Journal of Telecommunication, Electronic and Computer Engineering
ISSN: 2180 – 1843 e-ISSN: 2289-8131 Vol. 17 No. 4 (2025) 1-6
jtec.utem.edu.my
DOI: https://doi.org/10.54554/jtec.2025.17.04.001

Enhanced Local Disparity Map Algorithm Segment-Side Window-based Cost Aggregation and Refinement

M. Y. A. Mohammed Al-Asbahi[1], Ahmad Fauzan Kadmin[1,2]*, Rostam Affendi Hamzah[1], Khairul Azha A. Aziz[1], Nasharuddin Zainal[3], Nabil Jazli Jazlan[4]

[1]*Fakulti Teknologi dan Kejuruteraan Elekronik dan Komputer, Universiti Teknikal Malaysia Melaka, Durian Tunggal, 76100 Melaka, Malaysia.*
[2]*Centre for Telecommunication Research and Innovation (CETRI), Universiti Teknikal Malaysia Melaka, Durian Tunggal, 76100 Melaka, Malaysia.*
[3]*Faculty of Engineering and Built Environment, Universiti Kebangsaan Malaysia, Bangi, 43600 Selangor, Malaysia.*
[4]*IT Support Department, Amcorp Services Sdn Bhd, Petaling Jaya, 46050 Selangor, Malaysia.*

Article Info

Article history:
Received Jun 25th, 2025
Revised Aug 14th, 2025
Accepted Oct 21st, 2025
Published Dec 24th, 2025

Index Terms:
Disparity map
Segment cost
Side window
Stereo matching
Stereo vision

Abstract

Accurate disparity map estimation is crucial for applications such as 3D reconstruction, autonomous navigation, and object detection. Local window-based cost aggregation often suffers from edge fattening and texture inconsistency. This paper introduces a Segment-Side Window-based (SSW) stereo matching algorithm that combines Truncated Absolute Difference (TAD), Gradient Magnitude (GM), and Census Transform (CT) to build a robust cost volume. In the proposed approach, SLIC superpixels guide adaptive aggregation, while Side Window Filtering (SWF) preserves edges and enhances texture consistency. Winner-Takes-All optimization and SWF refinement further improve depth accuracy. On the Middlebury dataset, the proposed method achieves 13.3% (Nonocc) and 21.8% (All) bad pixel errors, outperforming BF, GF, iGF, and MF in both edge preservation and texture robustness.

*Corresponding Author: fauzan@utem.edu.my

I. INTRODUCTION

Stereo vision plays a pivotal role in computer vision, enabling the extraction of three-dimensional (3D) information from two-dimensional (2D) stereo images through depth estimation using disparity maps [1]. A disparity map encodes the pixel-wise displacement between corresponding points in a stereo image pair, which directly correlates with depth perception. Accurate disparity estimation is fundamental to a wide range of advanced applications, including 3D image reconstruction, autonomous navigation, medical imaging, and immersive 3D entertainment systems [2].

The process of disparity map estimation typically follows a four-stage framework: (1) matching cost computation, (2) cost aggregation, (3) disparity selection, and (4) disparity refinement, as shown in Figure 1 [3]. Despite the structured framework, several challenges persist in stereo vision, particularly with local window-based cost aggregation methods. These methods often suffer from edge fattening and texture inconsistency [4]. Edge fattening occurs when the cost aggregation window crosses object boundaries, causing a blurring effect that merges disparities from different depth layers, resulting in inaccurate depth estimations. Texture inconsistency, on the other hand, arises in regions with low or repetitive textures, where matching becomes ambiguous, leading to noisy or incorrect disparity values. These limitations are exacerbated by fixed window sizes that fail to adapt to varying scene structures, further compromising the reliability of the resulting disparity map.

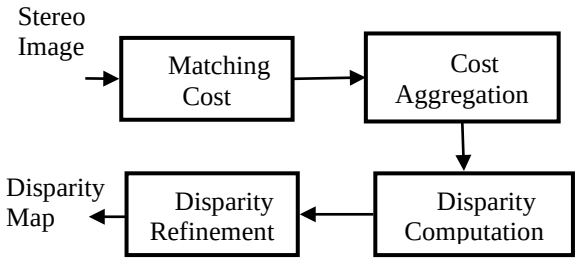

Figure 1. Four-stage framework of disparity map estimation

To address these issues, this study proposes an enhanced local disparity map algorithm using segment-side window-based cost aggregation and refinement. The method integrates a side window approach to dynamically adjust the orientation according to local edge structures. Additionally, segment-based aggregation preserves object boundaries while enhancing texture consistency, thereby mitigating the common pitfalls of local methods. The primary aim of this research is to improve disparity estimation accuracy while

maintaining computational efficiency, making it suitable for real-time applications.

The remainder of this paper is structured as follows: Section 2 reviews related works on cost aggregation techniques. Section 3 details the proposed methodology. Section 4 presents experimental results and performance evaluations. Finally, Section 5 concludes the study and suggests directions for future research.

II. RELATED WORKS

This section focuses on the study that has been proposed with current methods that deals with local window aggregation and refinement in achieving optimal accuracy for bad pixel error, texture inconsistency and edge fattening.

Huang et al.[5] introduced a channel-based 2D cost aggregation method, achieving high efficiency with minimal computational resources. Its lightweight design suits real-time applications, but the reliance on 2D aggregation may struggle with texture inconsistency, as finer details in complex scenes may not be preserved. Similarly, Zahari et al.[7] addressed depth edge preservation by using a global non-local approach. Although it improves accuracy on standard benchmarks like Middlebury, the method is computationally intensive and may still exhibit edge fattening due to the smoothing effect inherent in non-local operations. One 3D cost aggregation method rethinks stereo matching through a histogram-based approach, reducing computational redundancy [9]. While this method offers complexity reduction, handling high disparity variability remains a challenge, exacerbating texture inconsistencies in regions with abrupt depth changes.

Deng et al.[10] incorporated multi-scale information to enhance accuracy and efficiency. However, the increased computational demand may still struggle to resolve texture inconsistencies, especially in fine-detail regions, while abrupt depth changes can cause edge fattening. Gómez[11] reduced computational load by combining semi-global and local guided aggregation modules, balancing efficiency and accuracy. Despite this, scenes with low texture or high noise might compromise performance, with the risk of blurring edges. Huang et al.[12] employed 3D CNNs to achieve high accuracy in disparity estimation. While robust across datasets, the method's computational intensity poses challenges for real-time applications, and complex scenes with texture inconsistencies may still result in blurred or fattened edges. Rahim et al.[13] utilized neural architecture search to optimize 3D cost aggregation. Despite its flexibility, the computational expense involved in training may limit effectiveness in fine-textured regions and edge fattening during deployment.

Bangunharcana et al.[14] used guided cost volume excitation to improve real-time stereo matching accuracy. While fast, the use of 3D CNNs limits its adaptability for edge devices, and it may not optimally handle dynamic environments where texture inconsistency becomes a significant issue. Liu et al.[15] constructed a 3D cost volume and introduced intra-scale and cross-scale 2D cost aggregation modules. Although efficient, its multi-scale approach may still experience computational bottlenecks, particularly resulting in edge fattening when dealing with abrupt depth variations. Shamsafar et al.[16] reduced computational costs using MobileNet blocks, offering a lightweight solution. However, its performance, reported as lower than other methods, may not handle fine-textured regions effectively, contributing to texture inconsistency and edge blurring. Zeglazi et al.[17] introduced a novel dissimilarity measure to improve robustness in complex scenes, yet the method's additional computational needs may result in subtle edge fattening, especially in high-noise environments.

Yang et al.[19] enhanced robustness by detecting and preserving edges. While it performs well with edges and textures, computational intensity may still cause edge widening, particularly when noise is present. Lastly, Yuan et al.[20] proposed using 1D support windows based on spatial and gradient information. Though efficient, this method requires fine-tuning to prevent edge fattening and may struggle with intricate textures. Overall, the reviewed methods highlight the ongoing struggle to balance accuracy, efficiency, and computational complexity in cost aggregation for stereo matching. While advances have been made, tackling texture inconsistency and edge fattening remains an open challenge, necessitating further innovations to achieve robust and real-time performance.

III. METHODOLOGY

Fundamentally, the proposed stereo matching algorithm is illustrated in Figure 2. To accomplish the objective of this work, several steps are carried out. The proposed methodology consists of five primary stages, pre-processing, matching cost computing , cost aggregation, disparity optimization and final disparity refinement.

A. Matching Cost Computation

In this framework, the matching cost is computed using a multi-cost matching approach that integrates three different cost functions: Truncated Absolute Difference (TAD), Gradient Magnitude (GM), and Census Transform (CT). This combination enhances disparity estimation by addressing various challenges such as illumination changes, textureless regions, and edge preservation. The TAD function minimizes sensitivity to outliers by truncating large intensity differences between corresponding pixels in the left and right images [21].

The Absolute Difference (AD) measures the pixel intensity difference between the left (I_L) and right (I_R) images as in Equation (1):

$$AD(p,d) = [I_L(p) - I_R(p-d)] \quad (1)$$

However, to reduce sensitivity to illumination variations and outliers, TAD is computed as shown in Equation (2):

$$TAD\ (p,d) = \min([I_L(p) - I_R(p-d)], \tau) \quad (2)$$

where τ is a predefined truncation threshold that limits the influence of large intensity differences. This prevents high-cost values from dominating the disparity estimation.

Meanwhile, GM captures structural variations by computing gradient differences, ensuring robust performance in regions with weak textures [22]. Textureless regions (e.g., walls, sky) may have similar pixel intensities, making them hard to distinguish. To address this, the GM is used to capture edge and structural details as shown in Equation (3):

$$GM(p,d) = [G_L(p) - G_R(p-d)] \quad (3)$$

where $G_L(p)$ and $G_R(p\text{-}d)$ are the gradient magnitudes of the left and right images at pixel p, calculated using the Sobel operator. The CT method further strengthens the framework by encoding local spatial patterns into binary descriptors, making it highly resistant to illumination variations. The CT is defined as shown in Equation (4):

$$CT(p,d) = H(\oplus q \in N_p \xi(I_L(p),(I_L(q)) \oplus \xi(I_R(p-d), I_R(q-d))) \quad (4)$$

where N_p is the neighborhood window around pixel p and $(I_L(p),(I_L(q)$ is a binary function. This converts local intensity variations into a binary string. The Hamming distance H(·) measures the difference between the binary strings in the left and right images.

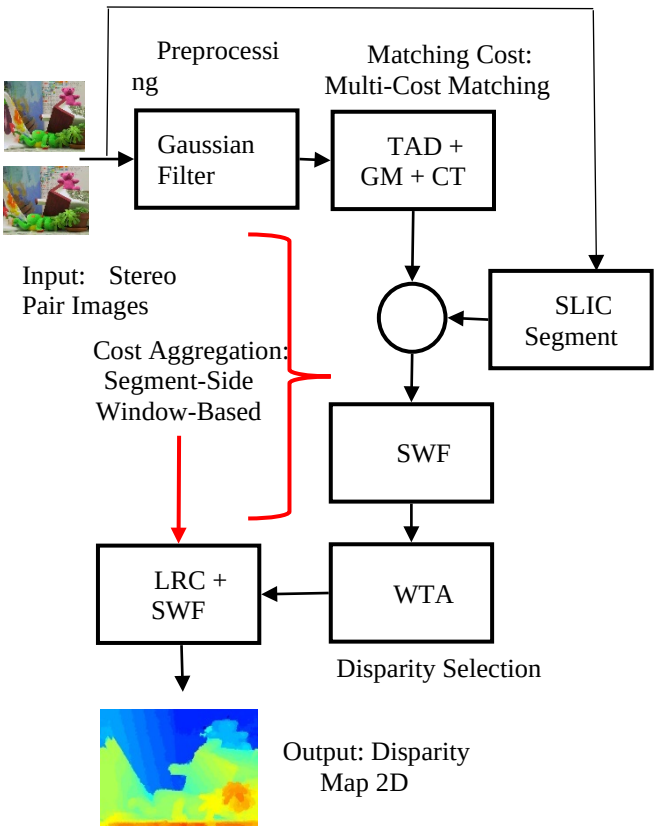

Figure 2. Proposed disparity map algorithm

The final matching cost is obtained by combining these three measures in a weighted sum, where the contributions of TAD, GM, and CT are balanced using tunable parameters. This approach enhances stereo matching performance by ensuring higher accuracy in textureless areas. The integration cost is given in Equation (5):

$$C(p,d) = \alpha TAD(p,d) + \beta GM(p,d)\ \gamma CT(p,d) \quad (5)$$

where α, β, γ are tunable weight parameters that balance the contributions of TAD, GM, and CT. The appropriate selection of these parameters is critical, as they directly influence the quality and robustness of the generated cost volume. By optimally balancing the complementary strengths of TAD, GM, and CT, the framework ensures improved handling of texture inconsistencies, depth discontinuities, and illumination variations, ultimately enhancing disparity estimation accuracy. Typically, higher weights are given to CT and GM in low-texture areas, while TAD is more reliable for detailed regions.

B. Cost Aggregation

Cost aggregation is employed to reduce correspondence ambiguity by applying a smooth filter to minimize noise in the initial raw matching cost. Since information obtained from a single pixel during cost computation is insufficient for accurate matching, aggregation is necessary to enhance reliability. Instead of traditional filtering approaches, the proposed method employs a segment-side window-based aggregation technique, which consists of Simple Linear Iterative Clustering (SLIC) and Side Window Filtering (SWF).

SLIC is a superpixel segmentation algorithm that clusters pixels into compact, perceptually meaningful regions based on color and spatial proximity. By grouping pixels with similar characteristics, SLIC ensures more reliable cost aggregation within homogeneous regions while preserving boundaries at object edges. SLIC performs superpixel segmentation to group pixels with similar color and intensity characteristics, ensuring more reliable aggregation within meaningful regions [23]. The SLIC algorithm minimizes the following distance measure to assign pixel p to a superpixel cluster k. SLIC iteratively refines clusters by updating centroids and reassigning pixels until convergence as given in Equation (6):

$$D = \sqrt{\left(\frac{d_c}{m}\right)^2 + \left(\frac{d_s}{S}\right)^3} \quad (6)$$

where d_c is the color distance in CIELAB space between pixel p and superpixel center k while d_s is the Euclidean spatial distance between pixel p and superpixel center k. S is the nominal superpixel size and m is the compactness parameter controlling trade-off between color similarity and spatial proximity.

SWF is then applied to refine the cost volume by reducing outliers and preserving object boundaries, which improves accuracy in disparity estimation. The SWF filtering operation is given in Equation (7):

$$SWF(p,d) = \sum_q \in w_p . w(p,q)[C(p,d)) + S(p)] \quad (7)$$

where $SWF(p,d)$is the filtered intensity at pixel p and disparity range of d. w_p represents the side window, which consists of pixels positioned at the edge of the window rather than centered around p. w(p,q) is the weight function, determined based on spatial and intensity similarities while $C(p,d))$ is the matching cost volume and S(p) is the SLIC value. SWF ensures that filtering is applied along object boundaries rather than across them, leading to better edge preservation.

C. 3.3. Disparity Selection

Disparity selection identifies the optimal disparity value for each pixel by minimizing the aggregated cost. In this process, a Winner-Takes-All (WTA) strategy is applied, as represented in Equation(8):

$$DS(p) = \arg\min_{d \in d_{r1}} SWF(p,d) \quad (8)$$

where DS(p) is the disparity value with the minimum cost, SWF(p, d) is the aggregated cost, and D represents the set of all possible disparity values. The selection of the disparity map is guided by the best match obtained through the multi-cost matching computation. While local approaches accumulate support areas by averaging or distributing them evenly, their accuracy is sensitive to noise and ambiguous regions. By integrating the segment-side window-based aggregation, the proposed method ensures more stable disparity estimation, particularly in textureless areas.

D. Disparity Refinement

The final stage involves postprocessing and disparity refinement to enhance the accuracy of the disparity map. In this framework, left-right consistency checking (LRC) is applied to identify and eliminate mismatches. Invalid pixel filling is performed using median interpolation, ensuring that missing or unreliable disparities are effectively reconstructed as given in Equation (9):

$$D_{median}(p) = median\{DS(p) \mid q \in N(p)\} \quad (9)$$

where DS(p) is the disparity of neighboring pixels q in the local window N(p) around pixel p, ensuring robust filling of invalid or noisy disparity values while preserving depth edges. Additionally, SWF is used to refine the disparity map by preserving edges while reducing noise. This refinement process effectively removes artifacts, enhances edge definition, and ensures high-quality depth estimation.

IV. RESULT AND DISCUSSION

The quantitative evaluation of the proposed SSW stereo matching method compared to BF, GF, iGF, and MF across the Middlebury dataset is presented in Table 1. The assessment is based on the bad pixel error for all pixels (All) and the non-occluded region (Nonocc). The performance of SSW demonstrates significant improvements over conventional approaches in both metrics, indicating enhanced accuracy in disparity estimation.

Table 1
Middlebury datasets training quantitative performance based on average all and nooccluded of bad pixel errors

Image Middlebury	BF		GF		iGF		MF		SSW	
	All	Nonocc	All	Nonocc	All	Nonocc	All	Nonocc	All	Nonocc
Adirondack	26.9	23	23.8	19.6	23.8	19.6	17.6	13.3	14.2	9.7
Artl	31.8	16.7	29.6	13.8	29.6	13.8	29.4	11	23.6	8.71
Jadeplant	80.6	61.7	75.5	55.1	75.5	55.1	65.2	42.7	54.8	34.5
Motorcyle	23.4	16.6	20.3	13.3	20.3	13.3	17.6	9.82	13.6	6.6
MotorcycleE	22.1	15.2	19.4	12.3	19.4	12.3	17.6	9.79	13.1	6.08
Piano	24	19.7	21.7	17.3	21.7	17.3	17.1	11.4	16.3	11.9
PianoL	36.6	33.3	34.3	30.9	34.3	30.9	30	25.4	25.9	22.3
Pipes	32.8	20.5	30.2	17.6	30.2	17.6	26.7	12.9	22.6	9.97
Playroom	40.8	25.9	37.8	22.5	37.8	22.5	36.3	17.6	29.6	14
Playtable	36.8	31.3	34.8	29	34.8	29	27.6	20.6	25.3	19.2
PlaytableP	25.8	19.7	23.3	17.1	23.3	17.1	19.5	11.4	15.6	9.23
Recycle	22.7	19.3	20.2	16.7	20.2	16.7	14.5	11.3	12.8	8.93
Shelves	26.4	24.4	24.8	22.6	24.8	22.6	18.6	15.8	17.1	14.7
Teddy	21.8	13.7	19.3	10.9	19.3	10.9	17.2	7.22	14.6	5.95
Vintage	73.7	70.2	68.5	64.5	68.5	64.5	51.2	43.1	44.3	39.2
Average	**33.5**	**25.5**	**30.7**	**22.3**	**30.7**	**22.3**	**25.9**	**16.2**	**21.8**	**13.3**

The proposed SSW method achieves the lowest average bad pixel error across all tested images, with an overall error of 21.8% (All) and 13.3% (Nonocc). This represents a substantial improvement compared to BF (33.5%, 25.5%), GF (30.7%, 22.3%), iGF (30.7%, 22.3%), and MF (25.9%, 16.2%). The trend is consistent across individual images, with SSW consistently outperforming the other methods. Notably, for complex texture regions, such as Jadeplant and Vintage, SSW reduces the error margin significantly, achieving 54.8% (All) and 34.5% (Nonocc) for Jadeplant, compared to MF's 65.2% (All) and 42.7% (Nonocc). Similarly, for Vintage, SSW records 44.3% (All) and 39.2% (Nonocc), in contrast to MF's 51.2% and 43.1% respectively.

These results demonstrate that SSW effectively mitigates texture inconsistency, which is a common problem in disparity estimation. Traditional approaches such as BF and GF suffer from excessive smoothing in texture-rich regions, leading to loss of structural details. The proposed method, by contrast, preserves finer details, reducing disparity errors in areas with complex textures.

In edge-sensitive regions, SSW shows superior performance in maintaining sharp transitions. For instance, in the Motorcycle image, SSW achieves 13.6% (All) and 6.6% (Nonocc), demonstrating better preservation of object boundaries compared to BF (23.4%, 16.6%) and GF (20.3%, 13.3%). This improvement stems from SSW's ability to minimize edge fattening, a common problem in traditional filter, in which sharp edges are excessively smoothed, resulting in incorrect depth estimation.

The qualitative results further confirm the advantages of SSW over other methods as shown in Figure 3. Visual comparisons indicate that SSW produces sharper disparity

maps with clearer object boundaries and fewer erroneous depth assignments in occluded regions. In high-frequency texture regions, such as Piano, traditional approaches introduce blurring, leading to inaccurate depth variations. However, SSW maintains fine structures with reduced texture bleeding, thereby ensuring greater accuracy in disparity reconstruction. For images such as Teddy, where occlusion handling plays a crucial role, SSW demonstrates superior robustness. The proposed method successfully suppresses noise in occluded regions while preserving depth transitions, leading to a smoother and more coherent disparity map. The non-occluded error for Teddy is 5.95%, the lowest among all methods, whereas BF records 13.7% and GF 10.9%, indicating that SSW significantly reduces disparity inconsistencies caused by occlusion.

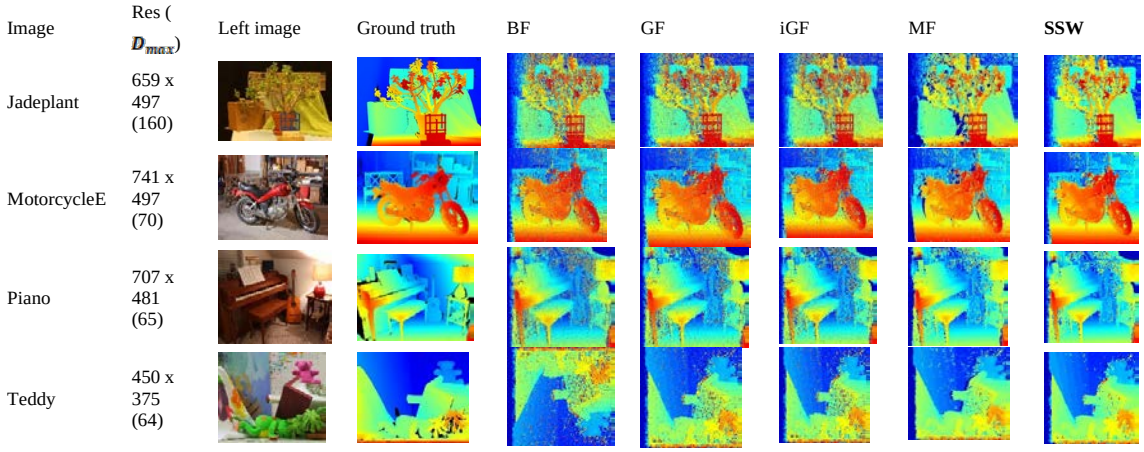

Figure 3. Middlebury datasets training qualitative performance

A comparison of different filtering techniques highlights the strengths and limitations of each approach. BF performs poorly in texture-dense regions due to its tendency to over-smooth high-frequency details. GF and its iterative variant, iGF, offer marginal improvements but still exhibit significant texture leakage, leading to depth inconsistencies. Meanwhile, MF demonstrates better edge preservation but remains susceptible to noise and occlusion artifacts. SSW outperforms these methods by incorporating a structure-sensitive weighting mechanism that selectively refines disparities based on local depth variations.

This approach successfully mitigates texture inconsistency by adapting to local intensity variations, ensuring that textures are accurately preserved without introducing artificial smoothing. Furthermore, SSW effectively addresses edge fattening by maintaining sharp transitions, leading to enhanced depth accuracy.

The experimental results confirm the proposed SSW method offers a significant improvement in disparity estimation accuracy compared to existing techniques. By effectively handling texture inconsistency and edge fattening, SSW achieves lower bad pixel errors and produces higher-quality disparity maps.

V. CONCLUSION

The proposed methodology introduces a five-stage local disparity map estimation framework that integrates multi-cost matching, segment-side window-based cost aggregation, and structured disparity refinement to address long-standing issues of edge fattening and texture inconsistency in local stereo matching. The first contribution lies in the use of multi-cost matching, where Truncated Absolute Difference (TAD), Gradient Magnitude (GM), and Census Transform (CT) are combined to capture complementary intensity, gradient, and structural information, resulting in a more robust and discriminative cost volume. The second contribution is the Segment-Side Window (SSW) aggregation approach, which combines SLIC superpixel segmentation with Side Window Filtering (SWF) to preserve object boundaries and enhance texture robustness while reducing noise in low-texture areas.

The third contribution is an enhanced disparity refinement process that incorporates left-right consistency checking and adaptive SWF smoothing to remove residual noise, recover fine details, and improve accuracy in occluded regions. Experimental results on the Middlebury benchmark show that the proposed framework consistently achieves lower error rates than existing techniques, while qualitative evaluations demonstrate sharper edges, smoother depth transitions, and reduced disparity errors in complex scenes. These findings confirm that the method offers a superior balance between accuracy and computational efficiency, making it suitable for high-precision stereo vision applications. Future work will focus on optimizing the framework for real-time GPU implementation and exploring hybrid deep learning integration to enhance scalability across applications such as industrial inspection, robotics, and 3D scene reconstruction tasks.

ACKNOWLEDGEMENTS

The authors wish to express their gratitude to Ministry of Higher Education (MoHE) Malaysia and Universiti Teknikal Malaysia Melaka (UTeM) for providing the funding necessary to accomplish this study via Fundamental Research Grants Scheme No: FRGS/1/2024/ICT09/UTEM /03/1/F00561.

CONFLICT OF INTEREST

The authors have no conflicts of interest to declare.

AUTHOR CONTRIBUTION

M. Y. A. Mohammed Al-Asbahi: Conceptualization, investigation, experimenting, data collection, writing – review and editing.

Ahmad Fauzan Kadmin: Supervision, problem statement, theory development, writing – review and editing.

Rostam Affendi Affendi: Writing – review and editing.

Nasharuddin Zainal: Data collection, examine and correct the manuscript.

Khairul Azha: Conceptualization, Investigation, Examine and correct the manuscript.

Nabil: Investigation and experimenting.

DATA AVAILABILITY

The data considered in this study were collected and accessed from Middlebury Stereo Dataset Standard Benchmarking under Middlebury College, Microsoft Research, and the National Science Foundation. The data is publicly available.

REFERENCES

[1] R. A. Hamzah, M. M. Roslan, A. F. Bin Kadmin, S. F. B. A. Gani, and K. A. A. Aziz, "JPG, PNG and BMP image compression using discrete cosine transform," TELKOMNIKA Telecommunication, Computing, Electronics and Control. 19(3), pp. 1010-1016, 2021.

[2] P. Maken and A. Gupta, 2D-to-3D: "A Review for Computational 3D Image Reconstruction from X-ray Images", Archives of Computational Methods in Engineering. 30(1), pp. 88-114, 2023.

[3] R. A. Hamzah et al., "A study of edge preserving filters in image matching", Bulletin of Electrical Engineering and Informatics, 10(1), pp. 111-117, 2021.

[4] M. S. Hamid, N. A. Manap, R. A. Hamzah, A. F. Kadmin, S. F. A. Gani, and A. I. Herman, "A new function of stereo matching algorithm based on hybrid convolutional neural network," Indonesian Journal of Electrical Engineering and Computer Science. 25(1), pp. 223 - 236, 2022.

[5] J. Huang, X. Fu, Z. Xiao, F. Zhao, and Z. Xiong, "Low-Light Stereo Image Enhancement", IEEE Transactions on. Multimedia, 25(1), pp. 2978-2992, 2023.

[6] S. F. A. Gani, R. A. Hamzah, R. Latip, S. Salam, F. Noraqillah, and A. I. Herman, "Image compression using singular value decomposition by extracting red, green, and blue channel colors," Bulletin of Electrical Engineering and Informatics, 11(1), pp. 168-175, 2022.

[7] M. Zahari, R. A. Hamzah, N. A. Manap, and A. I. Herman, "Stereo matching algorithm for autonomous vehicle navigation using integrated matching cost and non-local aggregation," Bulletin of Electrical Engineering and Informatics, 12(1), pp. 328-337, 2023.

[8] W. Gan, W. Wu, S. Chen, Y. Zhao, and P. K. Wong, "Rethinking 3D cost aggregation in stereo matching", Pattern Recognition Letters, 167, pp. 75-81, Mar. 2023.

[9] M. N. Zarina, and Z. Madiha, "Disparity Map from Stereo Images for Three-dimensional Surface Reconstruction", Engineered Science, 19, pp. 167-174, Jan. 2022.

[10] C. Deng, D. Liu, H. Zhang, J. Li, and B. Shi, "Semi-Global Stereo Matching Algorithm Based on Multi-Scale Information Fusion", Applied. Science, 13(2), pp. 1027, Jan. 2023.

[11] A. Gómez, "An Overview of GANet – Guided Aggregation Net for End-to-end Stereo Matching", Image Processing on Line, 13 (2023), pp. 215–226, Jul. 2023.

[12] Z. Huang, J. Gu, J. Li, and X. Yu, "A stereo matching algorithm based on the improved PSMNet", PLoS One, 16(8), pp. 1-16, Aug. 2021.

[13] R. Rahim, S. Woerz, and A. Zell, "LeanStereo: A Leaner Backbone based Stereo Network", 2023 International Joint Conference on Neural Networks (IJCNN), pp. 1-8, June 2023.

[14] A. Bangunharcana, J. W. Cho, S. Lee, I. S. Kweon, K. S. Kim, and S. Kim, "Correlate-and-Excite: Real-Time Stereo Matching via Guided Cost Volume Excitation", 2021 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), Dec. 2021.

[15] Z. Liu, Z. Li, W. Ao, S. Zhang, W. Liu, and Y. He, "Multi-Scale Cost Attention and Adaptive Fusion Stereo Matching Network", Electronics, 12(7), pp. 1594, Mar. 2023.

[16] F. Shamsafar, S. Woerz, R. Rahim, and A. Zell, "MobileStereoNet: Towards Lightweight Deep Networks for Stereo Matching", 2022 IEEE/CVF Winter Conference on Applications of Computer Vision, pp. 2417-2426, Feb. 2022.

[17] O. Zeglazi, M. Rziza, A. Amine, and C. Demonceaux, "Structural similarity measurement based cost function for stereo matching of automotive applications", Journal of Imaging, 6(8), pp. 77, Aug. 2020.

[18] W. J. Yang, Z.-S. Tsai, P.-C. Chung, and Y.-T. Cheng, "An Adaptive Cost Aggregation Method Based on Bilateral Filter and Canny Edge Detector with Segmented Area for Stereo Matching", International Workshop on Advanced Image Technology (IWAIT), pp. 110491J, Dec. 2019.

[19] W. Yuan, C. Meng, X. Tong, and Z. Li, "Efficient Local Stereo Matching Algorithm Based on Fast Gradient Domain Guided Image Filtering", Signal Processing: Image Communication, vol. 95, pp. 116280, Jul. 2021.

[20] R. A. Hamzah, M. N. Z. Azali, Z. M. Noh, M. Zahari, and A. I. Herman, "Development of depth map from stereo images using sum of absolute differences and edge filters", Indonesian Journal of Electrical Engineering and Computer Science, 25(2), pp. 875-883, Feb. 2022.

[21] R. A. Hamzah, A. F. F. Kadmin, M. S. S. Hamid, S. F. A. 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.

[22] J. J. Liaw, C. P. Lu, Y. F. Huang, Y. H. Liao, and S. C. Huang, "Improving census transform by high-pass with haar wavelet transform and edge detection", Sensors, 20(9), pp. 2537, Apr. 2020.

[23] D. Le and L. Nguyen, "Simple linear iterative clustering based low-cost pseudo-LiDAR for 3D object detection in autonomous driving", Multimedia Tools and Applications, 82, pp. 25253–25269, Feb. 2023.