High Accuracy Tracking with an Active Pan-Tilt-Zoom Camera

W. Fu, L. Gao

School of Electronics and Information Engineering
Xi'an Technological University
Xi'an, China

***Abstract*—Traditional PTZ tracking system focus on tracking algorithm, but PTZ camera control is not taken seriously and the control method has the large deviations. The algorithm for PTZ camera control is defined according to the target position which is achieved by the tracking algorithm in the image, calculates the rotation angles of the two motors to control the PTZ camera. As a result, the target appears in the centre of the image. This paper proposes a model-based algorithm for PTZ camera control, taking into account the camera distortion and the deviation is within 5 pixels. Experimental results show that the algorithm can be effectively applied to PTZ tracking system.**

Keywords-PTZ camera control; tracking algorithm; camera calibration

I. INTRODUCTION

PTZ(Pan, Tilt and Zoom) tracking mainly means that the PTZ cameras is under the automatic control for tracking of moving objects of visual guide to ensure tracking target always appearing in the center of the lens in the range of the camera monitoring scene. In the video surveillance, the distance of the target from the camera is generally 10 meters or more. If the object is a pedestrian, the characteristics of the pedestrian cannot be seen clearly since the image is low quality. In this case, we can manually zoom in the camera to obtain the high-resolution images of pedestrians. However, the method is only effective when the pedestrian is not moving. If the pedestrian is moving, it is easy to get out of the view of camera's field. It is very difficult to manual control the PTZ camera to make the pedestrians appear in the center of view. But PTZ tracking can make ensure that the moving target appears in the center of view. After suitable optical zoom in, the target will not be out of the scope of view and then we can obtain the high-resolution images of pedestrians.

In the PTZ tracking system, the most important is the tracking algorithm and the control algorithm. Tracking algorithm detects and tracks the target in the image. The control algorithm controls the rotation angle of the motor according to the target position in the image, and then to make the target appear in the center of the image. Chang et al [1]. Suggested that the center of the image can be divided into eight directions, namely east, west, south, north, southeast, southwest, northeast and northwest. If the target is detected at the center of the southeast, then make the control to let motor go to southwest. Xiang et al [2]. proposed closed-loop control method basing on the concept of feedback.

Owing to the larger deviation of the traditional algorithm, this paper directly establishes the relationship between the target position and the motor control angle for achieving a high accuracy based on pinhole camera model and camera distortion model.

II. PTZ TRACKING SYSTEM OVERVIEW

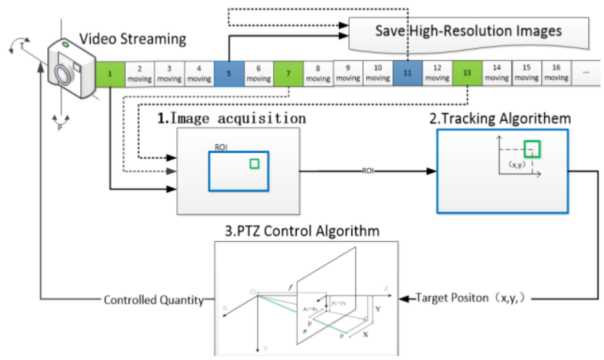

FIGURE I. THE STRUCTURE OF PTZ TRACKING SYSTEM.

As is shown in Figure 1, the whole PTZ tracking system consists of three parts, the image acquisition part, the tracking algorithm part and the PTZ control algorithm part. The system firstly obtains video streaming from the PTZ camera, and then uses the tracking algorithm to obtain the position of the targets. Then the PTZ camera control algorithm calculates the rotation angle and gives the orders to the PTZ camera. Drive the PTZ camera moving and make the target appear in the center of the image.

Image acquisition part mainly acquires images from the video stream as an input of tracking algorithm. In order to meet the requirements of the tracking algorithm, this part also consists of image feature enhancement and pre-image processing algorithm. It is noted that the poor image quality and the image blurring problems because of the moving of camera. So during the video capture, we capture the image interval, the image is captured at a frame 1, frame 7.

There are many different kinds of tracking algorithm. There is not a tracking algorithm which is popular now can perfectly solve the pose variation, illumination, occlusion and blur in tracking. And specific applications, the tracking algorithm is strict with real-time capability. This paper used compressive tracking algorithm proposed by Zhang Kaihua et al [3]. The compressive tracking algorithm runs in real-time and performs favorably against state-of-the-art algorithms on challenging sequences in terms of efficiency, accuracy and robustness. Our experiments show that the speed can reach about 50 frames per second under the 720P resolution and the algorithm can meet requirements of the PTZ tracking system. So this paper focus on PTZ control algorithm.

III. THEORY

PTZ control algorithm means that how to control the camera pan and tilt to make the tracking target always has been the central position of the image. There are generally two kinds of thoughts to solve the problem of control. First， we can stresses in precise mathematical mode .In other words, we can get accurate mathematical model of system through inputs and outputs based on a certain theory. Along with the simple controller, we can achieve very good effect. Second, we can design the controller of excellent performance. However, generally it is difficult to get the precise mathematical model, so we always choose the second solution to solve the control problem, but is not better than the first solution in control precision.

Specific to the PTZ control, also have the above two kinds of thought. We can use classical control algorithms such as the PID. We can detect error (the distance between detected positon and the center of the image) at every frame then put it into PID algorithm and get outputs to control camera. The PID has three parameters, proportional coefficient, the integral coefficient and the differential coefficient and this three parameters can be adjusted based on experiments. When selecting the appropriate parameters, the system will be in the steady-state after n frames within the permissible range of error. But the best control algorithm cannot reach steady state in the k+1 time based on the position of k time. If we want to do this, we must select the first method and know the exact mathematical model of the system. At the same time, the system has no integral parts and great noise. The tracking system of this paper also happens to meet these conditions, so it can be done in one step and make target appears in the center position of the image.

In the 3.1 section, the pinhole camera model and lens distortion model will be given a brief description. This part is the theoretical basis to get accurate mathematical model. In 3.2 sections, a model-based algorithm for PTZ camera control was described in detail.

A. *The Pinhole Camera Model and the Lens Distortion Model*

o-xyz is defined as the camera coordinate where *o* is the optical center of the camera. *o-uv* is defined as camera image plane coordinates in pixels units, where origin point is the upper left corner. *o-xy* is defined as the physical image plane coordinate in millimeters unis where origin point is defined the intersection of the optical axis and the image plane. It is also called the principal point of the image in the *o-uv* and coordinates is $[u0, y0]^T$. As is show in figure 2, a point $P=[X, Y, Z]^T$ in the camera coordinate is projected to a point $\overline{P}=[xc,yc,1]^T$ in the camera image plane coordinates. According to the pinhole camera model, the relationship of projection is

$$\lambda \begin{bmatrix} x_c \\ y_c \\ 1 \end{bmatrix} = \begin{bmatrix} f_x & 0 & u_0 \\ 0 & f_y & v_0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \end{bmatrix} \tag{1}$$

where *fx* and *fy* denote focal length (pixels) of the camera. Since the optical system of the camera has some flaws in the process of machining and assembly, when a point is projected onto the image plane, there is an offset between the actual point and the ideal point. In this paper, we only consider the radial distortion and eccentric distortion [4]. The normalized coordinates of *P* is $Pn=[X/Y, Y/Z]^T=[x, y]^T$. Let r represents the distance from point P to the principle point. We can get offset

$$\begin{aligned} \varepsilon_x &= x(k_1 r^2 + k_2 r^4 + k_3 r^6) + 2p_1 xy + p_2(r^2 + 2x^2) \\ \varepsilon_y &= y(k_1 r^2 + k_2 r^4 + k_3 r^6) + p_1(r^2 + 2y^2) + 2p_2 xy \end{aligned} \tag{2}$$

In the eqn(2), *k1*, *k2* and *k3* is the radial distortion coefficients, *p1*and *p2* is the eccentric aberration coefficients. Let $P_d=[x_d, y_d]^T$ denotes the idea positon. It can be expressed as

$$x_d = x + \varepsilon_x, \quad y_d = y + \varepsilon_y \tag{3}$$

There are 4 intrinsic parameters and 5 distortion coefficients above the formula.we can get 9 parameters based on zhang[5].

B. *A Model-Based Algorithm for PTZ Camera Control*

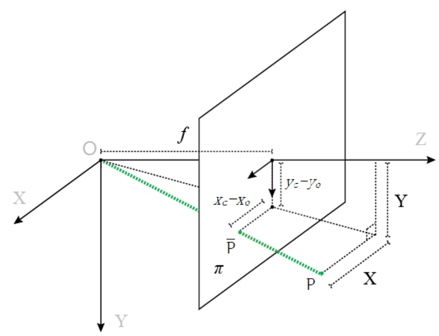

FIGURE II. THE DIAGRAM OF PTZ CONTROL ALGORITHM BASED ON THE MODEL.

It is necessary to find the mathematical relationship between the control amount and the amount of error and to build mathematical model to achieve the steady state at time k based on time k-1 the position of the target. A model-based algorithm for PTZ camera control proposed in Figure 2. $[x_0, y_0]^T$ is the point that expected position of the target appearing. We can make *o-xyx* passes through expected point by transform. If we rotate *o-xyx* by specific angles around X-axes and Y-axes respectively and let the Z-axis passes through the point *P*, the point *P* can be imaged the center position of the CCD sensor. According to the basic principles of geometry, if the rotation angle follows the eqn(4), the point P can appear at the center position of the image.

$$\Delta p = \arctan\left(\frac{(x_c - x_0)}{f_x}\right), \quad \Delta t = \arctan\left(\frac{(y_c - y_0)}{f_y}\right) \tag{4}$$

In the eqn(4), in order to calculation of the correct point (x_c, y_c), We need five distortion parameters. Among them, f_x and f_y are camera intrinsic parameters. We can get them from 3.1 sections. Since the PTZ camera is the zoom camera, we need to find a function of Z and focal length. We can use the least-squares fitting method to obtain the relationship between Z and the focal length

$$\begin{aligned} f_x &= a_{x0} + a_{x1}Z + a_{x2}Z^2 + \ldots + a_{xn}Z^n \\ f_y &= a_{y0} + a_{y1}Z + a_{y2}Z^2 + \ldots + a_{yn}Z^n \end{aligned} \tag{5}$$

We can obtain the corresponding f_x, f_y and Z after camera calibration at different Z values, then use a least squares fit the function of Z and focal length. Five distortion parameters not only can be treated in a similar method but also can use piecewise linear fit method. In this paper, the focal length is fitted with six-order least-squares polynomial. Five distortion parameters are fitted with piecewise linear.

From 3.1, we know that the position of the target that is detected by tracking algorithm cannot be directly substituted into the eqn(4) due to the presence of distortion. If you want to get high accuracy, the distortion correction is needed. Let $[u,v]^T$ is the position of the target in *o-uv*, provided by tracking algorithm. $P=[X,Y,Z]^T$ is defined as the position of target in *o-xyz*, the normalized coordinates is defined as $P_n=[X/Y,Y/Z]^T=[x,y]^T$. We can get

$$x=\frac{(\mathrm{u}-u_0)}{f_x},\quad y=\frac{(\mathrm{v}-v_0)}{f_y} \tag{6}$$

If the eqn(6) is substituted into eqn(3), we can get the ideal projection point in *o-xy*. Let $P_d=[x_d,y_d]^T$ is the ideal projection point, so the correct point in *o-uv* is

$$x_c=x_d f_x+u_0,\quad y_c=y_d f_y+v_0 \tag{7}$$

If the eqn(7) is substituted into eqn(4), the rotation angle after distortion correction can be got.

IV. EXPERIMENTS

As is shown in Figure 3, the experimental platform is based on Hikvision DS-2DF5286 Camera. Pan part can rotate for degree 360.Tilt part can rotate from degree -5 to degree 90. Its control interface uses the RS485 interface and video capture uses RJ45 cable interface.

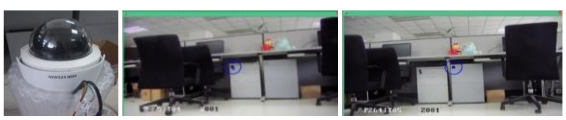

FIGURE III. THE LEFT IMAGE IS THE HIKVISION DS-2DF5286 PTZ CAMERA.

Other two images are the screenshot of the running test application. When you click any point on the screen, the point will be able to move to the canter of the screen. The middle image is the result of clicking the keyhole of the left cabinet. The right image is the result of clicking the keyhole of the right cabinet.

The entire program is implemented in the Windows operating system, using the C ++ programming language. The software environment is the Visual Studio 2013 and Hikvision development kits. Experiments on a PC equipped with Windows 8.1, where the processor is Intel Xeon E5-1603, clocks at 2.8GHz and memory is 8GB.

In the figure 3, there is a dot of radius of 5 pixels and a circle of radius of 25 pixels. The two marks can be used to estimates approximately error of PTZ control. The parameters of current PTZ values are displayed on the lower half of the screen. Figure 3 selects two cases from the experimental results. the left PTZ parameters are P275 and T04 and the others are P264 and T05.

TABLE I. THE RESULTS OF PTZ CONTROL ALGORITHM TEST

	0-5 pixels	5-25 pixels	>=25 pixels
Count	267	37	4
Percent	86.7%	12%	1.3%

We randomly click on the screen at 308 times and table 1 shows the results. Experimental results show that the algorithm proposed can be effectively applied to PTZ tracking system and the error is within 5 pixels in most cases.

FIGURE IV. THE RESULTS OF PTZ TRACKING TEST. THE TIMESTAMP AND PTZ PARAMETERS CAN BE SEEN ON THE TOP AND BOTTOM OF THE APPLICATION

As is shown in figure 4, this paper implements a simple PTZ tracking system based on the proposed algorithm and compressive tracking algorithm. The experiment shows the final results of the human face PTZ tracking system. From the change of PTZ parameters and the positon of face, it can be seen that the calculated rotation angle according to the proposed PTZ control algorithm can make the face keep on the center of the screen.

V. CONCLUSION

The most important is the control algorithm and the tracking algorithm in the PTZ tracking system. The tracking algorithm finds the target position in the image and the control algorithm controls the rotation angle of the PTZ camera based on the target position and finally let the target appear in the center of the image. This paper focuses on the PTZ control algorithms. We propose a model-based algorithm for PTZ camera control using a pinhole camera model and camera distortion model. The experiments have shown that the error of the control algorithm is within 5 pixels. But because the image acquisition is not continuous, resulting in the loss of the many information of the relationship between frame and frame, so the compressive tracking algorithm is not as better as continuous acquisition, and then resulting in poor PTZ tracking effect(such as fast-moving targets). Therefore, the optimal PTZ tracking system should be a static camera for detecting and a PTZ camera for tracking targets that use the target positon from static camera and solve the problem perfectly.

REFERENCES

[1] Chang, Faliang, et al. "PTZ camera target tracking in large complex scenes." Intelligent Control and Automation (WCICA), 2010 8th World Congress on. IEEE, 2010.

[2] Xiang, Guishan. "Real-time follow-up tracking fast moving object with an active camera." Image and Signal Processing, 2009. CISP'09. 2nd International Congress on. IEEE, 2009.

[3] Zhang Kaihua, Lei Zhang,and Ming-Hsuan Yang."Real-time compressive tracking." Computer Vision–ECCV 2012. Springer Berlin Heidelberg, 2012. 864-877.

[4] Fryer, John G., and Duane C. Brown. "Lens distortion for close-range photogrammetry." Photogrammetric engineering and remote sensing 52.1 (1986): 51-58.

[5] Zhang, Zhengyou. "A flexible new technique for camera calibration." Pattern Analysis and Machine Intelligence, IEEE Transactions on 22.11 (2000): 1330-1334.