www.s2is.org
International Journal on Smart Sensing and Intelligent Systems
ISSN 1178-5608

NEW STEREO MATCHING METHOD BASED ON IMPROVED BP ALGORITHM

Qian. Zhang, Shaomin Li, Y. Zhang, P. Wang, and JF. Huang

College of Information, Mechanical and Electrical Engineering, ShanghaiNormalUniversity, Shanghai 200234, China

Submitted: Oct. 6, 2014 ***Accepted: Jan. 20, 2015*** ***Published: Mar. 1, 2015***

Abstract—As stereo matching methods are widely used in computer vision and stereo reconstruction, from the perspective of improving the matching accuracy, this paper focuses on the global optimization algorithm. An improved stereo matching method based on 8-neighbor Belief Propagation method is proposed in this paper, by involving more pixels into information transmission, our method improves the accuracy of stereo matching. The experimental results verify the efficiency and reliability of our method.

Index Terms—Disparity map, stereo vision, *Belief Propagation*, MRF Model, Disparity map, Markov Random.

I. INTRODUCTION

With the development in the technology of computer and communication, it's believed that the three-dimension video display will be the next natural and logical development in the video display society[1-2], which provide a more natural visual experience in immersion and three-dimensional sense than the traditional two-dimension display. FVV (Free Viewpoint Video) or FTV (Free Viewpoint Television) and 3D-TV are the main applications of three-dimension video. Other than the three-dimensional sense, another characteristic of FTV is interactivity, which let users can choose the viewpoint freely[3,4] and enable users a more natural visual home entertainment experience. Multi-view video is captured by a set of cameras arranged on different positions of the same scene, which leads to the generation of mass video streams .However, restricted to transmission bandwidth, we cannot transmit all viewpoints video. To efficiently storage and transmission of the data, the common solution is the IBR (Image-Based Rendering) technology[5-9, 34],which generates video of new viewpoint only by some reference views, and can reduce the data needed for storage and transmission.

Stereoscopic vision plays an important role in the areas of DIBR applications, such as 3D stereoscopic television, arbitrary viewpoint video, traffic regulation, surgery and so on. Image processing algorithm of the objective world model is built on the basis of the theory proposed by Marr and Poggio in the 8th century. The most fundamental issue of stereo vision is to build three-dimensional diagram between a pair of matching relations in a group after correction of three-dimensional diagram, find the parallax of each pixel in the target figure in the reference image.

How to get an accurate dense disparity map is the core of image matching algorithm. Most difficult are often occlusions, object boundaries and fine structures, which can appear blurred. Matching is also challenging for structured environment. Additional practical problems originate from recording and illumination differences. Furthermore, fast calculations are often required, either because of real-time applications or because of large images or many images that have to be processed efficiently. Most of these methods concentrated on stereoscopic image pairs [5]-[7], e.g., Zhang proposed the reconstruction for intermediate views from stereoscopic images, using the block-wise maximum-likelihood (ML) disparity estimation method[8-9]. The area-based sum of SSD(sum of squared difference) is calculated over all view in [10], and an algorithm for disparity estimation with energy-based variational regularization using three image views was presented in [11].

On the basis of forming dense disparity map and different matching strategy, the stereo matching is divided into local image matching, global image matching and semi-global image matching. Local matching algorithm is also called the method based on window. Including area based matching algorithm, feature based matching algorithm, phase based matching algorithm,and gradient based matching algorithm [12-16], etc. It usually calculates the size, shape and weight of each pixel to find a suitable support area, and then calculate the weighted mean of disparity values in support area. The ideal support area should cover textureless area and stops at depth discontinuities. It usually adopt local optimization methods to improve the disparity, such as SAD, SSD and other algorithms. As well as global stereo matching algorithm, it's using energy function minimization method to estimate the disparity. However, there's only data item not smooth item in the energy function.

Global matching algorithm adopts the theory of global optimization method to estimate the disparity via building the global energy function [17-18]. Then achieve the minimum the energy function to get the optimal global disparity. This method could lead to better and more accurate results, but its computation complexity is too high and not suitable for real-time operation. Most popular algorithms are graph cuts, belief propagation and dynamic programming algorithm.

Semi-Global matching algorithm is an improved dynamic programming method essentially. It's based on the idea of pixel-wise matching of Mutual Information and approximating a global, 2D smoothness constraint by combining many 1D constraints. (It ensures the smooth constraint by adding different penalty according to different changes of depth. Besides, it ensures the uniqueness constraints by checking the consistency between the reference image and target image.) It has higher efficiency and accuracy compared with the other two methods.

The segment based stereo matching algorithm framework is proposed by Bleyer etc [11, 12] in recent year, and received extensive attention. Figure.1 shows the framework of stereo image matching based on segmentation. First, input two pieces of image, set the left image as the reference image and right one the match image. Both images are shot under the ideal environment which means there's only horizontal disparity. Secondly, proceed the image matching process with the similarity of the window region according to the image color information (grayscale information). That is achieving the initial disparity map adopted the method base on local window. Then segment the reference image. Finally, the initial disparity map will be optimized.

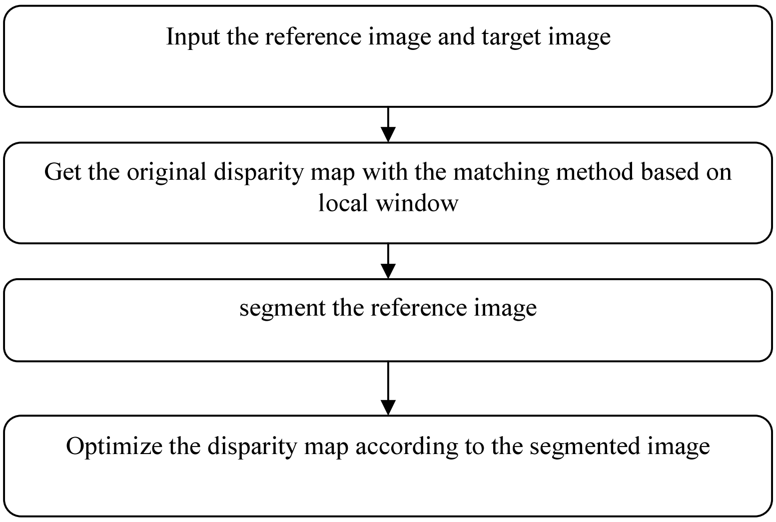

Figure 1. Framework of the segmentation algorithm

As we all known that local methods using winner-takes-all (WTA) method to select the disparities of image, which are usually used in real-time system. However, the limit of the Local methods which could only find the local optimal solution. While the global methods can achieved a excellent matching accuracy by modeling the disparity map surface as a Markov random, and various of optimization methods are used to minimize the cost solutions, such as belief propagation (BP) [22,23]and graph-cut[24-25], dynamic programming [26], non-linear diffusion [27], plane-fitting algorithm is adopted to guarantee matching accuracy [28] and so on.

There are many methods to improve the BP, mainly focus on ameliorate the energy function or decrease computation complexity by using segments instead of pixel to pass message[29].However in real-time stereo system, it is usually difficult to segment object especially in low texture area such as the wall.

Therefore, in this paper we still pass belief messages through pixels and try to improve the way of message passing. This proposed method can achieves good results without using any image segmentation, occlusion handling, or disparity refinement. The experiment results indicate that our method improve the performance.

The rest of this paper is organized as follows: Firstly, we describe the Overview of View Synthesis Scheme, in Section II. Then, MRF Model is proposed in Section III. The 8-neighbor based Belief Propagation method is described in section IV. Finally, experimental results and conclusion are given in Section V and VI, respectively.

II. THE CAMERAS GEOMETRY

Ideally, two cameras located in the same horizontal position, whose optical axises are parallel. Binocular stereo vision in this paper is also based on parallel camera model under the ideal circumstances. The left and right cameras can be imaginary as the eyes, the captured image can be regarded as the picture obtained by the left and right eyes. The depth image could be calculated according to the disparity between the images set.

Showed in Figure 2, Z represents the depth of point *S*; *f* is the focal length of camera; *Xr* represents the distance between the image *P* of *S* which regard the O_r as the optical center and the left optical axis; *Xt* is the distance between the image *P'* of *S* which regards the O_t as the optical center and the right optical axis; *T* is the distance of camera baseline.

According to the concept of disparity, we can know that the disparity formed between the left and right projection planes is the distance between *P* and *P'*. According to the disparity calculation formula (1), under the ideal case where $y_L = y_R$

$$d(p_L, p_R) = \sqrt{(x_L - x_R)^2 - (y_L - y_R)^2} \tag{1}$$

We could have $(p, p') = | x_L - x_R |$, which is the horizontal disparity.

Besides, from the figure (1), $\Delta SSP'$ and $\Delta SO_R O_L$ are similar and can easily get the formula showed as below:

$$\frac{z}{T} = \frac{T - x_R + x_L}{z - f} \tag{2}$$

$$z = \frac{fT}{x_R - x_L} = \frac{fT}{d} \tag{3}$$

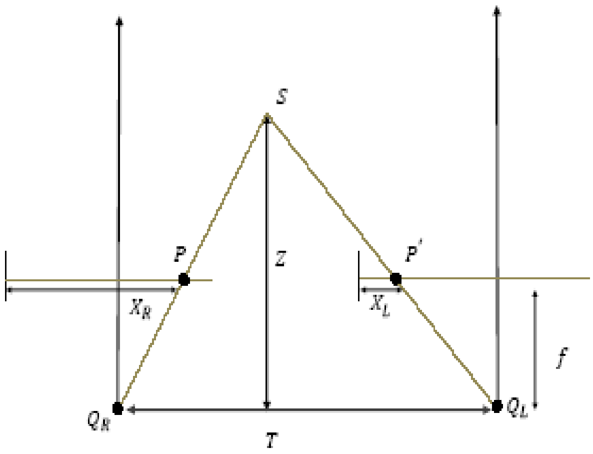

Figure.2 Binocular Disparit

According to the formula (2), we can see there is a relationship between disparity and depth, who can be converted to each other. The smaller the disparity is, the bigger the depth is; vice verse.

III MRF MODEL

In this section, we modelling the disparity map as a Markov Random Field. Denoted, is the variable at the location p,q. F={f_p}, G={g_p} , so we can get the posterior P(X|Y) as:

$$P(F \mid G) \propto \Pi_{p,q} \Psi(f_p, f_q) \Pi_p \Phi(f_p, g_q) \tag{4}$$

Where $\Psi(\cdot)$ expresses the compatibility between neighboring points, $\Psi(\cdot)$ is called compatibility function between f_p, g_q, which relates how compatible the disparity value with the intensity differences in the image. Bigger intensity correspond to lower compatibilities, vice-versa.

Taking the log of equation (4), we can get:

$$E(f) = \sum_{(p,q)\in N} -\log\Psi(f_p, f_q) + \sum_{p\in N} -\log\Psi(f_p, g_q) \tag{5}$$

Boykov[30] express this equation as:

$$E(f) = \sum_{(p,q)\in N} V(f_p, f_q) + \sum_{p\in N} -D_p(f_p, g_q) \tag{6}$$

Where, E(*f*) is the energy function, so maximizing the probability in (1) is equivalent to minimizing the energy Where, E(*f*) is the energy function, so maximizing the probability in (1) is equivalent to minimizing the energy function. While $\Psi(f_p, f_q)$ is corresponding to the function $V(f_p, f_q)$ relates to the smoothness constraint. Among most global methods, $V(f_p, f_q)$ is generally computed by Potts model.

Let $m_{pg}^{t}(f_p, f_q)$ be the message that the node f_p send to node f_q at the time t . They are defined as follows:

$$m_{pg}^{t}(f_p, f_q) = \min_{f_p} \left\{ \begin{array}{l} D_p(f_p, f_q) + V(f_p, f_q) \\ + \sum_{s \in \frac{N(p)}{q}} m_{sp}^{t-1}(f_p, f_q) \end{array} \right\} \tag{7}$$

m_{pq}^{0} is initialized 0. After T_{th} iterations ,we can compute belifes:

$$b_q(f_q) = D_q(f_q, g_q) \sum_{p \in N(q)} m_{pq}^{T}(f_p, f_q) \tag{8}$$

IV 8-NEIGHBOR BASED BELIEF PROPAGATION METHOD EXPERIMENT RESULT

In traditional belief propagation (BP)-based algorithm, local cost aggregations of different pixels are repeated travel from whole image to each pixel using 4-neighbor system[31]. A new 8-neighbor based Belief Propagation method is proposed in this section.

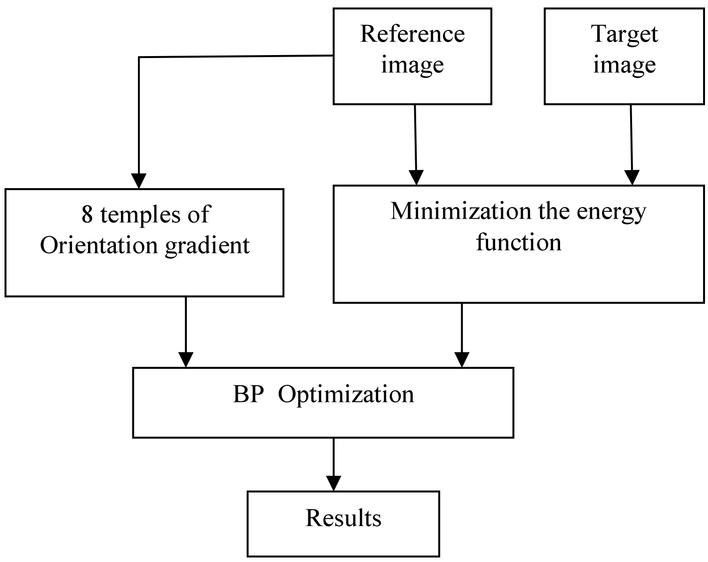

Figure.3 The flow-chart of our algorithm

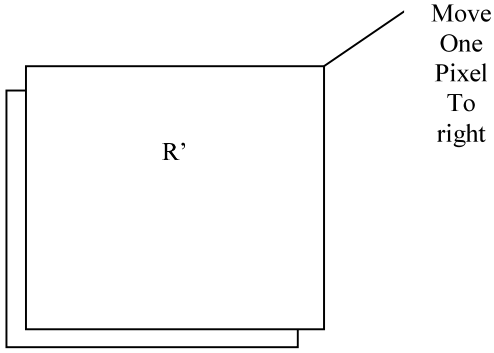

Figure.4. schematic diagram of left gradient

8 temples of orientation gradient including: up gradient, left gradient, down gradient, right gradient, up-left gradient, up-right gradient, down-left gradient and down-right gradient. The schematic diagram of left gradient is depicted in Figure.4.

$$b_n = a_n - a_{n-1} \tag{9}$$

$$Gradient = \begin{cases} 0, b_n < T \\ 1.b_n \geq T \end{cases} \tag{10}$$

If we define:

$$ans(x,y,z) = \begin{cases} \Pr aw_{\cos t}(x,y,z), \Pr aw_{\cos t}(x,y,z) < T' \\ T', \qquad \Pr aw_{\cos t}(x,y,z) \geq T' \end{cases} \tag{11}$$

$$\Pr aw_{message} = e^{-ans(x,y,z)/D} \tag{12}$$

The smoothness terms：

$$smooth_1 = e^{-s/D}, smooth_2 = e^{-P*s/D} \tag{13}$$

T is the Local threshold, which is confirmed by the experience value. The solution of others gradients are similar to the left gradient.

For example, if we pass message as follows:

$$\begin{aligned} &UpDownLeft\Pr aw_{message}(Z) = \\ &Up_{message}(z) * Down_{message}(z) * Right_{message}(z) * \\ &Left_{message}(z) * DownRight_{message}(z) * \\ &UpRight_{message}(z) * \Pr aw_{message} \end{aligned} \tag{14}$$

$$\begin{aligned} temp_{vector\,1}(z) = DownLeft_{message}(z) * \\ UpDownLeft\Pr aw_{message}(z) \end{aligned} \tag{15}$$

$$smooth = \begin{cases} smooth_1, Gradient(x, y) = 1 \\ smooth_2, Gradient(x, y) = 0 \end{cases} \tag{16}$$

$$temp_{vector\,2}(z) = smooth * temp_{vector\,1}(z) \tag{17}$$

$$temp_{\max} = \max(temp_{vector\,2}(z)) \tag{18}$$

$$temp_{vector2}(z) = \max(temp_{vector1}(z), temp_{\max}) \tag{19}$$

After normalizition, then choose the max belief as the final disparity.

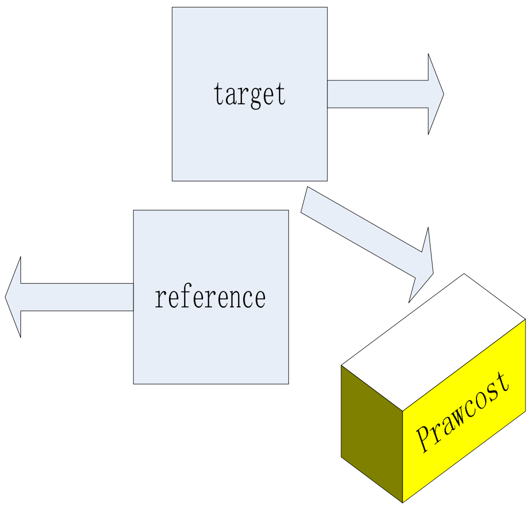

Figure 5. The cost function to obtain information

V EXPERIMENT RESULT

In this section, we evaluate the proposed method and the efficiency is verified by experimental results under Matlab2012 platform. We use four pairs of image sets: "Map", "Tsukuba", "Sawtooth" and"Venus". "Tsukuba" is a indoor environment with frontal surfaces,others consist of mainly slanted planes, and provided by http://vision.middlebury.edu/stereo/ synthetic images , that ignore the noise caused by the factors of camera. This website also provides the standard disparity map which could be used for qualitative comparison evaluation.

We uploaded the results of four image sets, which contains initial disparity map and optimized ones, to the test platform provided by the website http://vision.middlebury.edu/stereo/. We got the data showed in table I after quantitative evaluation. In the table, Nonocc means the occlusion area matching error rate; All means overall matching error rate; Disc means disparity boundary matching error rate.

Figure.6 and Figure.7 show the results using our algorithm mentioned in this paper. The brighter the area is, the greater disparity there is, the smaller the depth it is. We can find that most results are overly smooth because the 8 templets of orientation gradient including more pixels which could result in good quality. Thus, the algorithm in this paper is effective.

TABLE Ⅰ RESULT COMPARISON

	Our disparity map		
Image set	Nonocc	All	Disc
tsukuba	2.27	2.95	10.4
venus	7.84	8.49	24.1
teddy	14.6	21.7	23.7
cones	6.81	14.9	16.8

reference | Result of our method | the standard disparity map

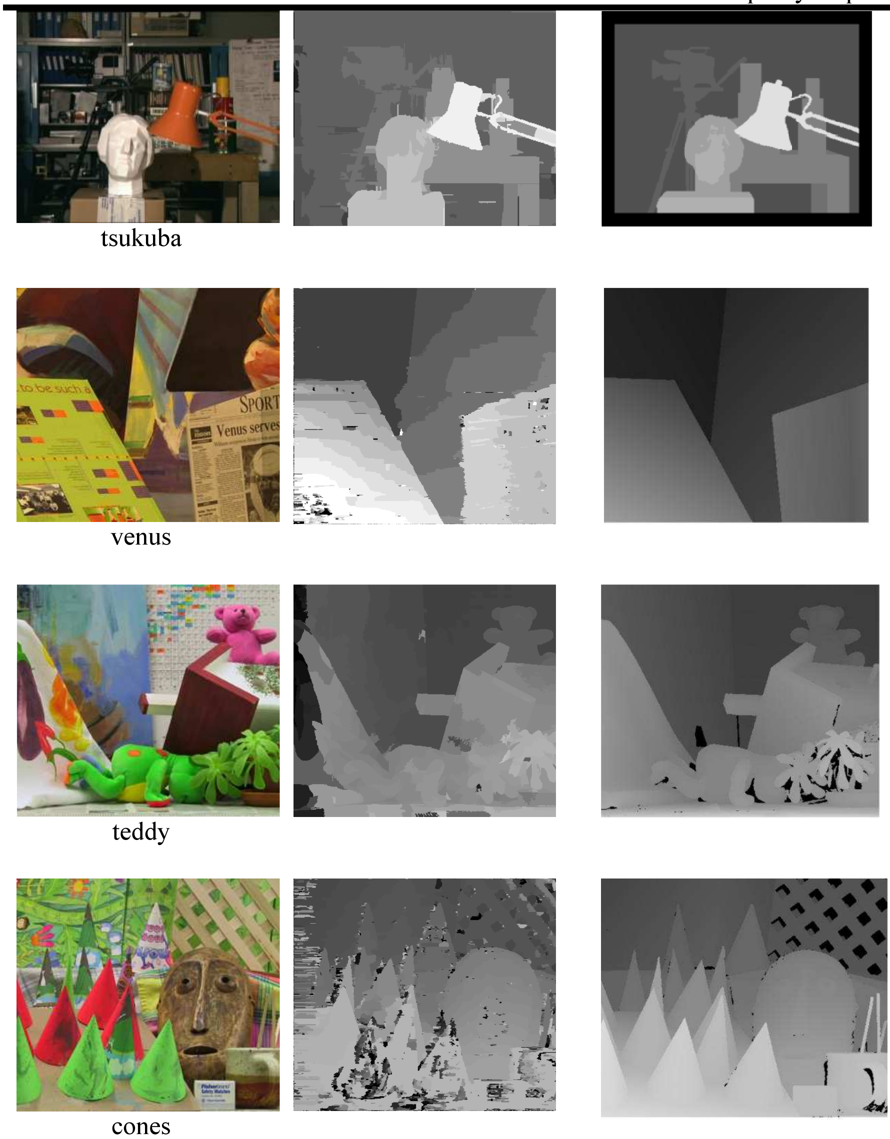

Fig.6 Results using our algorithm

In Tsukuba, the disparity map with our algorithm has better result with depth discontinuity area. The shape of the lamp is clearly visible though the light pole has a little false match.

In Venus, the disparity map used our algorithm regard the areas which are depth discontinuous as continuous area due to the optimization based on segmentation. This leads to a small area of false matching in the lower left corner. In Teddy, the disparity with our algorithm looks similar to the standard disparity map.

In teddy and Cones, Our result has good boundary matching precision, especially the details in bottom right hand. Our result has nice matching within the depth discontinuous areas and also keeps the details of the original image.

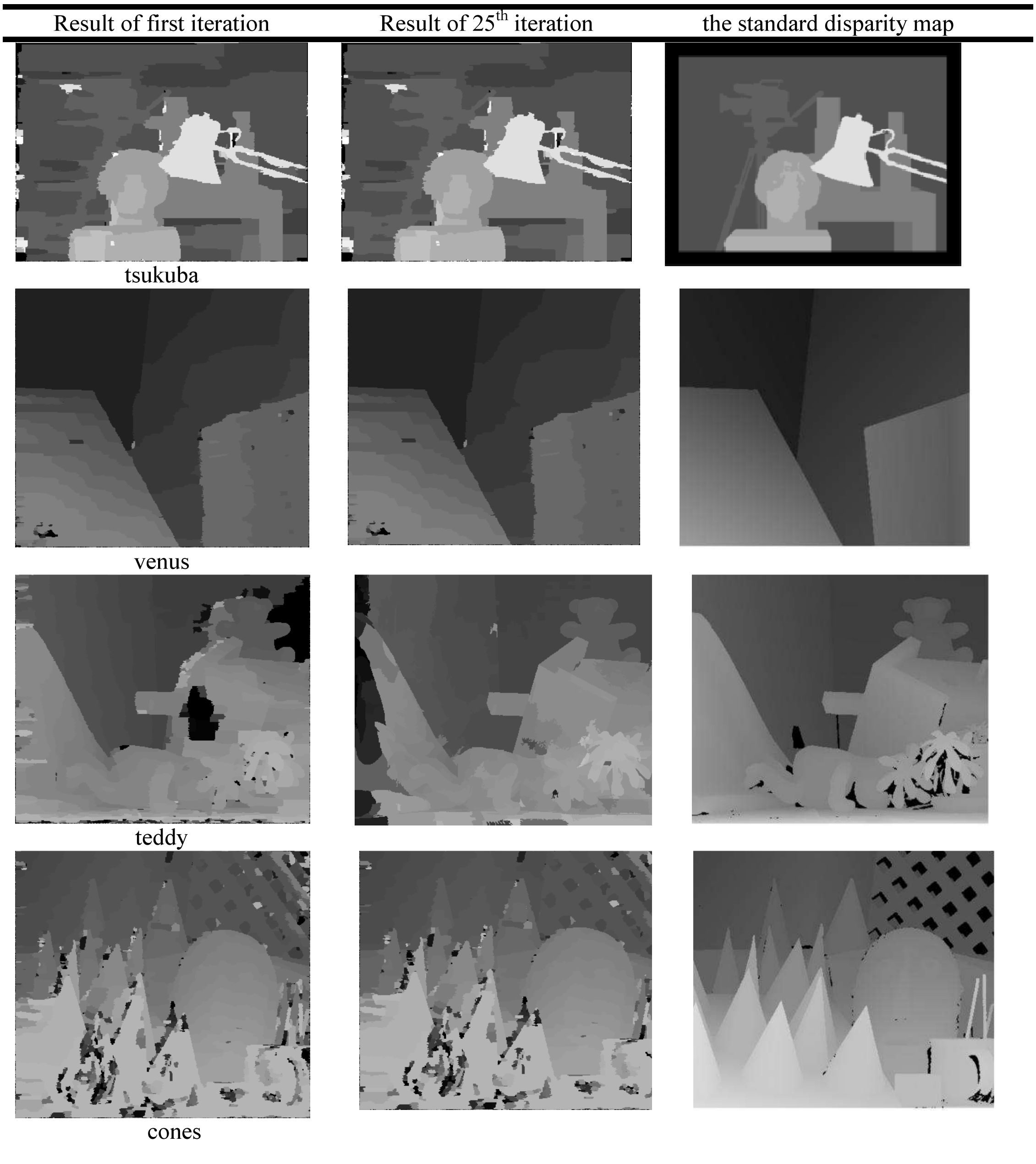

Fig.7 Results using our algorithm

VI CONCLUSION

Stereo matching is the difficulty and key in stereo vision, the research achievements of stereo matching can promote medical, military, aerospace and industry, geography, etc. we studied the binocular stereo matching method in this paper, proposed a new bp method based on 8 templets of orientation gradient and evaluate it by using middlebury image sets. Experimental results show that our proposed method is a good way to improve the performance of global stereo methods. Although this algorithm works well especially with the depth discontinuous and low texture area, it still has some disadvantage, and we need to accelerate the online computation in the future work. Later, the researchers can undertake further research on this issue, also can develop multi-view stereo matching research combined with image fusion. Try to get better dense disparity map with high quality.

ACKNOWLEDGMENT

This research was jointly sponsored by Natural Science Foundation of Shanghai (No. 15ZR1431500), Key Laboratory of Advanced Display and System Applications (Shanghai University), Ministry of Education, China (Project No. P200803), General research project of Shanghai Normal University (Grant No. SK201225), Shanghai municipal education commission of scientific research innovation projects (Grant No. 14ZZ125), Funding scheme for training young teachers in Shanghai Colleges.

REFERENCES

[1] Jianhua Wang ; Qingxiang Zeng ; Fei Xie ; Weiyi Sun.Research on 3D reconstruction of ATV's driving environment based on binocular vision.Electronics, Computer and Applications, 2014 IEEE Workshop,2014 , pp: 556- 559.

[2] Zi-wei Zhou, Ge Li, Ji-zhuang Fang, Jie Zhao, A new stereo matching algorithm based on image segmentation. Information and Automation (ICIA), 2012 International Conference on 6-8 June 2012. pp: 861 – 866

[3] Fan, Xinjian ; Wang, Xuelin ; Xiao, Yongfei.A shape-based stereo matching algorithm for binocular vision Security, Pattern Analysis, and Cybernetics (SPAC), 2014 International Conference ,2014 , pp: 70-74

[4] Nagy, A.-E. ; Szakats, I. ; Marita, T. ; Nedevschi, S.Development of an omnidirectional stereo vision system.Intelligent Computer Communication and Processing (ICCP), 2013 IEEE International Conference ,2013 , Page(s): 235- 242.

[5] Rueihung Li ; Ham, B. ; Changjae Oh ; Kwanghoon Sohn.Disparity search range estimation based on dense stereo matching.Industrial Electronics and Applications (ICIEA), 2013 8th IEEE Conference ,2013 , Page(s): 753- 759.

[6] A. -R. Mansouri and J. Konrad, Bayesian winner-take-all reconstruction of intermediate views from stereoscopic images, IEEE Trans. Image Process., 2012,vol. 9, no. 10, 1710-1722.

[7] Lina Yi, Guifeng Zhang, Zhaocong Wu, A Scale-Synthesis Method for High Spatial Resolution Remote Sensing Image Segmentation, Geoscience and Remote Sensing, IEEE Transactions on Oct. 2012, pp:4062 – 4070.

[8] Sakthivel, P. ; Balakrishnan, G.Measurement of stereo matching on images using dissimilarity estimation.Devices, Circuits and Systems (ICDCS), 2014 2nd International Conference 2014 , Page(s): 1-9.

[9] Shuai Zhang ; Chong Wang ; Chan, S.C.A new high resolution depth map estimation system using stereo vision and depth sensing device.Signal Processing and its Applications (CSPA), 2013 IEEE 9th International Colloquium.2013 , Page(s): 49- 53

[10] Gupta, R.K. ; Siu-Yeung Cho.Window-based approach for fast stereo correspondence Computer Vision, IET ,Volume:7, Issue: 2,2013 , Page(s): 123- 134

[11] X. Huang and E. Dubois, Three-view dense disparity estimation with occlusion detection, in Proc. IEEE Conf Image Process. Vol. III, Genoa, Italy, Sept. 2005, 393-396.

[12] Xing Mei ; Xun Sun ; Weiming Dong ; Haitao Wang ; Xiaopeng Zhang.Segment-Tree Based Cost Aggregation for Stereo Matching.Computer Vision and Pattern Recognition (CVPR), 2013 IEEE Conference ,2013 , Page(s): 313-320

[13] Puerto-Souza, G.A. ; Mariottini, G.-L.A Fast and Accurate Feature-Matching Algorithm for Minimally-Invasive Endoscopic Images Medical Imaging, IEEE Transactions on .Volume:32, Issue: 7,2013 , Page(s): 1201- 1214

[14] Jiaqi Liu ; Qiang Wu ; Xuwen Li.Research on Image Matching Algorithm Based on Local Invariant Features Intelligent Information Hiding and Multimedia Signal Processing, 2013 Ninth International Conference, 2013 , Page(s): 113- 116.

[15] Joglekar, J. ; Gedam, S.S. ; Mohan, B.K.Image Matching Using SIFT Features and Relaxation Labeling Technique—A Constraint Initializing Method for Dense Stereo Matching Geoscience and Remote Sensing, IEEE Transactions on ,Volume:52 , Issue: 9, 2014 , Page(s): 5643- 5652.

[16] Kun Jia.Image matching algorithm based on grayscale and its improvement,Mechatronic Sciences, Electric Engineering and Computer (MEC), Proceedings, 2013, pp: 1203- 1207

[17] Zamir, A.R. ; Shah, M.Image Geo-Localization Based on MultipleNearest Neighbor Feature Matching UsingGeneralized Graphs,Pattern Analysis and Machine Intelligence, IEEE Transactions on ,Volume:36, Issue: 8,2014 , pp: 1546- 1558.

[18] Shuang Yang ; Ling-Yu Duan ; Jie Lin ; Tiejun Huang.A novel pair-wise image matching strategy with compact descriptors,Image Processing (ICIP), 2013 20th IEEE International Conference 2013 , Page(s): 2572- 2576.

[19] Bleyer M, Gelautz M. A Layered Stereo Algorithm Using Image Segmentation and Global Visibility Constraints[J].ISPRS Journal of Photogrammetry and Remote Sensing ,2005,59(3): 128-1 50.

[20] Bleyer M, Gelautz M. Graph-Cut-Based Stereo Matching Using Image Segmentation with Symmetrical Treatment of Occlusions. Signal Processing: Image Communication, 2007, 22(2): 127-143.

[21] D. Scharstein and R. Szeliski. A taxonomy and evaluation of dense two-frame stereo correspondence algorithms, a - SSD + min-filter . IJCV 2002.

[22] Jin-Hyung Kim ; Kwon, J.W. ; Yun Ho Ko.Multi-baseline based texture adaptive belief propagation stereo matching technique for dense depth-map acquisition.Electronics, Information and Communications (ICEIC), 2014 International Conference,2014 , pp: 1- 2

[23] Eslami, H. ; Kasampalis, T. ; Kotsifakou, M.A GPU implementation of tiled belief propagation on Markov Random Fields,Formal Methods and Models for Codesign (MEMOCODE), 2013 , Page(s): 143-146.

[24] Xiaofang Wang ; Huibin Li ; Bichot, C.-E. ; Masnou, S. ; Liming Chen.A graph-cut approach to image segmentation using an affinity graph based on ℓ0-sparse representation of features

[25] ICIP, 2013 20th IEEE International Conference, 2013 , Page(s): 4019- 4023.

[26] Sujung Kim ; Seong Dae Kim.Surface normal vector force driven 3D object reconstruction via Graph-cut,Consumer Electronics (ISCE 2014), The 18th IEEE International Symposium on,2014 , Page(s): 1- 2

[27] Minh Nguyen ; Yuk Hin Chan ; Delmas, P. ; Gimel'farb, G.Symmetric dynamic programming stereo using block matching guidance,Image and Vision Computing New Zealand (IVCNZ), 2013 28th International Conference , 2013 , Page(s): 88- 93.

[28] Maiseli, B. ; Elisha, O. ; Jiangyuan Mei ; Huijun Gao.Edge preservation image enlargement and enhancement method based on the adaptive Perona–Malik non-linear diffusion model.Image Processing, IET ,Volume:8, Issue: 12,2014 , Page(s): 753- 760.

[29] G.Sengupta, T.A.Win, C.Messom, S.Demidenko and S.C.Mukhopadhyay, “Defect analysis of grit-blasted or spray printed surface using vision sensing technique”, Proceedings of Image and Vision Computing NZ, Nov. 26-28, 2003, Palmerston North, pp. 18-23.

[30] Sheng-jun Xu ; Jiu-qiang Han ; Liang Zhao ; Guang-hui Liu.Efficient Belief Propagation for Image Segmentation Based on an Adaptive MRF Model Dependable, Autonomic and Secure Computing (DASC), 2013 IEEE 11th International Conference, 2013 , Page(s): 324- 329

[31] Y. Boykov, O. Veksler, and R. Zabih. Fast approximate energy minimization via graph cuts. IEEE Transactions on Pattern Analysis and Machine Intelligence, November 2001, 23(11):1222–1239.

[32] G. Sen Gupta, S.C. Mukhopadhyay and M Finnie, Wi-Fi Based Control of a Robotic Arm with Remote Vision, Proceedings of 2009 IEEE I2MTC Conference, Singapore, May 5-7, 2009, pp. 557-562.

[33] K. Zhang, J. Lu and L. G, “Scalable Stereo Matching withLocally Adaptive Polygon Approximation,” ICIP, 2008, pp.313-316

[34] Chastine Fatichah, Diana purwitasari, Victor hariadi, Faried effendy, Overlapping White Blood cell segmentation and counting on microscopic blood cell images, vol. 7, no. 3, pp, 1271 – 1286, 2014.

[35] Lu Pengyu, Pu Jingchuan, et al., A Lexicon-Corpus-based Unsupervised chinese word Segmentation Approach, vol. 7, no. 2, pp, 263 – 282, 2014.

[36] Huanbing Gao, Shouyin Lu, Guohui Tian, Jindong Tan, Vision-integrated physiotherapy service robot using cooperating two arms, International Journal on Smart Sensing and Intelligent Systems, vol.7, no.3, pp.1024 – 1043, 2014.