Improved Adaptive Support-weight Algorithm Based on Guide Filtering and Disparity Calibration

XIAO Jing-wen

School of Electrical Engineering and Information
Southwest Petroleum University
Chengdu, China
e-mail: fox7xjw@163.com

WAN Min   WU Yao
School of Mechatronic Engineering
Southwest Petroleum University
Chengdu, China
e-mail:18940103@qq.com

Abstract—A novel local stereo matching method based on guided filtering and disparity calibration was presented in this paper. Firstly, the three dimensional cost volume was built by computing initial matching costs at given disparity values, and it was enhanced with guided filtering to preserve edge information effectively and improve operational efficiency; secondly, a disparity calibration approach was designed to refine the initial disparity, so the matching accuracy can be further improved. Experimental results on Middlebury dataset demonstrate that this proposed technique can obtain accurate and dense disparity map while eliminating effectively matching ambiguity, without affecting the running speed.

Keywords-Binocular vision system; Stereo matching; Adaptive support-weight; Guided filtering; Disparity calibration

I. INTRODUCTION

Binocular vision is an important branch of computer vision, Binocular vision system consists of image acquisition, camera calibration, image correction, stereo matching, depth acquisition and other steps. The stereo matching is the core technology and key steps of the binocular vision system, the accuracy and efficiency of the stereo algorithm directly affects the performance of the whole system.The literature [1] made a summary and classification of stereo matching algorithm. In which the local algorithms because of its simple structure, easy realization, faster, more practical application has been to get everybody's attention. The local algorithms seach the matching points through constructing the rectangular support window in the target image. But its matching accuracy is poor, because the assumption of smooth in depth discontinuities is destroyed, and leads to the 'foreground fattening effect'.

Adaptive support weight stereo matching method(AdaptWeight)[2], which was proposed In 2005, overcomes the edge-fattening to a certain extent by assign an individual weight to each pixel of the support window. A pixel's adaptive support-weight is related to the color and space distance relationship with pixels in the its neighborhood. And it led to the local algorithm in performance has been greatly improved. On this basis, the SegmentSupport[3] proposes a selective weighting function which dependent on color segmentation; GeoSup algorithm[4] defines the weights within a window by computing the geodesic distance to the central pixel. But the computational complexity of those algorithms above depends on the size of the support window, so that it consumes too much time. In order to improve the operation speed, [5] has been quick to realize adaptive weight algorithm using integral histogram to approximate the joint bilateral filter; variableCross algorithm [6] uses a cross-shaped filter; DCBGrid algorithm [7] uses a species approximate bilateral grid method; [8] propose a new "message permeability" filtering, which is a hybrid method based on GeoSup[4] and VariableCross [6]. However, these methods improve the calculation speed at the expense of matching precision.

According to local algorithms need to balance the accuracy and speed of the trend, and inspired by the above studies, proposed an improved stereo matching algorithm, which maintains the advantage in speed of local algorithm at the same time obtaining accurate disparity map.

II. ADAPTIVE WEIGHT ALGORITHMS

The main idea of AdaptWeight[2] is based on correspondence as the center to build a fixed support window, and to assign different weight to each pixel of the window. A pixel's weight inside window is associated with the color difference and the spatial distance between two points. Give high weight to the point which is similar with the center's color and has closer distance, and vice versa given less weight.

The algorithm consists of the following steps: (1) cost computation, (2) the weighted cost aggregation, (3) disparity selection. Using truncated AD to calculate the original matching cost between the pixel point p on the reference image and the matching point $\bar{p}_d$:

$$C\left(p,\bar{p}_d\right)=\min\left\{\sum_{i\in\{r,g,b\}}\left|I_l^i\left(p\right)-I_r^i\left(\bar{p}_d\right)\right|,T\right\} \quad (1)$$

Weighted support window pixel value expression as follow:

$$w\left(p,q\right)=\exp\left(-\left(\frac{\Delta c_{pq}}{\gamma_c}+\frac{\Delta g_{pq}}{\gamma_p}\right)\right) \quad (2)$$

Where Δc_{pq} and Δg_{pq} are respectively the similarity of color and space distance between pixels, γ_c and γ_p is respectively for the empirical parameters of color and distance. After the weighted calculation, the polymerization matching cost

between the pixels p and matching point $\bar{p}_d$ is expressed as follows:

$$C^{'}(p,\bar{p}_d) = \frac{\sum_{q \in N_p, \bar{q}_d \in N_{\bar{p}_d}} w(p,q)w(\bar{p}_d,\bar{q}_d)c(q,\bar{q}_d)}{\sum_{q \in N_p, \bar{q}_d \in N_{\bar{p}_d}} w(p,q)w(\bar{p}_d,\bar{q}_d)} \quad (3)$$

N_p and $N_{\bar{p}_d}$ is respectively support windows of the reference image and the target image, q and $\bar{q}_d$ is respectively two pixels inside window. The weighted cost aggregation steps can be equivalent to using the joint bilateral filter which was based on color and space distance to filter the original matching cost. After the polymerization, disparity of each pixel can be chosen by the simple WTA method:

$$d_p = \arg\min_{d \in S_d} C^{'}(p,\bar{p}_d) \quad (4)$$

$S_d = \{d_{\min}, \cdots, d_{\max}\}$ is a collection of all possible parallax.

Adaptive weight algorithm considers both the support weights reference support window and target window support, by combining two support window to calculate Because the weights of parallax. Which is largely increased the computation complexity. [9] propose the Relational Analysis Method(RAM) and points out that the size of support windows have great influence on the matching performance. And the size of window is larger, the running time is longer.

III. IMPROVED STEREO MATCHING ALGORITHM

Aiming at the disadvantage of AdaptWeight algorithm which computational complexity depends on the window size, the paper proposed an improved algorithm which is based on guided filtering and disparity calibration. The matching process can be divided into two steps: the initial match and disparity calibration. Firstly, calculating the original cost to construct the three-dimension cost volume, using guided filtering[5] filter it. Then, according to the color similarity and spatial proximity assign pixels' adaptive weight, to correct the initial disparity map in order to produce more accurate disparity maps.

A. the initial matching

1) Cost calculation

In this step, construct a three-dimensional cost volume (x,y,d). (x,y) is the coordinates of pixels p on reference image, d represents disparity search range of $[d_{\min}, d_{\max}]$. Cost is the original matching cost between the pixels p on reference image and pixels $(p-d)$ on the target image. This paper uses the truncated absolute difference of color and gradient which is robust to calculate it. In the range of disparity d, the color difference $M(p,d)$ between pixel p and its matching point is calculated as follows:

$$M(p,d) = \sum_{i=1}^{i=3} |I_l^i(p) - I_r^i(p-d)| \quad (5)$$

In the formula, $I^i(p)$ represents the value of the pixel p on the i-th color channels in the RGB space. The absolute difference $G(p,d)$ of the gradient between p and its matching point is expressed as:

$$G(p,d) = |\nabla_x(I_l(p)) - \nabla_x(I_r(p-d))| \quad (6)$$

In the function $\nabla_x(I(p))$ represents p's gradient in the x direction. The function $C(p,d)$ of finial cost is expressed as:

$$C(p,d) = \alpha \cdot \min(T_c, M(p,d)) + (1-\alpha) \cdot \min(T_g, G(p,d)) \quad (7)$$

α is the balance parameters of color and gradient terms, T_c and T_g are truncated value, which is in order to reduce occlusion pixel on matching results.

1) Weighted cost aggregation

In the adaptweight [2],the aggregation matching cost $C^{'}(p,d)$ can be represented as:

$$C^{'}(p,d) = \sum_{q \in W_p} W(p,q) \cdot C(q,d) \quad (8)$$

W_p represents a custom size rectangular window which make the pixels p as center.

The guided filtering [11] which was introduced by CostFilter[10], is based on local linear models, which can guarantee that there is a high correlation between the edge pixels of the output image and the guidance image after filtering. The guided filter has edge keeping property of joint bilateral filter, at the same time, its running time does not depend on the filter kernel size. It is one of the most fast edge preserving filters.

Therefore, in order to improve the computing speed while maintaining accuracy, guided filtering weights was used in the algorithm of this paper. Defined $W()$ as:

$$W_{p,q}(I) = \frac{1}{|\omega|^2} \sum_{k:(p,q) \in \omega_k} (1 + (I_p - \mu_k)^T (\Sigma_k + \varepsilon U)^{-1}(I_q - \mu_k)) \quad (9)$$

I was the reference image, and it also is the guide image of the guided filtering. Σ_k and μ_k are covariance matrix and mean vector in the window ω_k which make k as center and its size is $(2r+1) \times (2r+1)$. $|\omega|$ represents the number of pixels in the window, ε is a smoothness parameter.

After the cost aggregation, using WTA(King Winner-Take-All) strategy to select initial disparity.

$$d_p = \arg\min_{d \in D} C^{'}(p,d) \quad (10)$$

Obtain the initial disparity map.

B. disparity calibration

For local algorithms, the initial matching result exist some amend error, which will reduce their accuracy. Therefore, in this paper, reference AdaptDispCalib[12], add a disparity calibration steps to improve the accuracy of the initial disparity.

Due to the partial matching algorithm contains an implicit assumption, that is ,in a limited area, the point with similar color and short distance of space should have the same depth. So in the disparity calibration steps, it is needed to take p as the center, selecting a calibration window, use the distribution of initial disparity of all pixels to calibrate the initial disparity of point p. Distribute an adaptive support weight $W()$ to each pixel in the calibration window, expression weights as follows:

$$W(p,q) = \exp(-(\Delta c_{pq} / \gamma_i + \Delta g_{pq} / \gamma_g)) \quad (11)$$

Δc_{pq} and Δg_{pq} is respectively the RGB color difference and geometric distance of window pixel q and the center

pixel p, γ_i and γ_g is the corresponding weight coefficient. It is similar to the adaptive weight bilateral filtering algorithm weights, but the color similarity constraint is stronger, so γ_i is small.

After selected the calibration window for the pixel p, to analyze the distribution of the initial disparity window all pixels, that is to count the number of d_i among $d_i \in [d_{min}, d_{max}]$. Then, remember the maximum number of d_i as parallax pixel p.

Specific steps are as follows:

a) Define a three-dimensional zero matrix $C_d(x, y, d)$ which have the same width with reference image, x and y is the image coordinates, $d \in [d_{min}, d_{max}]$.

b) Traversing the initial disparity map. If there are some pixels whose initial disparity is d, mark the corresponding element of C_d as 1.

c) Weighted calculation of C_d.

d) For the pixel $p(x, y)$, select the d of maximum value as the pixel's disparity.

After calibration, disparity can be expressed as:

$$D(x, y) = \arg \max_{d \in [d_{min}, d_{max}]} \left\{ \sum_{i=(w-1)/2}^{(w-1)/2} \sum_{y=(w-1)/2}^{(w-1)/2} W((x, y), (x+i, y+j)) \times C_d(x+i, y+j) \right\} \quad (12)$$

w is the size of the calibration window.

IV. EXPERIMENTAL RESULTS AND ANALYSIS

Using the four standard images of the dataset which was provided by Middlebury website, do the experiment with the same parameter set. The calculation efficiency of CostFilter get from MATLAB2012b code which provided by [9, 10]. The algorithm based on the same post processing steps, and not through code optimization and parallel speedup processing.

Figure 1 is the running time of AdaptWeight and Guided filtering to calculate initial disparity for 'Tsukuba' under different size support windows. As can be seen, the running time of the guided filtering weights does not depend on the size of the window.
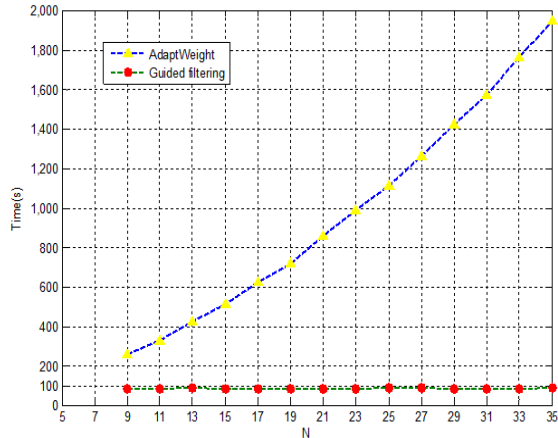

Figure 1.   Comparison of running time

In Figure 2, it visually presents the test result which shows the affection the algorithm on the four groups of image and the comparison of the actual disparity map. The first line is the standard test image, from left to right respectively is Tsukuba, Venus, Teddy, Cones; the second line is the real disparity map provided by Middlebury website; the third and fourth lines are the initial disparity map and the final disparity map; the last line is the error of the error threshold to one.
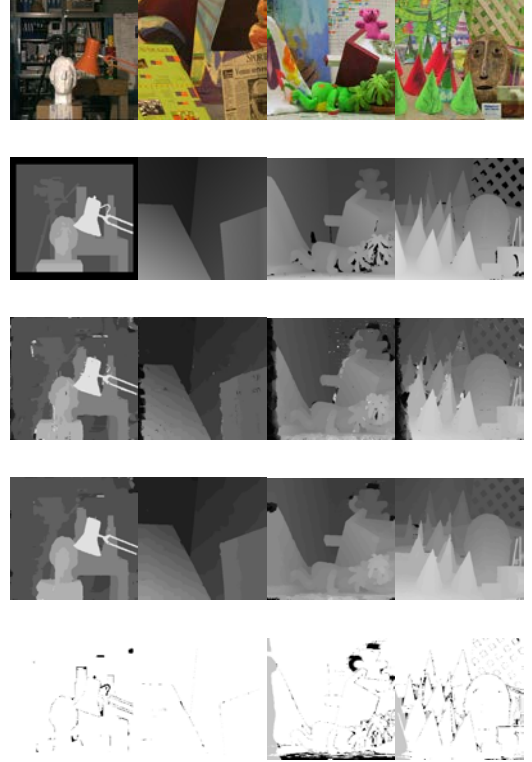

Figure 2.   Disparity maps for Middlebury dataset

Table 1 show the comparison of the initial disparity which obtained after guided filtering and the calibrated disparity. Thus, the disparity calibration steps can effectively remove the error matching points.

TABLE I.   THE COMPARISON OF INITIAL DISPARITY AND CALIBRATED DISPARITY

method	Tsukuba			Venus			Teddy			cones			APBP
	nonocc	all	disc	nonocc	all	disc	nonocc	all	disc	nonocc	all	disc	
initial disparity	2.51	3.29	8.75	2.02	3.18	16.10	8.47	17.00	19.00	3.61	12.60	9.60	8.85
calibrated disparity	2.02	2.53	7.65	1.00	1.70	8.27	7.89	15.70	18.40	3.51	12.40	9.35	7.54

The performance of the proposed method is shown in Table 2 to compare with others. The numbers represent the percentage of bad pixels(PBP) when the error threshold is 1. The 'nonocc', 'all', 'disc' are respectively represent the PBP of all pixels, pixels in non-occluded regions, and pixels near depth discontinuities. As can be seen from the table, the

matching effect of the proposed method for 'Venus' is better than CosFilter[9], and it is the best algorithms of the list.

Table 3 lists the comparison of the computational efficiency between the proposed method and CostFilter for Tsukuba. As can be seen, although increased the disparity calibration step, the computational efficiency has not been seriously affected.

TABLE II. THE MIDDLEBURY SITE ASSESSMENT RESULTS

方法	Tsukuba			Venus			Teddy			cones			APBP
	nonocc	all	disc	nonocc	all	disc	nonocc	all	disc	nonocc	all	disc	
AdaptWeight[2]	1.38	1.85	6.9	0.71	1.19	6.13	7.88	13.3	18.6	3.97	9.79	8.26	6.67
SegmentSupport[3]	1.25	1.62	6.68	0.25	0.64	2.59	8.43	14.2	18.2	3.77	9.87	9.77	6.44
GeoSup[4]	1.45	1.83	7.71	0.14	0.26	1.9	6.88	13.2	16.1	2.94	8.89	8.32	5.8
AdaptDispCalib[6]	1.19	1.42	6.15	0.23	0.34	2.5	7.8	13.6	17.3	3.62	9.33	9.72	6.1
VariableCross[8]	1.99	2.65	6.77	0.62	0.96	3.2	9.75	15.1	18.2	6.28	12.7	12.9	7.6
CostFilter[9]	1.51	1.85	7.61	0.2	0.39	2.42	6.16	11.8	16	2.71	8.24	7.66	5.55
OurS	1.63	1.89	7.29	0.15	0.26	1.91	6.69	12.1	16	2.78	8.22	7.91	5.57

TABLE III. RUNNING TIME

method	AdaptWeight[2]	CostFilter[9]	OurS
Time(s)	258.91	83.76	90.21

V. CONCLUSION

The proposed method puts forward effective improvement. Firstly, against the situation that the computational complexity of AdaptWeight algorithm is dependent on the size of the support window, which leading a longer running time, using the guided filtering to saves the running time while maintaining the edge information at the same time; secondly, after obtaining the initial disparity map, increased the disparity calibration step, according to the color similarity and spatial proximity assign adaptive weights, calibrate the initial disparity map, to further improve the matching accuracy. Experimental results show that the algorithm maintain the matching accuracy, effectively improves the speed, and has practical value.

REFERENCES

[1] SCHARSTEIN D, SZELISKI R. A taxonomy and evaluation of dense two-frame stereo correspondence algorithms [J]. International Journal of Computer Vision, 2002, 47(1/2/3):7-42.

[2] YOON K J, KWEON I S. Locally adaptive support-weight approach for visual correspondence search [J].IEEE Transactions on Pattern Analysis and Machine Intelligence, 2006, 28(4):650–656.

[3] TOMBARI F, MATTOCCIA S, STEFANO L D. Segmentation-based adaptive support for accurate stereo correspondence [C]. // Proc of IEEE Pacific Rim Symposium on Image and Video Technology, 2007 :427-438.

[4] HOSNI A, BLEYER M, GELAUTZ M, RHEMANN C. Local stereo matching using geodesic support weights [C]. // Proc of IEEE International Conference on Image Processing, 2009, 2093-2096.

[5] ZHANG K, LAFRUIT G, LAUWEREINS R, et al. Joint integral histograms and its application in stereo matching[C]. // Proc of IEEE International Conference on Image Processing, 2010, 817–820.

[6] ZHANG K, LU J, LAFRUIT G. Cross-based local stereo matching using orthogonal integral images [J]. Circuits and Systems for Video Technology, 2009, 19(7):1073–1079.

[7] RICHARDT C, ORR D, DAVIES I, et al. Real-time spatiotemporal stereo matching using the dual-cross-bilateral grid [C]. // Proc of European Conference on Computer Vision, 2010, 6313:510–523.

[8] CIGLA C, ALATAN A A. Efficient edge-preserving stereo matching [C]. // Proc of IEEE Conference on Computer Vision Workshops, 2011, 696-699.

[9] Guan T, Zhang G X. Performance enhancement of adaptive support-weight approach by tuning parameters[C]. Proceedings of the IEEE International Conference on Advanced Computational Intelligence, 2012: 206-211.

[10] RHEMANN C, HOSNI A, BLEYER M, et al. Fast cost-volume filtering for visual correspondence and beyond [J].In IEEE Conference on Computer Vision and Pattern Recognition. 2011,3017-3024.

[11] HE K M, SUN J, TANG X O. Guided image filtering [J]. In European Conference on Computer Vision. 2010,1-14.

[12] GU Z, SU X, LIU Y,et al. Local stereo matching with adaptive support-weight, rank transform and disparity calibration [J]. Pattern Recognition Letters, 2008, 29(9):1230-1235

[13] HOSNI A, GELAUTZ M, BLEYER M, Accuracy-efficiency evaluation of adaptive support weight techniques for local stereo matching[C]. // Proc of DAGM-OAGM, 2012, 7476:337–346.