International Journal of Emerging Technology in Computer Science & Electronics (IJETCSE)
ISSN: 0976-1353 Volume 20 Issue 3 – MARCH 2016.

AN ADAPTIVE FEATURE BASED STEREO MATCHING

R.Lavanya[1], S.Nirenjena[2]

[1]*Under graduate , Department of Computer Science andEngineering, IFET college of engineering, lavanyaravi50@gmail.com*
[2]*Assistant Professor, Department of Computer science and Engineering,, IFET College of Engineering, nirenjena25@gmail.com*

***Abstract*-A stereo matching scheme is widely used in computer vision and stereo reconstruction, from the perspective of improving the matching accuracy, this paper focuses on the global optimization algorithm. A novel high-accuracy Adaptive stereo matching scheme based on (based on image ---SURF-features. The approach is based on real time generation of facial disparity map, requiring neither expensive devices nor generic face model. An algorithm based on incorporating topological information of the face in the disparity estimation process is proposed to enhance the result of the 3D reconstruction. Some experimental results are presented to demonstrate the reconstruction accuracy of the proposed method.**

Key words- stereo matching, SURF feature, disparity map, global optimization.

I. INTRODUCTION

Face depth estimation is an important problem that was conjointly studied with face animation, facial analysis and face recognition. In the past few decades, many approaches have been proposed, including 3D from stereo matching (Furukawa and Ponce, 2010),3Dmorphable model based methods and (Choi et al.,2010), structure from motion (Chowdhury and Chellappa, 2003) and shape from shading techniques (Chow and Yuen, 2009). However, how to efficiently acquire facial depth information from stereo images is still a challenging problem especially for real time application.So far, several attempts have been made to deal with 3D face reconstruction from stereo images. (Lengagne et al., 2000) proposed a user interactive approach to deform an animated 3D mesh model from two stereo images. They use a priori knowledge and differential constraints on the 3D model to recover the surfaces of facial areas that are not reliably obtained from stereo alone. (Mallick and Trivedi, 2003) use parallel stereo images and a set of manually selected corresponding feature points to compute the rotation and translation matrix that are used to fit the 3D mesh model to the computed 3D feature points. (Cryer et al., 1995) proposed to merge the dense depth maps obtained separately from stereo and Shape From Shading (SFS) in the frequency domain. The merging process is based on the assumption that shape from stereo is good at recovering high frequency information and shape from shading is good at recovering low frequency information. Recently, many methods use improved SFS techniques to enhance the stereo results (Chow and Yuen, 2009). (Zheng et al., 2007),used a reference 3D face as an intermediate for correspondence estimation. The virtual face images with known correspondences are first synthesized from the reference face. The known correspondences are then extended to the incoming stereo face images, using face alignment and warping. In (Wu et al., 2008), authors do not use any external model, the feature correspondences between images are extracted and the disparity map is initialized. Then an iterative algorithm was used to refine automatically the disparity map using other images taken with different baseline. In this paper, we propose an improved framework for determining the disparity information of a human face from stereo matching in a binocular vision system using correlation based methods. While many stereo matching algorithms have been proposed in recent years (Scharstein and Szeliski, 2002), correlation-based algorithms still have an edge due to speed and less memory requirements (Heo et al., 2011). For the same reasons, we choose to use a correlation based method improved by incorporating the topological information specific to the face obtained by fitting an Active Shape Model (ASM) (Milborrowand Nicolls, 2008) on both images in the initialization step of the algorithm, while maintaining its real-time suitability. Our method demonstrated a satisfactory performance in terms of processing time and point matching accuracy.

II. RELATED WORK:

The notion of Ground Control Points (GCPs) was firstly introduced by Bobick and Intille [16]. GCPs were sparse points that can be matched reliably. The disparity values of the other points are estimated by merging the messages from GCPs. Thus, a semi-dense disparity map could be obtained directly [17]. Wei [18] combined GCPs messages with image segmentation. To deal with the half-occlusion problem, Xu and Jia [19] defined the Outlier Confidence by the probability of pixel occlusion. With the outlier confidence, they defined a penalty term for occlusion and added it to the energy function to be optimized. Good results were achieved in both occluded and non-occluded areas. Sun et al. [20] proposed an algorithm of reliable pixels' message propagation by line segments along 1-D direction. Different from [17], [18], their algorithm only propagated the reliable seed points' messages through the

scanning lines. Single-direction propagation was used to avoid the highly complex and unstable color segmentation. However, owning to the uncertainty of the endpoints of the line segments, line segments could be affected by the stripe defect problem.Wang [21] used MRF as the global optimization model with GCP as inputs.Their energy function consisted of three terms: the energy constraint with GCP, the smoothness constraint between neighboring pixels and the data term measured by absolute difference (AD) between the left view and the right view. The top rank in the Middlebury website of their method proved that GCP was a good expression of regional discrimination. GCP-based algorithms could strengthen the useful information of the scene.

III. PROPOSED METHODOLOGY

In order to estimate the disparity map, we adopta correlation based methods because of their low cost in processing time. However, photo-consistency measures used in these methods are not always suffi-cient to recover precise geometry, particularly in lowtexturedscene regions, in case of occlusion and large disparity. It can therefore be helpful to impose shape priors that bias the reconstruction to have desired characteristics.For this purpose, we use an Active Shape Model (ASM) to obtain prior topological information about the face. These information reduce the search area from the entire epipolar line to only a small segment. In other words, given a right point in the nose region of the face, we search only in the same regionin the left image. This guarantees the smoothness of the disparity map because a point in a topological region (nose, eye,etc.) in the left image will certainly matched with a point in the same region in the right image. As a consequence, disparities values will be continuous and a pixel in eye region will never exceed another in nose region. Fitting the ASM on both images determine the coordinates of the main features points in the right and left image which are subsequently used to compute the shift vectors of the corresponding feature points of the face without applying the classical methods discussed in Section . The shift vectors for nonfeaturepoints are then determined using a correlationbasedmethod using the features point's disparities.

Sparse disparity calculation

In order to establish the sparse matching, we start byapplying an ASM fitting algorithm on both images. The ASM algorithm aims to match a statistical face shape obtained by an offline training process, to a new face image by updating the model parameters to best match to all features points (see Figure 2). In our method, we used the ASM fitting, not only for detection facial features points, but also as an automatic stereo point matching. After fitting the ASM, we obtain the 2D coordinates of n face feature points in the right R=(xi;yi); I 2 [1;n], and the

Figure 1: ASM fitting on left and right stereo images.

left L = (x0i;y0 i); i 2 [1;n] images.Since we use a calibrated system and rectified stereo pairs, the y coordinates of each corresponding points are then normalized to their mean. For each face feature point pi, the Euclidian distance between its right and left coordinates is calculated to obtain its disparity d(pi) as follows:

$$d(p_i) = \sqrt{(x_i - x_i')^2 - (y_i - y_i')^2}. \qquad (3)$$

After the ASM fitting and the disparity calculation,we now have a set n of features points with 3coordinates:P= fpi(x;y;d); i 2 [1;n]g which will used in the dense disparity calculation step.

Dense disparity calculation

In this step, we calculate disparity of noncharacteristicpoints of the face using the obtainedsparserepresentation.The first step consists of projecting the face featurepoints in the 3D space (o;~x;~y; ~d) to obtain a 3D ASM for the face. The 3D ASM is then projected on (o;~x; ~d) and (o;~y; ~d) spaces. This step provides information about the disparity variation on the horizontal and the vertical profiles.Using the characteristic points projected on both 2D plans (o;~x; ~d) and (o;~y; ~d), we define the disparity interval of each point p(x;y) according to its neighbor characteristic points that are the left neighbor pLe f tNeighbor and the right neighbor pRightNeighbor.The disparity interval on ~x axis is defined as [DispMinpx ;DispMaxpx], where DispMinpx is the disparity of pLe f tNeighbor and DispMaxpx is that of pRightNeighbor. In the same way, we define the disparity interval according to y coordinate as [DispMinpy ;DispMaxpy]. Finally, the disparity interval of p(x;y) is given as shown in the following equation

$$[DispMin^p, DispMax^p] = [DispMin_x^p, DispMax_x^p] \cup [DispMin_y^p, DispMax_y^p]. \qquad (4)$$

In the second step, we calculate the disparity ofall non-characteristic points, using their disparity intervals to initialize the algorithm of the correlation, to obtain the dense disparity map.Given a left image point pl, a correlation window w and a

disparity interval [DispMinpl ;DispMaxpl], we aim at obtaining the disparity d 2 [DispMinpl ;DispMaxpl], which maximizes the correlation equation E(d) :

$$E(d) = Similarity(p_l(x,y), p_r(x+d,y)) \qquad (5)$$

For the similarity function, we have used the SAD measure that it is calculated bysubtractingpixel grey level values within an rectangular neighborhood window w between the reference image Il and the target image Irfollowed by the aggregation of absolute differences within the square window.

$$SAD_{I_l(x,y),I_r(x',y'))} = \sum_{u=0}^{m}\sum_{v=0}^{n} |I_l(x+u,y+v) - I_r(x'+u+d,y'+v)|. \qquad (6)$$

IV. RESULTS AND DISCUSSION

In this section, we describe our implementationand our results. We use, in our work, a Bumblebee stereoscopic system composed of two CDD precalibratedcameras mounted on a horizontal support.Inwe compare the disparity map estimated by the SAD method to our method that includes the prior knowledge by applying ASM. Results show that integrating prior knowledge about face can enhance the disparity map in terms of smoothness and also in terms of reducing the missing data named holes (or noise) occurring from uncertain disparities.

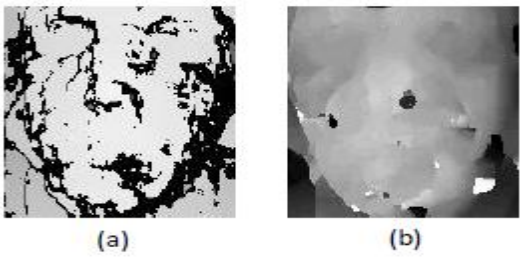

Figure 2: disparity map: (a) SAD Correlation based methods(window 11x11), (b) Our method (SAD+ASM).

In order to obtain the 3D model of the face, Thedepth map is generated using Equation (2) in Section 2 and preprocessed by applying an ellipsoid mask tocrop the face region. In order to fill holes, we applied a selective median filter (with a 7_7 kernel size), which is often used to preprocess depth data. A point cloud for the face is then generated and the texture mapping is performed using the OpenGL library of computer graphics. The results show that the proposed strategy, consisting of incorporating prior knowledge in the disparity estimation process, is robust and accurate. It

Figure 3: Texture mapping and 3D model generation.improves the result of general correlation based methods by considering the face shape and its topological regions, while maintaining its real-time suitability.

V. CONCLUSION

This paper presents an original attempt to a practicalface depth estimation in passive stereoscopic system. Unlike other general methods used for disparity calculation for any object, we introduced a specific method for depth estimation of face that uses the shape characteristics of the human face, obtained by adjusting the form of an active model, to improve result of the correlation-based method. Our method enhanced the classical correlation based method for disparitycalculation, in terms of depth estimation efficiency,with maintaining its real-time suitability. The experimental results show that the proposed algorithm produces a smooth and dense 3D point cloud model of human face, applicable to a wide range of real-time 3D face reconstruction situations.

REFERENCES

[1] Choi, J., Medioni, G., Lin, Y., Silva, L., Regina, O., Pamplona,M., and Faltemier, T. (2010). 3d face reconstruction using a single or multiple views. In Pattern Recognition (ICPR), 2010 20th International Conference on, pages 3959 –3962.

[2] Chow, C. and Yuen, S. (2009). Recovering shape by shading and stereo under lambertian shading model.International journal of computer vision, 85(1):58–100.

[3] Chowdhury, A. K. R. and Chellappa, R. (2003). Face reconstruction from monocular video using uncertainty analysis and a generic model. Computer Vision and Image Understanding, 91:188–213.

[4] Cootes, T., Edwards, G., and Taylor, C. (2001). Active appearance models. Pattern Analysis and Machine Intelligence,IEEE Transactions on, 23(6):681–685.

[5] Cryer, J., Tsai, P., and Shah, M. (1995). Integration of shape from shading and stereo* 1. Pattern recognition,28(7):1033–1043.

[6] Furukawa, Y. and Ponce, J. (2010). Accurate, Dense, and Robust Multiview Stereopsis.Pattern Analysis and Machine Intelligence, IEEE Transactions on, 32(8):1362–1376.

[7] Heo, Y. S., Lee, K. M., and Lee, S. U. (2011). Robust stereo matching using adaptive normalized cross-correlation.IEEE Transactions on Pattern Analysis and Machine Intelligence, 33:807–822. mixed pixels using adaptive over-segmentation," in *Proc. IEEE Conf.*

[8] C. Rhemann, A. Hosni, M. Bleyer, C. Rother, and M. Gelautz, "Fast costvolume filtering for visual correspondence and beyond," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2011, pp. 3017–3024.

[9] M. Bleyer and S. Chambon, "Does color really help in dense stereo matching?" in *Proc. 3DPVT10*, 2010, pp. 1–8.

[10] K. Wegner and O. Stankiewicz,"Similarity measures for depth estimation," in *Proc. 3DTV Conf., True Vis.-Capture, Transmiss., Display 3D Video*, May 2009, pp. 1–4.

[11] M. A. Fischler and R. C. Bolles, "Random sample consensus: A paradigm for model fitting with applications to image analysis and automated cartography," *Commun. ACM*, vol. 24, no. 6, pp. 381–395, 1981.

[12] A. Klaus, M. Sormann, and K. Karner, "Segment-based stereo matching using belief propagation and a self-adapting dissimilarity measure," in *Proc. 18th Int. Conf. Pattern Recognit.*, vol. 3. 2006, pp. 15–18.

[13] Y. Taguchi, B. Wilburn, and C. L. Zitnick, "Stereo reconstruction with *Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2008, pp. 1–8.

[14] M. Bleyer, C. Rother, P. Kohli, D. Scharstein, and S. Sinha, "Object stereo—Joint stereo matching and object segmentation," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2011, pp. 3081–3088.

[15] J. Sun, Y. Li, S. B. Kang, and H.-Y. Shum, "Symmetric stereo matching for occlusion handling," in *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2005, pp. 399–406.

[16] X. Mei, C. Cui, X. Sun, M. Zhou, Q. Wang, and H. Wang, "On building an accurate stereo matching system on graphics hardware," in *Proc. GPUCV*, 2011, 467–474.

[17] Y. Mizukami, K. Okada, A. Nomura, S. Nakanishi, and K. Tadamura,"Sub-pixel disparity search for binocular stereo vision," in *Proc. 1st Int. Conf. Pattern Recognit.*, Nov. 2012, pp. 364–367.
18.M. Bleyer, C. Rhemann, and C. Rother, "PatchMatch stereo-stereo matching with slanted support windows," in *Proc. Brit. Mach. Vis. Conf.*, 2011, pp. 1–11.

[19] R. Zabih and J. Woodfill, "Non-parametric local transforms for computing visual correspondence," in *Proc. Eur. Conf. Comput. Vis. (ECCV)*, 1994, pp. 151–158.

[20] X. Hu and P. Mordohai, "Evaluation of stereo confidence indoors and outdoors," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*,Jun. 2010, pp. 1466–1473.

[21] Q. Yang, R. Yang, J. Davis, and D. Nister, "Spatial-depth super resolution for range images," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2007, pp. 1–8.

[22] Z.-F. Wang and Z.-G. Zheng, "A region based stereo matching algorithm using cooperative optimization," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2008, pp. 1–8.

[23] M. Bleyer and M. Gelautz, "Simple but effective tree structures for dynamic programming-based stereo matching," in *Proc. Int. Conf. Comput. Vis. Theory Appl. VISAPP*, 2008, pp. 415–422.

[24] S. Zhu, L. Zhang, and H. Jin, "A locally linear regression model for boundary preserving regularization in stereo matching," in *Proc. 12th Eur. Conf. Comput. Vis.*, 2012, pp. 101–115.

[25] F. Besse, C. Rother, A. Fitzgibbon, and J. Kautz, "PMBP: Patch Match belief propagation for correspondence field estimation," in *Proc. BMVC*, 2012.

[26] M. Bleyer, S. Chambon, U. Poppe, and M. Gelautz, "Evaluation of different methods for using colour information in global stereo matching approaches," *Remote Sens. Spatial Inf. Sci.*, vol. 37, pp. 415–422, Jul. 2008.