A video stitching system based on mirror pyramids and non-overlapping calibration method

Ling Zhu, Wei Wang, Yu Liu, ShiMing Lai, Jing Li
College of Information System and Management
National University of Defense Technology
Changsha, China 410073
Email: orange4working@gmail.com

***Abstract*—The panoramic cameras on the market do not allow seamless stitching of the object close to the camera with the existence of non-zero parallax angle. Meanwhile, the traditional real-time stitching technology requires that the geometry of the camera group is known, and the external projection points are used to manually find feature points in the overlapped region. Therefore, this paper presents a method in creating a mirror reflection panoramic real-time splicing system by calibration. And because that the images taken by each camera have very narrow overlapping areas, a non-overlapping calibration method is proposed, requires only four fixed calibration patterns by rotations to complete a simple and convenient calibration work with evident effect. The ability of the camera being built on capturing a close-up shot during live video is of great significance for the future of panoramic video shooting.**

***Index Terms*—Mirror pyramids, panoramic camera, real-time stitching system, calibration,**

I. Introduction

The techniques for constructing panoramic cameras can be divided into the following categories: refractive methods, including only refractive elements (e.g., lenses), and catadioptric methods using a reflecting member (e.g., a mirror) in combination with the refractive element. The literature [1] provides a good overview on panoramic imaging techniques. The refractive system includes a camera cluster, a system using a fisheye lens, and a rotating camera. The market is generally dominated by a panoramic camera that has non-zero parallax angle, which does not allow close-ups shots, seamless splicing when the object is close to the camera, or other similar scenes.

Certain authors have attempted to solve the problem of parallax angles using a single camera and sensors with curved mirrors [2] [3] [4] [5]. However, a single camera is too low-resolution; it is unable to meet the current user requirements for high-definition video. Multiview panoramic camera using mirror pyramid is built by arranging individual cameras inclined to form a pyramid. The viewing angles of the individual cameras can be co-located at a single point within the pyramid, thereby effectively forming a virtual camera with a wide field of view.Majumder et al. [14] provided a method for real-time stitching of specular-reflection panoramic imaging systems. However, the method requires that the geometry of the camera group is known and the external projection points are used to manually find the feature points.

Therefore, this paper presents a method building the mirror reflection panoramic real-time stitching system based on calibration. The three main steps are as follows: first is the camera calibration, second is calculating coordinates that correspond to the template of the first frame of the video by projection transformation in the camera parameters, and third is filling the pixel of the rest of the video using the template for real-time stitching.

The general approach to resolve the problem of calibration among camera groups is to assume that the camera has overlapping regions, in which camera parameters can be estimated to use the points in the overlapping region in registering the position relative to the camera [15, 16]. However, because of the structure of the specular panoramic camera, the images taken by each camera have very narrow overlapping areas. To solve the non-overlapping area or small overlap area of the camera calibration problems, such an approach must increase extra cameras, thereby greatly increasing the experimental cost and the possibility of the cumulative error.

In this paper, Section 3 shows the concept of a panoramic real-time stitching system based on specular reflection. In section 4, a rigid constraint is proposed to solve the problem of non-overlapping region calibration of camera based on specular reflection for a uniform optical center. Section 5 shows the results of the proof of concept and a conclusion ends the paper.

II. RELATED WORK

A. Multiview panoramic camera using mirror pyramids

Panoramic cameras using mirror pyramids have many attractive features, including single-view imaging, high resolution, and video capture. Nalwa et al. [6] developed a four-sided specular reflector and proposed the four-faceted specular reflection device with stereoscopic imaging idea in [7]. Then, [8] added the upward camera, increased the vertical field of view angle, and theoretically discussed the situation of the pyramid for eight surfaces. Shimamura et al. [9] used a pair of six-sided mirror pyramid to achieve three-dimensional imaging and to accomplish experimental implementation. Aggarwal et al. [10] realized high dynamic imaging with a specular reflector.

Tan et al. [11] proposed a method for summarizing existing designs to create multiple viewpoints in a single mirror pyramid. Schreer, O. et al. [12] designed a six-sided mirror pyramid device that captures three-dimensional panoramic video using

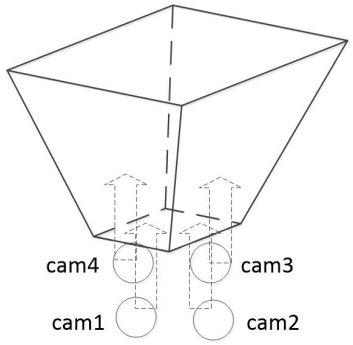


Fig. 1. Design of panoramic camera

the positional relationship of adjacent cameras. Their designs were validated in [13].

B. Calibration of cameras with non-overlapping regions

Solutions to non-overlapping area calibration are to keep the system stationary, attaching additional cameras to move it around, and observe feature points that overlap with each camera to be able to calibrate the camera group indirectly. Caspi and Irani proposed a method in registering non-overlapping region image sequences [17], but a constraint exists in which multiple camera groups have a common optical center point and do not apply to a general multi-camera system.

Certain authors use mirrors to create overlaps between scenes. The external calibration of a multi-camera system is done using a known [18] or unknown [19] geometric calibration pattern. However, the use of an additional mirror in this method is not convenient in the calibration process [20].

Lamprecht et al. [21] completed the calibration of the moving multi-camera system. In the calibration process, the camera tracks a static calibration object, and requires a priori knowledge of system speed. Esquivel et al. [22] used the rigidity constraints between camera groups to perform the calibration work by calculating the relative attitude of each camera track. The relative posture of the camera group is then deduced without any changes using the position relative to the camera. During the process, the quaternions of unit length replaced the rotation vector. Based on this study, Esquivel et al. [23] referred to the process of solving the rigid constraint parameters as sfm (re-projection process of reconstruction scale factor), introducing enhanced rotation (pole geometry constraints) and translation constraints (transformed into rotation phase with consistent size proportion). Recently, they proposed a new re-projection optimization parameter function to replace the geometric constraints [24], thereby consolidating the eye-to-eye calibration and sfm combined pipeline.

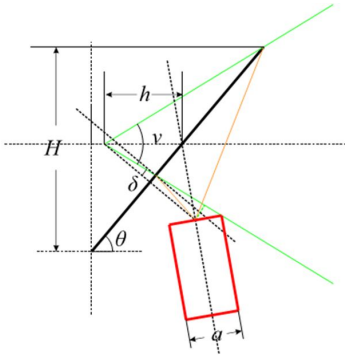


Fig. 2. Analysis of the camera

One calibration plate was used at the end of the experiment, whereas the other was extracted with features to reconstruct the scale factor. Following Esquivel et al., Dai et al. [25] proposed a solution for beam averaging of camera sets by direct calculation of posture averaging for the relative attitude instability. Leopardy et al. [26] suggested the sparse optimization beam adjustment in the iterative phase.

III. Design of Panoramic Real - Time Stitching System Based on Specular Reflection

A. Panoramic camera design

This camera structure design is similar to [7]. Four up-shot HD cameras are placed on the pallet for the parallel light to be reflected by the mirror in the camera. Fig. 1 shows the design of the camera. The vertical viewing angle of this virtual camera design is analyzed in Fig. 2 : where a is the width of the camera lens, H is the height of the mirror, θ is the horizontal angle for the mirror, δ is the field of view because of the camera loss, h is the distance between the virtual center and the mirror center, v is the vertical field angle. By setting h (i.e., the size of the device), the relationship between θ and the loss of field angle δ is as follows:

$$2h\sin\theta\sin\delta = \frac{a}{2}\sin(\theta-\delta) \tag{1}$$

$$\sin^2\delta = \frac{\frac{a^2}{4}\sin^2\theta}{4h^2\sin^2\theta + 2ha\sin\theta\cos\theta + \frac{a^2}{4}} \tag{2}$$

The vertical field angle can be obtained by θ and δ:

$$v = 2(\frac{\pi}{2} - \theta - \delta) \tag{3}$$

Obtaining the desired height of the device when the horizontal angle is $\theta$is possible

$$H = r\tan\theta + \frac{h}{\cot(\frac{v}{2}) - \cot(\theta)} \tag{4}$$

B. video stitching framework

This paper proposes the process of real-time stitching server based on calibration to solve the problem of the high accuracy of location as well as to manually identify the corresponding points in the overlapping area in the traditional real-time splicing server process. The design of real-time stitching server process based on the calibration is as follows (Fig. 3):

1. First, the intrinsic and extrinsic parameters of the camera is calibrated. In this case, because of the structure of the mirror-reflex reflection, the calibration problem is changed to multi-camera non-overlapping area calibration. The general method is to assume that the camera has overlapping areas

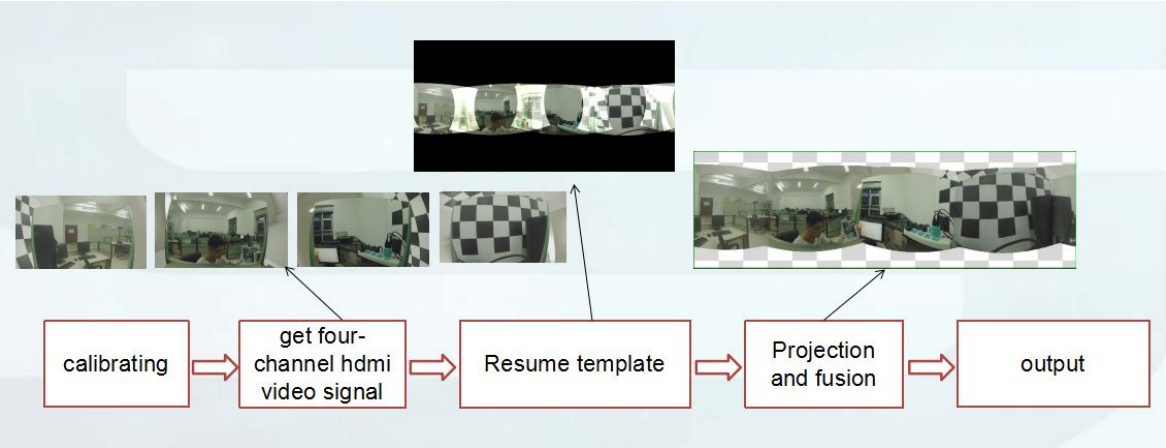


Fig. 3. Flow of real-time stitching system design

to estimate these camera parameters by using the points in the overlapping area in registering the relative position of the camera [15, 16]. Therefore, this paper presents a method based on the rigid constraints to complete the calibration work (Section 4).

2. The camera video signal is read, and processing is awaited.

3. The first frame of the video is taken to calculate the projection transformation. Considering the virtual refraction structure of the mirror that makes the virtual optical center of the multi-camera invariable, this paper then discusses the projection transformation in Section 5 to adapt to such conditions.

4. Following the coordinates of the corresponding template on filling the pixels of follow-up frame after projection and fusion, the output of latitude and longitude coordinates for the unified coordinate system panorama are obtained.

5. The video being spliced can output via HDMI or transferred via RTMP for network streaming.

IV. Calibration

The calibration phase plays an important role in the construction of the real-time splicing server, whereas the special structure of the specular reflection panoramic camera converts the calibration problem into non-overlapping area camera group calibration. Therefore, the calibration method of the same optical center camera based on rigid constraint is adopted.

A. Rigid coupled motion constraints with unified optical center

We assume that cam_0^0 and cam_0^k are the positions of the main camera at time 0 and time k respectively. cam_j^0 and cam_j^k are the position of camera j at time 0 and time k respectively. X_0^0 and X_j^k are the coordinate systems under the space coordinates of host camera 0 at time 0 and camera j at time k of the space coordinates (Fig. 4).Then two constraints can be obtained as follows[22]:

$$R_j^k \Delta R_j = \Delta R_j R_0^k \tag{5}$$

$$R_{\mathrm{j}}^k \Delta C_j + C_j^k = \Delta\lambda_j \Delta R_j C_0^k + \Delta C_j C_0^k \tag{6}$$

Each scalar ($\Delta\lambda_j > 0$) describes the isometric zoom between the camera J and the reference camera, and $\Delta R_j, \Delta C_j$ describes the position of the camera J relative to the reference camera. When the light center is very close, which is close to 0, then the constraint (5) will be:

$$C_j^k = \Delta\lambda_j \Delta R_j C_0^k \tag{7}$$

B. Using a pair of calibration plates

A point on the calibration board can replace the scene information. The benefits of using the calibration board are as follows. First, considering that the coordinates of the corner points on the calibration plate are $\begin{bmatrix} x & y & 0 & 1 \end{bmatrix}^T$ in the coordinate system and the real coordinate is transformed to the camera coordinate system by the external parameters, the scaling factor of the feature points in the camera coordinate system is 1. Therefore, the original constraints can be written as

$$R_j^k \Delta R_j = \Delta R_j R_0^k \tag{8}$$

$$C_j^k = \Delta R_j C_0^k \tag{9}$$

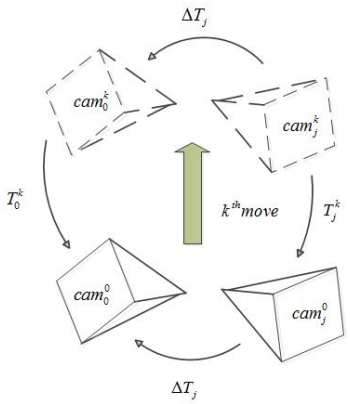


Fig. 4. Rigid constraints

Second, more than three calibration plate images can be obtained by transforming the camera position. This images can be directly obtained by the traditional calibration method [27]. R is the rotation vector and C is the translation vector. K is the camera internal matrix.

The relative attitude point in Fig. 5 that $T^{\mathrm{k}} = [R^{k0}|C^{k0}]$ of the camera reference frame to the k-time point can be obtained as follows: We suppose that the calibration plate coordinate system under the point x^0 and x^k, for the camera at 0 time and k time, respectively, are as follows:

$$R^0 x^m + C^0 = x^0 \tag{10}$$

$$R^k x^m + C^k = x^k \tag{11}$$

Thus, the following equation can be obtained

$$(R^0)^{-1}x^0 - (R^0)^{-1}C^0 = (R^k)^{-1}x^k - (R^k)^{-1}C^k \tag{12}$$

for:

$$R^{ko}x^k + C^{k0} = x^0 \tag{13}$$

Thus, the following formula can be obtained

$$R^{k0} = R^0(R^k)^{-1} \tag{14}$$

$$C^{k0} = C^0 - R^0(R^k)^{-1}C^k \tag{15}$$

C. Improved Solution of Rigid Constraints

In order to facilitate the solution, a unit quaternary array is used instead of the rotation matrix. The quaternion per unit length [28] is commonly used in computer vision and graphics for rotational estimation because the variable suffices to be just slightly over parametric in the rotation representation, not affected by singularity, and computationally convenient to use:

$$q = (\cos(\frac{\alpha}{2}), \sin(\frac{\alpha}{2})r) = (w,x,y,z) \tag{16}$$

The solution constraint (12) is transformed into a solution constraint

$$q_j^k \Delta q_j = \Delta q_j q_0^k \tag{17}$$

equivalent to solving

$$(T_{q_0^k} - T_{q_j^k}{}^*)\Delta q_j = 0 \tag{18}$$

Where T_q and T_q^* represent left multiplication and right multiplication of $q = (w, x, y, z)$ To solve the linear equation :

$$A * \Delta q = 0 \tag{19}$$

where

$$A = \begin{bmatrix} \mathrm{w}_0^k - w_i^k & -x_0^k + x_i^k & -y_0^k + y_i^k & -z_0^k + z_0^k \\ x_0^k - x_i^k & \mathrm{w}_0^k - w_i^k & -z_0^k - z_i^k & y_0^k + y_i^k \\ y_0^k - y_i^k & z_0^k + z_i^k & \mathrm{w}_0^k - w_i^k & -x_0^k - x_i^k \\ z_0^k - z_i^k & -y_0^k - y_i^k & x_0^k + x_i^k & \mathrm{w}_0^k - w_i^k \end{bmatrix} \tag{20}$$

When the constraint conditions are greater than the parameters and when it is needed to solve the following least squares problem, the problem can be solved by SVD [29]

$$\begin{aligned} &\min_{\Delta q_j} \sum_{k=0}^{m} ||A^k \Delta q_j||^2 \\ &s.t. ||\Delta q_j|| = 1 \end{aligned} \tag{21}$$

To use the constraint condition (9), refer as the initial value, then the following minimization problem is solved:

$$f = \sum_{k=1}^{K} ||q_j^k \Delta q_j - \Delta q_j q_0^k||^2 + ||C_j^k - \Delta R_j C_0^k||^2 \tag{22}$$

Based on the transformed four element array into the four-dimensional hemisphere space that can remove the constraint ($||\Delta q_j|| = 1$), the result of the above equation can be obtained. q, R are represented by $\psi = (x, y, z)$:

$$q = (\frac{2x}{\alpha^2+1}, \frac{2y}{\alpha^2+1}, \frac{2z}{\alpha^2+1}, \frac{1-\alpha^2}{\alpha^2+1}) \tag{23}$$

where

$$\alpha^2 = x^2 + y^2 + z^2 \tag{24}$$

The minimization problem is converted to an unconstrained nonlinear optimization problem with the parameters, which can be solved by the method of L-M.

V. Experimental Results and Proof-of-Concept

This experiment camera uses four GoPro4 Black cameras to shoot 4k video from bottom to top. The horizontal angle of the pyramid is set to 45 degrees. On the basis of the analysis in Chapter 3, the vertical angle of the camera is roughly 60

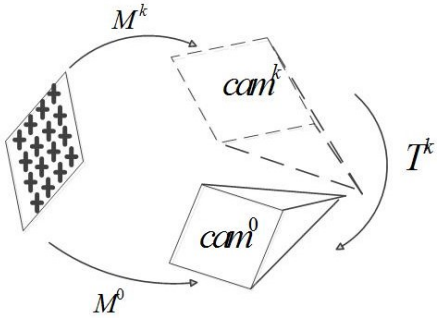
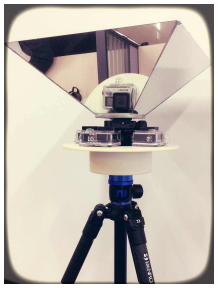


Fig. 5. Calculation of the relative pose



(a) Show of camera

(b) Calibration of the camera

Fig. 6. Actual experiment

Fig. 8. Image projected to the same coordinate system

degrees. Figure 6(a) shows the actual experimental camera designed.

Following the proposed calibration method, the following experimental environment is set up: four calibration patterns are placed around the camera, and the calibration is finished by rotating the camera several times. Figure 6(b) shows the experimental environment during the calibration phase.

Estimating the experimental error by taking the points in the overlapping region is difficult because the calibration plate is in the non-overlapping region. Therefore, the design of the sample frame by the above method to train the calibration results is calculated by using the test frame formula of the projection error. As shown in Fig. 4, the coordinate system cam_j^k randomly generated N points and through T_j^k, ΔT_j , $(T_0^k)^{-1}$ and ${\Delta T_j}^{-1}$ that can be re-projection of the points, the design of the projection function:

$$u_{ij}^k = {\Delta T_j}^{-1}(T_0^k)^{-1}\Delta T_j T_j^k x_{ij}^k \tag{25}$$

$$err = \sum_{j=1}^{N}\sum_{k=1}^{K}\sum_{i=1}^{M}||x_{ij}^k - u_{ij}^k||^2 \tag{26}$$

Where x_{ij}^k represents the ith point of the j-th camera calibration plate in attitude k,u_{ij}^k is the point after the projection. In this experiment, four cameras were used to compose the camera group. In each camera, the camera's coordinate system randomly generated 500 spatial points to test the error of the main camera relative to other servo cameras when the number of gestures k = 3 to 8. Experiments show that the increase of the attitude of the camera can obviously improve the relative attitude of the camera accuracy. The calibration data is accurate as the camera rotates (see Fig. 7).

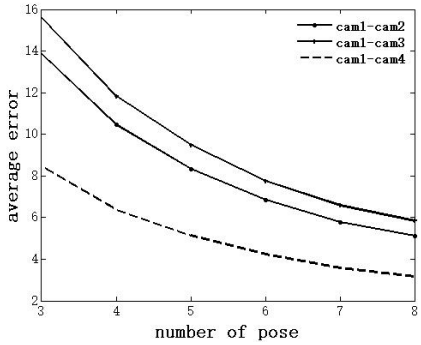


Fig. 7. The average error of calibration result

Figure 8 shows images that are projected to the same coordinate system by the intrinsic and extrinsic parameters of the four cameras. We can see that there is no obvious seam when the object is close to the camera. The method can complete the calibration of the mirror-reflection panoramic camera, and the effect is evident.

VI. Conclusions

In this paper, we propose a method of building a panoramic real-time stitching system based on calibration. The main steps of the method are divided into three steps, the first step is the camera calibration. Aiming at the extremely narrow overlapping area of the mirror-reflex type panoramic camera, a calibration method based on rigid constraint is proposed. After the four calibration patterns being fixed, the calibration can be completed only by rotating the camera. Using the calibration method of this paper, the intrinsic parameters of virtual camera can be calibrated directly at the same time. It is simple and convenient and the effect is obvious. The second step is calculating the coordinate corresponding template of the first frame of the video using the calibrated camera parameters. The parameters calibrated are used to generate the panoramic image by using the projection transformation algorithm which is adapted to the single viewpoint imaging. The third step is to apply the template to fill the remaining frames of the video to realize real-time splicing.The panoramic video produced by this method can be directly previewed on-line,or for broadcast by network streaming.

Although the method proposed in this paper is simple and effective, there is still a lot of work to be done in the future. First, there is still a slight registration error at the splice junction. Secondly, the panoramic imaging system based on specular reflection has a smaller vertical field angle, which will reduce the effect of live video. We will try to add up the additional camera or mirror pyramid in pairs in the future, combined with the proposed real-time stitching system building methods to achieve better live results.

References

[1] Tan K H, Hua H, Ahuja N. Multiview Panoramic Cameras Using Mirror Pyramids[J]. IEEE Transactions on Pattern Analysis Machine Intelligence, 2004, 26(7):941-6.

[2] Nayar S K. Catadioptric omnidirectional camera[C]// Conference on Computer Vision and Pattern Recognition. IEEE Computer Society, 1997:482-488.

[3] Kawanishi T, Yamazawa K, Iwasa H, et al. Generation of high-resolution stereo panoramic images by omnidirectional imaging sensor using hexagonal pyramidal mirrors[J]. 1998, 1(3):485-489 vol.1.
[4] Baker S, Nayar S K. A Theory of Single-Viewpoint Catadioptric Image Formation[J]. International Journal of Computer Vision, 1999, 35(2):175-196.
[5] Peleg S, Benezra M, Pritch Y. Omnistereo: Panoramic Stereo Imaging[J]. IEEE Transactions on Pattern Analysis Machine Intelligence, 2001, 23(3):279-290.
[6] V. Nalwa. A true omnidirectional viewer. Bell Laboratories Technical Report,1996.
[7] Nalwa V S. Stereo panoramic viewing system: US, US6141145[P]. 2000.
[8] Nalwa V S. Panoramic viewing system with offset virtual optical centers: U.S. Patent 6,111,702[P]. 2000-8-29.
[9] Shimamura J, Yokoya N, Takemura H, et al. Construction of an Immersive Mixed Environment Using an Omnidirectional Stereo Image Sensor[C]// Omnidirectional Vision, 2000.
[10] Aggarwal M, Ahuja N. Split Aperture Imaging for High Dynamic Range[J]. International Journal of Computer Vision, 2001, 2(1):7-17.
[11] Aggarwal M, Ahuja N. Split Aperture Imaging for High Dynamic Range[J]. International Journal of Computer Vision, 2001, 2(1):7-17.
[12] Schreer O, Kauff P, Eisert P, et al. Geometrical design concept for panoramic 3D video acquisition[J]. 2012:2757-2761.
[13] Weissig C, Schreer O, Eisert P, et al. The Ultimate Immersive Experience: Panoramic 3D Video Acquisition[C]// International Conference on Advances in Multimedia Modeling.
[14] Majumder A, Seales W B, Gopi M, et al. Immersive teleconferencing: a new algorithm to generate seamless panoramic video imagery[C]// ACM International Conference on Multimedia. 1999:169–178.
[15] Hartley, R., Zisserman, A.: Multiple View Geometry in Computer Vision. Cambridge University Press, Cambridge (2000)
[16] Zisserman, A., et al.: Metric Calibration of a Stereo Rig. In: Proc. WRVS (1995)
[17] Y. Caspi and M. Irani, Alignment of non-overlapping sequences,in Computer Vision, 2001. ICCV 2001. Proceedings. Eighth IEEEInternational Conference on, vol. 2. IEEE, 2002, pp. 7683.
[18] R. Kumar, A. Ilie, J.-M. Frahm, and M. Pollefeys, Simple calibration of non-overlapping cameras with a mirror, IEEE Conference onComputer Vision and Pattern Recognition, 2008. CVPR 2008.
[19] P. Lebraly, C. Deymier, O. Ait-Aider, E. Royer, and M. Dhome,Flexible Extrinsic Calibration of Non-Overlapping Cameras Using a Planar Mirror: Application to Vision-Based Robotics, IROS 2010.
[20] Y. Dai, J. Trumpf, H. Li, N. Barnes, and R. Hartley, Rotation Averaging with Application to Camera-Rig Calibration, Computer VisionACCV 2009, pp. 335346, 2010.
[21] B. Lamprecht, S. Rass, S. Fuchs, and K. Kyamakya, Extrinsic camera calibration for an on-board two-camera system without overlapping field of view, in Intelligent Transportation Systems Conference, 2007. ITSC 2007. IEEE, sept. 2007, pp. 265 270.
[22] Esquivel S, Woelk F and Koch R, Calibration of a multi-camera rig from non-overlapping views, Pattern Recognition, Springer Berlin Heidelberg, 82-91 (2007).
[23] Esquivel S, Koch R. Structure from Motion Using Rigidly Coupled Cameras without Overlapping Views[M]// Pattern Recognition. Springer Berlin Heidelberg, 2013:11-20.
[24] Esquivel S, Koch R. Multi-Camera Structure from Motion with Eye-to-Eye Calibration[M]// Pattern Recognition. 2015:29-40.
[25] Dai Y, Trumpf J, Li H, et al. Rotation Averaging with Application to Camera-Rig Calibration[C]// Computer Vision - ACCV 2009, Asian Conference on Computer Vision, Xi'an, China, September 23-27, 2009, Revised Selected Papers. 2009:335-346.
[26] Lebraly P, Royer E, Ait-Aider O, et al. Fast calibration of embedded non-overlapping cameras[C]// IEEE International Conference on Robotics Automation. 2011:221-227.
[27] Zhang Z. A Flexible New Technique for Camera Calibration[J]. IEEE Transactions on Pattern Analysis Machine Intelligence, 2000, 22(11):1330-1334.
[28] Kuipers J B. Quaternions and rotation sequences[M]. Princeton: Princeton university press, 1999.
[29] Esquivel S, Woelk F, Koch R. Calibration of a Multi-camera Rig from Non[M]// Pattern Recognition. Springer Berlin Heidelberg, 2007:82-91.
[30] Terzakis G, Culverhouse P, Bugmann G, et al. On quaternion based parameterization of orientation in computer vision and robotics[J]. Journal of Engineering Science Technology Review, 2014, 7(1):82-93.