신뢰도 공간과 선형 제어를 통한 스테레오 정합 기법

안효위°, 한영준*, 한헌수*
°숭실대학교 전자공학과
*숭실대학교 정보통신전자공학과
e-mail:{aaxww, young, hahn}@ssu.ac.kr

Stereo Matching Algorithm Based on Line Constraint and Reliability Space

Xiao-Wei An°, Young-Joon Han*, Hern-Soo Hahn*
°Dept. of Electronic Engineering, Soong-sil University
*Dept. of Electronic Engineering, Soong-sil University

● 요 약 ●

A new method is proposed for stereo vision where solution to disparity map is presented in terms of Line constraint and Reliability space -- the first constraint proposes a progressive framework for stereo matching which applies local area pixel-values from corresponding lines in the left and right image pairs. The second states that reliability space based on corresponding points records the disparity and then we are able to apply the median filter in order to reduce the noises which occur in the process. A coarse to fine result is presented after the median filtering, which improves the final result qualitatively. Experiment is evaluated by rectified stereo matching images pairs from Middlebury datasets and has proved that those two adopted strategies yield good matching quantitative results in terms of fast running speed.

키워드: Stereo matching, Corresponding lines, Lines constraint, Reliablity space

I. Introduction

Stereo vision has been one of the most extensively investigated topics in the area of computer vision. Simulating the human vision system in order to estimate 3D position with the modern computer technology is the final goal. As stereo techniques are able to convert 2D images to 3D model, they have been applied to computer graphics and virtual reality or estimation of relative positions of objects in understanding semantic relationships among different environments [7]. Reliable depth map shows the distance of different objects in the reality, which has importance in robotic applications.

Many approaches have been applied for improving the matching quality and efficiency in the past years, as Scharstein and Szeliski presented a taxonomy of dense matching methods which presented almost all the classic correspondence algorithms in their paper[1]. Usually the main applications focus on the two categories---Local (area) methods and Global (feature) methods. Local methods mainly applied the adaptive window which could get the dense stereo results, but commonly the expensive computation.

Obviously, it is trivial to balance the tradeoff between dense results and cheap computation time. Especially if window containing more pixels may make matching results unclear in the local areas. In order to smooth this obscure, some relative-costs must be computed in low-texture areas such as Normalized Cross correlation (NCC), Sum of Squared Difference (SSD) and Sum of Absolute Difference (SAD). Those calculations limit the real time synchronization of the stereo vision system, and some of them produce coarse results without any details. However, methods based on global features are applicable to those issues. The framework based on the graph cuts provided necessary energy minimization techniques in vision [8]. The dynamic programming for stereo like Baker and Binford paper [9] uses Viterbi algorithm with the edge information. Both of them match the standard image pairs with cheap computation. According to those issues, our motivation is to develop an algorithm for fast stereo matching that is able to produce dense maps with edge details and reduce computation time.

II. 3D Stereo Vision System

The vision system for capturing the 3D position of objects was based on stereo geometry. Usually vision system consists of calibration part, matching part and reconstruction part.

1. Calibration

In order to determine the detailed geometric configuration of the two parallel cameras, it is necessary to take a calibration before the matching algorithm in terms of system inner geometric parameters (focal length, image center and lens distortion). Calibration refers to the act of evaluating and adjusting the precision and accuracy of measurement equipment in the image pairs processing. Application of Epipolar geometry (Fig. 1) produces the epipolar line which reduces the complexity of matching correspondence points between images. The calibration between binocular images finally benefits the matching of correspondence points.

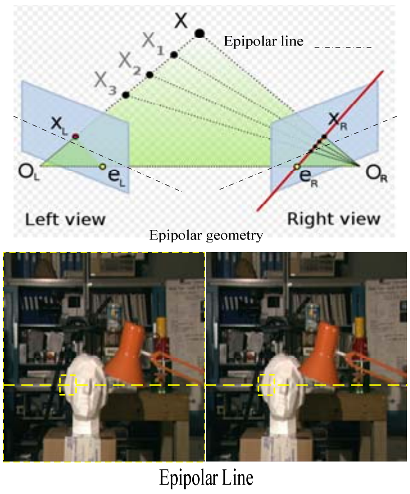


Fig.1. Epipolar geometry.

2. Triangulation

After setting the Epipolar line in the left and right image pairs, relationship between pairs is easily determined under the Triangulation constraint. (Fig.2.)

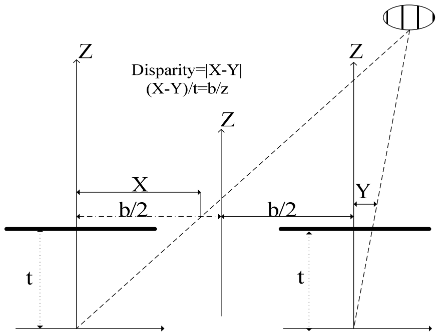


Fig.2. Triangulation.

Figure 2 shows that according to the triangulation, the distance of object (Depth) is inversely proportional to the disparity variance.

3. Disparity Space Image

The most crucial problem in stereo vision system analysis is to estimate the correspondence problem. In another words, to determine which pixels in the left and right images correspond to the one point in 3D world. After obtaining the correspondences, map of disparity was constructed, from which the position of the object was calculated by triangulation.

The set of initial matching costs that are fed into a stereo matcher's optimization stage is often called the disparity space image (DSI) [2]. Here we just apply the horizontal line constraint for correspondences and occlusions simultaneously [3]. Equation (1) shows the mathematical definition of DSI.

$$DSI_{d_n} = | I_L(x_l + k, d_n) - I_R(x_R, d_n) | \quad (1)$$

I_L I_R ---- Left image and Right image

d_n ---- Disparity distance

3.1 Line Constraint

The basic idea of line constraint is that only comparing row information based on epipolar row, it is able to get the low computational cost. We know that the information between the adjacent odd rows and even rows is similar to each other in the standard image pairs. That is because intensities among neighbor pixels existing similar relationship. As shown in the Fig. 3 the adjacent row pixels in gray level. Calculating the odd rows in the matching stage will make the computational time half of that checking all the rows in the both images.

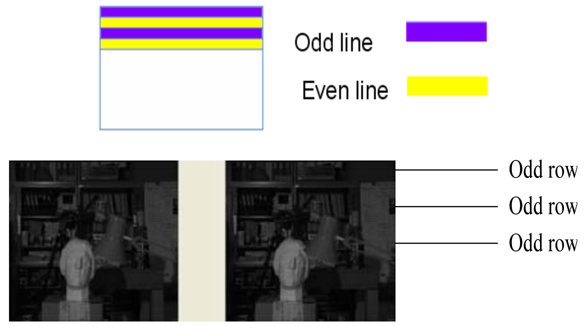


Fig.3. the adjacent row pixels in gray level

3.2 Reliability Space

Reliability space is the data set which records correspondence informations between two images. It is a kind of good evaluation about matching estimation. Now, the original image pairs have been transformed into a new space---including only odd row informations. In this paper, new image pairs under the condition of the fixed step substracted each other in a " pixel to pixel " way. We can apply Eq. (2) for constructing Reliability Space.

$$\begin{cases} R_{reliability} = R_{reliability} + \beta, if(DSI_{d_n}(I_L, I_R) < thr) \\ R_{reliability} = 0, otherwise \end{cases} \qquad (2)$$

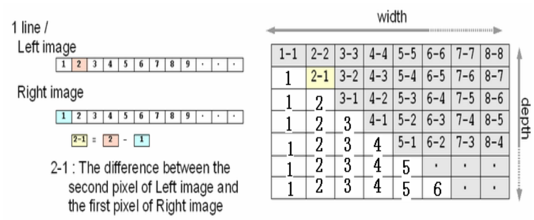
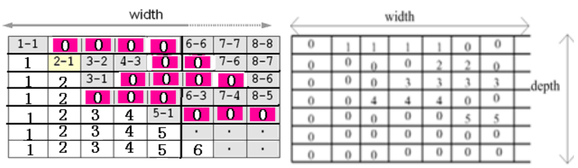


Fig.4. pixels substraction

Here thr is the threshold which evaluates the degree of correspondence. Given the Reliability space, we just need calculate the biggest lengths under the different depths.



Length arrangement pixels arrangement

Fig.5. pixels arrangement

After obtaining the length value arrangement, we set the position where is "0" in the length arrangement gets the value of its depth. And then compute the longest "0" positions in the pixels arrangement space. Here the longest "0" positions include the best informations about the best depth.

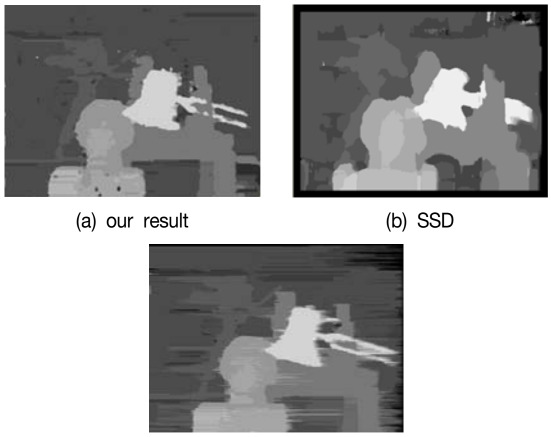
3.3 Median Filter

Experiments show that many noises which are produced by occlusion and other factors such as shadows. In order to reduce the effect of noise, median filter (3*3template) is applied in the final process. Two cases exist in the filtering which can improve the results:

i) In weakly textured region [4,5], the signal to noise ratio is low and often some pixels are rejected although the disparity can correctly be estimated in the neighborhood.

ii) Isolated matches were always false ones.

III. Experiment



(a) our result (b) SSD

(c) Dynamic programming

Fig.6. Result of Algorithms

To verify the performance of the algorithm, experiments on standard stereo image pairs were conducted on PC. For384*288(pixels) image pairs, the disparity was computed in 0.17 seconds. The Tsukuba image is a famous scene from Ohta's lab. We compared our own implementations with several known stereo matching algorithms: WTA with SSD (time cost 3Seconds), DP(2.4 Seconds). We gets the better dense result especially the edge informations and obtains a good time cost in our algorithm.

Ⅳ. Conclusion

As shown in the Fig. 6, a new stereo model was proposed, which demonstrated that improved quality result could be obtained. Future work will be aimed at reducing effects of distortions and noises.

This research was supported by the MKE(The Ministry of Knowledge Economy), Korea, under the ITRC(Information Technology Research Center) support program supervised by the NIPA(National IT Industry Promotion Agency)

References

[1] D. Scharstein and R. Szeliski, “A Taxonomy and Evaluation of Dense Two-Frame Stereo Correspondence Algorithms”, International Journal of Computer Vision, 47(1/2/3):7–42,April–June2002

[2] R. Szeliski and D. Scharstein, “Sampling the Disparity Space Image”, Introduction to Robotics, Publisher, NewYork,2006.

[3] Birchfield and C. Tomasi. “A pixel dissimilarity measure that is insensitive to image sampling”. IEEE Trans. PAMI, 20 (4) : 401–406,1998. 2.

[4] S. Mattoccia. “A locally global approach to stereo correspondence”. 3DIM 2009.

[5] Samuel Morillas and Valent´ın Gregori, “A New Vector Median Filter Based on Fuzzy Metrics”, Image Analysis and Recognition Second International Conference, ICIAR 2005, Canada,September28-30,2005.

[6] Barıýs Baykant ALAGÖZ, “Obtaining Depth Maps From Color Images By Region Based Stereo Matching Algorithms”, OncuBilim Algorithm And Systems Labs. Vol.08, Art.No:04,(2008)

[7] J. Cai, ‘'Fast Stereo Matching: Coarser to Finer with Selective Updating”, Proceedings of Image and Vision Computing New Zealand 2007, pp. 266–270, Hamilton, NewZealand, December 2007.

[8] Yuri Boykov and Vladimir Kolmogorov, “An Experimental Comparison of Min-Cut/Max-Flow Algorithms for Energy Minimization in Vision”, In IEEE Transactions on PAMI, Vol. 26, No. 9, pp. 1124-1137, Sept. 2004

[9] H.H. Baker, T.O. Binford, Depth from edge and intensity based stereo, in Proc. Int. Joint Conf. Artif. Intell. (1981) 631–636.