that you use a text box t

Estimation of a Relative Camera Orientation with Few Correspondences using Unsynchronized Viewpoints

Jieun Kim
Graduate School of Information
Yonsei University
Seoul, Republic of Korea
lilly9928@yonsei.ac.kr

Deokwoo Lee
Dept. of Computer Engineering
Keimyung University
Daegu, Republic of Korea
dwoolee@kmu.ac.kr

***Abstract*— Recovery of 3-dimensional(3D) coordinates information from 2-dimensional(2D) image has long been of interest in computer vision. Geometric relationship between 2D and 3D coordinates is one of the most key to accurate 3D reconstruction result. Prior to performing 3D reconstruction, internal and external parameters of cameras need to be estimated and the task is called camera calibration. In this paper, we propose the approach to camera calibration with few correspondences leading to difficulty in accurate calibration work. To alleviate the limitation of few correspondences, unsynchronized images from each viewpoint are utilized so that geometric relationship between 3D coordinates of a calibration object and 2D coordinates of a projected image can be fully exploited. To substantiate the proposed approach, we provide experimental result**

Keywords—camera calibration; camera geometry; fundamental matrix; homography;

I. INTRODUCTION

Recovery of 3D information from 2D images always has been one of the most fundamental and crucial problem in computer vision and computational 3D imaging while it remains challenges to achieve high accuracy in practical applications [1,2]. Numerous methods have been proposed to achieve accurate 3D information from a set of 2D images, and there have been two main approach to 3D reconstruction, the one is called passive method and the other one is called active method [3]. Passive method employs multiple cameras each of which is located in different positions. The most widely used method is stereo vision system, two camera capture the same 3D scene from different locations leading to generating parallax between two viewpoints [4]. Geometric relationship between 3D coordinates of a target 3D scene and the ones of optical cameras provides rich information about 3D coordinates of a target scene. Active method replace one camera by a light source that generate a specific light pattern or projects a light ray onto a 3D target object [5]. In the active method, three main approaches have been popularly used, the one is using known light pattern, e.g., grid patterns, line(horizontal or vertical) patterns, circular pattern, coded patterns, etc., the second one is using point cloud that is generated by light source that projects dot light patterns onto the target object to be reconstructed [6]. The third one is using a light ray that is projected onto the target object, and the round trip time of the light ray and the speed of the ray are used to calculate a distance between the light source and the object.

Prior to the reconstruction work, calibrating cameras is mandatory procedure to achieve accurate reconstruction result. In particular, passive method requires camera parameters, the intrinsic and the extrinsic ones. Numerous work has been introduced in the areas of camera calibration, but it has gained less attention compared to the work of accurate depth estimation and 3-dimensional(3D) reconstruction work. Calibration plays a key role to estimate depth and 3D coordinates of a target object by estimating the intrinsic and the extrinsic parameters. Depth, distance between the optical center and the 3D object point is calculated using the following equation.

$$D = \frac{bf}{d},$$

where D is a depth, b and f are baseline and the focal length of a camera, respectively. d is a disparity. The focal length is an element of the intrinsic parameter. Disparity is acquired by vertical alignment of a pair of stereo images. To achieve the alignment, rectification is required. Relative orientation angle is needed to rectify a pair of stereo images, and the angle is from the extrinsic parameters.

In the past few decades, calibration has been one of the challenges in computer vision, because the calibration requires high accuracy and needs iterative optimization process. There have been two popular calibration methods, the one is Tsai's method and the other one is Zhang's method [7]. Readers can find their work or can read review article by Salvi [8]. The methods aforementioned are based on the geometric relationship between the 3D calibration object and the projected image in 2D space. The calibration has been evaluated by calculating re-projection error. Recently, deep neural network model is applied to the calibration work by training a pair of stereo images with calibration parameters they are estimated *a priori* [9]. Although deep neural network based calibration can provide high accuracy, it consumes long time leading to inefficiency especially for mobile camera or small embedded

insert a graphic (which is i

that you use a text box t

system that need real-time calibration. In this paper, we provide calibration approach for three-view camera system, that has few correspondence between neighboring viewpoints. Some previous work has proposed calibration methods in case of few correspondence, however, it has used additional instrument or has been helped by moving camera. This paper uses unsynchronous view to generate sufficient correspondence followed by performing correspondence matching. The next section briefly introduces camera model and parameters before we propose the approach for our calibration in section 3. Section 4 provides experimental result before we conclude this paper in section 5..

II. Camera Models and Parameters

In computer vision, camera is approximately modeled to a pinhole camera model. Although the model is approximately established, it has been popularly used in the calibration work. The calibration basically aims to achieve the internal and the external parameters of cameras that are located in different positions. The former ones are composed of the focal length of a camera, principal points of an image plane and skew factors. Once the internal parameters are calculated, relationship between the 2D and 3D coordinates of a target scene. The latter ones define relative orientation of the cameras, rotation and translation (distance between optical centers). As similar to the traditional calibration, this work also obeys a traditional calibration model. Calibration work aims to achieve the parameters aforementioned as follows.

$P = p[R\backslash t]$, (1)

where $\boldsymbol{P}$ is 3D coordinates of a target scene and its projection onto 2D image plane is represented by $\boldsymbol{p}$. $\boldsymbol{R} \subset \mathbb{R}^{3\times3}$ represents relative orientation(rotation angle) between viewpoints, and a distance between optical centers of the viewpoints(cameras) is denoted by $\boldsymbol{t} \subset \mathbb{R}^{3\times1}$. Typically, camera parameters (or calibration parameters) are composed of 5 variables of the internal parameters and 12 variables of the extrinsic parameters. Thus, the calibration work is to calculate 17 unknown variables, and the linear model can solve the calibration problem. Zhang's method has been one of the most widely used one since late 1990s. The planar object with known patterns is captured from different locations, and the internal and the external parameters are calculated with optimization process. Deep learning based calibration has gained attention, but the results does not show significant improvement. Once calibration is performed, re-projection error provides quantitative accuracy of the calibration result. In the next section, calibration with few correspondence between neighboring viewpoints is introduced, and the calibration using unsynchronous viewpoint is proposed.

III. Proposed Method

This section details calibration of three-view camera system with few correspondences between neighboring viewpoints. Since insufficient number of correspondences is provided, typical procedure of the calibration cannot be achieved. In this work, three-view thermal camera captures the 3D object that is manufactured for the calibration work. The object is composed of three planar object as shown in Fig. 1.

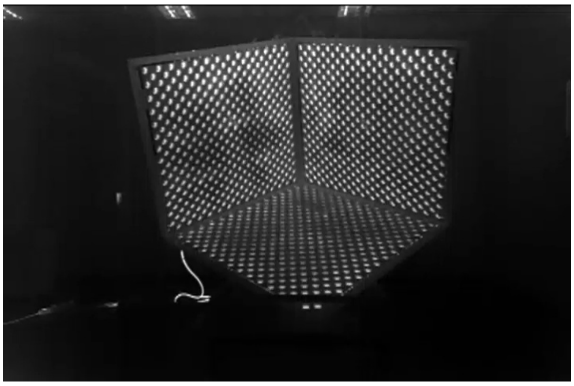

Fig 1. 3D planar object with known pattern for the calibration

Since the camera does not provide sufficient correspondences, traditional feature matching followed by estimating fundamental matrix or by usual capture from different locations does not provide rich information for the calibration. A structure of a camera system is shown in Fig 2.

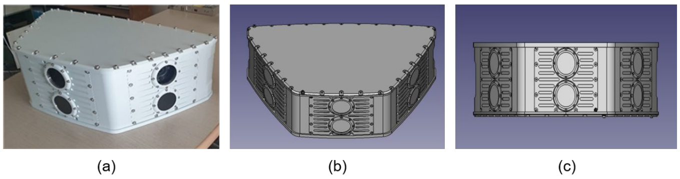

Fig 2. Structure of a three-view camera system.

As shown in Fig 2, if the camera is located at close position (usually closer than 5 meters), three viewpoints share very few feature points. Thus, we try to generate corresponding feature points by unsynchronizing the viewpoints as shown in Fig 3.

insert a graphic (which is i

that you use a text box t

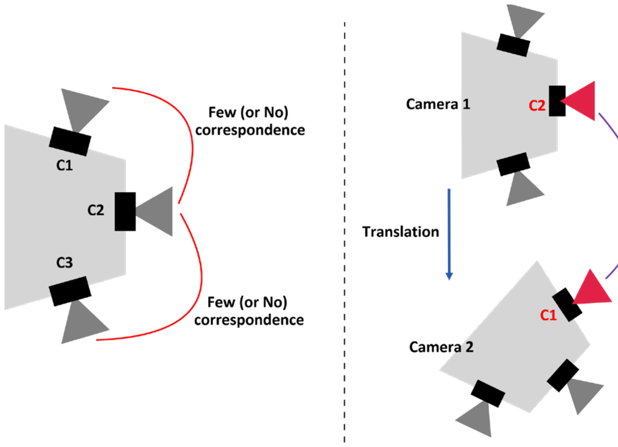

Fig 3. Correspondence is acquired by generating unsynchronous viewpoints in artificial way.

Let $C = (C_1, C_2, C_3)$ denotes a vector of the original viewpoints C_1 and C_2 and C_3 are optical centers of each viewpoint, respectively. Let $C = (C_{1N}, C_{2N}, C_{3N})$ (each element of C_N is an each optical center of a rotated camera)is denoted by a vector of new viewpoints, respectively, then C and C_N have the following relationship,

$$C_N = RC \quad (2)$$

Rotation matrix $\boldsymbol{R} \subset \mathbb{R}^{3\times3}$ ($\mathbb{R}$ is defined as a set of real numbers) is known a priori, then corresponding feature points are extracted from C_2 in camera1 and C_1 in camera2(Fig 3). Let C_2 of camera1 denoted by C_2 and C_1 of camera2 denoted by C_{1N}, then we can extract sufficient corresponding points between C_2 and C_{1N}.. Once correspondence is extracted, fundamental matrix($\boldsymbol{F}$) is estimated. Based on the decomposition of $\boldsymbol{F}$, rotation between C_2 and C_{1N}, can be estimated, denoted by$\tilde{R}$. Since we already have initial rotation angle between C and C_N (which is $\boldsymbol{R}$), the ultimate value of C_1 of camera1 and C_2 of camera1 can be calculated to $\tilde{R}$. Rotation between C_2 and C_3 and C_1 and C_3 also can be calculated using similar approach aforementioned.

Once $\boldsymbol{R}$ is determined, initial rotation value that is for generating unsynchronous viewpoint is used to find a final solution of the relative orientation between viewpoints, written as

$$R_{final} = R + \tilde{R} \quad (3)$$

IV. Experimental Results

To substantiate the proposed approach, this section provides experimental results with detailed procedure of the calibration work. Fig 4 shows a flow diagram of the experimental procedure for the calibration.

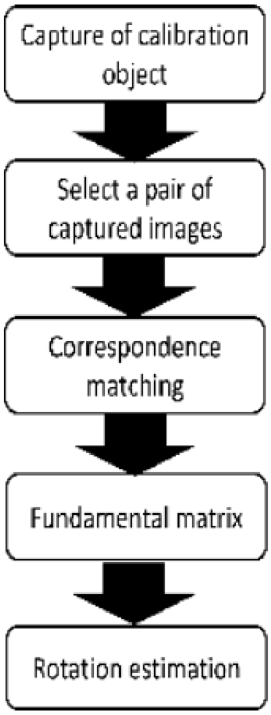

Fig. 4. Experimental procedure for the calibration

In the first, three-view camera captures a 3D object which has a known pattern. Since the angle between neighboring viewpoints is bigger than 180-degree, it is difficult to acquire corresponding feature points (e.g., corners of a checkerboard patterned object). To solve the difficulty, the camera captures the calibration object with various rotation angle that is a known a priori. In this experiment, we rotate the camera with a multiple of 2-degree, i.e., 2, 4, 6, 8-degree, etc. Rotation is carried out until the angle reaches 90-degree. Fig 5 shows the result of correspondence matching between neighboring viewpoints. In the fig 5, a pair of images is captured unsynchronous way. In other words, the left image is captured by rotating the camera to 26-degree. In this work, calibration object is composed of circles, and the distance between centers of neighboring circles is 10centineters. Provided known size of a circle, relationship between 2D pixel coordinates and 3D real coordinates is established. Correspondence matching between neighboring images is achieved. If the rotation is set to 90-degree, as shown in fig 6, a number of corresponding points is significantly increased, however, based on the experiment, fundamental matrix is reliably acquired with 26-degree rotation.

insert a graphic (which is i

that you use a text box t

Fig. 5. Correspondence matching between viewpoints (26-degree rotation)

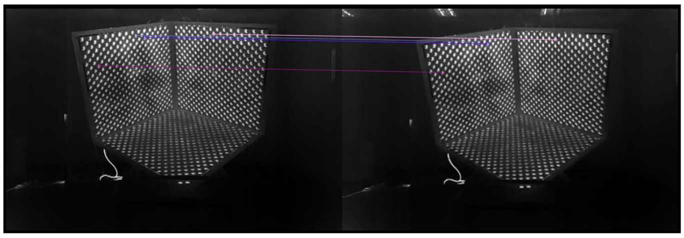

Fig. 6. Correspondence matching between unsynchronized viewpoints (90-degree rotation)

Rotation (represented using 3D angle in x, y and z direction and in radian) and translation (represented using a 3D vector) values of a C1 and C2 are provided in Table 1.

Camera C1		Camera C2	
Rotation(rad)	Translation	Rotation(rad)	Translation
-2.1502	+124.4072	-1.2281	-7.3168
+0.5293	-35.4247	-1.3521	-11.9284
-0.0793	+375.6424	+0.7950	+26.6254

Homography between neighboring viewpoints is estimated, and the result is about 88-degree. As aforementioned, 26-degree is initially rotated, the final result is 114-degree. The groundtruth value of a rotation is about 120-degree (see fig 1). Thus, the estimated value of a rotation using the proposed approach is closed to the groundtruth value.

V. Conclusion

This paper proposes estimating the extrinsic parameters, particularly the relative rotation between neighboring viewpoints. Since a traditional approach does not apply to the camera setup of this work, unsynchronous views are generated to acquire sufficient corresponding points. Unsynchronous views are generated by rotate a camera with known angles, leading to acquisition of sufficient number of corresponding feature points. In the calibration work, circles are used instead of checkerboard patterns. The intrinsic parameters are estimated using usual approach, and the extrinsic parameters are estimated by decomposing fundamental matrix. The rotation value is added to the initial angle which is set to generate unsynchronous viewpoint. Future work will focus on modeling more accurate and reliable mathematical and geometric relationship for the calibration.

Acknowledgment

This work was supported by the National Research Foundation of Korea(NRF) grant funded by the Korea government (Ministry of Education) (No. 2022R1I1A3069352).

References

[1] Q. Xu, W. Wang, D. Ceylan, R. Mech and U. Neumann, "DISN: Deep Implicit Surface Network for High-quality Single-view 3D Reconstruction", Advances in Neural Information Processing Systems 32 (NeurIPS 2019), pp. 1-11, 2019.

[2] R. Guler and I. Kokkinos, "HoloPose: Holistic 3D Human Reconstruction In-The-Wild", Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2019, pp. 10884-10894.

[3] I. Lai, B. Yang, B. ma, M. Liu, Z. Yin, L. Yin and W. Zheng, "An Improved Stereo Matching Algorithm Based on Joint Similarity Measure and Adaptive Weights", Applied Sciences, vol. 13, no. 1, pp. 1-21.

[4] N. O'Mahony, S. Campbell, L. Krpalkova, D. Riordan, J. Walsh, A. Murphy and C. Ryan, "Computer Vision for 3D Perception", Part of the Advances in Intelligent Systems and Computing book series (AISC,volume 869), pp. 788-804, 2018.

[5] H. Nguyen, Y. Wang and Z. Wang, "Single-Shot 3D Shape Reconstruction Using Structured Light and Deep Convolutional Neural Networks, Sensors, vol. 20, no. 13, pp. 1-1, 2020.

[6] S. Zhang, "High-speed 3D shape measurement with structured light methods: A review", Optics and Lasers in Engineering, vol. 106, pp. 119-131, 2018.

[7] Z.Zhang, "A flexible new technique for camera calibration", IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 22, no. 11, pp. 1330-1334, 2000.

[8] J. Salvi, "A comparative review of camera calibrating methods with accuracy evaluation", Pattern Recognition, vol. 35, no. 7, pp. 1617-1635, 2002.

[9] J. Zhang, B. Luo, Z. Xiang, Q. Zhang, Y. Wang, X. Xu, J. Liu and W. Wang, "Deep-learning-based adaptive camera calibration for various defocusing degrees", vol 46, no. 22, pp. 5537-5540, 2021.

[10] G. Eason, B. Noble, and I.N. Sneddon, "On certain integrals of Lipschitz-Hankel type involving products of Bessel functions," Phil. Trans. Roy. Soc. London, vol. A247, pp. 529-551, April 1955. (*references*)

[11] J. Clerk Maxwell, A Treatise on Electricity and Magnetism, 3rd ed., vol. 2. Oxford: Clarendon, 1892, pp.68-73.

[12] I.S. Jacobs and C.P. Bean, "Fine particles, thin films and exchange anisotropy," in Magnetism, vol. III, G.T. Rado and H. Suhl, Eds. New York: Academic, 1963, pp. 271-350.

[13] K. Elissa, "Title of paper if known," unpublished.

[14] R. Nicole, "Title of paper with only first word capitalized," J. Name Stand. Abbrev., in press.

[15] Y. Yorozu, M. Hirano, K. Oka, and Y. Tagawa, "Electron spectroscopy studies on magneto-optical media and plastic substrate interface," IEEE Transl. J. Magn. Japan, vol. 2, pp. 740-741, August 1987 [Digests 9th Annual Conf. Magnetics Japan, p. 301, 1982].M. Young, The Technical Writer's Handbook. Mill Valley, CA: University Science, 1989.

insert a graphic (which is i