

An Improved Adaptive Support-weight Sparse Region-based Algorithm for Stereo Matching

Di Lu and Xue Lin

College of Electrical and Electronic Engineering, Harbin University of Science and Technology, Harbin 150080, China
ludizeng@hrbust.edu.cn

Abstract

Stereo matching is to obtain disparity map from two or more images whose pixels existing disparity between each other. Aim at the difficulty of meeting both accuracy and speed when using the adaptive support-weight algorithm, an improved adaptive support-weight sparse region-based algorithm accorded with HVS is proposed. First, improve the traditional support-weight formula, and the support-weight of the points can be calculated according to the improved support-weight formula. Second, the dense disparity map is obtained by sparse region-based matching. Third, left-right consistency check and blocking filling is performed for the obtained disparity map. Finally, median filter is used to remove isolated mismatching points and noise points. Experiment results show that, using the presented method, the matching efficiency is above 90 times faster than that of the adaptive support-weight algorithm proposed by Yoon; and the matching accuracy is 12.34% higher than that of SSD. So the improved algorithm is verified for obtaining accurate disparity map fast and meeting the requirements of system practicability.

Keywords: *stereo matching; improved adaptive support-weight; sparse region-based matching; left-right consistency check*

1. Introduction

Binocular stereo vision is an important issue in computer vision, which is widely used in 3D reconstruction, poses detection and target tracking [1], [2], [3]. As one of the focuses and difficulties of binocular stereo vision, the accuracy of stereo matching has an important effect on practical application. Therefore, it is significant to get high-precision disparity map rapidly for binocular stereo vision [4].

According to the difference of optimization mode, region matching is divided into global region stereo matching algorithm and local region stereo matching algorithm. High-precision disparity map can be obtained by using global algorithm at the cost of more time taken, so it is hard to meet the requirement of practical application [5]. The accuracy of local algorithm is influenced by the size of support window. As the increase of the window, the matching accuracy of textureless area becomes higher because of more information contained in the window, but, at the same time, the matching accuracy in texture area and occluded area becomes lower; on the contrary, the smaller the window is, the higher becomes the matching accuracy of texture area and occluded area, but the matching accuracy of textureless area is lower, because information can't be acquired enough in textureless area when using small window.

The fixed window stereo matching algorithm [6] is an algorithm that establishes a rectangle support window centered on the waiting-for-matched point. Although the computing speed of the algorithm is fast, it can not get high precision disparity map because of less consideration of disparity smoothness. Fusiello [7] uses 9 shiftable windows to proceed stereo matching, instead of limiting the waiting-for-matched points to the center of support window. Compared with fixed window algorithm, its matching

effect gets better, but the calculation amount is multiplied, and whether the neighbor points and the waiting-for-matched point in support window have the same or similar disparity can not be ensured as well. Yoon [8] proposes the adaptive support-weight algorithm, whose nature is to compute the support-weights of neighbor points based on color similarity and geometric proximity, so the neighbor points which have strong similarity with the center point can get larger support-weight value. The algorithm proposed by Yoon has high accuracy, but its matching efficiency is poor.

Aiming at the adaptive support-weight algorithm can not meet the matching accuracy and matching speed simultaneously, an improved adaptive support-weight sparse region stereo matching algorithm is proposed. The improved algorithm contains three steps: 1) use the adaptive support-weight formula which accorded with HVS to compute support-weight values of each neighbor points in support window; 2) after computing support-weights, do sparse region stereo match in order to acquire dense disparity map; 3) make a left and right consistency check on the acquired disparity map and fill the occluded points, then conduct median filter on the disparity map to eliminate isolated noise points and mismatching points.

2. Adaptive Support-weight Algorithm

2.1. Yoon's Adaptive Support-weight Algorithm

In literature [8], Yoon proposed the adaptive support-weight algorithm, whose core concept is to set a rectangle window centered on the waiting-for-matched point, then compute support-weight values of each neighbor point based on color similarity and geometric proximity between neighbor points and the center point in the support window. Thus, a neighbor point can be gotten a larger support-weight value, when it has stronger similarity with the center point in color and geometric. The formula of support-weight is shown below:

$$w(p, q) = \exp\left(-\left(\frac{\Delta c_{pq}}{\gamma_c} + \frac{\Delta g_{pq}}{\gamma_p}\right)\right) \quad (1)$$

where p is the center point, q is the neighbor point; Δc_{pq} represents the color similarity of p and q , Δg_{pq} is the geometric distance between p and q ; γ_c and γ_p respectively represents the parameter of color and distance. Use the formula below to compute the similarity measure of the two input images after the support-weights is calculated:

$$C(p, \bar{p}) = \frac{\sum_{q \in N_p, \bar{q} \in N_{\bar{p}}} w(p, q) w(\bar{p}, \bar{q}) e(q, \bar{q})}{\sum_{q \in N_p, \bar{q} \in N_{\bar{p}}} w(p, q) w(\bar{p}, \bar{q})} \quad (2)$$

p , \bar{p} respectively represents the center point in left image and right image; N_p is the support window centered on p in left image, q is the neighbor point in N_p ; $N_{\bar{p}}$ is the support window centered on \bar{p} in right image, \bar{q} is the neighbor point in $N_{\bar{p}}$, $e(q, \bar{q})$ is the matching cost between q and \bar{q} . Yoon uses the sum of absolute differences(SAD) to compute the initial matching cost. The adaptive support-weight algorithm proposed by Yoon has high matching precision, but its calculation amount is large and it is time consuming, so it can not meet the requirement of practicality. The support-weight feature curve of Yoon's algorithm is shown in Figure 1.

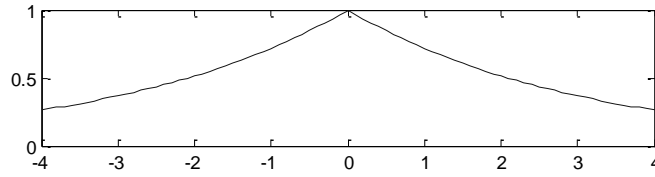


Figure 1. The Feature Curve of Yoon's Algorithm

In Figure 1, 0 is the center point. The support-weight values decrease slowly. Whether the distance between neighbor point and center point is large or small, the down trend of curve is all the same, which is not accorded with HVS. Thus, in order to make support-weight value more accorded with HVS, an improved adaptive support-weight is proposed.

2.2. Improved Adaptive Support-weight

The general stereo matching algorithms assume the points in support window have the same disparity, which can lead to mismatching in depth discontinuous area, literature [9] considers that the disparity obey the Gaussian distribution. HVS [10] shows that the closer the distance from attention center, the slower attenuation the attention value, the farther the distance from attention center, the faster attenuation the attention value. Based on the above two issues, an improved adaptive support-weight is proposed. Use the geometric distance between neighbor point and the center point computing support-weight value to accomplish stereo match. The formula of support-weight W is shown below:

$$W(x+i, y+j) = ae^{-kf^2} \quad (3)$$

W obeys Gaussian distribution. x and y is the coordinate of the center point, $f = \sqrt{i^2 + j^2}$ is the geometric distance between neighbor point and center point, a , k are regulating parameters, a is used to regulate amplitude, k is used to adjust the width of feature curve of W . The feature curve of W is shown in Figure 2.

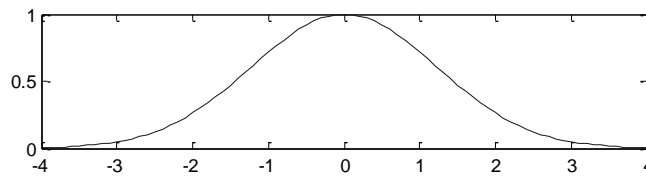


Figure 2. The Feature Curve of Improved Adaptive Support-Weight W

In Figure 2, 0 is the center point, comparing Figure 2 with Figure 1, the feature curve of improved support-weight W is slowly declined when neighbor point is close to the center point, the reference value of this neighbor point is large, the curve of W is fast declined when the distance between neighbor point and the center point is large, the reference of this neighbor point is small, so the feature curve of W is accorded with HVS aforementioned.

3. Sparse Region Stereo Matching

This thesis mainly study on grayscale images stereo match, which takes the truncated absolute differences (TAD) as the matching cost, the TAD formula is shown below:

$$TAD(x, y, d) = \min \{|I_L(x+i, y+j) - I_R(x+i+d, y+j)|, T\} \quad (4)$$

I_L , I_R respectively represents the gray value of left image and right image, d represents disparity, $TAD(x, y, d)$ is the absolute value of gray value difference of corresponding points in the left and right image, T is the truncated threshold. In depth discontinuous area, the gray value difference of corresponding points is large in the left and right image, and it would lead to mismatching if using this result for subsequent calculation directly. Thus, T is used to make up the shortage of using absolute differences (AD).

The matching accuracy can be improved effectively by using the improved adaptive support-weight and left-right consistency check (which will be mentioned below), at the same time, the calculation amount of algorithm will be increased. Therefore, a sparse aggregation method of matching cost is proposed. The improved algorithm proposed in this thesis needs to calculate the disparity map of left image to right image and that of right image to left image, and there are some information redundancies. Thus, select points every other row and column in support window, and compute support-weight values. The points on the cross direction of the center point are more able to characterize the center point. Therefore, after selecting points every other row and column, select all points on the cross direction of the center point, and compute their support-weight. Take a 5×5 window for example, the specific aggregation shape is shown in Figure 3.

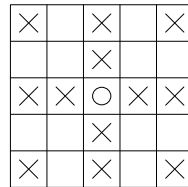


Figure 3. Schematic Diagram of the Sparse Aggregation Method

In Figure 3, ‘o’ represents the center point, ‘×’ represents the neighbor points which needed to compute support-weight and participate in aggregation. Figure 3 shows that the sparse aggregation method for matching cost can reduce calculation amount, especially when the support window is large.

Combined with the improved adaptive support-weight, the matching cost aggregation formula can be acquired. It is shown below:

$$SSWTAD = \sum_{(i,j) \in w_L} W(x+i, y+j) \times TAD(x, y, d) \quad (5)$$

w_L represents the support window of left image. The formula (5) is used as similarity measure function of region matching, and then winner take all (WTA) strategy is selected to get initial disparity map. That is, the point in target image which has the minimum SSWTAD with the waiting-for-matched point in reference image is selected as the corresponding point of the waiting-for-matched point.

4. Disparity Refinement

Occlusion is unavoidable in the process of stereo matching. Occlusion means that some points only appear in one image, so they don’t have corresponding point in the other image, but in the process of stereo matching the algorithm must assign a disparity value to these points, which can result in mismatching and reduce the matching accuracy. So left and right consistency check (LRC) is used to check the occluded points in image, and then these occluded points would be filled. After the above processing, there still could be isolated noise points and mismatching points in disparity map. Median filter is extremely effective on filtering out impulse interference and salt and pepper noise. Therefore, in

order to remove these isolated noise points and mismatching points, median filter is selected to optimize the disparity map.

5. Improved Stereo Matching Algorithm Process

The process of the improved stereo matching algorithm is as below:

- 1) Use the improved algorithm to do the stereo matching of left image to right image, and the left-to-right disparity map D_L can be acquired;
- 2) Proceed to do right image to left image stereo matching, and the right-to-left disparity map D_R can be gotten;
- 3) There is one point P in left image and the disparity value of P is $D_L(P)$. The corresponding point in right image of P is $P - D_L(P)$, the disparity value of $P - D_L(P)$ is $D_R(P - D_L(P))$, and then check whether $|D_L(P) - D_R(P - D_L(P))| < \delta$ or not, if not, P will be marked as the occluded point;
- 4) Find the first non-occluded point on both left and right side of the occluded point, their disparity value is respectively marked as PL and PR , select the minimum of the two values as the occluded point disparity value, namely, $D(P) = \min(D(PL), D(PR))$. The reason why selecting the smallest value is that as for left image, occluded points are usually background, and the depth of background is always bigger than the depth of foreground. As well, the depth is inversely proportional to the disparity, so the minimum disparity value is selected as the disparity of occluded point;
- 5) Use median filter to process the disparity map derived from 4), after that, the final disparity map can be acquired.

The flow diagram of the improved algorithm is shown in Figure 4.

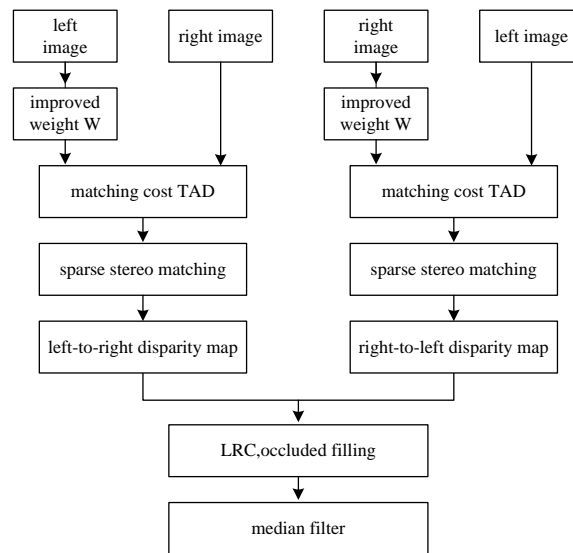


Figure 4. The Improved Algorithm Flow Diagram

6. Experiment and Analysis

In order to analysis the improved algorithm in both part of matching accuracy and computing time, the tsukuba image pairs is selected in the experiment, which is provided by the Middlebury [11] platform of Carnegie Mellon University, all tests are executed on

a computer with AMD A6-3670 2.70GHz CPU and 2G internal storage, the software environment of experiment is Matlab R2010b.

In order to evaluate the improved algorithm's performance, use the four algorithms respectively on the tsukuba image pairs. They are the sum of squared differences (SSD), the adaptive support-weight algorithm proposed by Yoon in literature [8], the method that combines the improved adaptive support-weight with un-sparse region matching (SWTAD for short), and the improved algorithm proposed in this thesis to get the disparity map.

First, the optimal window size of SSD is choiced. The experiment results (in Figure 5) show that the optimal window size is 15×15 , and Average PBM (Percentage of bad matching pixels) is 21.23%, which is the lowest than the others.

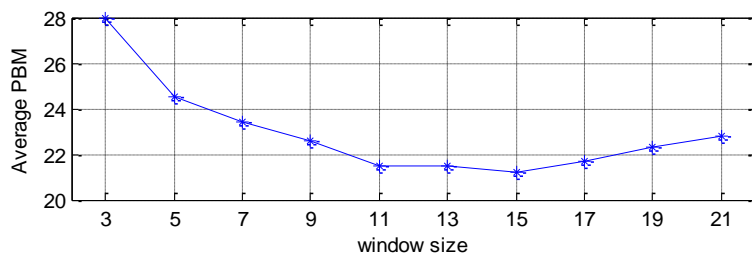


Figure 5. Average PBM and Window Size of SSD

Second, the optimal window size and parameters of the improved algorithm is determined, the results (in Figure 6) show that the optimal window size is 11×11 . Truncated threshold $T=6$, and in the improved support-weight mode $a=0.3, k=1/172$, LRC threshold $\delta=1$. The parameters of the improved algorithm are acquired from several experimental analyses. The results comparisons of different algorithms are shown in Figure 7.

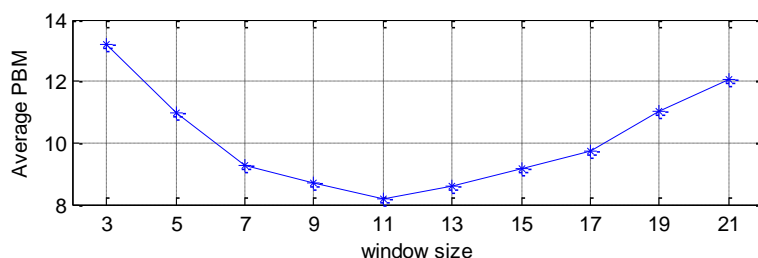
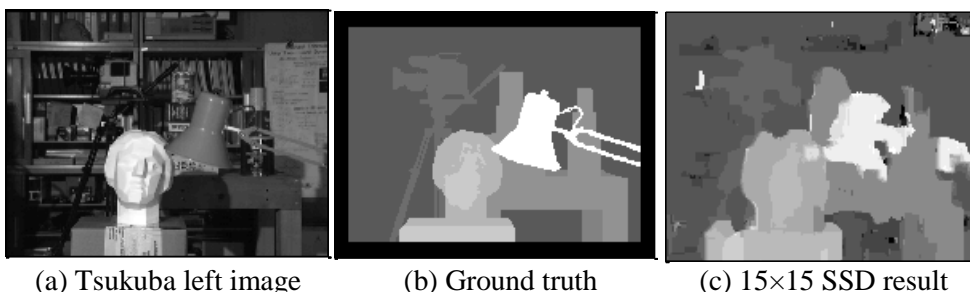


Figure 6. Average PBM and Window Size of the Improved Algorithm



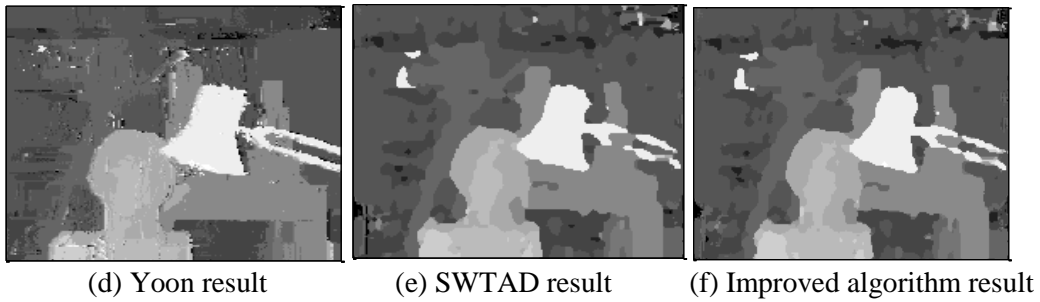


Figure 7. Comparison of Experiment Results

Figure 7(a) is the left tsukuba image, Figure 7(b) is the ground truth of tuskuba image, which is provided by Middlebury platform. The disparity map acquired by 15×15 SSD is shown in Figure 7(c), the outline of Figure 7(c) is indistinct, because the process of matching cost aggregation in a window of SSD is equivalent to a low pass filtering(LPF), the bigger the window , the smaller the cut-off frequency and the more indistinct the outline of object. The outline of Figure 7(d) is the most shipshape and clear. The outline of Figure 7(f) is more obvious than the outline of Figure 7(c), which explains that the improved algorithm can weaken the outline blurring problem result from the window size increasing and it can keep the details of the image better.

The Middlebury test results and the computing time of the four algorithms above are shown in Table 1.

Table 1. Middlebury Test Results and Computing Time

Tsukuba	n-occ(%)	all(%)	disc(%)	Average PBM(%)	Calculating time(s)
SSD	10.1	11.9	41.7	21.23	18.38
Yoon algorithm	1.38	1.85	6.90	3.38	1719.88
SWTAD	5.22	6.15	15.3	8.89	30.71
Improved algorithm	4.94	5.85	13.8	8.20	18.18

(Note: n-occ is the percentage of bad pixels in non-occluded area, all is the percentage of bad pixels in all area, disc is the percentage of bad pixels in depth dissentious area)

It can be seen from Table 1, Figure 7(e) and Figure 7(f) that, though the improved algorithm uses sparse aggregation method, the matching accuracy is slightly higher than the method that does not use sparse aggregation (namely, SWTAD), and the time cost is reduced by nearly half. The matching accuracy is not decreased, but slightly increased; this is because in the square support window not all points have the same or similarity disparity with the center point. The sparse aggregation method can reduce these points taking part in aggregation to some degree, thus the matching accuracy is improved slightly. Therefore, using the sparse aggregation method is significant. The percentage of bad pixels of the improved algorithm is 8.20%, compared with SSD is 21.23%, and in non-occluded area, all area and depth discontinuous area the matching accuracy increases obviously. As it can be seen from Table 1, the time cost of SSD is almost identical to that of the improved algorithm, while the accuracy of the improved algorithm is 12.34% higher than that of SSD, so the improved algorithm is more competitive than SSD. Although Yoon's algorithm has higher matching accuracy, it costs 1719.88s about 29 minutes to accomplish the matching, so it can not meet the requirement of practical application in time cost. High-precision disparity map can be acquired rapidly by the improved algorithm, thus, on several occasions of high matching accuracy and matching efficiency, the improved algorithm has great advantages.

Using the improved algorithm to process other image pairs provided by Middlebury platform can also get good matching results, as shown in Figure 8.

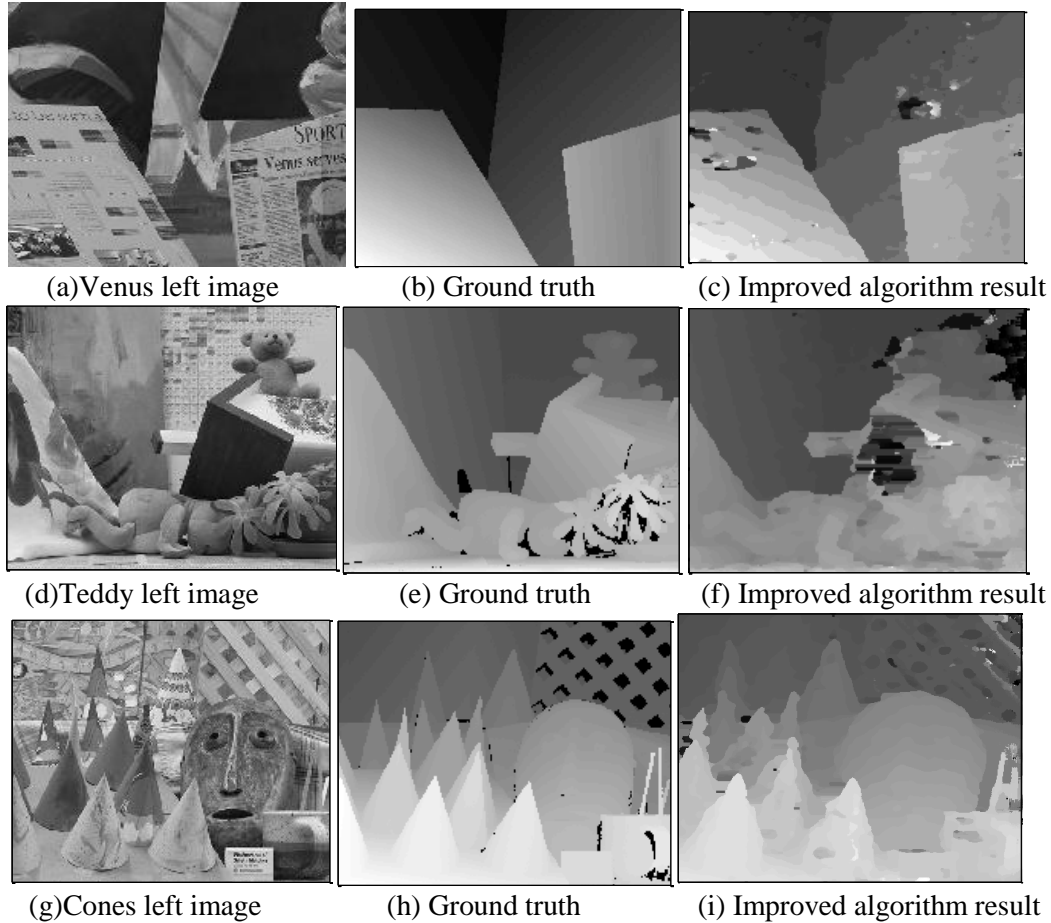


Figure 8. Other Stereo Image Pairs Experiment Results

7. Conclusion

Against to the problem of adaptive support-weight algorithm, an improved adaptive support-weight sparse region stereo matching algorithm accorded with HVS is proposed. First, calculate the improved support-weight value, then carry through the sparse region matching to input images, and the initial disparity map of the images can be acquired. Finally, in order to optimize disparity map, LRC and median filter is done to the initial disparity map. The main conclusions are as follows:

1) The Gaussian function is used as the adaptive support-weight, which can make the support-weight value of each neighbor point accorded with HVS, accordingly better results can be acquired by the improved algorithm;

2) A sparse aggregation method is proposed, which can reduce the information redundancy of LRC and improve the matching efficiency;

3) Compared with Yoon's adaptive support-weight algorithm and SSD, the improved algorithm can improve matching accuracy largely on the basis of matching efficiency requirement. Therefore, the improved algorithm is more suitable for practical application than Yoon's algorithm and SSD.

4) The weakness of improved algorithm is that support window is fixed. In further study, the support window will be adaptive to improve algorithm performance.

References

- [1] W. Xin, Y. Kun, Y. Xiao and Z. Mingzhao, "Design of Binocular Vision 3D Reconstruction System Based on Motion Recovery", *Optics and Precision Engineering*, vol. 22, no.5, (2014).
- [2] Zh. Liang, X. Jinfan and X. Qingyuan, "Pose Estimation Algorithm and Verification Based on Binocular Stereo Vision for Unmanned Aerial Vehicle", *Journal of Harbin Institute of Technology*, vol.46, no.5,(2014).
- [3] Zh. Juan, P. Jianshou, W. Yapeng and L. Jiyan, "Tracking and Measurement of Moving Object in Binocular Stereo Vision", *Computer Engineering and Applications*, vol. 45, no.25, (2009).
- [4] H. Fu and D. Feipeng, "Stereo Matching Using Belief Propagation and Local Edge Construction-based Cost Aggregation", *Journal of Image and Graphics*, vol.16, no.11,(2011).
- [5] D. Scharstein and R. Szeliski, "A Taxonomy and Evaluation of Dense Two-frame Stereo Correspondence Algorithms", *International Journal of Computer vision*, vol.47, no.11, (2002).
- [6] L. Jing and Y. Hong, "Regional Real-time Stereo Matching Algorithm", *Journal of Xi'an Technological University*, vol.32, no.3,(2012).
- [7] A. Fusiello, V. Roberto and E. Trucco, "Efficient Stereo with Multiple Windowing", *Computer Vision and Pattern Recognition Conference*, (1997), pp. 858-863.
- [8] K. J. Yoon and I. S. Kweon, "Adaptive Support-weight Approach for Correspondence Search", *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol.28, no.4, (2006).
- [9] T. Kanade and M. Okutomi, "A Stereo Matching Algorithm with an Adaptive Window: Theory and Experiment", *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol.16, no.9, (1994).
- [10] G. Nur, V.D. Silva, H.K. Arachchi, A. Kondoz, W.Fernando, M.O.Martinez-Rach and S.Dogan, "Sensitivity of the HVS for Binocular Disparity Cue in 3D Displays Under Different Ambient Illumination Conditions", *IEEE International Conference on Consumer Electronics(ICCE)*, (2012),pp.459-460.
- [11] <http://vision.middlebury.edu/stereo>.

