Proceedings of Symposium on Electrical, Mechatronics and Applied Science 2018 (SEMA'18), pp. 153-154, 8 November 2018

Local stereo matching algorithm based on pixel difference adjustment and weighted median filter

Yeouwei Gan[1*], R.A. Hamzah[2], N.S. Nik Anwar[1]

[1)] Fakulti Kejuruteraan Elektrik, Universiti Teknikal Malaysia Melaka, Hang Tuah Jaya, 76100 Durian Tunggal, Melaka, Malaysia
[2)] Fakulti Teknologi Kejuruteraan Elektrik & Elektronik, Universiti Teknikal Malaysia Melaka, Hang Tuah Jaya, 76100 Durian Tunggal, Melaka, Malaysia

[*]Corresponding e-mail: melvin.gyw92@gmail.com

Keywords: Disparity map; Stereo matching; weighted median

ABSTRACT – This paper presents a new algorithm for stereo vision system to obtain depth map or disparity map. The proposed algorithm for stereo vision system consist of three stages. The first stage starts with matching cost computation where pixel based differences methods are used. The methods are Absolute Difference (AD) and Gradient Matching (GM). Next, second stage, where disparity optimization focuses on Winner-Takes-All (WTA) approach to normalize the disparity values of each pixel of the image. Finally, weighted median (WM) filter is added to reduce the noise on the disparity map.

1. INTRODUCTION

The purpose of this work is to proposed a new algorithm for stereo vision system to obtain depth map or disparity map. For this proposed algorithm, a three step taxonomy is used. These steps are matching cost computation, disparity optimization and disparity refinement. Disparity map optimization method can be classified into three; global, semi global and local method. Their classification relies on the method the disparity is calculated [1].

Global methods define the disparity map based on the energy minimization method. This method is computationally complex, expensive and require longer processing time. Local method requires less computational time as this method is usually less complex. Therefore, it is able to obtain faster execution time. While for the semi global method is combination of both.

This paper proposes a new local-based stereo matching algorithm to obtain disparity map. The propose work utilized Absolute Differences (AD) and Gradient Matching (GM) at it initial stage to combine the pixel of the image from the left and right. Then, local method, Winner-Take-All (WTA) is used for optimization or normalization of the disparity map. Finally, the noise is reduced using Weighted Median (WM) filter. The fifteen samples of stereo images are from a standard benchmarking dataset, that is Middlebury Stereo. The samples are then processed by the proposed algorithm. Finally, the processed sample are uploaded to Middlebury again. The quantitative result will be given by Middlebury, where they are given based on the percentage of error for each sample of pictures.

2. METHODOLOGY

The proposed work is categorized as a three steps process and the following is the block diagram of the proposed methods, presented in figure 1.

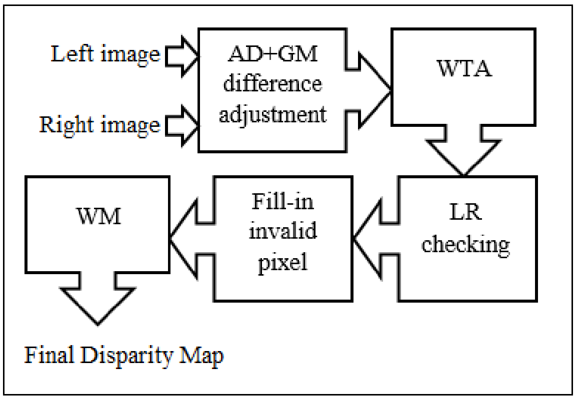

Figure 1 Block diagram for proposed algorithm.

2.1 Matching Cost Computation

The matching process was started by using the absolute difference (AD), and gradient matching (GM) implemented by Hamzah et al. [2]. This is where the pixels of the left image correspond to the right image form the initial disparity map. The equation for AD is presented in Equation (1):

$$AD(p,d) = \left| I_l(p) - I_r(p,d) \right| \quad (1)$$

where $p = (x, y)$, represent the coordinate of the pixel of interest, d is disparity or depth value and I_l represent left image and I_r represent right image. The *GM* method was implemented together with *AD* to improve the accuracy by overcoming the high distortion on disparity map such as error at boundaries and low texture area. This is based on the *AD* algorithm implemented by Tan and Monasse [3]. The equation for *GM* is presented in Equation (2):

$$GM(p,d) = \left| m_l(p) - m_r(p,d) \right| \quad (2)$$

where m_l represent the gradient from left image and m_r, represent gradient of the right image The combination of both method resulting the matching cost function $M(p,d)$. The equation is given in Equation (3):

$$M(p,d) = AD(p,d) + GM(p,d) \quad (3)$$

© Centre for Robotics and Industrial Automation, Faculty of Electrical Engineering

2.2 Disparity optimization

The disparity map obtained from the first process is optimized by utilizing Winner-Take-All (WTA) strategy. Minimum disparity value is selected on each pixel of the disparity map. The formulation is given on Equation (4):

$$d(x,y) = \arg\min_{d \in D} M(p,d) \qquad (4)$$

where D denotes the range of disparity on an image, $d(x,y)$ is the selected disparity value at the location of (x,y) and $M(p,d)$ is the value of the first stage.

2.3 Disparity Refinement

This stage started with the post processing, LR checking, implemented by Mattoccia et al. [4] and Kordelas et al. [5] to detect outliner. Next, Fill-in invalid pixels was done with left image as reference. The filling process started with left to the right valid pixel replacement. The process was continued by utilizing weighted median filter with bilateral filter. The equation of bilateral filter $B(p,q)$ is shown in Equation (5):

$$B(p,q) = \exp\left(-\frac{|p-q|^2}{\sigma_s^2}\right)\exp\left(-\frac{|d(p)-d(q)|^2}{\sigma_c^2}\right) \qquad (5)$$

where the (p,q) represent the pixel of interest , and $|p-q|$ refer to spatial Euclidean and $|d(p)-d(q)|^2$ refer to Euclidean. σ_s^2 and σ_c^2 are the spatial distance and color similarity parameters. Bilateral filter is a type of edge preserving filter which aid in disparity map accuracy. A higher weighted is use on this filter and weighted median [6,7] was implemented. The weighted of $B(p,q)$ is transform into summation of histogram $h(p,d_r)$ which result in Equation (6).

$$h(p,d_r) = \sum_{q \in w_p | d(q) == d_r} B(p,q) \qquad (6)$$

where d_r denotes the disparity range and w_p is the window size with the radius $(r \times r)$ at centered pixel of p. The final disparity value WM is determined by the median value of $h(p,d_r)$ given by Equation (7):

$$WM = med\{d \mid h(p,d_r)\} \qquad (7)$$

3.0 RESULT AND DISCUSSION

A standard benchmarking dataset was used in this work, which is fifteen set of stereo images from Middlebury Stereo [9]. The experiments are run through a personal computer with the feature or CPU i7 and 8G RAM. The *nonocc error* represent the error of invalid disparity values on the non-occluded regions and *all error* represents the error of invalid disparity values on all pixels of the disparity map image.

Table 1 shows the quantitative result of the propose framework which is compared with AVERAGE_ROB and MEDIAN_ROB. If compared with these current method, the proposed algorithm is ranked top of the table with value of 38.0% and 41.9% of nonocc and all errors respectively. When compare with AVERAGE_ROB, the propose algorithm reduces nonocc and all error by 34.7% and 28.9%. Meanwhile, compare with MEDIAN_ROB, the proposed algorithm reduces nonocc and all error by 41.9% and 36.1%.

Table 1 The average results of the Middlebury training dataset.

Algorithm	nonocc error (%)	all error (%)
Proposed work	38.0	41.9
AVERAGE _ROB [9]	72.7	70.8
MEDIAN_ROB [9]	79.9	78.0

4.0 CONCLUSIONS

A new stereo matching algorithm is presented in article. Furthermore, the proposed work is able to yield better accuracy compared with AVERAGE_ROB and MEDIAN_ROB in Table 1.

ACKNOWLEDGEMENT

Authors are grateful to Universiti Teknikal Malaysia Melaka for the financial support through PJP/2018/FTK(13C)/S01632.

REFERENCES

[1] Scharstein, D., & Szeliski, R. (2002). A taxonomy and evaluation of dense two-frame stereo correspondence algorithms. *International Journal of Computer Vision*, *47*(1-3), 7-42.

[2] Hamzah, R. A., Ibrahim, H., & Abu Hassan, A. H. (2017). Stereo matching algorithm based on per pixel difference adjustment, iterative guided filter and graph segmentation. *Journal of Visual Communication and Image Representation 42*, 145-160.

[3] Tan, P., & Monasse, P. (2014). Stereo disparity through cost aggregation with guided filter, Image Process. *On Line (IPOL) 4*, 252–275.

[4] Mattoccia, S., Tombari, F., & Di Stefano, L. (2007). Stereo vision enabling precise border localization within a scanline optimization framework. *Asian Conference on Computer Vision, Springer*, 517–527.

[5] Kordelas, G. A., Alexiadis, D. S., Daras, P., & Izquierdo, E. (2016). Content-based guided image filtering, weighted semi-global optimization, and efficient disparity refinement for fast and accurate disparity estimation. *IEEE Trans. Multimedia 18 (2),* 155–170.

[6] Min, D., Lu, J., & Do, M. N. (2012). Depth video enhancement based on weighted mode filtering. *Pattern Recogn. Lett. 21 (3)* 1176–1190.

[7] Ma, Z., He, K., Wei, Y., Sun, J., & Wu, E. (2013). Constant time weighted median filtering for stereo matching and beyond, *IEEE International Conference on Computer Vision.* 49–56.

[8] Hamzah, R. A., & Ibrahim, H. (2016). Literature survey on stereo vision disparity map algorithms, *J. Sensors 8742920.*

[9] Szeliski, R. (2011). Computer Vision: Algorithms and Applications. *Springer London Ltd.*

[10] Scharstein, D., & Szeliski, R. (2018). Middlebury stereo evaluation (accessed date: July 2018, http://vision.middlebury.edu/stereo/eval3/).

© Centre for Robotics and Industrial Automation, Faculty of Electrical Engineering