

A study of edge preserving filters in image matching

Rostam Affendi Hamzah, A. F. Kadmin, S. F. A. Gani, KA Aziz, T. M. F. T. Wook,
N. Mohamood, M. G. Y. Wei

Fakulti Teknologi Kejuruteraan Elektrik & Elektronik, Universiti Teknikal Malaysia Melaka, Malaysia

Article Info

Article history:

Received Aug 2, 2019

Revised Dec 12, 2019

Accepted Jun 26, 2020

Keywords:

Bilateral filter
Edge preserving filter
Image matching
Median filter
Stereo vision

ABSTRACT

This article presents a study on edge preserving filters in image matching which comprises a development of stereo matching algorithm using two edge preserving filters. Fundamentally, the framework is reconstructed by several sequential processes. The output of these processes is a disparity map or depth map. The corresponding points between two images require accurate matching to make accurate depth map estimation. Thus, the propose work in this article utilizes sum of squared differences (SSD) with dual edge preserving filters. These filters are used due to edge preserved properties and to increase the accuracy. The median filter (MF) and bilateral filter (BF) will be utilized. The SSD produces preliminary results with low noise and the edge preserving filters reduce noise on the low texture regions with edge preserving properties. Based on the experimental analysis using the standard benchmarking evaluation system from the Middlebury, the disparity map produced is 6.65% for all error pixels. It shows an accurate edge preserved properties on the disparity maps. To make the proposed work more reliable with current available methods, the quantitative measurement has been made to compare with other existing methods and it displays the proposed work in this article perform much better.

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 & Elektronik,
Universiti Teknikal Malaysia Melaka,
Taman Tasik Utama, 75450 Ayer Keroh, Malacca, Malaysia.
Email: rostamaffendi@utem.edu.my

1. INTRODUCTION

Image matching is a corresponding process from one point on reference image to another point on the target image. The target image can be two or more images. In this article, the images are using from the stereo camera input which is also known as stereo images. The matching algorithm from the matching process produces disparity map. This map consists of depth information which is valuable for many applications such as robotics automation [1-2], face recognition [3], 3D surface reconstruction [4] and virtual reality [5]. The stereo baseline can be setup in a wide or short baseline [6] distance which depends on the applications. To determine the range or distance estimation, the triangulation function is applied to each of the pixel on the disparity map. Therefore, to get an accurate result, the matching process requires complex and challenging solution for depth or distance estimation. It requires precise function on the propose framework. Basically, from the Szeliski and Scharstein [7], stereo matching process can be developed by multiple stages. First, matching cost computation to get the preliminary disparity map. Second, the cost aggregation stage to filter the preliminary disparity map. Third stage, the optimization process to normalize the disparity values on the reconstructed map. The last stage is pre-processing which contains several consecutive sub processes to filter the final disparity map.

The development of stereo matching, there are two major approaches available in developing the algorithm framework. It is local methods as published in [8-10] and global method [11]. Mostly local methods use local properties or local contents using windows-based technique such as fixed windows implemented in [12-13], adaptive window [14], convolution neural network [15] and multiple windows [16]. In common, Winner-Takes-All (WTA) strategy is applied for local based optimization. It is low computational complexity and fast execution time [17-19]. Local method such implemented in [20] that used fitting the plane to increase the accuracy at the final stage. This method also known as RANSAC that efficiently works on the low textured areas. However, the error still occurred on the object edges. Their method requires several iterations for plane fitting process. If wrong iterations, then it will affect the results. Commonly, local methods show fast running time, but low accuracy on the edges due to improper selection of windows sizes. Hence, to get an accurate result for the local approach is a challenge to the researchers.

Another approach in stereo matching algorithm to produce the disparity map is global optimization method. Fundamentally, this method uses energy-based function which is known as Markov random field (MRF). The method in global optimization such as belief propagation (BP) [21] and graph cut (GC) [22] produce accurate results. Each pixel of interest calculation requires all pixel's energy in disparity map. It calculates neighboring or nearby pixels using maximum flow and the selection is made based on the minimum cut-off energy on the disparity map. The algorithms implemented using global optimization approach normally involve high computational requirement due to all pixel's energy calculation and absorption. Global methods involve iterations which increase the execution time each disparity map reconstruction. This article aims to produce accurate results and competitive with some established methods. The first function or stage will be implemented using improved Sum of squared differences (SSD) [23]. Then, the second stage utilizes the edge preserving filter which is known as bilateral filter (BF) [24]. This filter is capable to remove noise and preserved object edges. The third stage is optimization based on WTA strategy. Last stage, the median filter (MF) is applied to remove unwanted or remaining invalid pixels. The MF is also one of the edges preserving filter where it is capable to increase the accuracy at the object edges.

2. RESEARCH METHOD

The framework of the proposed algorithm is displayed on Figure 1. The stereo matching algorithm starts with STEP 1 to get the preliminary disparity map. The improved SSD has been proposed which the weighted technique is used on the block matching process. The technique should improve the effectiveness of the corresponding process. Then at STEP 2, the BP is utilized to reduce the noise and preserved the object edges. The BP is capable to efficiently remove noise on the low texture regions and sharpening the object boundaries. The optimization uses WTA strategy which this method normalizes the floating points numbers and selects minimum disparity values on the disparity map. Final stage at STEP 4 is also using one of the edges preserving filters known as MF. This filter is a type of nonlinear filter and capable to improve final disparity map.

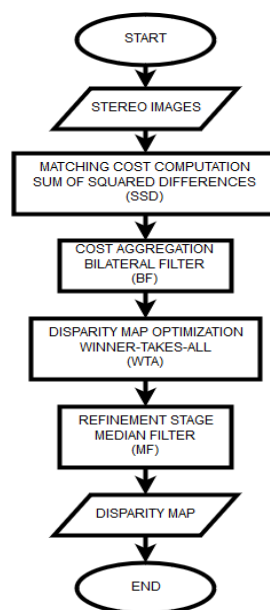


Figure 1. A flowchart of the proposed algorithm

2.1. Matching cost computation

The first stage of the proposed framework is using the weighted SSD. The preliminary disparity map is produced at this stage. Hence, robust function must be used to increase the effectiveness on the disparity map. The problem on matching process at this stage on the low texture regions must be at minimum. The weighted is proposed at SSD to improve the values on the low texture. Thus, the consistency of the weight at the low texture region is capable to make the matching process accurate and reduces the mismatch or invalid pixels. In (1) shows the weighted SSD which the RGB images are used as an input;

$$SSD(x, y, d) = \frac{1}{M} \sum_{(x,y) \in M} |I_l^i(x, y) - I_r^i(x - d, y)|^2 \quad (1)$$

where the left and right images are I_l and I_r respectively, the coordinates of disparity d is represented by (x, y, d) , i denotes the RGB channels of left and right images and the SSD window with M is the weighted with the window size at (13×11) .

2.2. Cost aggregation

This second stage more likely to filter the preliminary disparity map from stage one. Normally the preliminary disparity map contains high noise and it must be removed. Some of invalid and uncertainties pixels are constructed during the matching process. Hence, at this stage the filter must be robust and is capable to remove high noise of invalid pixels and preserved the object boundaries. The BP is used due to strong preserving object edges and at the same time efficient to remove high noise especially on the plain color and low texture regions. In (2) is the BF function used in this article.

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

where p is the location pixel of interest at (x, y) , w_B and q are window size of BF and neighboring pixels respectively. The σ_s denote a factor of spatial adjustment and σ_c equals to similarity factor for the color detection. The $p - q$ is spatial Euclidean interval and $I_p - I_q$ denotes the Euclidean distance in color space. Hence, (3) is the cost aggregation function of BF with the matching cost computation input.

$$C(x, y, d) = WM_{x,y,q}^{BF} SSD(x, y, d) \quad (3)$$

2.3. Disparity optimization

This stage optimizes the disparity values on disparity map. The normalization is based on the minimum disparity values with the floating-point number which the WTA is selected in this article. The WTA is normally being used in the local based methods due to fast implementation [17, 25]. The WTA function is given by (4).

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

where D represents a set of valid disparity values for an image and $C(x, y, d)$ denotes the second stage of aggregation step. Fundamentally, after this stage the disparity map still contains noise or invalid pixels. Thus, this map needs to be improved and the last stage is will remove remaining noise.

2.4. Refinement stage

The last stage of the algorithm framework is known as refinement or post processing stage. It has several continuous processes which start with handling the occlusion regions, filling the invalid pixels and filtering final disparity map. The left-right consistency checking process is conducted to identify occlusion areas and some invalid pixels. Then, these invalid pixels are restored with valid pixel values through the filling process. Some of artifacts and unwanted pixels will be removed using the MF and at the same time preserved the object boundaries. The MF smoothes the final disparity map as indicates by (5).

$$d_f = \text{median}\{d_{x,y} | (x, y) \in w_{MF}\} \quad (5)$$

where w_{MF} denotes the MF window size and d_f is the final disparity value at the location of (x, y) .

3. RESULTS AND DISCUSSION

The results of disparity maps are evaluated based on the grayscale values. The darker object or intensity on the disparity map determines the far away that object to the sensor. Higher grayscale values indicate the objects are closer to the sensor. The experimental analysis has been executed on a computer with Windows 10, 3.2GHz and 8G RAM. The input images are from the Middlebury stereo evaluation dataset [25] which contains 15 standard images and must be submitted online. These images are very complex, and each image consists of different characteristics and properties such as light settings objects depth, incoherence regions, different resolutions and low texture areas. The values of $\{M, \sigma_s, \sigma_c, w_B, w_{MF}\}$ are $\{13 \times 9, 17, 0.3, 9 \times 9, 13 \times 13\}$.

Figure 2 shows a sample disparity map result of the Adirondack image. It has without and with the proposed framework. Image with the proposed work produces more accurate result with reduced noises. The shapes of the object are sharper with preserved edges. Based on this Figure 2, accurate disparity map is formed and reconstructed with clear contours level of depth estimation. Figure 3 displays the disparity map results from the Middlebury. It contains 15 images which the accuracy is determined by all (all) error and non-occluded (nonocc) error. The all error is evaluated based on all pixels on an image and non-occluded represented by the pixel without the occluded region of disparity map within these 15 images, Pipes and Jadeplant images are the most difficult images to be matched. These images comprise several piping lines and leaves with different sizes respectively. Yet, the propose algorithm can reconstruct almost accurate disparity map with clear discontinuities regions.

Fundamentally, real images from the Middlebury are difficult and very challenging to get an accurate corresponding point. It was developed to test the robustness of an algorithm where same corresponding point maybe contains different pixel values. Additionally, each image contains difference characteristics such as plain color objects, shadow, discontinuity regions and occluded areas. With referring to Figure 3; the disparity maps of low texture surfaces such as Motorcycle, MotorcycleP, Playtable and PlaytableP are well recreated with different depth and disparity contour. Other regions difficult to be matched are plain colour objects and shadow such as images of ArtL, Recycle, Piano and PianoL. These regions consist of similar pixel values and possibility to get wrong matching is very high.

The disparity maps from the proposed work display almost accurate matching for these images. It shows that the proposed work is able to get correct matching pixels on these regions and robust against the plain colour areas. The quantitative measurement is tabulated in Table 1 and 2. These results are produced by the Middlebury online benchmarking evaluation system with two errors attributes as explained above. Some established methods are also included in these Tables to show the competitiveness of the proposed work. Overall, an average error measurement is assessed to rank the best results. For Table 1, the proposed method is ranked at top of the table with 6.65%, and Table 2 with 9.32%. It shows that the proposed work is competitive with other recently published methods and can be used as a complete algorithm.

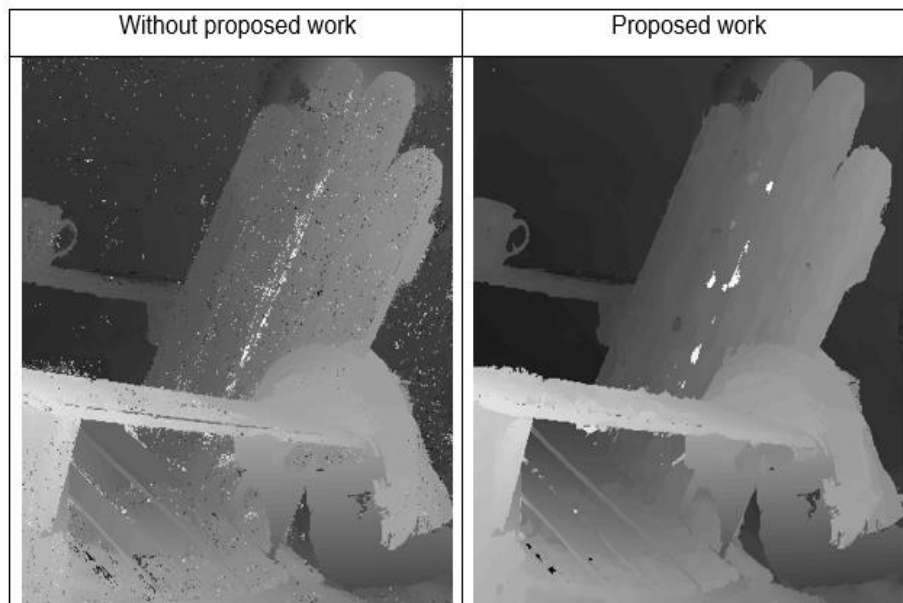


Figure 2. A sample disparity map result to show the efficiency of edge preserving filter on the Adirondack

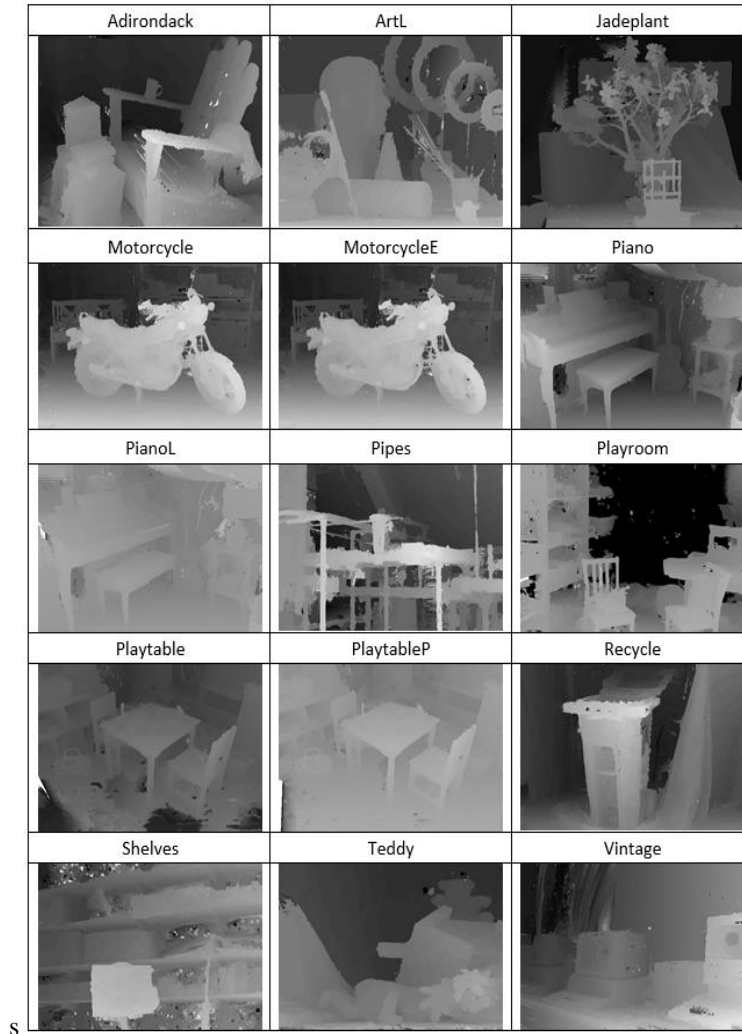


Figure 3. These 15 images are the results from the proposed algorithm using the Middlebury standard benchmarking evaluation system

Table 1. Performance comparison of *nonocc* error from the Middlebury images

Algorithms	Adiron	ArtL	Jadepl	Motor	MotorE	Piano	PianoL	Pipes	Playrm	Playt	PlayP	Recy _c	Shelvs	Teddy	Vintge	Ave
Proposed Algorithm	3.55	4.39	12.45	2.98	2.91	5.34	10.10	6.88	5.57	24.76	5.89	3.77	8.84	2.32	8.87	6.65
SNCC [26]	2.89	4.05	18.10	2.68	2.52	3.52	7.08	6.14	5.64	45.40	3.13	2.90	7.59	1.58	13.50	6.97
ELAS [27]	3.09	4.72	29.70	3.28	3.29	4.30	8.31	5.61	6.00	21.80	2.84	3.09	9.00	2.36	10.90	7.22
MPSV [17]	3.83	6.00	19.70	5.85	5.53	5.68	34.30	9.59	5.86	15.30	4.20	4.59	13.00	3.70	14.30	8.81
ADSM [18]	13.30	6.10	15.00	3.67	5.67	7.08	20.60	6.57	13.20	23.10	3.55	5.76	17.20	3.05	10.10	8.95
DoGGuided [19]	15.20	9.57	27.10	5.64	8.31	8.09	32.40	9.67	14.00	24.50	5.32	5.56	16.20	4.15	15.00	12.00
BSM [28]	7.27	11.40	30.50	6.67	6.52	10.80	32.10	10.50	12.50	24.40	12.80	7.42	16.40	4.88	32.80	13.40

Table 2. Performance comparison of *all* error from the Middlebury images

Algorithms	Adiron	ArtL	Jadepl	Motor	MotorE	Piano	PianoL	Pipes	Playrm	Playt	PlayP	Recy _c	Shelvs	Teddy	Vintge	Ave
Proposed Algorithm	4.54	7.65	31.90	6.23	4.56	5.65	10.99	12.34	7.75	25.99	6.34	4.78	9.87	3.54	9.83	9.32
SNCC [26]	3.63	6.78	39.80	5.12	5.11	4.65	8.23	11.80	8.05	45.60	4.36	3.29	8.10	2.55	14.80	10.40
ELAS [27]	4.08	7.18	52.80	5.39	5.45	4.96	9.00	10.70	7.94	23.20	3.83	3.78	9.46	3.34	11.60	10.60
MPSV [17]	5.87	9.43	40.20	9.11	8.80	7.03	34.20	15.80	8.58	16.90	5.89	6.78	13.70	4.82	16.80	12.70
ADSM [18]	14.30	10.60	34.10	6.00	8.00	7.37	20.40	12.10	16.90	25.50	5.84	5.83	17.20	4.11	11.10	12.30
DoGGuided [19]	20.10	28.00	56.50	13.80	16.80	13.40	37.30	23.80	30.30	30.80	13.00	9.13	19.00	13.40	23.60	22.30
BSM [28]	12.70	28.70	58.70	14.80	14.70	16.00	35.80	24.50	29.40	31.00	20.20	12.10	19.20	14.30	39.30	23.50

4. CONCLUSION

This article presented the stereo matching algorithm based on edge preserving filters. These filters were capable to remove the noise and preserve the object edges as shown in the results above. The framework started with weighted SSD at the matching cost function. Then, the BF was utilized at aggregation stage to filter high noise and conserve the object boundaries of the preliminary disparity map. The WTA strategy was implemented at the optimization stage to normalize the floating point numbers to the disparity values. The second edge preserving filter was used at the last stage of the proposed work using MF. This nonlinear filter removed remaining noise and increases the efficiency of final disparity map. Overall, these edge preserving filters used in the proposed framework were able to remove noise especially on the low texture regions and able to preserve the object edges as shown by Figure 2. The quantitative measurement from the standard benchmarking Middlebury system also demonstrated low average errors were produced by the proposed framework at 6.65% and 9.32% of non-occluded and all pixel errors respectively. The training images are shown by Figure 3.

ACKNOWLEDGEMENTS

This project was sponsored by a grant from the Universiti Teknikal Malaysia Melaka with the Number: JURNAL/2018/FTK/Q00008.

REFERENCES

- [1] W. Wu, H. Zhu, S. Yu and J. Shi, "Stereo matching with fusing adaptive support weights," *IEEE Access*, vol. 7, pp. 61960-61974, 2019.
- [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.
- [3] E. Winamo, 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.
- [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)*, Yogyakarta, pp. 119-124, 2016.
- [5] I. Vedamurthy, D. C. Knill, S. J. Huang, A. Yung, J. Ding, O. S. Kwon, D. Bavelier and D. M. Levi, "Recovering stereo vision by squashing virtual bugs in a virtual reality environment," *Philos Trans R Soc Lond B Biol Sci.*, vol. 371, no. 1697, p. 20150264, 2016.
- [6] 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.
- [7] D. Scharstein and R. Szeliski, "A taxonomy and evaluation of dense two-frame stereo correspondence algorithms," *International Journal of Computer Vision*, vol. 47, no. 2, pp. 7-42, 2002.
- [8] Q. Yang, "A non-local cost aggregation method for stereo matching," *2012 IEEE Conference on Computer Vision and Pattern Recognition*, Providence, RI, pp. 1402-1409, 2012.
- [9] A. Hosni, C. Rhemann, M. Bleyer, C. Rother and M. Gelautz, "Fast Cost-Volume Filtering for Visual Correspondence and Beyond," in *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 35, no. 2, pp. 504-511, 2013.
- [10] R.A. Hamzah and H. Ibrahim, "Improvement of stereo matching algorithm based on sum of gradient magnitude differences and semi-global method with refinement step," *Electronics Letters*, vol. 54, no. 14, pp. 876-878, 2018.
- [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)*, Stanford, CA, pp. 276-285, 2016.
- [12] Q. Liang, Y. Yang and B. Liu, "Stereo matching algorithm based on ground control points using graph cut," *2014 7th International Congress on Image and Signal Processing*, Dalian, pp. 503-508, 2014.
- [13] Q. Yang, P. Ji, D. Li, S. Yao, M. Zhang, "Fast stereo matching using adaptive guided filtering," *Image and Vision Computing*, vol. 32, no. 3, pp. 202-211, 2014.
- [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, 2013.
- [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)*, Boston, MA, pp. 1592-1599, 2015.
- [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, pp. 229-246, 2002.
- [17] S. F. A. Ghani, A. F. Kadmin, M. S. Hamid, R. A. Hamzah and K. H. Fen, "Development of a portable community video surveillance system," *International Journal of Electrical and Computer Engineering*, vol. 9, no. 3, pp. 1814-1823, 2019.
- [18] 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.

- [19] 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)*, Nagoya, pp. 258-261, 2017.
- [20] R. A. Hamzah, M. G. Y. Wei and N. S. N. Anwar, "Stereo matching based on absolute differences for multiple objects detection," *Telecommunication Computing Electronics and Control*, vol. 17, no. 1, pp. 261-267, 2019.
- [21] 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)*, Dallas, TX, pp. 236-241, 2016.
- [22] H. Mohammadi, F. Samadzadegan and P. Reinartz, "2D/3D information fusion for building extraction from high-resolution satellite stereo images using kernel graph cuts," *International Journal of Remote Sensing*, vol. 40, no. 15, pp. 5835-5860, 2019.
- [23] 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.
- [24] R. A. Hamzah, S. F. A. Ghani, A. Din and K. A. A. Aziz, "Visualization of image distortion on camera calibration for stereo vision application," *2012 IEEE International Conference on Control System, Computing and Engineering*, Penang, pp. 28-33, 2012.
- [25] S. Daniel and S. Richard, "Middlebury Stereo Evaluation-Version 3," 2019. [Online]. Available: <http://vision.middlebury.edu/stereo/eval/references>. [Accessed: Nov 2019].
- [26] N. Einecke and J. Eggert, "Anisotropic median filtering for stereo disparity map refinement," *2013 International Conference on Computer Vision Theory and Applications*, Barcelona, pp. 189-198, 2013.
- [27] A. Geiger, M. Roser and R. Urtasun, "Efficient Large-Scale Stereo Matching," *2010 Asian Conference on Computer Vision*, Queenstown, pp 25-38, 2010.
- [28] 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)*, Tsukuba, pp. 356-359, 2012.