Feature-based Prevention of Error Propagation in Real-Time Stereo Matching Method

Ji-Hun Mun and Yo-Sung Ho
Gwang-ju Institute of Science and Technology (GIST)
123 Cheomdan-gwagi-ro, Buk-gu, Gwangju 61005, Republic of Korea
{jhm, hoyo}@gist.ac.kr

***Abstract*—Disparity estimation from stereo image is a critical step in 3D video processing. Consecutive image sequences have a similarity with each frame in terms of continuity. Based on the sequential frames, we efficiently restrict a disparity search range when estimating a disparity image by using a previously obtained disparity information. Although the previously estimated disparity values are correct, a small part of error was propagated in successive stereo matching result. In this paper, we propose the prevention of error propagation method in temporal domain stereo matching. Based on extracted feature points, disparity image is periodically refreshed. Since the feature points represent the characteristic of image, feature points based disparity map regeneration improves the correctness of result image. The proposed method shows a better BPR results than generally used stereo matching method.**

Keywords—Real time stereo matching; Error propagation; feature points; Temporal domain

I. INTRODUCTION

Depth information is used for providing a reality of world in 2D coordinate. Although the human can feel the perspective without any supporting devices, it is very complicate to express on 2D space. In order to generate a 3D images, the depth information is combined with 2D image. As a necessity of depth information was boosted, many kind of depth information acquisition methods have been developed. Mainly depth map acquisition methods are divided into two types. We can extract the depth information from a camera which containing a depth sensors. For example, ToF (Time-of flight) and Kinect camera can take a depth information using depth detecting sensors. Acquiring a depth map from sensor is called *active depth sensing method* [1], and the other method is called *passive depth sensing method* [2]. The other depth map generation method uses stereo matching. From the stereo image, we can compute a disparity of each object. The disparity values change into real depth information by using a relation of camera geometric system.

Although the depth information is easily obtained by depth sensors, it has many kind of restriction such as geometrical problem or field of view issue. Estimating a disparity map from stereo image can get a depth information without any physical restriction. However, it need more time to get a result than depth sensing method. Many kinds of research about solving a high time complexity problem have been conducted. Zhang et al. [3] propose a GPU acceleration method to reduce the time complexity. Since GPU can divide a problem and simultaneously computes an each problem using parallel processing, time complexity is impressively decreased. One of the state-of-art disparity estimation method is proposed by Yoon et al. [4]. They use support window which is called adaptive supporting-weight (ADSW) for correctness of disparity value, however it also need more computation time than general stereo matching methods. Chang et al. [5] simplify the ADSW by implementing that system on a hardware. Although the conventional method reduces the time complexity about depth map calculation, it needs other hardware assistance like GPU; CUDA or on board implementation.

Our proposed method basically use temporal domain stereo matching method. A video sequence is composed of many image frames, so that it is not possible to estimate the disparity value in real time. When computing a disparity value in stereo image, search range has an effect to the computation time. In temporal domain stereo matching, estimated disparity map which coming from previous stereo frame was used for current frame stereo matching. If we iteratively perform that procedure, the estimated disparity map result includes an error or incorrect values. In order to prevent the error propagation in sequential disparity results, firstly we extract the feature point by using FAST (Feature from accelerated segment test) method. Based on the extracted feature points, we refresh the disparity map to improve the accuracy of reference disparity map. If we frequently refresh the disparity map, it also cause a computational complexity problem. We refresh the disparity map 1 per 3 frames to improve the reference disparity map accuracy while reduce the time complexity.

Temporal domain stereo matching method with different initial disparity image and that concept will be explained in Section 2. In Section 3, feature based disparity map refresh method will be introduced in terms of time complexity and efficiency of reference disparity value. Using a various sequence images, we testify the performance of proposed method in Section 4 and analyze the simulation results in Section 5.

II. TEMPORAL DOMAIN STEREO MATCHING

A. Iterative temporal stereo matching

The stereo matching method typically divided into two types; *Local* and *Global*. Since global stereo matching method need high computational time complexity, we apply local stereo matching method in sequential image stereo matching. Iterative stereo matching in temporal domain represents that previously obtained disparity map is used in following frames stereo matching procedure. The proposed iterative stereo matching method in temporal domain is represented in Fig. 1. As shown in Fig. 1, the initial disparity map is obtained by using the 'Left(0)' and 'Right(0)' image. The zero frame stereo matching result image which 'Disparity(0)' is used for consecutive frame stereo matching procedure.

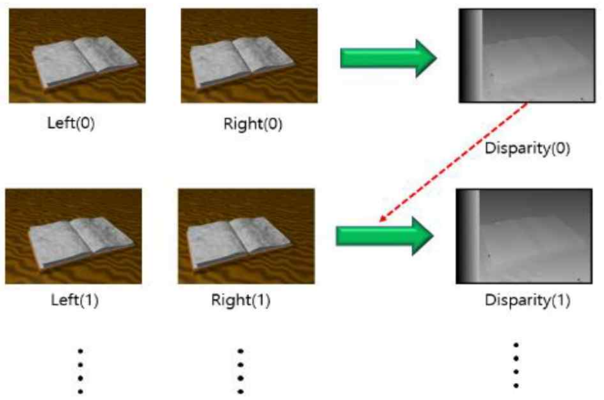

Fig. 1. Iterative stereo matching.

B. *Global based temporal stereo matching*

In temporal stereo matching method, we also use global stereo matching result as an input reference image. The global stereo matching method considers overall image pixels, instead of using a window kernel in matching procedure.

Typically dynamic programming, graph cuts and belief propagation methods are used for optimization function in global stereo matching. Since the global stereo matching method can more accurately estimate the disparity value than the local stereo matching method, the result of global stereo matching image is used for initial reference image. We use improved global stereo matching result image which treats the occlusion area based on discontinuity preserving method [6]. Fig. 2 demonstrate the improved disparity map based iterative stereo matching method in temporal domain.

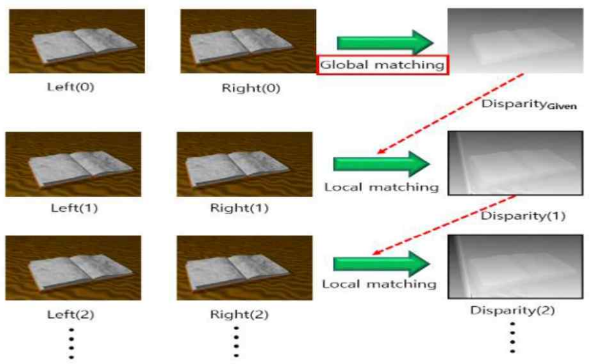

Fig. 2. Global matching based iterative stereo matching.

In Fig. 2, the global stereo matching is performed only one time at first stereo image. The global stereo matching result has an accurate disparity value than local stereo matching result, so that we can preserve the accuracy of disparity value in following image frames. In the same way, previously acquired disparity value restricts the disparity search ranges for computing a cost function in local stereo matching as represented in (1). Where $disp_{pre}$ represents the value which coming from previously obtained disparity map and factor n expand or reduce the disparity search range.

$$\begin{aligned} \mathrm{Min}_{\mathrm{disparity}} &= disp_{pre} - n \\ \mathrm{Max}_{\mathrm{disparity}} &= disp_{pre} + n \end{aligned} \tag{1}$$

Although the global stereo matching result is used as an input image, a small part of error is accumulated while performing the iterative stereo matching. Fig. 3 shows an error propagation effects in consecutive frames. Since we use the global stereo matching method at starting frame, Fig. 3 shows results from frame 1 to frame 15. As we can check in the sequential stereo matching results, the later frame of stereo matching results have an improper pixel value in terms of disparity map accuracy.

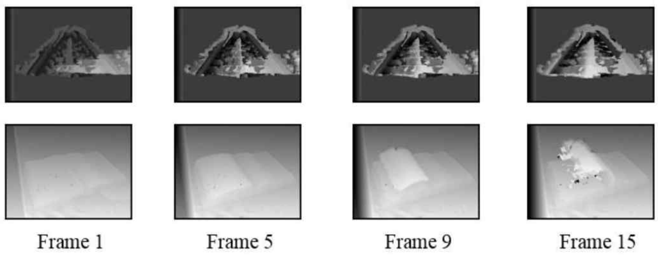

Fig. 3. Error propagation in sequential image.

Although we use global stereo matching result as an initial reference image, the error effect on the disparity result image as progressing the sequential stereo matching. Especially, the edges of the object and corner regions are seriously effected by error propagation.

III. FEATURE BASED DISPARITY MAP RENEWAL

Iterative stereo matching in temporal domain derives an unexpected disparity map as the number of frames are increased. To prevent the error propagation, we refresh the reference disparity map based on extracted feature points. Since the feature point extraction procedure is performed within millisecond, it does not effect to the computational time. However, we need a time for renewal of reference image.

A. *Feature points extraction on object region*

Many kinds of feature extraction methods are developed to detect the characteristic of image. Since the time complexity problem is very critical issue in temporal domain stereo matching, we adopt the FAST feature detection. FAST feature detection method considers neighbor region of current pixel to check the luminance value. Decision tree can help to determine whether feature points or not, so that we can rapidly detect the feature points from an image. Since we need feature points at edge or corner of objects, finding a proper threshold value is critical issue. Fig. 4 exhibits different results with varying threshold values. As we can check in Fig. 4, lower threshold value detects unnecessary feature points.

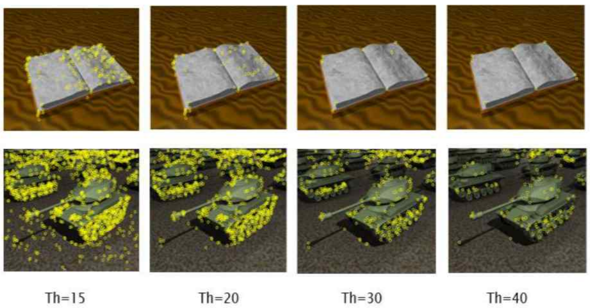

Fig. 4. Feature detection results with different threshold value.

The coordinate of extracted feature points in stereo image are not matched each other, since we find out correctly matched feature points. First, the extracted features in left and right image can match along the horizontal line by using an epipolar constraint. Based on horizontal matching, we compare order of neighbor corresponding feature points. Lastly, we measure the distance of two corresponding feature points as represented in Fig. 5. From initial feature point of epipolar line, we compare all of feature points until the end of the epipolar line. If the distance of two feature point is same, then we determine that these feature points are correctly matched.

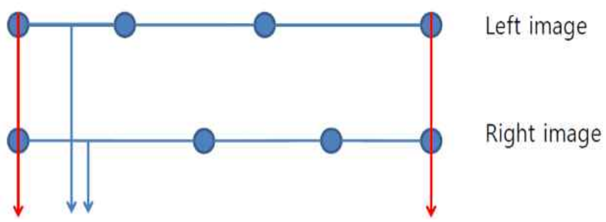

Fig. 5. Feature matching based on distance measuring.

Based on the correctly matched feature points, we periodically refresh the reference disparity image to improve the accuracy of disparity pixel value. When performing the stereo matching, the current pixel match with the extracted feature point, then we consider maximum and minimum disparity search range for cost calculation.

The renewal is performed 1 per 3 frames, since the feature point based renewal need more time to compute a cost value for matching. Fig. 6 demonstrates a feature point based iterative stereo matching procedure. In Fig. 6, when computing the disparity value of *Left(4)* and *Right(4)*, we use extracted feature point. Since at the object edge and corner points has feature points, the quality of disparity result image *Disparity(4)* improves than generally used temporal domain stereo matching result.

If we reduce the frequency of reference image renewal ratio, the overall time consuming is decreased and vice versa. Since we consider the disparity search range of extracted feature point in stereo matching procedure, non-feature points area still provide an inaccurate disparity search range. As we can see in Fig. 4, the lower threshold value can take more feature points than bigger one. However the purpose of temporal domain stereo matching method is reducing a time complexity, a higher threshold value is used for feature detection.

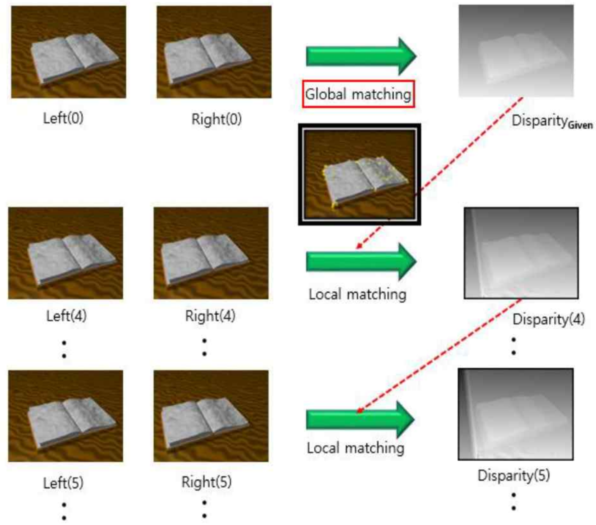

Fig. 6. Feature based temporal stereo matching.

IV. Experiment Results

The proposed algorithm has been implemented on a 3.0GHz Xeon i7 PC platform. Typically, the disparity is computed within 5 second for 20 consecutive image frames and window size is fixed to 5x5. In this section, we present the experiment results of proposed algorithm by using sequential test image.

Temporal domain stereo matching method is optimized to sequential images, so that we use Cambridge computer graphics test sequences. Totally 4 test sequences which *Book*, *Tank*, *Street* and *Temple* are used to testify our proposed method. The resolution of each test sequence is 400x300 and contain a ground truth image for each sequence. We adopt the FAST feature extraction method and use fixed threshold value 30.

Ratio of reference disparity map renewal effect on computational time and result of disparity image quality, so that we perform the test by changing the frequency of disparity map renewal. Fig. 7 exhibits the test results about feature based temporal domain stereo matching.

In Fig. 7, each row represents the renewal frequency for stereo matching and below number indicates the number of frames. As we can see in Fig. 7, when we frequently refresh the disparity value, the disparity accuracy was increased in following frames. The meaning of frequency 3 indicates that we extract the feature point and searching whole disparity range for cost calculation per 3 frame.

Especially in sequence *Temple*, it has very improper disparity value on the result image. Since that sequence has similar texture region on the background and object boundary area, when performing the cost calculation, improper pixel values are used for comparison. Although we frequently renew the reference disparity map, it still includes a small part of error pixel value. Note that the feature based stereo matching method partially improves the reference disparity value not whole disparity values.

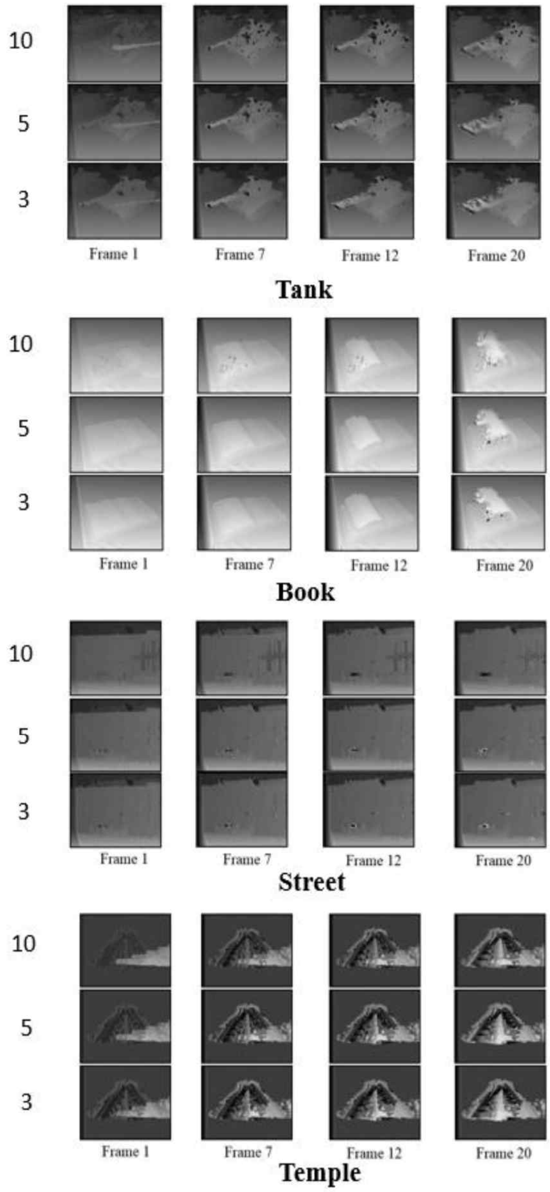

Fig. 7. Experiment results of proposed method.

The Cambridge test sequences are provided with ground truth images. To verify performance of our proposed method, we compare the result images with ground truth images. Usually the disparity images are compared by using a BPR (Bad pixel rates) value. If the differences between result image pixel and ground image pixel is bigger than 1, that pixel is determined as a bad pixel. Table 1 reports the detailed BPR value. That results demonstrate the difference of BPR with respect to disparity map renewal ratio. As demonstrated in Table 1, the time complexity was increased when we frequently refresh the reference disparity map.

TABLE I. BPR AND COMPLEXITY COMPARISON RESULTS

Method	Sequence	BPR(%)			Time complexity(sec)		
		1 per 10	*1 per 5*	*1 per 10*	*1 per 10*	*1 per 5*	*1 per 3*
General	Tank	11.3			4.34		
	Book	10.4			4.08		
	Street	13.2			4.34		
	Temple	10.9			3.87		
Feature based	Tank	11.1	10.7	9.8	3.42	3.68	4.42
	Book	10.0	9.8	9.5	3.12	3.17	3.27
	Street	12.8	11.6	11.4	4.09	4.14	4.21
	Temple	10.1	9.4	8.9	4.21	4.33	4.41

V. CONCLUSION

We have demonstrate a new algorithm for temporal domain stereo matching based on feature points. Using the FAST feature detection, we find out feature points and renew the reference disparity map. Since the feature based stereo matching has high time complexity than general method, we predetermine the ratio of disparity map renewal frequency. From the experiment results, we can check that the proposed method shows a better BRP performance about 1.43% than generally used matching method. However, the time complexity was increased about 0.37sec. Even though we use proposed method, still a small part of error was propagated to disparity result of consecutive frames. In the future, we plan to automatically refresh the disparity map with conditional feature detection method.

ACKNOWLEDGMENT

This research was supported by the 'Cross-Ministry Giga KOREA Project' of the Ministry of Science, ICT and Future Planning, Republic of Korea(ROK). [GK15C0100, Development of Interactive and Realistic Massive Giga-Content Technology]

REFERENCES

[1] "3dv systems, z-cam," http://www.3dvsystems.com/home/index.html

[2] D. Poussart and D. Laurendeau, "3-D Sensing for industrial computer vision, in advances in machine vision," chapter 3, 1989, pp. 122-159.

[3] K. Zhang, J. Lu, Q. Yang, G. Lafruit, R. Lauwereins and L. V. Gool, "Real-time and accurate stereo: A scalable approach with bitwise fast voting on CUDA," *IEEE Transaction on CSVT*, vol. 21, no. 7, pp. 867-878, July 2011.

[4] K. Yoon and I. Kweon, "Adaptive support-weight approach for correspondence search," *IEEE Transaction on Pattern Recognition and Maching Intelligence*, vol. 2, no. 4, pp. 650-656, Feb. 2006.

[5] N. Chang, T. Tsai, P. Hsu, Y. Chen and T. Chang, "Algorithm and archietecture of disparity estimation with mini-cencus adaptive support weight," *IEEE Transaction on Circuit and Systems for Video Technology*, vol. 20, no. 6, pp. 792-805, June 2010.

[6] W. Jang, Y. Ho, "Discontinuity preserving disparity estimation with occlusion handling," *Journal of Visual Communication and Image Representation*, vol. 25, no. 7, pp. 1595-1603, Aug. 2014.