48.5	42.1	71.9	36.8	34.1	40.9
Recyc	42.6	39.1	44.9	47.3	38.1	40.5	32.2	39.1	50.9	47.6	30.7	35.7
Shelvs	45.2	60.9	60.1	61.2	51.8	53.0	24.5	55.2	33.9	53.4	57.1	59.7
Teddy	42.8	24.5	17.5	24.2	22.0	27.3	36.3	17.9	78.2	24.7	15.5	24.3
Vintge	66.6	59.9	68.1	70.0	55.0	57.7	79.1	61.8	80.8	81.5	52.4	56.2
Avg	41.5	44.8	37.7	43.3	39.5	44.3	37.1	41.8	48.4	45.1	33.0	41.0

5. CONCLUSION

A new local framework of stereo matching based on improved modified dynamic cost was explained in this paper, mainly on the stereo correspondence illumination variations issue, the methodology and disparity map performance. The algorithm consists of five steps: Pre-processing using a Gaussian filter and a modified dynamic histogram with a census signature of matching cost with adaptive bilateral filter as cost aggregation. WTA optimisation is used as disparity selection, while consistency check for left and right with combination adaptive bilateral median filter is used as the final disparity refinement. The disparity map obtained shown better performance around the uniform and border surface. The accuracy accomplishment is competitive compared with several other state-of-the-art frameworks from the Middlebury evaluation platform. In the future, a local texture histogram analysis would be adopted to the proposed algorithm, which will improve the low texture region effectively.

ACKNOWLEDGEMENTS

We are grateful to the ministry of higher education (MOHE), Malaysia, Universiti Teknikal Malaysia Melaka (UTeM), and the centre for research and innovation management (CRIM) through the research grant of FRGS/1/2020/FTKEE-CACT/F00451 for their kind and helpful to complete this work.

REFERENCES

[1] H. Liu, R. Wang, Y. Xia, and X. Zhang, "Improved cost computation and adaptive shape guided filter for local stereo matching of low texture stereo images," *Applied Sciences*, vol. 10, no. 5, pp. 1869(1)-1869(17), 2020, doi: 10.3390/app10051869.

[2] R. A. Hamzah and H. Ibrahim, "Literature survey on stereo vision disparity map algorithms," *Journal of Sensors*, vol. 2016, pp. l-22, 2016, doi: 10.1155/2016/8742920.
[3] I. Cabezas, V. Padilla, and M. Trujillo, "A measure for accuracy disparity maps evaluation," in *Iberoamerican Congress on Pattern Recognition*, vol 7042, pp. 223-231, 2011, doi: 10.1007/978-3-642-25085-9_26.
[4] D. Scharstein, *et al.*, "High-resolution stereo datasets with subpixel-accurate ground truth," in *German conference on pattern recognition*, vol 8753, pp. 31-42, 2014, doi: 10.1007/978-3-319-11752-2_3.
[5] B. Jia, S. Liu, and Z. Du, "A progressive framework for dense stereo matching," *Pattern Recognition and Image Analysis*, vol. 26, no. 2, pp. 294-301, 2016, doi: 10.1134/S1054661816020036.
[6] E. Bebeselea-Sterp, R. R. Brad, and R. R. Brad, "A Comparative Study of Stereovision Algorithms," *International Journal of Advanced Computer Science and Application (IJACSA)*, vol. 8, no. 11, pp. 359-375, 2017.
[7] R. R. Basaru, C. Child, E. Alonso, and G. Slabaugh, "Quantized Census for Stereoscopic Image Matching," *2014 2nd International Conference on 3D Vision*, 2014, pp. 22-29, doi: 10.1109/3DV.2014.83.
[8] M. Mahato, S. Gedam, J. Joglekar, and K. M. Buddhiraju, "Dense Stereo Matching Based on Multiobjective Fitness Function-A Genetic Algorithm Optimization Approach for Stereo Correspondence," in *IEEE Transactions on Geoscience and Remote Sensing*, vol. 57, no. 6, pp. 3341-3353, June 2019, doi: 10.1109/TGRS.2018.2883483.
[9] G. Li, X. Zhang, C. Li, H. Jin, and J. Zhao, "Design and application of parallel stereo matching algorithm based on CUDA," *Microprocessors Microsystems*, vol. 47, pp. 142-150, Nov. 2016, doi: 10.1016/j.micpro.2015.09.006.
[10] Yong-Jun Chang and Yo-Sung Ho, "Adaptive Pixel-wise and Block-wise Stereo Matching in Lighting Condition Changes," *Journal of Signal Processing Systems*, vol 91, no. 11, pp. 1305-1313, February 2019, doi: 10.1007/s11265-019-1442-7.
[11] Yong-Jun Chang and Yo-Sung Ho, "Disparity map enhancement in pixel based stereo matching method using distance transform," *Journal of Visual Communication and Image Representation*, vol. 40, pp. 118-127, October 2016, doi: 10.1016/j.jvcir.2016.06.017.
[12] J. Lee, D. Jun, C. Eem, and H. Hong, "Improved census transform for noise robust stereo matching," *Optical Engineering*, vol. 55, no. 6, pp. 063107(1)-063107(10), 2016, doi: 10.1117/1.OE.55.6.063107.
[13] J. Lim and S. Lee, "Patchmatch-Based Robust Stereo Matching Under Radiometric Changes," in *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 41, no. 5, pp. 1203-1212, 1 May 2019, doi: 10.1109/TPAMI.2018.2819662.
[14] 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.
[15] I. Gallo and E. Binaghi, "A new algorithm for dense two-frame stereo correspondence," *Int. Conf. Comput. Vis. Theory Appl.*, 2005, pp. 1-9.
[16] M. Abdullah-Al-Wadud, M. H. Kabir, M. A. A. Dewan, and O. Chae, "A Dynamic Histogram Equalization for Image Contrast Enhancement," in *IEEE Transactions on Consumer Electronics*, vol. 53, no. 2, pp. 593-600, May 2007, doi: 10.1109/TCE.2007.381734.
[17] B. Subramani and M. Veluchamy, "Quadrant dynamic clipped histogram equalization with gamma correction for color image enhancement," *Color Research & Application*, vol. 45, no. 4, pp. 644-655, April 2020, doi: 10.1002/col.22502.
[18] S. K. P. Kaur, "A Novel Brightness Preserving Histogram Equalization Technique for Image Contrast Enhancement," *International Journal of Science and Research (IJSR)*, vol. 5, no. 9, pp. 1305-1309, 2016.
[19] J. Lim, Y. Kim, and S. Lee, "A census transform-based robust stereo matching under radiometric changes," *2016 Asia-Pacific Signal and Information Processing Association Annual Summit and Conference (APSIPA)*, 2016, pp. 1-4, doi: 10.1109/APSIPA.2016.7820841.
[20] C. Ttofis, C. Kyrkou, and T. Theocharides, "A Low-Cost Real-Time Embedded Stereo Vision System for Accurate Disparity Estimation Based on Guided Image Filtering," in *IEEE Transactions on Computers*, vol. 65, no. 9, pp. 2678-2693, 1 Sept. 2016, doi: 10.1109/TC.2015.2506567.
[21] Y. Gan, R. A. Hamzah, and N. S. N. Anwar, "Local Stereo Matching Algorithm Based on Pixel Difference Adjustment, Minimum Spanning Tree and Weighted Median Filter," *2018 IEEE Conference on Systems, Process and Control (ICSPC)*, 2018, pp. 39-43, doi: 10.1109/SPC.2018.8704131.
[22] Y. Du and K. Jia, "Neighborhood correlation and window adaptive stereo matching algorithm," *Journal of Information Hiding and Multimedia Signal Processing*, vol. 10, no. 4, pp. 509-516, December 2019.
[23] G. A. Kordelas, D. S. Alexiadis, P. Daras, and E. Izquierdo, "Enhanced disparity estimation in stereo images," *Image and Vision Computing*, vol. 35, pp. 31-49, March 2015, doi: 10.1016/j.imavis.2014.12.001.
[24] W. Wu, L. Li, and W. Jin, "Disparity refinement based on segment-tree and fast weighted median filter," *2016 IEEE International Conference on Image Processing (ICIP)*, 2016, pp. 3449-3453, doi: 10.1109/ICIP.2016.7533000.
[25] D. Scharstein and R. Szeliski, "Middlebury Stereo Evaluation-Version 3," 2020. [Online]. Available: https://vision.middlebury.edu/stereo/eval3/.
[26] 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.
[27] M. Shahbazi, G. Sohn, J. Théau, and P. Ménard, "REVISITING INTRINSIC CURVES for EFFICIENT DENSE STEREO MATCHING," vol. 3, pp. 12-19, July 2016, doi: 10.5194/isprs-annals-III-3-123-2016.
[28] R. A. Hamzah, A. F. Kadmin, M. S. Hamid, S. F. A. Ghani, and H. Ibrahim, "Improvement of stereo matching algorithm for 3D surface reconstruction," *Signal Processing Image Communication*, vol. 65, pp. 165-172, July

2018, doi: 10.1016/j.image.2018.04.001.
[29] K. Lingyin, Z. Jiangping, and Y. Sancong, "Stereo Matching Based on Guidance Image and Adaptive Support Region," *Guangxue Xuebao/Acta Optica Sinica*, vol. 40, no. 9, pp. 0915001-1-0915001-13, May 2020.
[30] Y. Li and S. Fang, "Removal-based multi-view stereo using a window-based matching method," *Optik*, vol. 178, pp. 1318-1336, February 2019, doi: 10.1016/j.ijleo.2018.10.126.

BIOGRAPHIES OF AUTHORS

Ahmad Fauzan Kadmin currently a Chartered Engineer and Professional Technologist attached with UTeM as a researcher; he has over 16 years of experience in the electronic & computer engineering field with a technical expert in R&D engineering, computer vision & medical electronics. He graduated with a Bachelor's Degree in Electronics Engineering from Universiti Sains Malaysia and a Master's Degree in Computer & Communication Engineering from Universiti Kebangsaan Malaysia. Previously, he worked with Megasteel Sdn. Bhd., Samsung SDI (M) Sdn. Bhd. and Agensi Angkasa Negara. He published several technical and engineering paperwork in computer vision and medical processing.

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.

Nurulfajar Abd Manap earned a Master of Electrical Engineering in image processing (2002, Universiti Teknologi Malaysia) and Bachelor of Electrical Engineering (2000, Universiti Teknologi Malaysia). He obtained his PhD (Image and Video Processing) at the University of Strathclyde, Glasgow, in 2012. He teaches various electronics engineering subjects and courses in his affiliation for undergraduate and postgraduate levels. His subjects include Advanced Digital Signal Processing, Computer Vision and Pattern Recognition, Data Structures, Multimedia Technology & Applications and Distributed & High-Performance Computing. He also one of the MIET Chartered Engineer and Apple Distinguished Educators (ADE). His research interests are 3D image processing, stereo vision and video processing.

Ts. Mohd Saad Hamid received his B. Eng degree majoring in Computer from Multimedia University. He completed his M. Eng majoring in Computer and Communication from The National University of Malaysia in 2014. Previously, he worked at Telekom Malaysia Berhad and Continental Automotive Components (M) Sdn Bhd. Currently, he is pursuing his PhD at Universiti Teknikal Malaysia Melaka. His research interests are computer vision, deep learning, and image processing and embedded systems.