Adapted Semiglobal Dynamic Cost Volume Solution for Stereo Matching Framework

A. F. Kadmin[1*], R. A. Hamzah[1], M. N. Abd Manap[2], M. S. Hamid[1], S. F. Abd Gani[1], K. A. A. Aziz[1]

[1]Fakulti Teknologi Kejuruteraan Eletrik dan Elektronik, Universiti Teknikal Malaysia Melaka, Hang Tuah Jaya, 76100 Durian Tunggal, Melaka, Malaysia
[2]Fakulti Kejuruteraan Elektronik dan Komputer, Universiti Teknikal Malaysia Melaka, Hang Tuah Jaya, 76100 Durian Tunggal, Melaka, Malaysia

[*]Corresponding author's email: fauzan@utem.edu.my

ABSTRACT: Stereo correspondences establishment for disparity map measurement is quite challenging in stereo matching due to textureless pixels and illumination variations in stereo images. This paper presents the solution to these issues with improved stereo matching framework based on adapted dynamic cost volume of Census Transform (BT) and dynamic histogram. The framework produced a precise cost volume structure at matching cost stage. An eight-directions Semiglobal (SGM) scanline and Anisotropic Diffusion (AD) filter employed in the cost aggregation to aggregate the cost volume. A final refinement method applied to increase the quality of the disparity map. Based on the Middlebury dataset platform, the framework has enhanced accuracy and surpassed several established frameworks.

Keywords: *Stereo matching, Semiglobal, Disparity map.*

1. INTRODUCTION

The stereo vision system aims to acquire depth measurement from 2D images called a disparity map. The main focus in stereo vision research is to improve the disparity map accuracy [1]. The traditional framework is developed based on four stages of the Schartein and Szeliski framework; 1. Matching cost, 2. Aggregation of cost, 3. Disparity optimisation, and 4. Final refinement [2]. It is challenging to produce precise stereo correspondences in the stereo matching due to several noises such as textureless pixels, radiometric differences, blurry edges reflected causes from poor segmentation methods [3].

This paper proposed an improved framework of stereo matching employing adapted dynamic cost volume of the semi-global method. A modified Census Transform (CT) combined with dynamic histogram to produce volume cost. An eight-directions Semiglobal (SGM) and Anisotropic Diffusion (AD) filter used in the cost aggregation to aggregate the cost volume. The remainder of this paper is organised as follows. Section 2 briefed the proposed framework. The experimental result and the discussion framework are explained in section 3. Section 4 will present the conclusion.

2. PROPOSED FRAMEWORK

The proposed improved stereo matching framework has been developed through four main stages. The framework blocks are presented in Fig. 1. A cost volume determined in the cost computation of matching stage which the matrix of cost volume is the stereo corresponding of the coordinate pixel to each other. The cost volume is the structure of pipeline cost data that will be directed in the cost aggregation and optimisation stage as shown in Fig. 2. The dimension is h x w x d where h = height, w = weight and d = disparity search space.

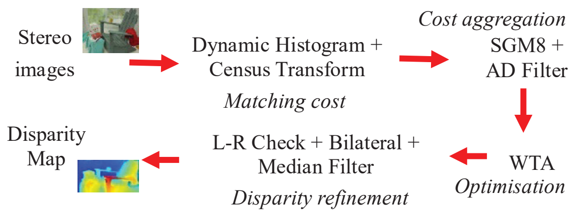

Figure 1 The framework of stereo matching

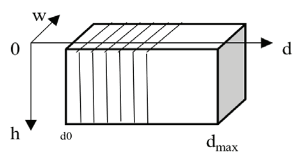

Figure 2 Cost volume structure

This stage partitioned the sub-histogram, splitting each sub-histogram for gray level portion and employing equalisation as in Eq. (1) [4]. It follows with Census Transform (CT) employed to focus on mismatched between the corresponding pixels [5] as in Eq. (2) and (3).

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

$$(x,y) = \underset{i=-u}{\overset{u}{\otimes}} \underset{i=-v}{\overset{v}{\otimes}} \xi(I(Y(x,y)), I(Y(x+i,y+i))) \quad (2)$$

$$\xi(q_1,q_2) = \begin{cases} 1, if\ q_1 > q_2 \\ 0, if\ q_1 \le q_2 \end{cases} \quad (3)$$

In the cost aggregation stage, an eight-directions scanline of SGM and AD filter approach are utilised to aggregate the cost volume. This framework computes the cost volume for disparity at this stage and minimised the aggregated cost using the Winners-Take-All (WTA) to acquire the accurate disparity map in the optimisation stage. The WTA improves the computational cost as applied by [6]. In the last stage of the framework, this technique is called disparity refinement, which corresponds to the target depth map employing the same method employed by [7]. The framework performance is evaluated using the Middlebury Vision Evaluation Platform by measuring the average error percentage for nonoccluded and all pixels.

3. RESULT AND DISCUSSION

All conducted experiments are accomplished utilising the Windows 10 platform with Intel Xeon 2.6GHz processor with 64GB memory. The algorithm executed all the stereo pair images in about 337 seconds with an average of about 22.1 seconds per image using the MATLAB environment. Fig. 3 shows the results of disparity map from the framework on the Middlebury images with the reference and ground truth images produced from standard 2D and 1D disparities [2]. It indicates that the framework worked well around the border and uniform surfaces, as marked in the red circle. However, the disparity map produces only minimum streak artefact effects compared with local methods that commonly appear in the matching process.

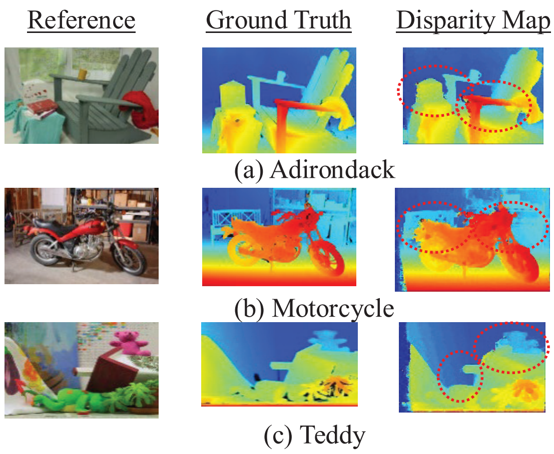

Figure 3 Reference image versus the disparity map (a) Adirondack (b) Motorcycle (c) Teddy

Table 1 reports the Middlebury quantitative evaluation of the framework for average errors assessed compared with several established frameworks. The other stereo matching framework is the adaptation of Hirschmuller's SGM method (SGBM1), Binary Stereo Matching (BSM) and Multiview Stereo (R-NCC). It shows that the established technique is produced average results in the comparison table, producing 12.9% and 8.9% of all and nonocc for average errors outperforming other frameworks. The adapted dynamic cost volume of cost computation also produces fine contour of disparity levels with minimum noises presented in the map.

4. CONCLUSION

An improved stereo matching framework based on adapting SGM dynamic histogram cost volume is formulated in this work to improve the disparity map accuracy. The experimental result shows the framework worked well around the surfaces and achieved impressive level of disparity accuracy with acceptable execution time in the Middlebury dataset platform compared with several other stereo matching framework.

Table 1 Comparison of the Middlebury results average error performance.

Image	SGBM1		BSM		R-NCC		Proposed	
	Non %	All %	Non %	All %	Non %	All %	Non %	All %
Adiro	18.3	21.1	7.27	12.7	20.5	21.2	**4.59**	**7.84**
Motor	3.48	11.0	6.67	14.8	9.59	11.5	**5.61**	**12.2**
Piano	6.51	11.6	10.8	16.0	9.12	9.59	**5.72**	**8.95**
Shelvs	15.1	17.6	16.4	19.2	11.5	11.7	**12.1**	**14.6**
Avg	10.8	15.3	10.3	15.7	12.8	13.7	**8.9**	**12.9**

ACKNOWLEDGEMENT

We are grateful to Centre for Research Innovation Management (CRIM), Universiti Teknikal Malaysia Melaka (UTeM) and Ministry of Higher Education (MOHE), Malaysia for their assistance and funding to support this work.

REFERENCES

[1] R. A. Hamzah and H. Ibrahim, "Literature survey on stereo vision disparity map algorithms," *Journal of Sensors*, vol. 2016, pp. 1–22.

[2] D. Scharstein *et al.*, "High-resolution stereo datasets with subpixel-accurate ground truth," in *German Conference on Pattern Recognition*, 2014, pp. 31–42.

[3] T. T. San and N. War, "Local stereo matching under radiometric variations," in *18th IEEE/ACIS International Conference on Software Engineering, Artificial Intelligence, Networking and Parallel/Distributed Computing (SNPD)*, 2017, pp. 245-249.

[4] M. A. Al Wadud *et al.*, "A dynamic histogram equalization for image contrast enhancement," *IEEE Transactions on Consumer Electronics, 53(2), 2007, pp. 593-600.*

[5] E. Bebeselea-Sterp, R. R. Brad, and R. R. Brad, "A Comparative Study of Stereovision Algorithms," *International Journal of Advanced Computer Science and Applications, vol. 8, no. 11, 2017, pp. 359–375.*

[6] 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," in *2018 IEEE Conference on Systems, Process and Controls (ICSPC)*, 2018, pp. 39-43.

[7] 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, 2019, pp. 509–516.