REAL GEOMETRIC EDGES AIDED TO PRESERVING THE DEPTH DISCONTINUITIES IN DENSE STEREO MATCHING

[1] *Ting Fang Chiu* (邱鼎方), [2] *Chiou-Shann Fuh* (傅楸善)

[1] Department of Civil Engineering,
[2] Department of Computer Science and Information Engineering,
National Taiwan University, Taiwan,
E-mail: r05521116@ntu.edu.tw

ABSTRACT

The integrity of geometric edge is vital to the high-quality 3-D scene representation. The way to conquer depth discontinuities usually need to precisely distinguish foreground and background that causes depth discontinuities, so the matching cost aggregation or the penalty parameters of depth discontinuities can be corrected. Therefore, geometric edge detection is an important clue to strengthen accuracy of depth continuities. This research develops an approach to combine local method and semi-global method, and the most importantly, simulates real geometric edges to combine with local method and semi-global method. The approach develop to quantify the maximum quality increase of edge detection aided to dense stereo matching.

Keywords: *Dense stereo matching, Depth discontinuities, Local method, Semi-global method.*

1. INTRODUCTION

In dense stereo matching, problems arising from image noise, textureless region, depth discontinuities and occlusions are quite common [1][2][3]. This research focuses on the problem of depth discontinuities. In global algorithms, the methods of conquering depth discontinuities usually allocate different matching cost weighting for distinguishing foreground and background that causes depth discontinuities, so the matching cost or the penalty parameters of depth discontinuities can be eased [4][5][6][7]. In local method, matching cost aggregation influences the depth result. The fixed window easily combines different depth then results wrong aggregation value, especially in depth discontinuities. In order to handle this problem, building a suitable support region that fixes object geometric shape is very important. There are many local algorithm modifying their support region, like: Shiftable Window[8], Multiple Window [9], Image Segmentation [10], Adaptive Weight [11], Cross-based [12]. Therefore, we can find that geometric edge detection is a necessary stage in giving clues of depth continuities, both in local methods and global methods. In this paper, we simulate real geometric edges from real depth images to find the maximum edge's benefit in stereo matching. The matching algorithm uses local method to find initial depth then uses semi-global method to refine the initial depth. The real geometric edges constraint both local method and semi-global method.

2. RESEARCH METHOD

This paper uses cross-based support region to compute the matching cost aggregation. The cross-based support region is determined by the real geometric edges. After finding the initial depth, we use median filter to determine the depth refined range. The semi-global method considers the position of the real geometric edges to give special penalty parameters. The modified semi-global method refines initial depth especially in depth discontinuities. In the end, this paper uses modified median filter, which reserves depth discontinuities, to reduce noise and avoid depth discontinuities region being over smooth. The research flowchart is shown as Figure 1.

2.1 Extract real geometric edges

We check the pixels' real depth and their neighbors' depth (Up, Down, left and right). If one of the directional difference is larger than one pixel, we consider this pixel to be geometric edge, like Figure 2.

2.2 Computation of matching cost

This paper combines RGB-band gradient's Absolute Difference (AD) and RGB-band Census Transform (Census) to be the similarity index between two pixels. The equation of AD, Census and combination is shown as (1).

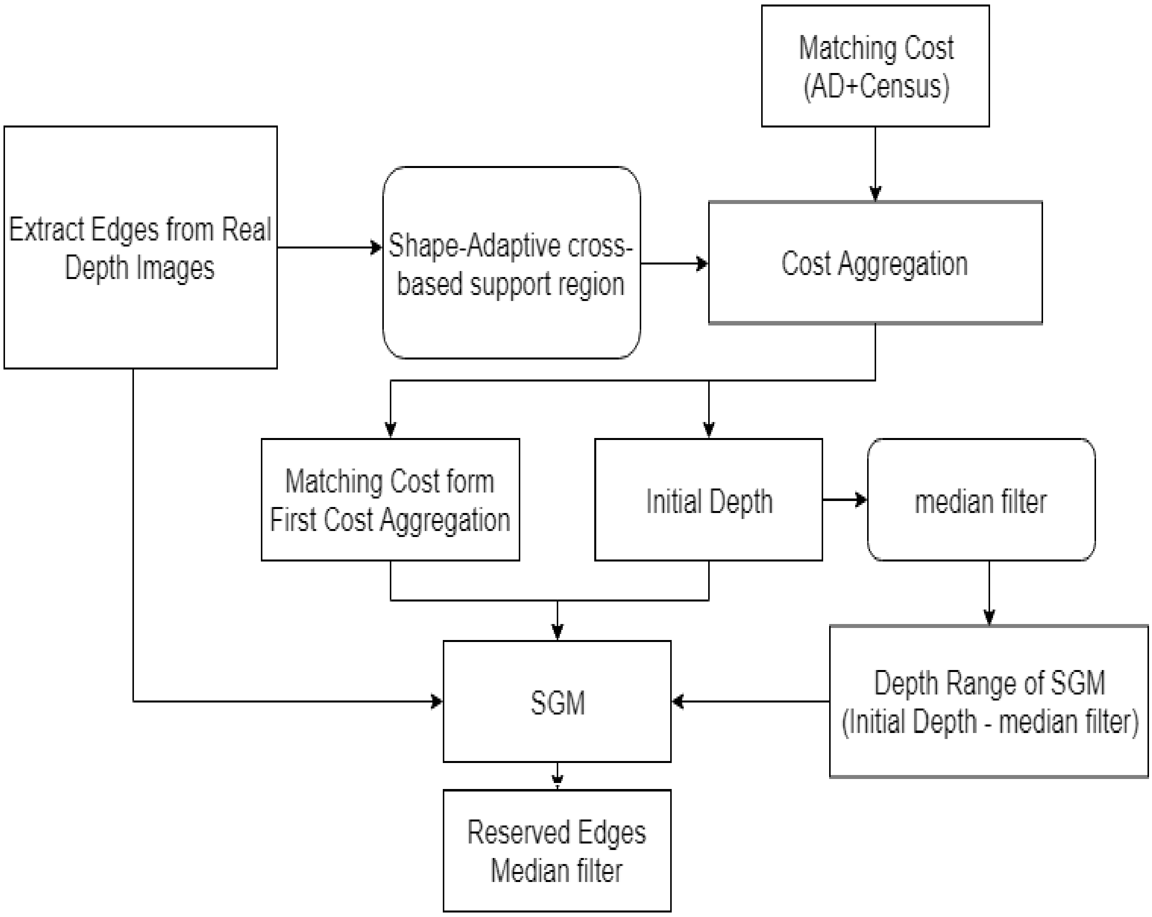


Figure 1: The research flowchart.

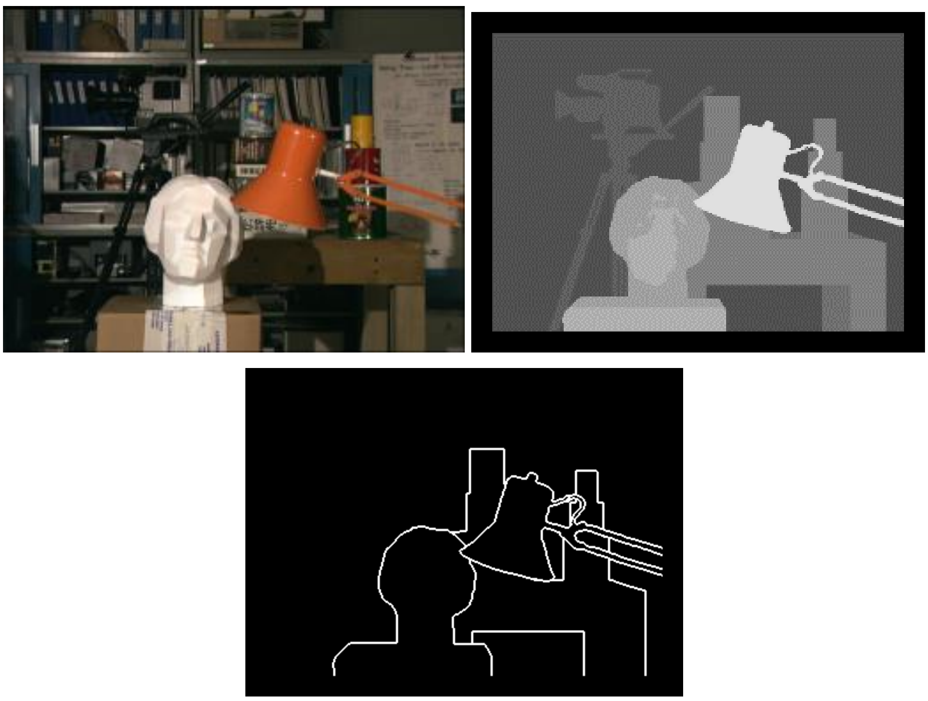


Figure 2: The color stereo image, real depth image and geometric edges.

$$C_{AD}(p,d) = \frac{1}{3}\sum_i \left| I_i^{left}(p) - I_i^{right}(p,d) \right|$$
$$\forall i \in \left\{ \frac{\partial I}{\partial x}, \frac{\partial I}{\partial y}, RGB - based \right\} \quad (1\text{-}1)$$

$$C_{Census}(p,d) =$$
$$\frac{1}{3}\sum_i Ham(Census_{left}(p) - Census_{right}(p,d))$$
$$\forall i \in \{RGB - based\} \quad (1\text{-}2)$$

$$\mathrm{C(p,d)} = \rho(C_{AD}(\mathrm{p,d}), \lambda_{AD}) + \rho(C_{census}(\mathrm{p,d}), \lambda_{census})$$
$$\rho(\mathrm{c}, \lambda) = 1 - \exp(-\frac{c}{\lambda}) \quad (1\text{-}3)$$

The *Ham* is the Hamming Distance like Figure 3. The purpose of using combining method is strengthening the image's textureless features and geometric features.

2.3 Aggregation of matching cost

In this paper, we use real geometric edge to constraint cross-based support region. The algorithm of edge aided to cross-based support region starts from the central pixel then search up, down, left and right. If the direction meets the edges, stop searching. After that, we build the horizontal *H*(*p*) and vertical branches *V* (*p*) . Then start again on the vertical branches *V* (*p*) and search horizontal regions. The entire steps are shown as Figure 4.

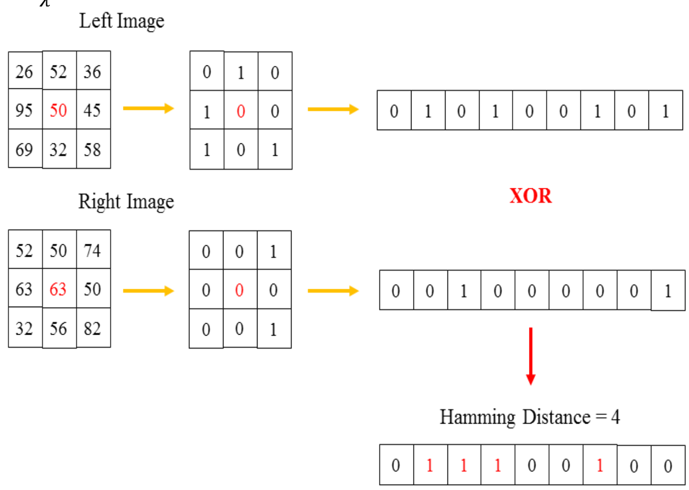


Figure 3: The illustration of Hamming Distance [7].

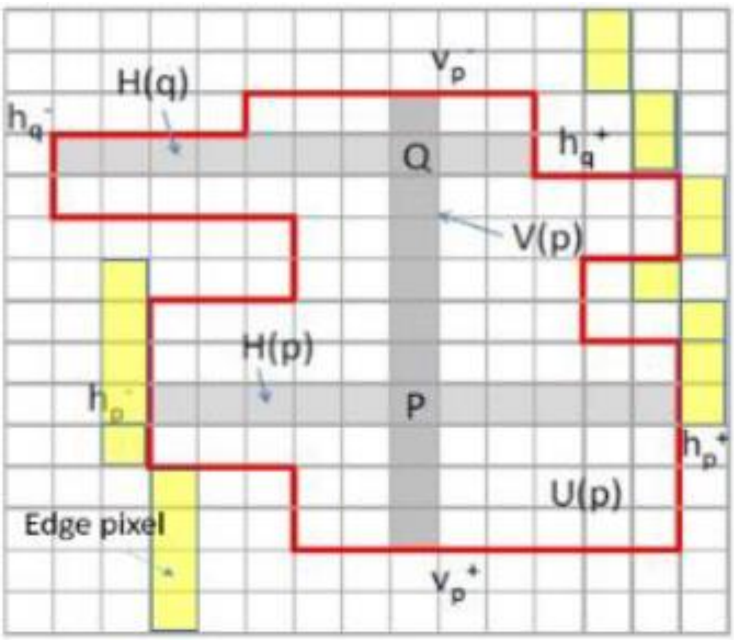


Figure 4: The illustration of cross-based method [6].

After defining support region, the approach gives weight to the neighbor pixels. The weight is based on the L, a, b difference in CIELab color space and geometric distance. The equation is shown as (2).

$$\Delta c_{pq} = \sqrt{(L_p - L_q)^2 + (a_p - a_q)^2 + (b_p - b_q)^2} \quad (2\text{-}1)$$

$$\Delta g_{pq} = \sqrt{(x_p - x_q)^2 + (y_p - y_q)^2} \quad (2\text{-}2)$$

$$\mathrm{w(p,q)} = \exp\left(-\frac{\Delta c_{pq}}{\gamma_c}\right) \cdot \sqrt{\exp\left(-\frac{\Delta g_{pq}}{\gamma_p}\right)} \cdot w_q \quad (2\text{-}3)$$

2.4 Depth refinement

After cross-based matching to find the initial depth, we use median filter to smooth initial depth then compute the difference between initial depth and smooth depth. The difference will be the refined range. In refining steps, we first use semi-global matching (SGM) to improve the depth based on the initial depth and refined range. The matching cost that SGM uses inherit the aggregation cost. In the end, the approach uses the reserved edge median filter to refine the last depth result.

2.4.1 Modified SGM

[13] proposed a semi-global matching method, which uses 8 or 16 directional one-dimension energy functions to replace original two-dimension energy function. The equation is shown as (3)

$$\mathrm{E(x,y,d)} = \mathrm{C(x,y,d)} + \min\begin{cases} \mathrm{E(x-r,y,d)} \\ \mathrm{E(x-r,y,d-1)} + P_1 \\ \mathrm{E(x-r,y,d+1)} + P_1 \\ min_i(\mathrm{E(x-r,y,i)} + P_2) \end{cases} - min_k(\mathrm{E(x-r,y,k)}) \quad (3\text{-}1)$$

$$\mathrm{S(x,y,d)} = \sum_r E_r\,(x,y,d) \quad (3\text{-}2)$$

$$\mathrm{d(x,y)} = argmin_d S(x,y,d) \quad (3\text{-}3)$$

$\mathrm{E(x-r,y,d)}$ is previous minimum sum of matching cost in the path r. $min_k(\mathrm{E(x-r,y,k)})$ is used to avoid over accumulate. It doesn't affect the final result. This method sums 8 or 16 directional matching cost $\mathrm{S(x,y,d)}$ and take the depth $\mathrm{d(x,y)}$, which has minimum matching cost to be the result. P_1 is the smaller penalty parameter. It is used to smooth depth for consistency. P_2 is the larger penalty parameter. It is used to sharpen depth discontinuities for describing object's contour. According to [6], this paper adds new penalty parameter P_3 to replace P_2 in depth discontinuities. The equation is shown as (4).

$$\begin{cases} P_2 = \frac{1}{N_{cp}} \sum_{i \in N_{cp}} (Ms_i - Mf_i) \\ P_1 = P_2/2 \\ P_3 = (P_2 + P_1)/2 \\ N_{cp} = number\ of\ confidence\ pixel \end{cases} \quad (4)$$

Ms_i is the second minimum matching cost. Mf_i is the minimum matching cost. N_{cp} is the sum of pixel i that $\frac{Mf_i}{Ms_i} > \tau$. The matching cost in modified SGM is inherited by aggregation matching cost in local method. It can have better performance than original similarity index. The depth difference between initial depth and smooth depth by median filter is the refined range. The SGM searching range is shown as (5).

$$\text{SGM search range} = [inid - refined\ range, inid + refined\ range]$$
$$\text{refined range} = \mathrm{abs}(\mathrm{inid} - median_{inid}) \quad (5)$$

2.4.2 Reserved edge median filter

First, we check the edge's depth difference between neighbors if larger than one pixel. If it is larger than one pixel, we reserve this edge's pixel. Next, according to the kernel size, we use cross-based method to cut support region based on the reserved edges. In the end, we take the median depth in the support region to be the final result.

3. EXPERIMENTAL RESULTS

In this paper we used [14] 4 stereo images with real disparity (Tsukuba, Venus, Teddy, and Cones) to be our experimental data. The related variables setting are shown as Table 1. The way of evaluating quality is based on [15][16]. We divided two area, depth discontinuities and non-occlusion area, to quantify the accuracy in different characteristic areas. The index of accuracy is percentage of pixel whom error is larger than one pixel.

Table 1: The related variables setting

Test image	Tsukuba	Venus	Teddy	Cones
Image				
Resolution	384×288	434×383	450×375	450×375
Depth range	5~16	4~20	8~55	17~55
Gradient setting ($\sigma, h \times w$)	(0 , 5×5)	(1 , 5×5)	(1 , 5×5)	(0 , 5×5)
Census kernel	3×3	3×3	3×3	3×3
Aggregation kernel	31 ×31	31 ×31	31 ×31	31 ×31
$(\lambda_{AD}, \lambda_{Census})$	(25 , 25)	(25 , 25)	(25 , 25)	(25 , 25)
τ	0.75	0.75	0.75	0.75

The results in every steps are shown as Figure 5. We compared results with original SGM by [13]. In Figure 5, we can see that the real geometric edges can really raise the entire depth's accuracy, especially in depth discontinuities. It can prove that the real geometric edges can precisely divide background and foreground in cost aggregation and depth optimization. The accuracy in every steps also shows that the entire researching flow refines the depth steps by steps. This approach has a good initial depth firstly then refines initial depth by depth optimization and depth refinement

In depth refinement, we compare result of the modified median filter and original median filter in Figure 6 and Figure 7. We can see that the modified median filter preserved more depth discontinuities features than original median filter. So it can have better results in depth discontinuities even in depth continuities area. It is because the precise depth border can separate background and foreground to make both areas smoother.

The results compared with other algorithms are shown as Table 2. We choose three stereo matching methods, an adaptive support-weight approach (ASW) [11] with the SGM depth optimization, a cross-based method (CB) [12] with the SGM depth optimization, and Hybrid method [6] to evaluate the virtues of the proposed method. In Table 2, the proposed method has the best result no matter in depth discontinuities or non-occlusion area. It can prove that this method if has good enough edge information, just like real geometric edges, we can have very accurate result by this approach especially in depth discontinuities.

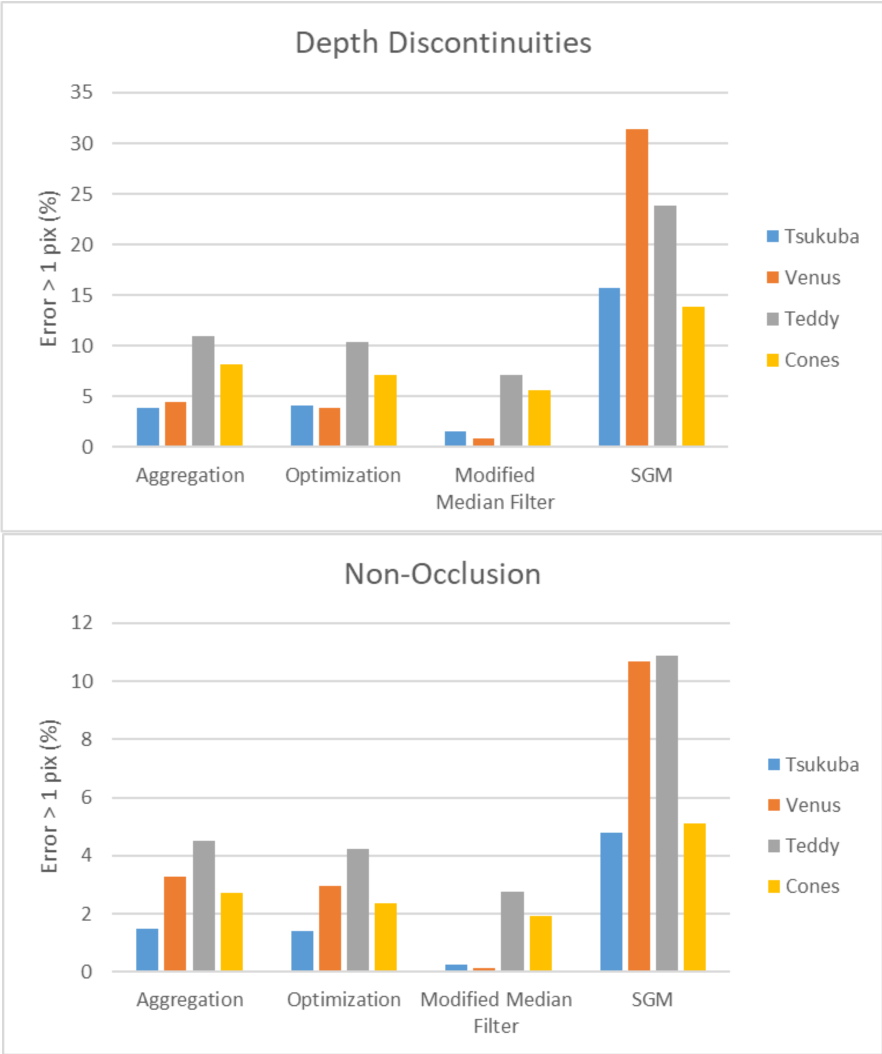
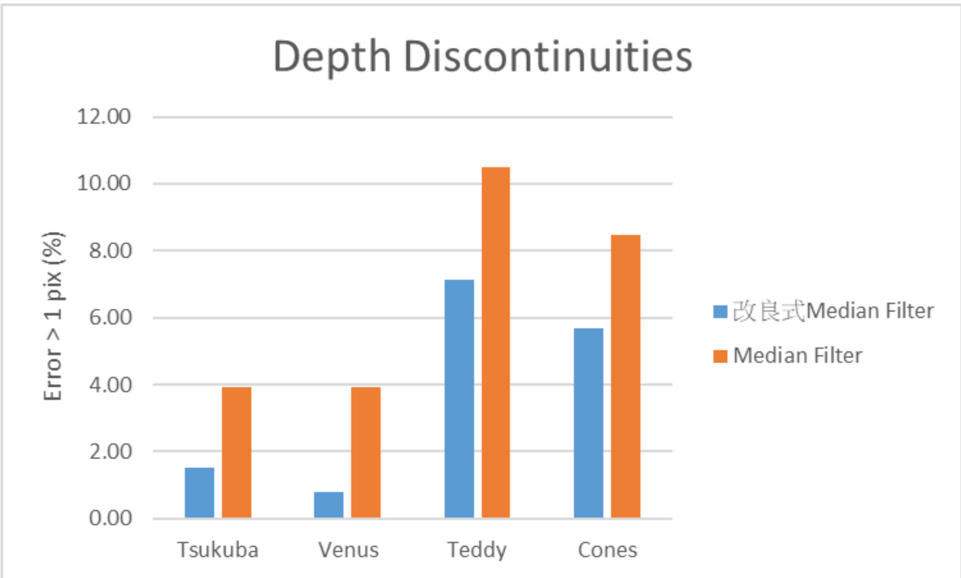


Figure 5: The accuracy in every steps, which compared with SGM.



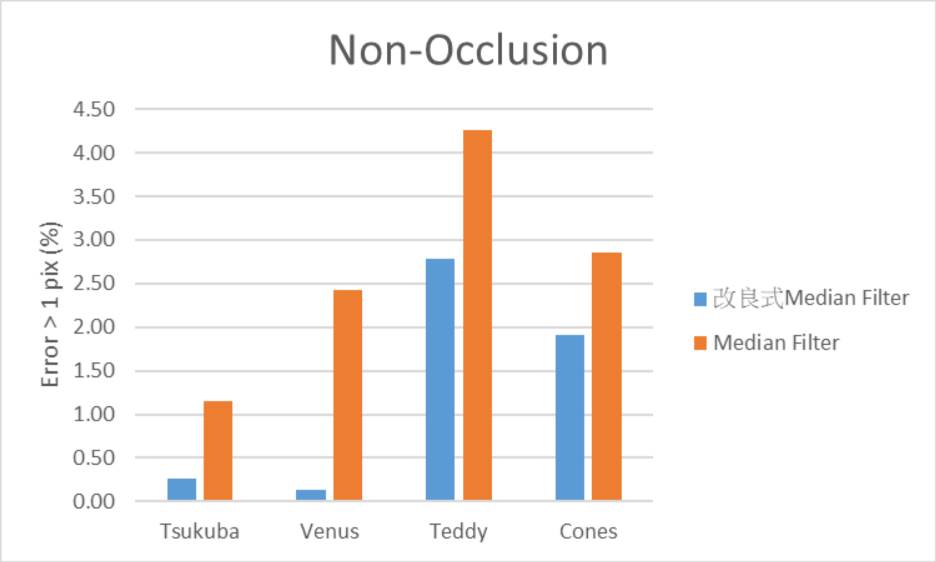
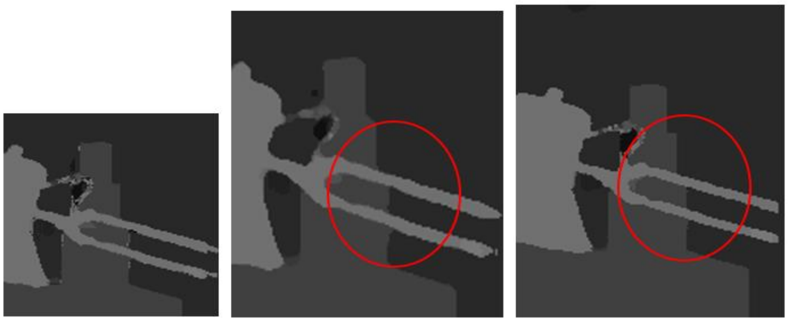


Figure 6: The accuracy of median filter and modified median filter.



(a) original depth (b) median filter (c) modified median filter

Figure 7: The comparison with median filter and modified median filter.

Table 2: The comparison with other methods.

Test image	Tsukuba	Venus	Teddy	Cones
	Non-occlusion (%)			
SGM	4.8	10.7	10.9	5.12
ASW	2.27	1.93	9.42	4.14
CB	1.48	0.62	7.74	3.17
Hybrid	1.44	0.59	7.71	3.13
Proposed method	0.26	0.13	2.78	1.91
	Disparity Discontinuity (%)			
SGM	15.7	31.4	23.9	13.9
ASW	11.1	20.6	22.7	11.9
CB	6.52	6.64	18.7	10.8
Hybrid	4.63	5.83	15.6	8.9
Proposed method	1.51	0.81	7.14	5.67

4. CONCLUSIONS AND FUTURE WORK

In this paper, the effectiveness of the proposed approach has been verified with the Middlebury stereo benchmark. This paper proposes the method that combines real geometric edges in cost aggregation, depth optimization and depth refinement. The results present that this approach has precise stereo correspondence, especially in depth discontinuities. The shape-adaptive cross-based aggregation approach with the edge constraint generates a precise initial disparity map and gives the precise penalty estimation. As a consequence, the sensibility of the SGM cost aggregation towards the penalty parameters can be alleviated. In summary, the study imposes the edge constraint onto the energy function of SGM and cost aggregation to improve the accuracy in depth discontinuity regions. We can see that if we have precise geometric edges, this approach can generate a very precise disparity map. So in the future, how to take these precise and complete geometric edges is the first task need to do. Moreover, the optimization in reducing computational complexity and further assessment will be studied in future work.

REFERENCES

[1] W. Wei, and K. N. Ngan, "Disparity estimation with edge-based matching and interpolation," *International Symposium on Intelligent Signal Processing and Communication Systems*, pp. 153-156, 2005.

[2] Y. J. Chang and Y. S. Ho, "Disparity estimation using edge preserving method." *TENCON 2015-2015 IEEE Region 10 Conference*, pp.1-4, 2015.

[3] J. Chen, D. Ian, S. Li, Z. Li, M. Madden, J. Mills, N. Paparoditis, F. Rottensteiner, M. Sester, C. Toth, J. Trinder, and C. Heipke, "Information from imagery: ISPRS scientific vision and research agenda." *ISPRS Journal of Photogrammetry and Remote Sensing*, 115: 3-21, 2016

[4] C. ÇIĞLA, and A.A. Alatan, "Efficient edge-preserving stereo matching." *Computer Vision Workshops (ICCV Workshops) 2011 IEEE International Conference*, pp.696-699, 2011.

[5] X. Sun, X. Mei, S. Jiao, M. Zhou, Z. Liu, and H. Wang, "Real-time local stereo via edge-aware disparity propagation." *Pattern Recognition Letters*, 49: 201-206, 2014.

[6] T. Y. Chuang, H. W. Ting, and J. J. Jaw, "Hybrid-Based Dense Stereo Matching." *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, Vol. XLI, Part B3, pp.495-501, 2016.

[7] H.C. Lee, "Semi-Global Matching by Integrating Edge Feature Information.", Thesis, National Taiwan University, Taipei, Taiwan. (in Chinese), 2016.

[8] D. Scharstein, and R. Szeliski, "A taxonomy and evaluation of dense two-frame stereo correspondence algorithms", *International journal of computer vision*, 47(1-3): 7-42, 2002

[9] H. Hirschmüller, P. R. Innocent, and J. Garibaldi, Real-time correlation-based stereo vision with reduced border errors, *International Journal of Computer Vision*, 47(1-3), 229-246, 2002.

[10] M. Gerrits, and P. Bekaert, "Local stereo matching with segmentation-based outlier rejection", *The 3rd Canadian Conference on Computer and Robot Vision*, pp. 66, 2006

[11] K. J. Yoon, and I. S. Kweon, "Adaptive support-weight approach for correspondence search", *IEEE Transactions on Pattern Analysis & Machine Intelligence*, (4): 650-656, 2006.

[12] K. Zhang, J. Lu, and G. Lafruit, "Cross-based local stereo matching using orthogonal integral images", *IEEE Transactions on Circuits and Systems for Video Technology*, 19(7): 1073-1079, 2009.

[13] H. Hirschmüller, "Accurate and efficient stereo processing by semi-global matching and mutual information", *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, Vol. 2, pp. 807-814, 2005.

[14] D. Scharstein, and R. Szeliski. Middlebury Stereo Vision Page. http://vision.middlebury.edu/stereo/ (17 Apr. 2018), 2008.

[15] D. Scharstein, and R. Szeliski, "A taxonomy and evaluation of dense two-frame stereo correspondence algorithms", *International journal of computer vision*, 47(1-3): 7-42, 2002.

[16] D. Scharstein, and R. Szeliski, "High-accuracy stereo depth maps using structured light", IEEE Computer Society Conference on Computer Vision and Pattern Recognition, Vol. 1, pp. 195-202, 2003