Bulletin of Electrical Engineering and Informatics
Vol. 12, No. 1, February 2023, pp. 328~337
ISSN: 2302-9285, DOI: 10.11591/eei.v12i1.4122

Stereo matching algorithm for autonomous vehicle navigation using integrated matching cost and non-local aggregation

Madiha Zahari[1,2], Rostam Affendi Hamzah[2], Nurulfajar Abd Manap[1], Adi Irwan Herman[3]
[1]Fakulti Kejuruteraan Elektronik dan Kejuruteraan Komputer, Universiti Teknikal Malaysia Melaka, Melaka, Malaysia
[2]Fakulti Teknologi Kejuruteraan Elektrik dan Elektronik, Universiti Teknikal Malaysia Melaka, Melaka, Malaysia
[3]Product and Test Engineering, Texas Instruments Electronics Malaysia, Melaka, Malaysia

Article Info

Article history:

Received May 24, 2022
Revised Sep 5, 2022
Accepted Oct 12, 2022

Keywords:

Autonomous
Disparity
Matching cost
Stereo matching

ABSTRACT

Stereo matching algorithm plays an important role in an autonomous vehicle navigation system to ensure accurate three-dimensional (3D) information is provided. The disparity map produced by the stereo matching algorithm directly impacts the quality of the 3D information provided to the navigation system. However, the accuracy of the matching algorithm is a challenging part to be solved since it is directly affected by the surrounding environment such as different brightness, less texture surface, and different image pair exposure. In this paper, a new framework of stereo matching algorithm that used the integration of census transform (CT) and sum of absolute difference (SAD) at the matching cost computation step, non-local cost aggregation at the second step, winner take all strategy at the third step, and a median filter at the final step to minimize disparity map error. The results show that the accuracy of the disparity map is improved using the proposed methods after some parameter adjustment. Based on the standard Middlebury and KITTI benchmarking dataset, it shows that the proposed framework produced accurate results compared with other established methods.

This is an open access article under the CC BY-SA license.

Corresponding Author:

Madiha Zahari
Fakulti Teknologi Kejuruteraan Elektrik dan Elektronik, Universiti Teknikal Malaysia Melaka
Hang Tuah Jaya, Melaka, Malaysia
Email: madiha@utem.edu.my

1. INTRODUCTION

In recent years, autonomous vehicle navigation becomes a popular research topic in computer vision since it is able to navigate without human interaction [1]. To ensure accurate decision-making made by the autonomous vehicle, an accurate three-dimensional (3D) representation is required to ensure the systems analyze the correct environment information. The surrounding environment information can be detected by many methods such as using radar detection, laser imaging, detection and ranging (LIDAR), and global positioning system (GPS) [2]. However, this method required an expensive new device to be installed in the vehicle instead of a computer vision that used the normal vehicle camera. Normal vehicle cameras commonly produce two-dimensional (2D) images which are required to be converted to 3D images in order to ensure the autonomous vehicle system navigation able to function accurately. Depth valuation is the critical step in the process of converting a 2D view to a 3D view. The depth is determined by using a stereo vision system based on the triangulation principle where the depth value is directly influenced by the disparity value of the stereo images [3].

The depth is determined by calculating the disparity value at two equivalent points between pairs of images. The stereo matching is the process of calculating the disparity between two images [4]. In this process, a stereo matching algorithm is developed to calculate the disparity value of each corresponding point in the images where the output of this process is known as a depth map. The depth map generated by the

developed algorithm makes a significant contribution to the quality of the 3D representation of the image, and as the accuracy of the depth map improves, so does the accuracy of the depth data. The matching algorithm can be developed based on three main classifications which are global method, semi-global method and local methods [5], [6]. The global method calculated the disparity value by calculating the overall energy function of all pixels in the images. Global methods able to produce high accuracy of depth map but due to the computational complexity, this method will lead to a low-efficiency matching algorithm while local methods which calculate the disparity value based on the predetermined support window resulted a high efficiency of the matching algorithm [7]. Due to this matter, many real-time applications used the local method to develop their matching algorithm [8]. However, the drawback of local methods is the difficulty to find an accurate depth map for the image with less texture region and in the image with high radiometric distortion. The difficulty of locating the equivalent point between the image pair in both locations causes the likelihood of an error matching process occurs is significant. Although many studies had been done in the development of stereo correspondence algorithm framework to get an accurate output, but there is still no ideal answer to this problem. Aiming to improve the accuracy of the depth map, this paper proposed a new matching algorithm framework which focused on the local based methods.

The local method is based on the four basic steps which consist of cost-volume computation (CC), cost-volume aggregation (CA), cost-volume optimization (CO), and disparity refinement (DR) [6]. In CC, the basic methods such as sum of absolute differences (SAD), sum of squared difference (SSD) and normalized cross correlation (NCC) is commonly used since these methods are simple and straightforward. However, these approaches are acutely vulnerable to amplitude distortion, which has a negative impact on stereo matching precision [9]. The approach developed by Zabih *et al*. addressed as census transform (CT) and rank transform (RT) reported able to overcome the issue of radiometric distortion because that method are relying on the neighboring pixel intensity rather than the intensity value itself [9]. Therefore, SAD and SSD may be able to handle the matching uncertainty effectively for image areas with comparable local structures, whereas the non-parametric local transformations may handle the matching uncertainties well for image regions with similar colors [10]. Due to this, researchers began integrating multiple matching costs to increase the stereo matching algorithm's accuracy. The performance of the integrated matching cost provided improved results, which attracted the researcher's attention to integrate CT with other cost matching methods. The combination between improved CT and texture filtering have been proposed by Hou *et al*. improved the accuracy of the proposed algorithm in the area which the pattern texture is relatively dense [8]. The combination between CT and Absolute Different resulted a high quality depth map for high definition images [11]. The combination of CT and SAD reported able to improve the accuracy of the disparity map [12]. The combination of SAD and gradient matching proposed by [13] also reported to have a good accuracy of the disparity map.

The second step of the stereo matching algorithm is the CA step. The most common aggregation methods are by using filtering techniques. Guided filter (GF) is the common filtering technique used in the stereo matching algorithm due to the capability of this method in preserving edge of the image [14]–[17]. However, this method is window-based cost filtering in which the aggregation is happening based on the predefined window. Non-local cost aggregation has also gained attention due to the efficiency of the technique and the show a good performance in the low texture region [18]. The non-local cost aggregation is performed based on the overall pixel in the cost volume. The non-local technique such as minimum spanning tree (MST) and graph segmentation are reported to perform aggregation better than the filtering aggregation technique [19], [20]. Wu *et al.* [18] proposed a stereo matching algorithm that fuses between non-local and GF and reported that the proposed framework outperforms the other state of art. The non-local CA approach used by [21] also reported that the method is comparable with other CA methods. However, the ability of the non-local aggregation to preserve the edge of the image still required improvement.

Aiming to improve the accuracy of the disparity map, a new stereo matching algorithm is proposed in order to increase the robustness of the matching algorithm for images pair with different brightness, less texture surface and different image pair exposure while preserving the edge of the image. Taking the advantages of the combined CC methods and the non-local CA, this paper proposed a new stereo matching algorithm with the integration of CT and SAD as the CC steps followed non-local cost aggregation. The optimized parameter for CC steps and CA steps has been determined throughout the experiment where the performance of the proposed algorithm is analyzed.

2. METHOD

This paper proposed a new stereo matching algorithm that follows the basic step of the local matching algorithm development which involved four steps as explained in the previous chapter, while Figure 1 shows the block diagram for the proposed algorithm. The first step of the development of the proposed algorithm starts with matching cost computed by calculating the cost volume using CT and combined with the SAD. Then the second step aggregated the cost volume using non-local cost aggregation.

For each acceptable pixel, the WTA method selects the lowest cost accumulated at the corresponding location. Unwanted pixels still exist at this time, particularly in occlusion and less textured areas [21]. The left-right consistency (LR) procedure will detect these unwanted pixels [19]. Following that, the fill-in method is used to replace faulty pixels with valid minimum pixel values. Finally, the weighted median filter is used to perform the disparity enhancement step.

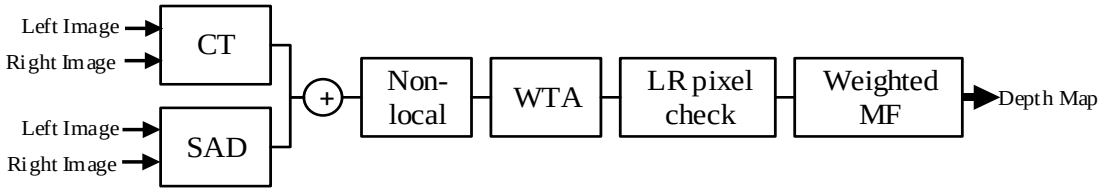

Figure 1. Block diagram of the proposed algorithm

2.1. Matching cost computation

SAD is the basic method in matching cost computation where all disparity value in the predefined small window is totalled up and the process repeated along the same horizontal line. The minimum value is considered as the best matching pixel region of the matching process. The algorithm of the matching cost using SAD is developed based on the expression shown in (1):

$$\mathrm{M_c}SAD(p,d) = \min \sum_{i \in \mathrm{w}} \left| I_l^i(p) - I_r^i(p-d) \right| \quad (1)$$

where p is symbolize the pixel at coordinates $(x;y)$, i symbolize the neighbouring pixel in the predefine window, $w.d$ represent the disparity value, while the pixel of left image and right image denoted by I_l and I_r respectively. CT process maps the surrounding pixel to a bit string that can be used to represent the pixel intensity value [9]. The process of CT is based on (2):

$$\mathrm{CT(p)} = \otimes_{\mathrm{i} \in \mathrm{w_{CT}}} \mathrm{cen(p,q)} \quad (2)$$

where p symbolize for the target pixel while q symbolize the neighboring pixels and $\otimes$ refer to the process of comparing the target pixel and its neighboring pixel with window size, w_{CT}. The binary function obtained from this process is represented as $cen\ (p,q)$. The transformation to the binary representation are based on condition stated in (3):

$$\mathrm{cen(p,q)} = \begin{cases} 1, & \mathrm{I(p)} \geq \mathrm{I(q)} \\ 0, & \text{otherwise} \end{cases} \quad (3)$$

where $I(p)$ is the target pixel and $I(q)$ is the surrounding pixels. The matching cost at the target point in the image pair is estimated using Hamming distance and represent as (4):

$$\mathrm{M_cCT(p,d)} = \mathrm{HammigDistance}(\mathrm{CT_l(p)} - \mathrm{CT_r(p-d)}) \quad (4)$$

where CT_l and CT_r are the bit string obtained by performing CT. Due to the unbalanced of the window size between $\mathrm{M_c}$SAD and M_cCT, the integration of the two matching costs is based on the normalized cost function proposed by [22]. The final matching cost which is known as integrated matching cost denoted by $i\mathrm{M_c(p,d)}$ is expressed in (5) where the α is the added parameter to control the sensitivity to the radiometric distortion.

$$\mathrm{iM_c(p,d)} = 2 - \left[\mathrm{e}^{(-\mathrm{M_cSAD(p,d)})} + \mathrm{e}^{\left(\frac{-\mathrm{M_cCT(p,d)}}{\alpha}\right)} \right] \quad (5)$$

2.2. Cost aggregation

A simple but effective non-local cost aggregation is absorbed in this work. This method is based on the non-local cost aggregation, which inspired by work done by [19] which used the left grayscale image as a guidance image. An undirected graph with four connected grids is developed based on the guidance image

where represent as G=[V;E] where V and E represent all the image pixels and all the edges between the neighbouring pixels. The output from the cost volume obtained by (5) is absorbed as input to the CA step and the cost volume function $\mathrm{iM_c(p,d)}$ is set as the vertices correspond to the data set for the spanning tree as represent in (6).

$$G = [\, iM_c(p,d), E] \tag{6}$$

From G, the spanning tree is developed by computing the weighted between the pair of neighbouring pixels by using (7) where $I(s)$ and $\mathrm{I}(r)$ is the neighbouring pixel value.

$$\mathrm{W(s,r)} = |\mathrm{I(s)} - \mathrm{I(r)}| \tag{7}$$

The sum of the possible edges of the spanning tree developed from G is calculated by summing up the weighted value. The edges with the large weighted value will be removed in the process of the spanning tree development which resulted an MST. The distance between two nodes in MST is define as (8):

$$\mathrm{D(p,q)} = \min \sum \mathrm{W(s,r)} \tag{8}$$

where $D(p,q)$ represents the distance between two nodes, p and q, in the connected edges. Then, the tree structure between two nodes is determined based on expression (9) [20], where σ represent a constant parameter to adjust the similarity between node *p* and node *q*.

$$\mathrm{S(p,q)} = \mathrm{e}^{\frac{-\mathrm{D(p,q)}}{\sigma}} \tag{9}$$

The final cost aggregation based on the MST structure is expressed as (10) by calculating the weighted summed of the matching cost for pixel *p* at disparity *d* and the tree structure of pixel *p* for all image pixels *q*.

$$\mathrm{CA(p)} = \sum_{\mathrm{q}} \mathrm{S(p,q)} * \mathrm{iM_c(p,d)} \tag{10}$$

2.3. Disparity optimization

The winner-takes-all (WTA) technique is used to select the lowest cost as the initial disparity value when the CA is completed [9]. In this step, the minimum cost volume at disparity, *d* is selected based on (11) where the initial disparity represents as d_i, *R* represents all potential disparity values and $CA(p,d)$ represent the aggregated cost volume at disparity.

$$\mathrm{d_i} = \arg\min_{\mathrm{d}\in\mathrm{R}} \mathrm{CA(p,d)} \tag{11}$$

2.4. Disparity refinement

Finally, the initial disparity obtained in the previous step will be refined in order to get a smoother disparity map. In this step, the invalid pixel is determined by checking tm he inconsistency of the left disparity map and the right disparity map. The valid disparity value is then used to replace the invalid values by taking the nearest valid pixels, and both the valid and invalid values must be located on the same scan line. After completing the filling process, the weighted median filter from [10] is implemented for final refinement. The final disparity map, d^F is refined as (12):

$$d^F(p) = \underset{q\in\Omega}{med}\{d(p)\} \tag{12}$$

2.5. 3D surface reconstruction

The 3D surface reconstruction is based on the triangulation theory where the depth value is obtained based on (13). The focal length denotes by *f* and *b* represents the baseline distance between the left camera and right camera and d is the disparity value of the image pair. By using the disparity map obtained in the stereo matching algorithm, the 3D representation of the image pairs is obtained by calculating the depth value of the image which represent as *Z* in (13).

$$Z = \frac{b*f}{d} \tag{13}$$

3. RESULTS AND DISCUSSION

The experiment was carried out to evaluate and analyze the performance of the proposed algorithm from various perspectives. It was conducted on a personal computer with Window 10, 3.2 GHz processor and 16 GB memory. The algorithm was developed using the C^{++} programming language and open CV library. The results were evaluated based on two standard benchmarking evaluation, which are Middlebury [23] and KITTI [24] stereo evaluation dataset. The quantitative results of the experiment evaluated using the online Middlebury database while the qualitative results are based on the KITTI images.

3.1. Parameter settings

In this experiment, all the constants are determined based on the evaluation of the error using the Middlebury training dataset. The errors are based on the percentage of absolute disparity error in two different masks which are whole image area pixel, all, and non-occluded area pixel, nonocc. In matching cost computation, the initial parameter α is determined by setting the initial constant parameter of the SAD as 1.0 while the value of α is set to minimum and gradually increases until the minimum point of average error is determined. Figure 2 shows the relationship between α and the percentage of the average error where the minimum average error is obtained at α is equal to 0.09. Based on the experiment using 15 training images of the Middlebury training dataset, the average error is 7.32% in the non-occluded region while the average error is at 10.6% in the all pixel regions.

In the cost aggregation step, the non-local method is optimized by determining the constant value of the non-local CA which is denoted by σ. The initial value of this parameter is set to 0.1 by referring to the work done by [19] and the value is decreasing until the minimum average error is obtained. Figure 3 showed the graph of the average error of the disparity map versus the constant parameter of non-local CA. Based on the results, it showed that the average error is decreasing while the value of the constant decreased. In this experiment, the nonlocal CA constant of 0.03 is used by considering the performance of the average error of all 15 training images of the Middlebury training dataset. Based on Figure 3, it is showed that the average error of image with different brightness, such as PionoL and ArtL start increasing after 0.03. The error of the Playtable also reduced by 50% compared with the error when sigma set at 0.1.

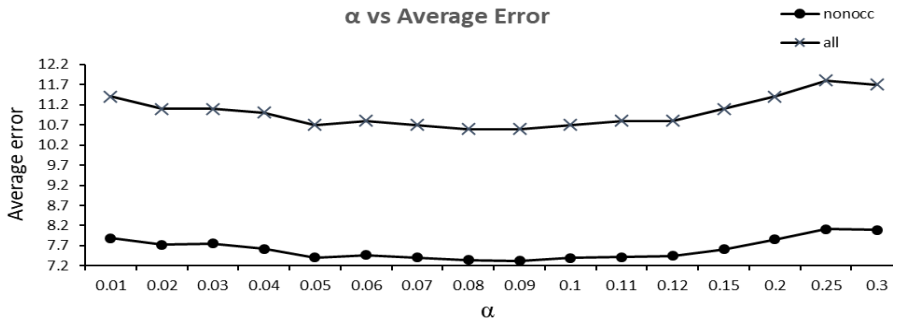

Figure 2. Average error in different parameter setting during MCC

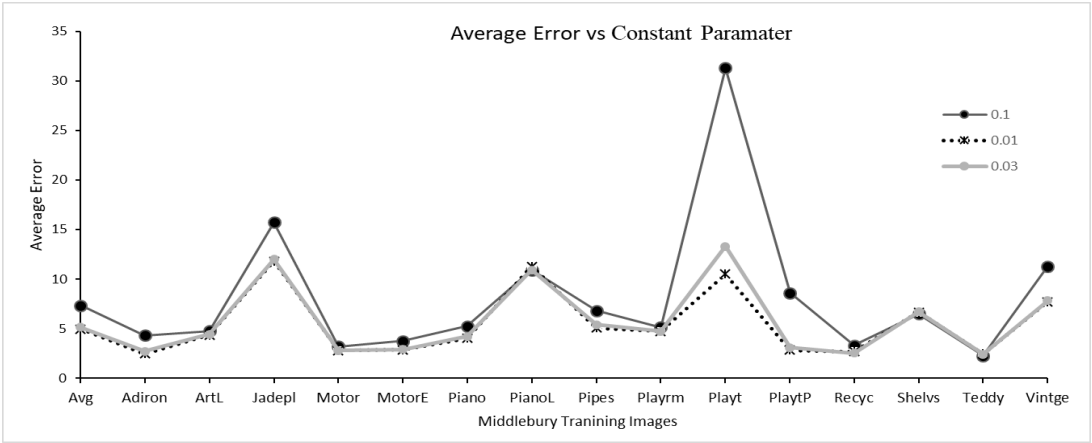

Figure 3. The average error for non-occluded region with different value of sigma

Figure 4 showed the comparison of Adiron and Playtable images before and after parameter adjustment. It showed that parameter optimization produced disparity maps with lower average error and the images showed improvement of edge preservation of the image after parameter adjustment. The disparity map produced using Adiron image pair showed that the edge of the image is more preserved by using sigma 0.03 compared with the sigma 0.1. The process continued with the WTA strategy after completion of parameter setting at CA and proceed with DR at stage 4.

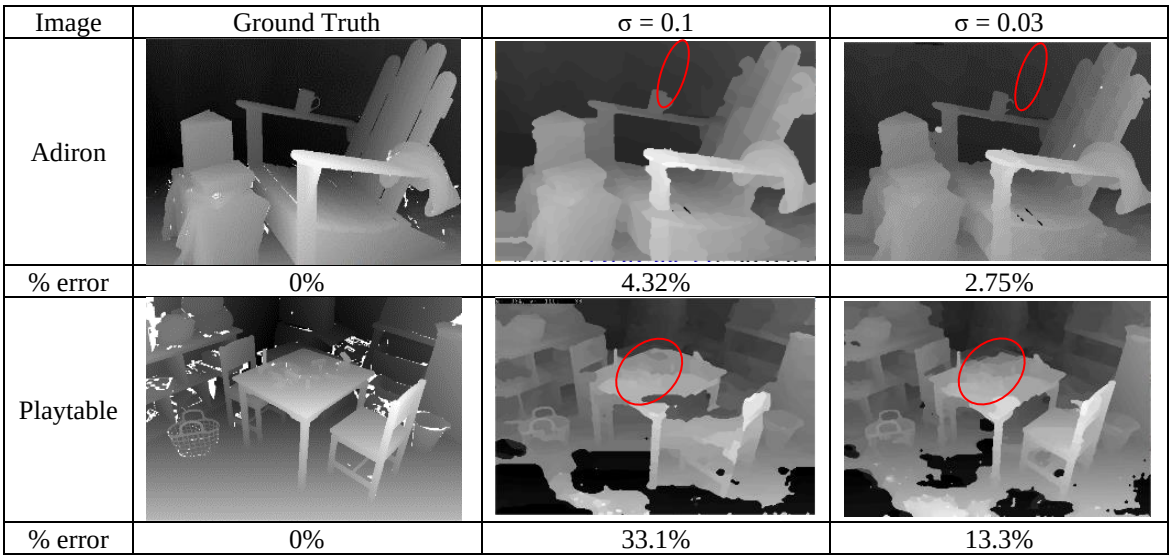

Figure 4. Example of disparity map with different setting at CA step

3.2. Middlebury dataset

Figure 5 showed the 15 training images and disparity map using the proposed algorithm. The comparison of the proposed algorithm and other published framework from online Middlebury evaluation are tabulated in Table 1. The comparison showed that the proposed methods are comparable with other published methods. The overall average error in all pixel regions showed that the proposed algorithm produced the lowest error compared with other frameworks while for overall average error for non-occluded region the proposed methods is at second rank. The comparison also showed that the proposed methods produced lowest average error in the image such as motor, MotorE, recycle, and shelves while at second lowest average error for most of the other images.

Table 1. The comparison of the percentage of average error for all pixel regions and non-occluded region between the proposed framework and other established framework

Framework	NEW %		tMGM-16 [25] %		FBW_ROB %		DDL [26] %		FASW [18] %		TCSCSM [27] %	
	all	*nonocc*	*all*	*nonocc*	*all*	*nonocc*	*All*	*nonocc*	*All*	*nonocc*	*All*	*nonocc*
Avg	8.39	5.14	9.48	5.78	8.65	3.96	8.63	5.44	8.59	5.18	8.47	5.12
Adiron	3.45	2.7	4.53	3.01	4.93	2.26	3.07	2.47	3.5	2.61	2.86	2.07
ArtL	8.89	4.43	8.41	3.91	7.7	2.72	7.83	4.25	7.84	5.09	8.03	4.44
Jadepl	28.7	12	22.1	11.2	37.1	14.2	32.8	15.5	35.4	15.1	34.7	14.4
Motor	5.08	2.79	7.93	2.81	6.01	1.87	5.83	3.37	6.04	3.45	5.44	3.01
MotorE	5.29	2.9	7.88	2.91	6.14	1.97	5.92	3.57	5.7	3.19	5.43	2.92
Piano	5.13	4.2	6.36	4.95	3.7	2.65	5.38	3.96	4.73	3.69	5.54	4.54
PianoL	11.7	10.9	27.7	27.1	6.92	5.98	8.13	6.97	9.65	8.93	10.8	10.2
Pipes	10.6	5.41	11	4.59	10.6	3.68	11.3	5.63	12	6.18	10.8	5.23
Playrm	8.24	4.71	8.51	5.49	7.39	3.71	5.66	3.82	6.57	4.38	7.31	5.17
Playt	16	13.3	16.1	12.3	5.59	2.4	13.4	10.8	9.45	5.89	14.5	11.7
PlaytP	4.81	3.1	6.6	2.58	4.89	2.2	4.26	3	4.33	3.06	3.32	2.5
Recyc	2.9	2.54	4.26	2.5	4.28	2.1	3.07	2.44	3.04	2.48	2.84	2.54
Shelvs	7.2	6.68	13.1	12.6	9.69	9.04	8.57	8.02	8.7	8.25	8.7	8.14
Teddy	3.86	2.41	2.86	1.86	3.51	1.87	2.76	2.02	3.42	2.52	2.83	1.99
Vintage	9.11	8.16	7.77	6.58	8.87	6.82	15.5	13.9	8.47	7.44	6.79	5.6

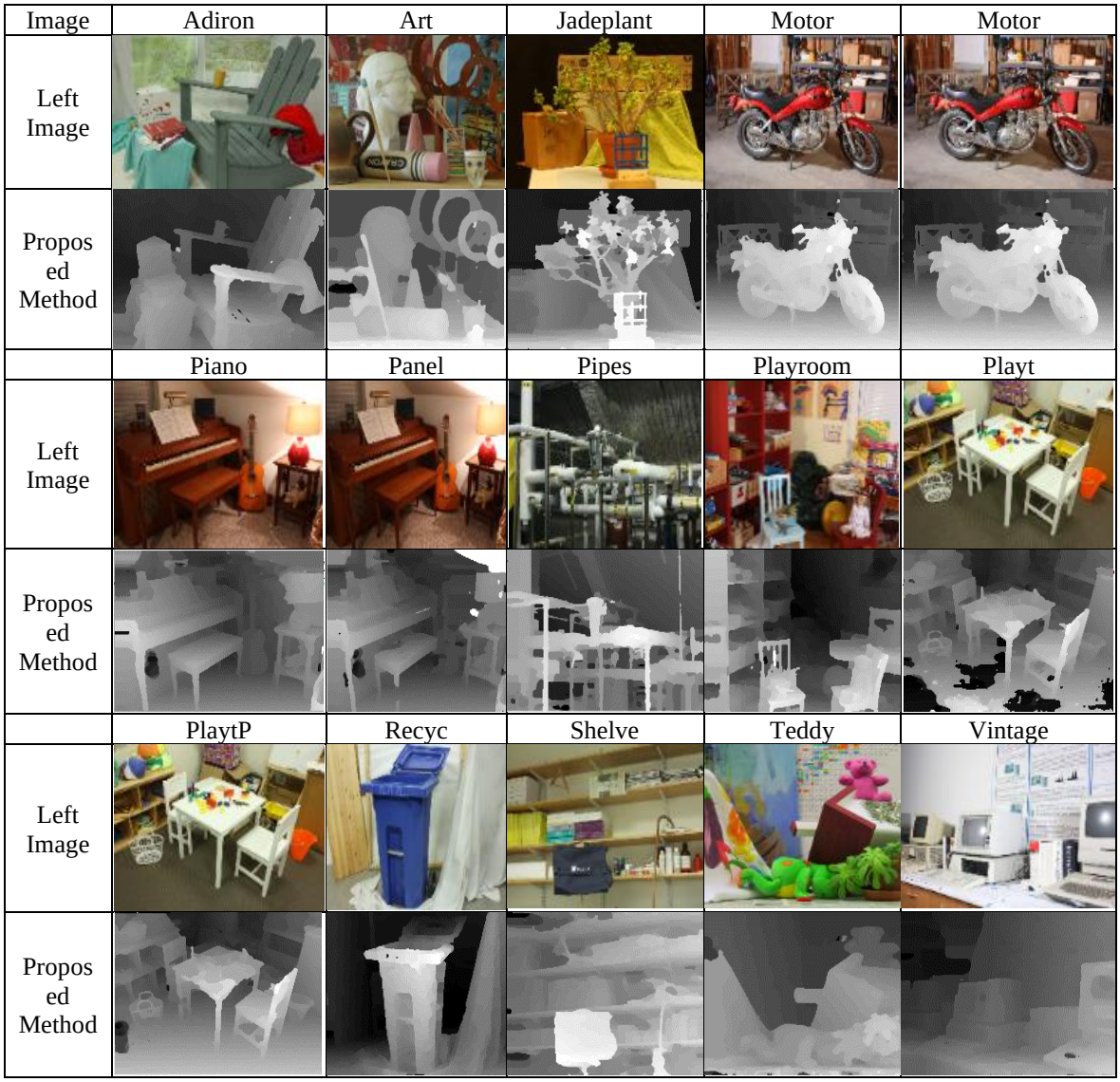

Figure 5. The left image and disparity map of the proposed method for middlebury training images

3.3. KITTI dataset

In order to evaluate the proposed algorithm to a more challenging environment and to ensure the matching algorithm is suitable for autonomous vehicle, KITTI dataset had been used as the input of the stereo images. KITTI provide a real world stereo image for autonomous vehicle which consists of various exposures such as less texture region and inconsistent brightness of image pairs. All the images with 1226×370 resolution is used with maximum disparity at 255. Figure 6 showed the left images, ground truth and disparity maps of the proposed algorithm for the image sequence from 000000_10 until 000006_10**.** The results showed that the proposed algorithm able to match the stereo images in real environment where the real environment had a large area of texture-less region, natural brightness and variety of environmental exposures.

3.4. 3D surface reconstruction

Based on the triangulation principle, the 3D surface is reconstructed from the disparity map obtained by using the cvkit development kit. Figure 7 showed an example of 3D surface reconstruction for image00000_10 using disp_occ_0 from KITTI training dataset and using the proposed method. The results showed that the proposed method is robust in outdoor environment and capable to determine small object.

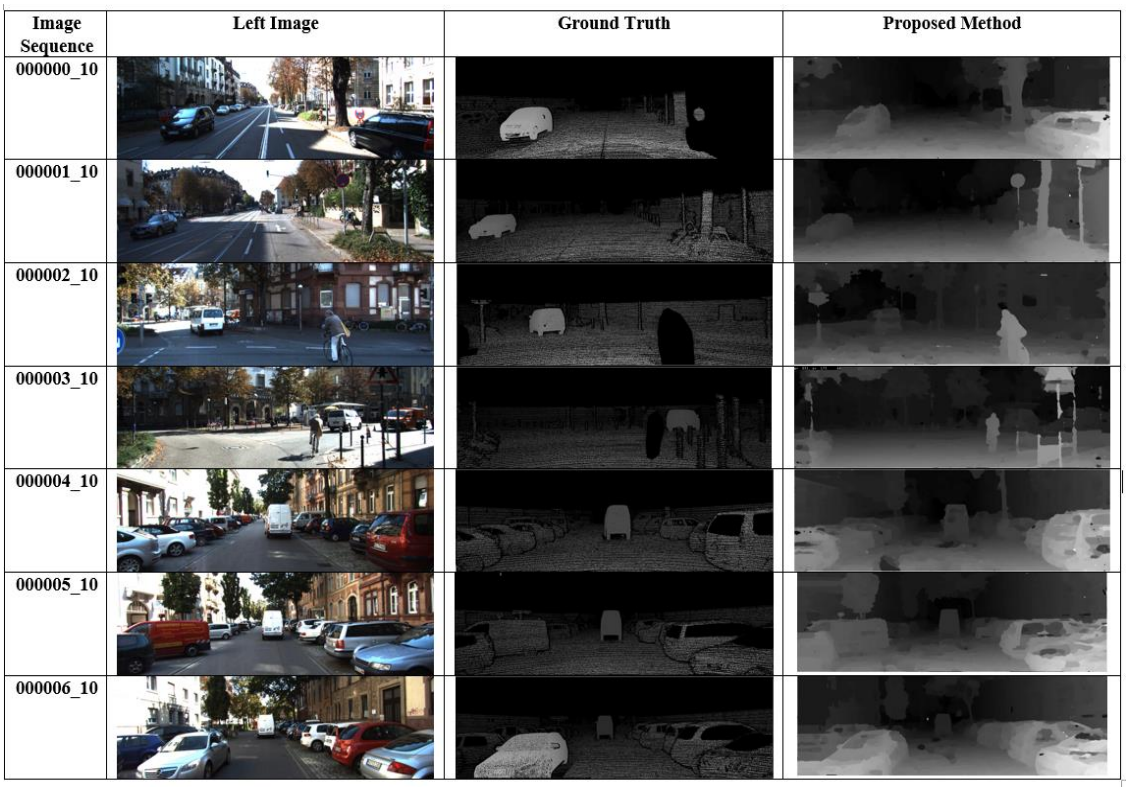

Figure 6. The disparity map produced by using proposed algorithm using KITTI training dataset

Image	Image 000000_10
disp_occ_0	
Proposed Method	

Figure 7. The 3D reconstruction of the image000000_10 from KITTI dataset

4. CONCLUSION

In this paper, a stereo matching algorithm based on the integration of two matching costs, which is CT and SAD is used with the non-local cost aggregation. The constant parameter had been introduced to the CT and non-local aggregation to improve the accuracy of the output. Through the experiment and comparison with other algorithm, it concludes that the proposed algorithm effectively reduces the percentage of error and by introducing the constant parameter to the CT and the nonlocal CA, better results are obtained especially in the less texture region and different illumination region. The evaluation using real environment dataset also shows that the proposed algorithm is able to produce good quality of disparity map. It is concluded that the proposed algorithm is suitable for 3D surface reconstruction and be implemented in autonomous vehicle navigation system.

ACKNOWLEDGEMENTS

This research project is supported by a grant from the Universiti Teknikal Malaysia Melaka and Ministry of Higher Education, Malaysia with the reference number FRGS/1/2020/FTKEE-CACT/F00451.

REFERENCES

[1] S. V. Aswin Kumer, L. S. P. Sairam Nadipalli, P. Kanakaraja, K. Sarat Kumar, and K. Ch. Sri Kavya, "Controlling the autonomous vehicle using computer vision and cloud server," *Materials Today: Proceedings*, vol. 37, pp. 2982–2985, Jan. 2021, doi: 10.1016/j.matpr.2020.08.712.

[2] B. Kim, D. Kim, S. Park, Y. Jung, and K. Yi, "Automated Complex Urban Driving based on Enhanced Environment Representation with GPS/map, Radar, Lidar and Vision," *IFAC-PapersOnLine*, vol. 49, no. 11, pp. 190–195, 2016, doi: 10.1016/j.ifacol.2016.08.029.

[3] G.-S. Hong and B.-G. Kim, "A local stereo matching algorithm based on weighted guided image filtering for improving the generation of depth range images," *Displays*, vol. 49, pp. 80–87, Sep. 2017, doi: 10.1016/j.displa.2017.07.006.

[4] Y.-J. Chang and Y.-S. 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, Oct. 2016, doi: 10.1016/j.jvcir.2016.06.017.

[5] R. A. Hamzah and H. Ibrahim, "Literature Survey on Stereo Vision Disparity Map Algorithms," *Journal of Sensors*, vol. 2016, pp. 1–23, 2016, doi: 10.1155/2016/8742920.

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

[7] Y. Chai and X. Cao, "Stereo Matching Algorithm Based on Joint Matching Cost and Adaptive Window," *2018 IEEE 3rd Advanced Information Technology, Electronic and Automation Control Conference (IAEAC)*, 2018, pp. 442-446, doi: 10.1109/iaeac.2018.8577495.

[8] Y. Hou, C. Liu, B. An, and Y. Liu, "Stereo matching algorithm based on improved Census transform and texture filtering," *Optik*, vol. 249, p. 168186, Jan. 2022, doi: 10.1016/j.ijleo.2021.168186.

[9] R. Zabih and J. Woodfill, "Non-parametric local transforms for computing visual correspondence," in *Computer Vision — ECCV '94*, Springer Berlin Heidelberg, 1994, pp. 151–158, doi: 10.1007/bfb0028345.

[10] H. Hirschmuller and D. Scharstein, "Evaluation of Stereo Matching Costs on Images with Radiometric Differences," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 31, no. 9, pp. 1582–1599, 2009, doi: 10.1109/TPAMI.2008.221.

[11] W. Wang, J. Yan, N. Xu, Y. Wang, and F.-H. Hsu, "Real-Time High-Quality Stereo Vision System in FPGA," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 25, no. 10, pp. 1696–1708, Oct. 2015, doi: 10.1109/tcsvt.2015.2397196.

[12] M. Zahari, R. A. Hamzah, N. A. Manap, and A. I. Herman, "Stereo matching algorithm based on combined matching cost computation and edge preserving filters," *Indonesian Journal of Electrical Engineering and Computer Science*, vol. 26, no. 3, p. 1415, Jun. 2022, doi: 10.11591/ijeecs.v26.i3.pp1415-1422.

[13] R. A. Hamzah, M. G. Y. Wei, and N. S. N. Anwar, "Development of stereo matching algorithm based on sum of absolute RGB color differences and gradient Matching," *International Journal of Electrical and Computer Engineering (IJECE)*, vol. 10, no. 3, p. 2375, Jun. 2020, doi: 10.11591/ijece.v10i3.pp2375-2382.

[14] K. Lingyin, Z. Jiangping, and Y. Sancong, "Stereo Matching Based on Guidance Image and Adaptive Support Region," *Acta Optica Sinica*, vol. 40, no. 9, p. 0915001, 2020, doi: 10.3788/aos202040.0915001.

[15] 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, Jul. 2018, doi: 10.1016/j.image.2018.04.001.

[16] A. Hosni, C. Rhemann, M. Bleyer, C. Rother, and M. Gelautz, "Fast Cost-Volume Filtering for Visual Correspondence and Beyond," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 35, no. 2, pp. 504–511, Feb. 2013, doi: 10.1109/tpami.2012.156.

[17] Y. Xu, Y. Zhao, and M. Ji, "Local stereo matching with adaptive shape support window based cost aggregation," *Applied Optics*, vol. 53, no. 29, p. 6885, Oct. 2014, doi: 10.1364/ao.53.006885.

[18] W. Wu, H. Zhu, S. Yu, and J. Shi, "Stereo Matching With Fusing Adaptive Support Weights," *IEEE Access*, vol. 7, pp. 61960–61974, 2019, doi: 10.1109/ACCESS.2019.2916035.

[19] Q. Yang, "A non-local cost aggregation method for stereo matching," *2012 IEEE Conference on Computer Vision and Pattern Recognition*, 2012, pp. 1402-1409, doi: 10.1109/cvpr.2012.6247827.

[20] X. Mei, X. Sun, W. Dong, H. Wang, and X. Zhang, "Segment-Tree Based Cost Aggregation for Stereo Matching," *2013 IEEE Conference on Computer Vision and Pattern Recognition*, 2013, pp. 313-320, doi: 10.1109/cvpr.2013.47.

[21] W. Wu, H. Zhu, and Q. Zhang, "Oriented-linear-tree based cost aggregation for stereo matching," *Multimedia Tools and Applications*, vol. 78, no. 12, pp. 15779–15800, Dec. 2018, doi: 10.1007/s11042-018-6993-2.

[22] K.-J. Yoon and I.-S. Kweon, "Locally Adaptive Support-Weight Approach for Visual Correspondence Search." doi: 10.1109/cvpr.2005.218.

[23] D. Scharstein *et al.*, "High-Resolution Stereo Datasets with Subpixel-Accurate Ground Truth," in *Lecture Notes in Computer Science*, Springer International Publishing, 2014, pp. 31–42, doi: 10.1007/978-3-319-11752-2_3.

[24] M. Menze and A. Geiger, "Object scene flow for autonomous vehicles," in *2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2015, pp. 3061–3070, doi: 10.1109/CVPR.2015.7298925.

[25] S. Patil, T. Prakash, B. Comandur, and A. Kak, "A Comparative Evaluation of SGM Variants (including a New Variant, tMGM) for Dense Stereo Matching." arXiv, Nov. 21, 2019, doi: 10.48550/arXiv.1911.09800.

[26] J. Yin, H. Zhu, D. Yuan, and T. Xue, "Sparse representation over discriminative dictionary for stereo matching," *Pattern Recognition*, vol. 71, pp. 278–289, Nov. 2017, doi: 10.1016/j.patcog.2017.06.015.

[27] C. Cheng, H. Li, and L. Zhang, "Two-Branch Convolutional Sparse Representation for Stereo Matching," *IEEE Access*, vol. 9, pp. 21910–21920, 2021, doi: 10.1109/access.2021.3056137.

BIOGRAPHIES OF AUTHORS

Madiha Zahari received the Bachelor degree in Electrical and Electronic Engineering from University Technology of Petronas in 2006 and Master of Engineering in Industrial Electronic and Control from University of Malaya in 2013. She had worked in the semiconductor industry between 2006 and 2009. She is presently a lecturer at Universiti Teknikal Malaysia Melaka's Fakulti Teknologi Kejuruteraan Elektrik dan Elektronik. She is also a Chartered Engineer with the MIET. Image processing and stereo vision are two of her main research interests. She can be reached at madiha@utem.edu.my.

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 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.

Nurulfajar Abd Manap received a Master of Electrical Engineering in image processing (2002, Universiti Teknologi Malaysia) and Bachelor of Electrical Engineering (2000, Universiti Teknologi Malaysia). He obtained his Ph.D (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. He can be contacted at email: nurulfajar@utem.edu.my.

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