Automatic Weak Calibration of Master-Slave Surveillance System Based on Mosaic Image

Li You, Song Li, Wang Jia
Institute of Image Communication and Information Processing
Shanghai Jiao Tong University
Shanghai, China
{*liyou, song_li, Jiawang*}@*sjtu.edu.cn*

***Abstract*—A master-slave camera surveillance system is composed of one(or more) wide FOV(field of view) static camera and one(or more) dynamic PTZ(Pan-Tilt-Zoom) camera. In such a system, master camera monitors a wide field and provides positional information of interesting objects to slave camera so that it can dynamically track them. This paper describes a novel method for the calibration of master-slave surveillance. The method uses a mosaic image created by snapshots of slave camera to estimate the relationship between static master camera plane and pan-tilt controls of slave camera. Compared with other ways, this solution provides an efficient and automatic way to calibration of a master-slave system.**

***Keywords*-mosaic image; master-slave camera system; calibration;**

I. Introduction

Video surveillance system is already prevalent in our daily life. We can see it in airports, railway stations, banks etc. However, with the development of surveillance system, monitored area is larger and larger while the demand of high resolution image about interesting event is also increased. One solution of that problem is master-slave surveillance system. In such a system, one(or more) fixed wide FOV sensor acts as the leader, and directs follow PTZ cameras to zoom into the targets of interest. Coupling these two types of sensors enable the exploitation of their their respective drawbacks. Then, an important question of the system is: how to control PTZ camera to aim at the same object detected by static camera? Calibration, as a pre-prepared step of fixing such system, enable information exchange between the two sensors. So, the object detected by master camera would be focused and tracked successfully by the slave camera. There are two approaches for system calibration.

Strong calibration completely constructs the real model of the system and enables 3D points in world coordinates project into 2D points in each image frame. However, such approach is almost impossible. In fact, transportation, installation, changes in temperature and humidity in outdoor environment would affect the estimated parameters. Weak calibration computes the relation between pixels in static camera and the pan-tilt parameters in active camera. The basic idea of weak calibration is to find a mapping between pixels coordinates and angles of rotation in the dynamic sensor. Through this way, a pixel in the static camera, which represents a 3D point in world, would be transformed into pan-tilt controls of PTZ camera, which make the camera focus to a point correspond to the same 3D point.

In the literature, several methods have been proposed to solve either strong or weak calibration of master-slave camera system. Hampapur et al[1] triangulated an object's position by two or more calibrated cameras and determine the steering parameters for a third, PTZ camera which is also calibrated. Zhou et al[2] selected a number of pixel locations in a static camera. For each pixel, manually move the slave camera to center the slave image at P_i and record the corresponding slave pan-tilt angles $S_i(P_i, T_i)$ to form lookup table(LUT) linking static camera coordinates with the pan and tilt angles. Though it is accurate enough to initialize the track of dynamic camera, the method is time consuming and inconvenient. Badri et al[3] described a method which can automatically sample grids in the master camera image, find proper pan-tilt parameters of slave camera to observe every grid, and extend lookup table by interpolation. Though image in static camera is automatically subdivided by this method, a manual process to find the proper parameters of active camera is inevitable.

In this paper, a novel method is proposed to calibrate a master-slave system. The solution is completely automatic and we show that its accuracy is sufficient in application. This paper is organized as follows: In section 2, the principle and details of this method are described. In section 3, we discuss accuracy of this method and show the experiment result.

II. Weak Calibration Based on Mosaic Image

A. Principle

As for a leader-follower system is composed of a static wide angle camera and a dynamic camera, let I_s and $I_d(\alpha, \beta, Z)$ represent the images of static and PTZ camera respectively. The parameters (α, β, Z) represent the pan, tilt and zoom control of a PTZ camera. The goal of our approach is: for any point in static camera image, a suitable camera control (α, β, Z) would be computed and adopted to make

1051-4651/10 $26.00 © 2010 IEEE
DOI 10.1109/ICPR.2010.450

dynamic camera focus the corresponding point. That is to say: learn a mapping ζ between the pixel position (x_s, y_s) in the static camera and the pan-tilt angles (α_Z, β_Z) at a fixed zoom level Z:

$$(\alpha_Z, \beta_Z) = \zeta(x_s, y_s, Z) \tag{1}$$

Inspired by Bimbo[4], a mosaic image and pair-wise geometry relationship are adopted to solve the problem. Fig.1[4] shows the geometry relationship among static image plane, mosaic image plane, and current dynamic camera plane.

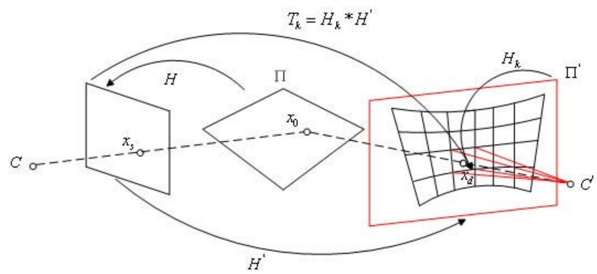

Figure 1. The geometry relationship in master-slave system

H denotes the homography of master camera from world to image plane. H' is the homography between the image plane of static camera and the mosaic image plane $\prod'$ of dynamic camera. H_k represent the homography between the mosaic image plane and actual image plane of dynamic camera. As long as we know H_k and H', the position of point x_s in static camera would be transformed to the slave camera C' by:

$$T_k = H_k * H' \tag{2}$$

$$x_d = T_k * x_s \tag{3}$$

where T_k denotes the transformation from static camera plane to current dynamic camera plane, and x_d is the corresponding point in dynamic camera. Meanwhile, T_k is mapped to the camera control parameters (α, β, Z), which means the corresponding point x_d of x_s would be in the center of dynamic camera by operating it with control $(\alpha, \beta, Z) : T_k \to \xi(\alpha, \beta, Z)$. Now, the calibration problem is transformed to:

$$T_k \to \xi(\alpha, \beta, Z) = \xi_k * \xi' \tag{4}$$

where ξ_k denotes the mapping from mosaic image plane to actual image plane in PTZ control parameters form, and ξ' is the mapping from image in static camera to mosaic image.

B. Automatically Create Image Mosaic

In our experiment, the mapping ξ_k is obtained through a process of creating mosaic image at a fixed zoom level Z_0. The mosaic image is created by regularly rotating dynamic camera and snapping sample images with certain rule. Usually, the regularly movement is chosen to cover the FOV of PTZ camera. Along with the movement of dynamic camera, source images are shot at specified locations and indexed by (α, β, Z).

Mosaic image, at some extent, is an integrated image created by warping a number of source images shot at different angles to one(or more) common image plane. Mosaic image is often regarded as a powerful method to enlarge the FOV(field of view) of a camera. The details of how to create a mosaic are described in [5]. In the experiment, we follow the feature based method to create mosaic image. For short, the key steps of stitch two source images are:

1. Computing correspondence: the features and its descriptor are found by SURF[6] algorithm, the best matched pairs of SURF feature points are also found.

2. Image registration: Based on the matched pairs of points, a RANSAC algorithm is used to find the best transformation matrix M between two images.

$$\begin{bmatrix} x' \\ y' \\ z' \end{bmatrix} = \begin{bmatrix} m_0 & m_1 & m_2 \\ m_3 & m_4 & m_5 \\ m_6 & m_7 & m_8 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} \tag{5}$$

3. Image blending: once we have found the best transformation M, the input images would be blended together. In order to integrate them seamlessly, we use a weighted image stitch algorithm.

The process is repeated between adjacent images in every row and column, finally, the mosaic image is created from 64 source pictures. Seen in Fig.2. The mosaic image is composed of 8*8=64 source images. The white lines are borders of every source image. Once we obtain mosaic image, the mapping ξ_k between mosaic image plane and source image could be found:

$$\xi_k \to (\alpha_i, \beta_i, Z) \tag{6}$$

Figure 2. The mosaic image mad from PTZ camera

C. Automatic Calibration Between Static Camera and Dynamic Camera

After the construction of mosaic image, the mapping between mosaic image plane and current actual image plane is obtained. The next step is to find the mapping from image in static camera to the mosaic image: ξ'. This problem is solved by interpolation of matched feature points between static image and mosaic image. Though not very precise, with the consideration of weak calibration and the FOV of dynamic camera, it is still effective through our experiment.

This procedure requires a method of detecting and matching visual features robust to scale, rotation, view-point, and lightning. The Speeded Up Robust Features(SURF)[6] exhibits great performance under these constraints. So, SURF feature points are selected and matched through mosaic image and static image. After this process, an approximate transformation ξ' is acquired, the final step of the method is combining two obtained mapping:ξ_k, ξ' together.

The final stage of weak calibration is composed of such steps:

1. For any point $P_s(x_s, y_s)$ in static image, a searching algorithm is adopted to find the neighbor SURF points of P_s. Let $N_R(P_s)$ denotes the nearest SURF points set in the range of $R : N_R(P_s) = \{(P_s^1, r_1), (P_s^2, r_2), ...\}$, where r_i is the distance between P_s and P_i.

2. For every point in the set $N_R(P_s)$, find its corresponding feature point P_d^i in the mosaic image. As the reason that P_d^i is in many source images at the same time, an algorithm based on the minimal distance from P_d^i to the border of source image is executed to find the "best" source image I_r^i which "best" contains P_d^i(the larger the minimal distance is , the closer the point is to the center of image).

3. After the steps above, the point $P_s(x_s, y_s)$ is related to several index numbers which represent the rotation parameters of PTZ camera. A liner interpolation solution is used to find the best rotation parameters of PTZ camera to observe the corresponding point of P_s in the dynamic camera. Let the indexes denote by $S_i(\alpha_i, \beta_i, Z_i), i = 1, 2, ..., n$. The interpolation process is followed by equation(7):

$$\begin{aligned} S &= S_1 * f_1(r_1, r_2, ..., r_n) + S_2 * f_2(r_1, r_2, ..., r_n) \\ &\quad + ... + S_n * f_n(r_1, r_2, ..., r_n) \end{aligned} \tag{7}$$

where f_i is a weighted interpolation function:

$$f_i(r_1, r_2...r_n) = r_i \Big/ \sqrt{r_1^2 + r_2^2 + ... + r_n^2} \tag{8}$$

Fig.3 illustrates all the steps mentioned above.

D. Experiment and Result

In order to evaluate the accuracy of the method, we estimate the error between the actual corresponding point in dynamic camera and the center in dynamic image. We firstly analyze the accuracy in the step of creating mosaic image, then, show some results in application. In the proposed method, camera control information is automatically obtained through creating mosaic image. Accuracy in this step directly influences the precision of calibration. We assume that the mosaic image covers the field of view in

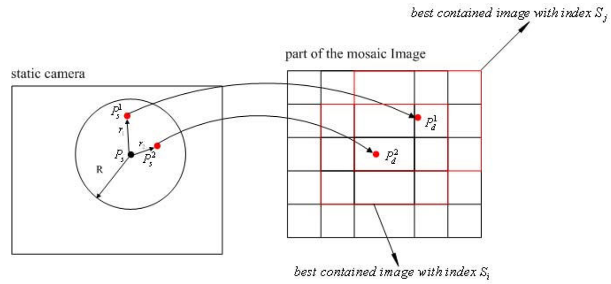

Figure 3. The final calibration process

static camera. Under this principle, the process of snapping and blending source images is similar to a sampling process. Accuracy in this step is determined by the density of sample points. The more the number of samples is , the more accuracy we obtained. Another important factor influencing the accuracy of calibration is global SURF matching. The number and distribution of matched points would influence the result of calibration. In order to enlarge the number of matched feature points and make its distribution more even, epipolar geometry constraint is adopted.

To estimate accuracy of the proposed method, we randomly select a serial of points in static camera, for every test point, run the calibration programme, calculate the distance between the center of dynamic camera image and its corresponding point in dynamic camera. The result are shown in Fig.4.

After the automatic calibration, we let a people appear in

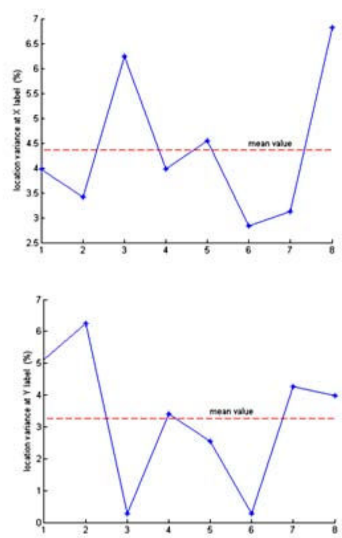

Figure 4. Results on master-slave system

the surveillance system and observe the result. Fig.5 shows some results of the surveillance system.

Figure 5. Results on test points

III. Conclusion

In this paper, we have proposed a novel calibration method in master-slave camera surveillance system based on mosaic image. the most distinct merit of this method is completely automatic compared with other solutions,and the results shown in this paper validate both availability and accuracy of the method.

However, there are still many researches and improvements to be explored. As for mosaic image plays a crucial role in the method, a deeper research in mosaic image could be found in [5]. For creating mosaic image in an area contains many moving object, a similar solution is also available. In[7], a method for composing mosaic image under such environment is proposed. When facing a larger environment, a circular panoramic mosaic image has to be adopted. In such situation, although approach of creating a circular panoramic mosaic image[5] is different to the approach depicted in this paper, the main principle is the same.

Acknowledgment

The work was supported in part under Natural Science Foundation of China grants No.60702044, 60625103, 60802020 and National Basic Research Program of China (2010CB731401 and 2010CB731406).

References

[1] A. Hampapur, L. Brown, J. Connell, S. Pankanti, A. W. Senior, Y. L. Tian, L. Brown, and R. Bolle. Face cataloger: Multi-scale imaging for relating identity to location.I wrote. *In: IEEE conference on Advanced Video and signal Based Surveillance*, (7):13–20, July 2003.

[2] X. Zhou, R. Collins, T. Kanade, and P. Metes. A master-slave system to acquire biometric imagery of humans at distance. *In: ACM Proc. Workshop on Video Surveillance*, 113-120, November 2003.

[3] J. Badri, C. Tilmant, J. Lavest, Q. Pham, P. Sayd. Camera-to-Camera Mapping for Hybrid Pan-Tilt-Zoom Sensors Calibration *In: 15th Scandinavian Conference, SVIA*, 132-141, July 2007.

[4] A. Bimbo, F. Dini, A. Grifoni, F. Pernici. Exploiting Single View Geometry in Pan-Tilt-Zoom Camera Networks. *In:IEEE Workshop on Multi-camera and Multi-model Sensor Fusion: Algorithms and Applications(M2SFA2)*, October 2008.

[5] R.Szeliski. Video Mosaics for Virtual Enviroments. *In: IEEE Computer Graphic and Application*, (16):22-30, Mar 1996.

[6] H.Bay, T. Tuytelaars, L. Gool. SURF: Speeded Up Robust Features. *In: IEEE European Conference in Computer Vision*, 404-417, May 2006.

[7] M. Irani, P. Anandan, S. Hsu. Mosaic Based Representations of Video Sequences and Their Applications. *In: IEEE International Conference in Computer Vision*, 1995.