SIFT-based Stereo Matching to Compensate Occluded Regions and Remove False Matching for 3D Reconstruction

Do Kyung Shin, Jeong Ho Lee and Young Shik Moon

Department of Computer Science and Engineeringek
Hanyang University
Ansan, Korea
E-mail: {dkshin,jhlee,ysmoon}@cse.hanyang.ac.kr

ABSTRACT

Generally, algorithms for generating disparity maps can be clssified into two categories: region-based method and feature-based method. The main focus of this research is to generate a disparity map with an accuracy depth information for 3-dimensional reconstructing . Basically, the region-based method and the feature-based method are simultaneously included in the proposed algorithm, so that the existing problems including false matching and occlusion can be effectively solved. As a region-based method, regions of false matching are extracted by the proposed MMAD(Modified Mean of Absolute Differences) algorithm which is a modification of the existing MAD(Mean of Absolute Differences) algorithm. As a feature-based method, the proposed method eliminates false matching errors by calculating the vector with SIFT and compensates the occluded regions by using a pair of adjacent SIFT matching points, so that the errors are reduced and the disparity map becomes more accurate.

Keywords: SIFT, stereo matching, 3D reconstruction

1. INTRODUCTION

Recently, many researchers are actively studying on the reconstruction of a 3-dimensional image using acquired 2-dimensional images from a stereo camera. The artificial 3-dimensional reconstruction for real-life environment is weakness that model creation is not realistic. But the virtual environment experience plays an important role in feeling a virtual reality.

In this paper, we propose an efficient method to generate a correct disparity map for a 3-dimensional reconstruction. Generally, algorithms[1] for generating disparity maps can be classified into two categories: region-based method[2] and feature-based method[3]. The proposed method uses a region based similarity measurement as a basis matching technique and improves correctness using characteristic method based on the extracted matching points from SIFT[4]. In order to automatically establish scan area for matching between two disparity maps, we extract a pair of matching points with the minimum and maximum value from SIFT. We can reduce the amount of calculation about useless by calculating the specified scan area and can create a sparse disparity map. As a matching method, we use the proposed MMAD algorithm which is a modification of the existing MAD algorithm and the SIFT feature points robust to changes of image scale and rotation. The proposed method reduces errors about depth information through eliminating false matching errors by calculating the vector with SIFT and compensates the occluded regions by using pair of adjacent SIFT matching points.

2. EXTRACTION OF MATCHING PAIR USING SIFT

In order to increase the accuracy of the matching points between stereo images, we extract feature points based on SIFT algorithm which robust to changes of the image scaling and rotating. In this paper, the matching calculation consist of four steps: i) extraction of the matching points by SIFT matching between stereo image, ii) acquisition of the minimum (X_{min}, Y_{min}) and maximum (X_{max}, Y_{max})coordinates from i), iii) decision of the scan range within candidate region acquired from two coordinates, iv) scan of the matching points based on a left image for the scan range from iii). Since a pair of the extracted matching points doesn't always locate in the most outer point, we define α as an error range in Eq. (1) The existing method may calculate the matching point for all the pixels, while the proposed method calculates the partial scan by automatically deciding the scan area.

$$\begin{aligned} Top &= \min(matchingPo\operatorname{int}Y) - \alpha \\ Bottom &= \max(matchingPo\operatorname{int}Y) + \alpha \\ Left &= \min(matchingPo\operatorname{int}X) - \alpha \\ Right &= \max(matchingPo\operatorname{int}X) + \alpha \end{aligned} \quad (1)$$

3. PROPOSED ALGORITHM

In order to generate a reliable disparity map, the proposed method which removes the false matching and compensates the occlusion region has procedures as followings.

Step1. Calculate the matching using MMAD which is the minimum cost function.

Step2. Separate the background and object from an image by checking edges.

Step3. Eliminate the false matching by calculating the vector using SIFT.

Step4. Compensate the occlusion regions by using SIFT and edges.

3.1 Matching Calculation of MMAD

The methods of the matching calculation are classified into region-based method and feature-based method. In this paper, we propose the algorithm considering both methods. In order to improve the matching accuracy, the proposed matching method uses region-based method for removing errors and utilizes feature-based method for compensating false area.
Region-based similarity decision function is classified into minimum and maximum cost function. The kinds of minimum cost function are SSD(Sum of Squared Differences)[6,7], SAD(Sum of Absolute Differences) and MAD[6,7], and the ones of maximum cost function are NSSD(Normalized Sum of Squared Differences), NC(Normalized Correlation), NCC(Normalized Cross Correlation), and ZNCC(Zero-mean Normalization Cross Correlation). Maximum cost function has a shortcoming that the processing speed is slow, though has high reliability. Therefore we propose MMAD in Eq. (3) that is modified from MAD in Eq. (2).

$$C_{MAD}(x,y,d) = \frac{\sum_{n=y-L}^{y+L}\sum_{m=x-K}^{x+K}|I_R(x+m+d,y+n)-I_L(x+m,y+n)|}{K\times L} \quad (2)$$

$$C_{MMAD} = \frac{\sum_{n=y-L}^{y+L}\sum_{m=x-K}^{x+K}|(I_R(x+m+d,y+n)-\mu_r)-(I_L(x+m,y+n)-\mu_l)|}{K\times L} \quad (3)$$

- I_R : right image
- I_L : left image
- d : disparity range
- $K \times L$: size of window

- μ_l : bright mean of left windows
- μ_r : bright mean of right windows

MMAD basically uses the existing MAD method, but it is applied to the difference between brightness averages (μ_l, μ_r) for the scan range of left and right image. Because MMAD is minimum cost function, a pixel with the lowest weight can be selected as a pair of matching points by using Eq.(4)

$$d_{MMAD}(x,y) = \arg\min C_{MMAD}(x,y,d) \quad (4)$$

Because MMAD is based on the correlationship normalized as independent zero-mean for the difference of brightness and the change of contrast, it has a shortcoming that the processing speed is slow. MMAD takes the effect similar to maximum cost function and reduces errors by compensating the difference of lightness for stereo image due to the method applied the normalized correlationship to MAD as minimum cost function. Besides, the proposed method has faster performance speed than existing maximum cost function.

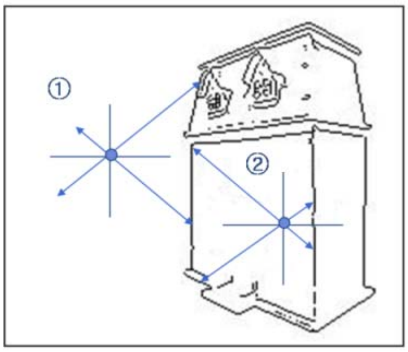

Fig. 1 The condition to classify an image into background and object region

$$\text{distance} = \sqrt{(Ex2-Ex1)^2} - \sqrt{(Ey2-Ey1)^{\ 2}} \quad (5)$$

$$\text{threshold} = \begin{vmatrix} \max(\text{siftLeftMin}, \text{siftRightMin}) \\ -\min(\text{siftLeftMax}, \text{siftRightMax}) \end{vmatrix} \quad (6)$$

$$\text{thresholdPoint}(x,y) = \begin{cases} \text{true} & \text{if threshold} \le \text{distance} \\ \text{flase} & \text{otherwhise} \end{cases} \quad (7)$$

3.2 Classification of Object and Background

Before compensating occlusion region by false matching, the proposed method examines whether the edges of four directions for all the pixels is existed. If at least one edge isn't existed as shown in the case ① of Fig. 1, then our method classifies this pixel into background area, else the pixel is classified into object area due to existing at least one edge as shown in the case ② of Fig. 1.
That is, the proposed method automatically classifies an image into background and object area for matching calculation. This method has an advantage that can reduce the calculation quantity due performing to the matching calculation only for object area except background area.

3.3 Error Compensation using SIFT

3.3.1 Remove of the False Matching

The proposed MMAD is region-based method and only considers brightness within a window. Accordingly, if there are no corresponding points, the reliability for the matching calculation gets lower because the result of matching is false. In order to remove errors for false matching, we propose the method to decrease the error rate by performing the vector calculation based on SIFT matching pairs and comparing brightness. Among SIFT matching pairs, we extract the matching pairs with minimum and maximum value for x-Axis on a right image in which the scan area is automatically set. The matching calculation has procedures as following: i) scan of all pixels from left-top to right-bottom within the scan area, ii) comparison of a pixel (x,y) on a right image with from a pixel (x,y) to a pixel (x+'disparity range') on a left image, iii) calculation of the vector for the extracted SIFT matching pairs with maximum and minimum x-axis value as shown in Fig. 3.

Fig. 4 shows procedures for the vector calculation. Firstly we move the vectors on a left image and a right image into origin on the same coordinate system. As show in Eq. (5), if Euclidean Distance between two vectors is less than threshold value as in Eq. (6), then we assume that these pixels may be true matching point as in Eq. (7).
Because the matching algorithm evaluates the similarity between a pixel (xL,y) on a left image and a pixel on a right image by shifting a pixel from a pixel (xR,y) to a pixel (xR+′disparity range',y), the value of threshold is |xR-xL|. The case of Fig. 2 and Fig. 3 are that each SIFT distance of left and right image is same and threshold is 3. In Fig. 2 the matching pairs are classified into true because threshold is less than 3. In Fig. 3 the matching pairs are classified into false because threshold is greater than 3. The cases of Fig. 4 and Fig. 5 are that each SIFT distance of left right image is different and threshold is 3. In Fig. 4 the matching pairs are classified into true because threshold is less than 3. In Fig. 5 the matching pairs are classified into false because threshold is greater than 3. Because the matching pairs with same SIFT implies that the rotation degree of the object is small, we suppose that fewer occlusion regions may be generated. Though threshold of two cases is all 3, in the case that each SIFT distance is same, the matching is classified into true until the disparity range is 2. In addition, in the case that each SIFT distance is different, the matching is classified into true only if the disparity range is 1. The latter case implies that the error range of the vector calculation is less than the former case.
Euclidean distance is calculated by SIFT vector, and then the weight value of MMAD is sorted by ascending order within disparity range.

Because pixels with minimum weight value are the most reliable matching pairs, we store the disparity between these pixels into final disparity map. However, the above method still has the problem for occlusion region because the corresponding points can't be found in the invisible area by distortion of image.

3.3.2 Compensation of Occlusion Region

Existing stereo matching still has unsolved problems for searching accurate corresponding points and compensating occlusion region generated by the absence of the corresponding points. In this paper, we propose the method to solve problems for occlusion region by using SIFT matching points.
SIFT-based matching algorithm is strong for rotating and resizing the image, and can improve reliability for searching a pair of the matching points. That is, this algorithm select a pair of the matching points with the closest distance among SIFT matching points existing in the quadrant of the 2D coordinate system which is made from a pixel point of occlusion region as the origin. If the quadrant without a pair of SIFT matching points exists, then we examine whether the edge in the quadrant exists. If the edge exists, then the proposed method classifies the region into foreground area and compensates the disparity with the average disparity of the SIFT matching points existing in each of the quadrant. 'S' in Fig. 6 indicates a pair of SIFT matching points and 'E' indicates the edge.

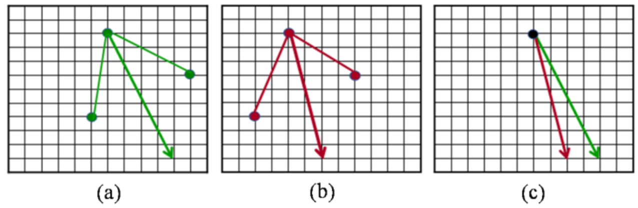

Fig. 2 When SIFT distance is equal – in case of true matching: (a)vector on left image (b)vector on right image (c)distance between two vectors

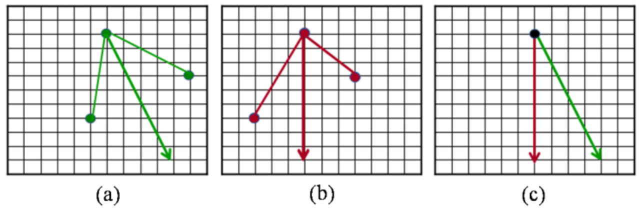

Fig. 3 When SIFT distance is equal – in case of false matching: (a)vector on left image (b)vector on right image (c)distance between two vectors

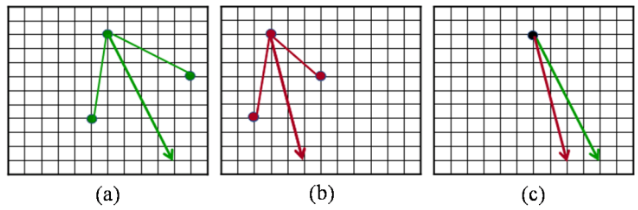

Fig. 4 When SIFT distance is not equal – in case of true matching: (a)vector on left image (b)vector on right image (c)distance between two vectors

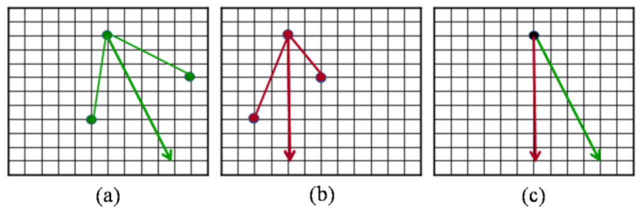

Fig. 5 When SIFT distance is not equal – in case of false matching: (a)vector on left image (b)vector on right image (c)distance between two vectors

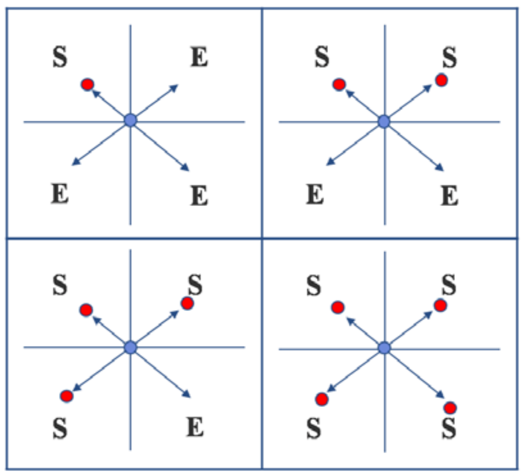

Fig. 6 Examination whether the edge exists in the quadrant

Tabel 1. The accuacy comparison of the proposed method with the existing methods

	SIFT	Minimum cost function			Maximum cost function		The proposed method
Matching algorithm	**SMNA**	**SSD**	**SAD**	**MAD**	**ZNCC**	**NC**	**MMAD**
The average of correct matching number(CMNA)	84.2	56.95	56.50	56.10	58.15	50.15	62.5
The mean matching ratio		67.63%	67.10%	66.63%	69.06%	59.56%	74.23%

Fig. 6 shows four examples of the compensation for occlusion area, but all the compensations are 15 cases.

Though Fig. 6 shows four examples of the compensation for occlusion area, all of the compensation has fifteen cases.

Condition 1 : If each of 1, 2, 3, and 4 quadrants doesn't have at least one edge and SIFT matching points, then we classify the pixel into background area.

Condition 2 : If 1, 2, 3, and 4 quadrants have at least one edge and SIFT matching pairs, then we compensate the disparity by using the searched SIFT matching pairs.

4. EXPERIMENTAL RESULTS

In the experiment, we used stereo images of a building model with 320×240 sizes and windows with 21×21 sizes to measure the area-based similarity. We used 20 stereo images as experimental data and Fig. 7 indicates original image for the result.

4.1 Results of Experiment

Fig. 8-(a) depicts the result with applying MAD after setting automatic scan area and Fig. 8-(b) depicts the result with applying MMAD. MMAD compensates the occlusion region by normalizing the difference between left and right image. However, as shown in Fig. 8-(b), false matching can be generated by absence of corresponding points.

Fig. 8-(c) indicates the result of generating another occlusion area by not matching the area of deciding on false matching after performing the vector calculation by using SIFT. Fig. 8-(d) indicates the result after compensating occlusion region by eliminating errors for false matching. The proposed method compensates the occlusion region using the similarity of adjacent pixels and the accuracy of SIFT matching pairs. Fig. 8-(e) indicates the result of removing noise and compensating the occlusion region for errors. Fig. 8-(f) shows the result that the image is classified into background and object area by reducing error range with an additional consideration for the existence possibility of the edge.

4.2 Evaluation of Performance

We need the reference disparity map to compare the accuracy of the disparity map measured from the proposed method. However, the accuracy measure is difficult because the reference disparity map isn't defined. In this paper, we use the SIFT matching pairs to measure the accuracy of the proposed method. Because the SIFT algorithm which is robust to the changes of the image scaling and rotating has high trust about the matching pair, we use the disparity extracted by the SIFT matching pair as a reference disparity map. Accordingly, in the disparity maps of the existing and the proposed method, we compare errors between two disparities which are extracted from the same location of SIFT matching pair (x,y). We call that the SIFT disparity is SD and the disparity of matching algorithm is MD. And, if the absolute value of the difference between SD and MD such as the condition in Eq. (8) is less than 10, we assume that the calculated disparity at the pixel is accurate. We call that the number of the SIFT matching pairs is SMN and the number of the matching pairs assumed to be correct is CMN. The accuracy rate is calculated as CMN ratio about SMN such as Eq. (9). Because the experiment in Table 1 uses ten test images, ACMN and ASMN as the average of each CMN and SMN are used as shown in Eq. (10). The accuracy rate of the proposed method is compared with one of the existing methods such as SSD, SAD, MAD, ZNCC and NC. In terms of the accuracy rate, the proposed method has been improved from 30% to 40% over the existing methods.

In addition, Fig. 9 indicates the results of the 3D reconstruction by using the generated disparity map. In Fig. 9, (a) ~ (e) indicate the results of the 3D reconstruction by using the existing methods and (f) indicates the result of the proposed method. In the existing methods, it can be confirmed that the result of reconstruction about edges is different from the structure of the original object because the disparity at edges is incorrect. On the other hand, we can know that the errors of the proposed method is less than the existing methods as showing that edges are distinguished.

$$|\mathrm{SD} - \mathrm{MD}| < 10 \tag{8}$$

$$\frac{\mathrm{AMN}}{\mathrm{SMN}} \times 100 \tag{9}$$

$$\begin{aligned} \mathrm{ASMN} &= \frac{\text{Add all SMN of the test image}}{\text{Total number of the test image}} \\ \mathrm{ACMN} &= \frac{\text{Add all CMN of the test image}}{\text{Total number of the test image}} \end{aligned} \tag{10}$$

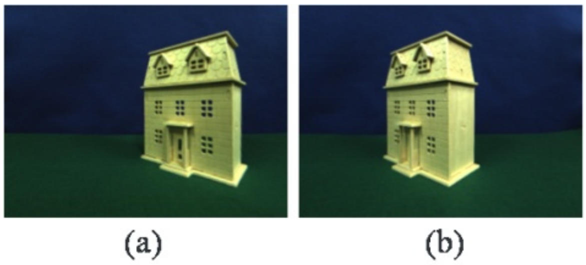

Fig. 7 Original stereo image used in the experiment (a)left image (b)right image

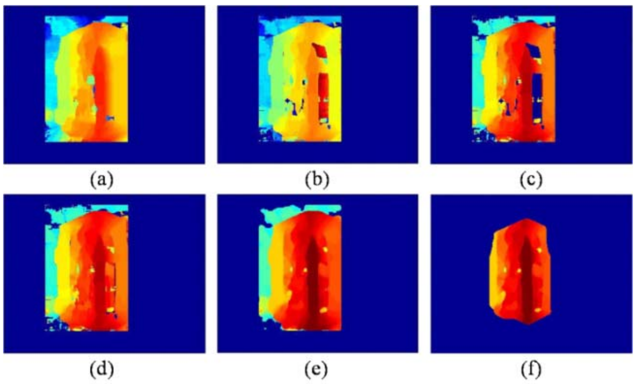

Fig. 8 The result of proposed disparity map (a)MAD (b)MMAD (c)Compensation by vector (d)Compensation of occlusion region (e)Exclusion of noise (f)Noise elimination

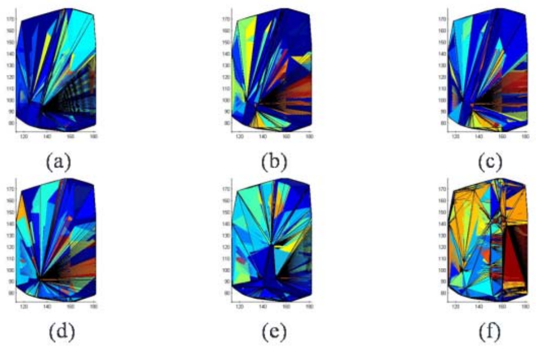

Fig. 9 The result of 3D reconstruction by using the proposed disparity map (a)SSD (b)SAD (c)MAD (d)ZNCC (e)NC (f)Proposed method

Acknowledgement

This research was supported by the MIC (Ministry of Information and Communication), Korea, under the ITRC (Information Technology Research Center) support program supervised by the IITA (Institute of Information Technology Assessment)

5. REFERENCES

[1] D.Scharstein and R.Szeliski, “High-accuracy stereo depth maps using structured light”, in IEEE Computer Society Conference on Compuer Vision and PatternRecognition(CVPR 2003), volume 1, pp.195-202, Madison, WI, june. 2003

[2] A. Fusiello, V. Roberto, E. Trucco, “Experiments with a new area-based stereo algorithm,” in International Conference on Image Analysis and Processing, 1997.

[3] Guo-Qing Wei, Gerd Hirzinger, “Intensity and feature based stereo matching by disparity parameterization”, ICCV, vol. 6, jan. 1998.

[4] D.G.Lowe, “Distinctive image features from fcale-invariant keypoints,” Intl. J. Computer Vision, vol. 60, no. 2 , pp. 91-110 , 2004.

[5] D. Y. Jung, S. Y. Park, J. H. Oh, C. S. Kang, K. G. Nam, “Stereo matching by dynamic programming involving occlusion and local bias,” in J. of Electrical Engineering and Information Science, vol. 3, no. 6, pp. 770-778, Dec. 1998.

[6] D.Sankoff and L. Haque, “Power boosts for cluster tests” Proc. RECOMB int’l Workshop Comparative Genomics(RCG’05), pp.121-130, Sept.2005.

[7] G.Blin, C.Chauve, G.Fertin, R.Rizzi, and S.Vialette, “Comparing genomes with duplications : A computational complexity point of view”, IEEE/ACM Transsactions on Computational Biology and Bioinformatics, vol.4, no.4, Dec. 2007