Fully Automatic PTZ Camera Calibration Method

Anton Obukhov, Konstantin Strelnikov, Dmitriy Vatolin
Department of Computational Mathematics and Cybernetics
Moscow State University, Moscow, Russia
{aobukhov, kstrelnikov, dmitriy}@graphics.cs.msu.ru

Abstract

In this paper we present a novel approach to fully automatic pan-tilt-zoom (PTZ) camera calibration. A calibration system based on the proposed method is easy to setup, is easily scalable and does not require any human participation during operation. We focus on calibration of extrinsic parameters while assuming that all intrinsic parameters are known *a priori*. Our calibration technique uses a set of measurements that are each represented by the correspondence between the Cartesian world coordinates and the camera's internal *pan-tilt* coordinates for a given point. Use of the proposed "direct measurement" approach makes the calibration process separable, which means that the camera position and rotation can be calculated independently. Although the internal coordinates are easily accessible in most contemporary cameras, the world coordinates of specific points must be obtained using a visual marker detection system. The output of the auto-calibration system for each camera is the extrinsic camera parameters that best fit the input measurements. The proposed method is designed for operation in areas under various lighting conditions or with complex wall topology and it exhibits adequate results in practice.

Keywords: *PTZ, pan-tilt-zoom, multiple camera calibration, auto-calibration, automation, geometric approach, nonlinear search.*

1. INTRODUCTION

Generally, calibration of a measurement device involves establishing a relationship between the device output and the units of measurement. The problem of camera calibration arose after scientists recognized cameras to be powerful measurement devices. Camera calibration processes are designed to produce a parameterized camera model by using several measurements that include frame information and, optionally, *a priori* data.

Before development of any camera calibration system, it is important to define the minimal sufficient subset of calibration parameters, because algorithms for estimation of complex model parameters are of high computational complexity. Traditional classification of camera calibration parameters includes external (position, rotation) and internal (focal length, radial distortion, etc.) parameters [1]. The last class contains both parameters that represent some physical properties or integral features of a device (e.g. focal length) and parameters that describe shortcomings of a device (e.g. radial distortion). The increase in quality of camera manufacturing processes, however, has helped eliminate some problematic intrinsic camera attributes such as radial distortion and mismatch between the rotation center and the focal point. Throughout this paper we assume that the subset mentioned above only includes extrinsic calibration parameters and that the intrinsic parameters are either known *a priori* or are not required. As is shown in the following discussion, this assumption does not introduce any additional errors into the calibration process.

This paper is organized as follows. Section 2 of this paper presents a review of related work. Section 3 provides some notation and mathematical background for the camera calibration problem. In Section 4 we describe the proposed method. Section 4.1 presents the outline of the auto-calibration method, followed by descriptions of position estimation in Section 4.2 and of rotation estimation in Section 4.3. Implementation notes, experimental results and a method evaluation are given in Section 5. Finally, conclusions are drawn in Section 6.

2. RELATED WORK

The problem of automatic calibration of a static camera was addressed in the early work of R. Tsai [1]. Tsai proposed various calibration techniques, such as calibration using a monoview coplanar set of points and multiple viewing position calibration. He suggested performing the calibration in several steps, utilizing calibration separability, and claimed that analytic methods of camera parameters evaluation, such as linear least squares, are preferable if no intrinsic parameters need to be calibrated. However, performing static recalibration of a PTZ camera after each rotation is not a desirable solution.

One group of methods for PTZ camera calibration includes those that use calibration objects such as visual markers or sets of coplanar points with known coordinates [2-4]. The calibration process can be described as follows: given the world coordinates and corresponding frame coordinates of a set of points, the calibration is performed by minimizing the error of mismatch between the projection of world coordinates and corresponding frame coordinates in accordance with a set of calibration parameters. Some of the techniques from this group of methods use unusual objects to obtain additional information from the environment. M. Agrawal *et al.* [3] used spheres as calibration objects to simplify the object detection process and to simultaneously calculate the frame scale factor along the horizontal direction (the uncertainty image scale factor [1]). Agrawal stated that every camera sees sphere contours as ellipses, and, thus, the uncertainty image scale factor can be calculated as the ratio of the radii of the ellipse. I.-H. Chen *et al.* [4] described a system that uses sheets of A4 paper lying on the horizontal plane. This approach yields information about the right angles of the paper, thus simplifying the calibration model. Common drawbacks of these clever methods, however, are the inconveniences that arise during the preparation stage or during the calibration.

A number of other approaches (self-calibration methods) [5-7] are based on inter-frame homography or on point correspondence between several frames. Most of these methods handle the nonlinearity of the calibration parameters calculation by making assumptions about scene factors and by taking into account *a priori* knowledge of intrinsic camera parameters. The methods in the first group mentioned above seem to be unstable in conditions with low lighting, and others are designed to perform mainly just intrinsic calibration. As a result, these methods are not of interest here. An additional difficulty associated with this group of methods is that they only operate in static environments, because any captured motion yields errors in homography and calibration results. This difficulty is rather valuable limitation of use cases.

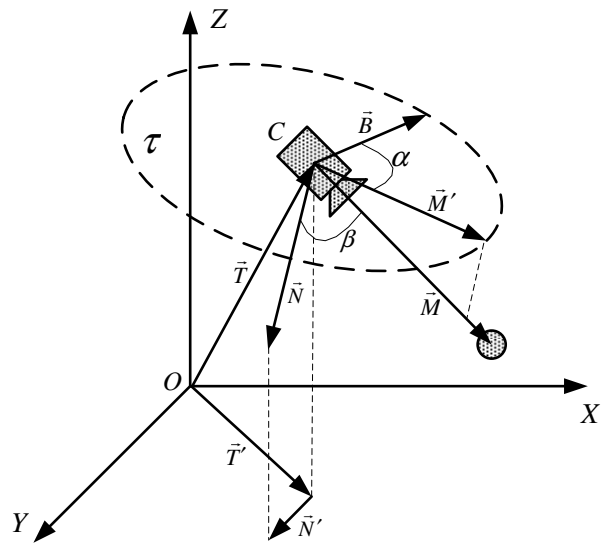

Figure 1: PTZ camera position and rotation in world coordinates

3. BACKGROUND AND NOTATION

Assume that a single PTZ camera is fixed at some location within a room. By convention, this location is described by the translation vector $\vec{T} = \overrightarrow{OC}$ in the Cartesian coordinates system using axes *X, Y and Z* (Fig. 1). As can be seen in the figure, we use a right-handed system, but this is not a necessary condition.

Every PTZ camera has internal polar coordinates that span a half-space that consists of all points *P* on one side of plane τ, such that $\left(\overrightarrow{CP}, \vec{N}\right) \geq 0$ and $\vec{N} \perp \tau$. $\vec{N}$ is the representation in world coordinates of the vector that specifies the origin of the camera's coordinate system. The two camera coordinates are *pan* and *tilt*. To specify the direction of the vector $\vec{M} = (\alpha, \beta)$ in the figure, *pan* is measured in the plane τ in a clockwise direction from vector $\vec{B}$, and *tilt* is measured from vector $\vec{N}$ in the plane defined by vectors $\vec{N}$ and $\vec{M'}$ ($\vec{M'}$ is the normalized projection of $\vec{M}$ onto τ). Since the camera has only two degrees of freedom, each 3D point in the world (*XYZ*) coordinates uniquely maps onto a 2D point in the camera's internal *pan-tilt* coordinates. The inverse coordinate transformation produces a unit vector in the direction defined by the 2D point in *pan-tilt* coordinates. These transformations are described in greater detail later.

The camera's extrinsic parameters for this notation are the vectors $\vec{T}$, $\vec{N}$ and $\vec{B}$. Every camera has six independent degrees of freedom: three for the translation vector and three for rotation [1]. The rotation can be presented using Euler's angles (φ, θ and ψ), as a series of three rotations around a set of known axes. This allows vectors $\vec{N}$ and $\vec{B}$ to be represented in terms of the scalar quantities φ, θ and ψ.

Although the translation calibration is clearly an essential means of camera calibration, rotation calibration is also critical, since it is practically impossible to install a camera such that $\vec{N} \parallel \vec{Z}$. Even if the camera was installed in such a precise manner (resulting in $\vec{N} \parallel \vec{Z}$, $\vec{N'} = \vec{0}$ and $\tau \parallel XY$, where $\vec{N'}$ is the projection of $\vec{N}$ onto the *X-Y* plane), it would still be necessary to either calibrate the rotation with one degree of freedom or to continue to tune the installation until $\vec{B} \parallel \vec{X}$. Here we consider the common rotation case with three degrees of freedom.

$$R_Y^{\omega} = \begin{bmatrix} \cos(\omega) & 0 & \sin(\omega) \\ 0 & 1 & 0 \\ -\sin(\omega) & 0 & \cos(\omega) \end{bmatrix}, \; R_Z^{\omega} = \begin{bmatrix} \cos(\omega) & -\sin(\omega) & 0 \\ \sin(\omega) & \cos(\omega) & 0 \\ 0 & 0 & 1 \end{bmatrix} \tag{1}$$

There are 12 possible ways to perform space rotation using Euler angles. Here we propose a series of rotations around the axes in the order *Z-Y-Z*. The rotations by angle ω around the *Y* and *Z* axes are given by (1).

Rotation of the camera using the matrix *R* is equivalent to the rotation of the scene using the inverse matrix. We denote the scene rotation matrix as the *rotation matrix*, which is shown in (2).

$$R_{Z,Y,Z}^{\varphi,\theta,\psi} = R_Z^{\psi} \cdot R_Y^{\theta} \cdot R_Z^{\varphi}. \tag{2}$$

The transformation of the world coordinates of some point P^W into the camera's Cartesian coordinates P^C, which are bound to the camera's internal *pan-tilt* coordinates, is the following:

$$P^C = R_{Z,Y,Z}^{\varphi,\theta,\psi} \cdot \left(P^W - C\right). \tag{3}$$

The transformation between the internal *pan-tilt* coordinates (α, β) of the camera and the Cartesian coordinates of point *P* is given in (4). The inverse transformation is given in (5).

$$\alpha = \arctan\left(\frac{P_Y}{P_X}\right), \; \beta = \arctan\left(\frac{P_Z}{\sqrt{P_X^2 + P_Y^2}}\right) + \frac{\pi}{2} \tag{4}$$

$$\begin{bmatrix} P_X \\ P_Y \\ P_Z \end{bmatrix} = \begin{bmatrix} \cos(\alpha)\cos\left(\beta - \frac{\pi}{2}\right) \\ \sin(\alpha)\cos\left(\beta - \frac{\pi}{2}\right) \\ \sin\left(\beta - \frac{\pi}{2}\right) \end{bmatrix} \tag{5}$$

4. PROPOSED METHOD DESCRIPTION

The proposed calibration method is similar to those of the first group of approaches described in Section 2. The algorithm receives a set of point correspondences and then determines the calibration parameters by minimizing the error of fitting input to model in accordance with the set of calibration parameters.

4.1 Outline of the auto-calibration method

To avoid getting stuck on errors that can be introduced by unknown (or uncalibrated) intrinsic parameters in case of their usage, it is preferable to use "direct measurements", which are data (3D points in world coordinates) collected by humans, as well as meta-information from the camera controller (using the camera's internal *pan-tilt* coordinates). Thus, each direct measurement contains five known values $\left(M^x, M^y, M^z, \varphi, \theta\right)$ which can be trusted as no others – this is the first key feature. The point $M = \left(M^x, M^y, M^z\right)$ must lie on the optical axis of the camera, this is called the "direct measurement requirement".

We use a nonlinear search algorithm to extract calibration parameters. As stated in [1], the dimensionality of the space in which the nonlinear search is applied should not increase beyond

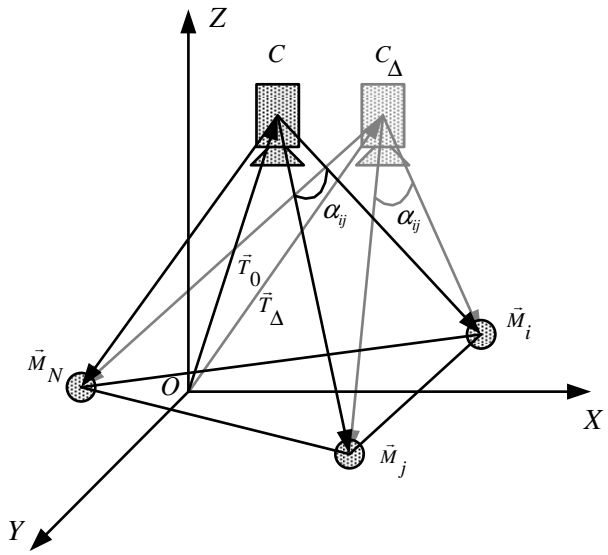

Figure 2: Position calibration scheme

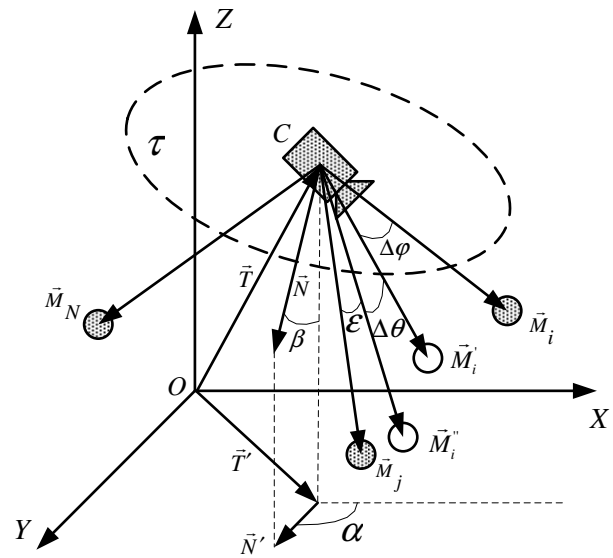

Figure 3: Rotation calibration scheme

five. At this point we come to the second key feature of direct measurements: they permit separation of the calibration process into several sequential stages:

1. Position estimation (transform vector $\vec{T}$);
2. Rotation estimation (rotation matrix *R*);
    a. Estimation of vector $\vec{N}$;
    b. Calculation of vector $\vec{B}$;
    c. Calculation of φ, θ and ψ.

The nonlinear search is used only in stages 1 and 2a and has a maximum dimensionality of three. The following algorithm is used for estimating both calibration parameters:

1. Set initial parameter value;
2. Pass the following subroutine into the numerical minimization algorithm:
    a. For each pair of unequal direct measurements (*i*,*j*), calculate the error functional (described below) for the parameter value in the current iteration and for direct measurements *i* and *j*;
    b. Calculate the sum of all error functional values obtained in the previous step;
3. Output the argument of the last iteration.

4.2 Camera position estimation

Given the set of direct measurements of length *N* (*N*>2), we perform the position calibration using the algorithm above. For the error functional, assume that a camera is placed at point *C* and that the markers from the ith and j^{th} measurements have coordinates M_i and M_j respectively (Fig. 2).

The points C, M_i, M_j are vertices of a triangle. The M_iM_j side of the triangle is known, and the lengths of the other two sides depend on the location of point *C*. The side M_iM_j can be calculated by means of the law of cosines using sides CM_i and CM_j and the angle α_{ij} between them. The angle α_{ij} can be calculated from internal camera coordinates using the following expression:

$$\cos\alpha_{ij} = (P_i, P_j), \tag{6}$$

where P_i and P_j are the 3D unit vectors that specify the camera's internal *pan-tilt* coordinates in the associated Cartesian system, as with (5).

Thus, the error functional for the pair of direct measurements can be written as the absolute difference between the actual length of M_iM_j and the calculated length. The final form of the functional is the following:

$$\begin{aligned} Err_{ij} &= \left|(M_iM_j)^2 - (CM_i)^2 - (CM_j)^2 + 2CM_iCM_j\cos\alpha_{ij}\right|, \\ Err &= \sum_{i=1}^{N}\sum_{j=i+1}^{N} Err_{ij}. \end{aligned} \tag{7}$$

As can be seen in Fig. 2, we are searching for point *C* such that we minimize the error of the wireframe model of our scene, thus minimizing the camera position offset $\|C_\Delta - C\|$.

During testing of the prototype for our calibration method, it was found that setting the initial camera position to zero may produce erroneous results in cases where either the input data is imprecise or the number of measurements is small. It was also discovered that setting the initial position to the average of input 3D points yields better results in problematic cases.

4.3 Camera rotation estimation

The process of rotation calibration requires both a set of *N* direct measurements (*N>1*) and the exact camera position in world coordinates. The same measurement set that was used in position calibration can also be used in rotation calibration. As stated at the beginning of Section 4, rotation estimation can be performed separately from position estimation.

The vector $\vec{N}$ (Fig. 3), which defines the camera's origin direction, is unity and can thus be represented by the two scalars (α and β) using (4) and (5). The main concept of the error functional is that the error of the rotation series that transform M_i into M_j exhibits a minimum value for the correct vector $\vec{N}$.

Speaking more precisely, we calculate the angle between $\overrightarrow{M_i''}$ and $\overrightarrow{M_j}$, where $\overrightarrow{M_i''}$ is the vector $\overrightarrow{M_i}$ rotated first around $\vec{N}$ by the *pan* difference angle (resulting in $\overrightarrow{M_i'}$) and then rotated in the plane containing vectors $\vec{N}$ and $\overrightarrow{M_i'}$ by the *tilt* difference angle (resulting in $\overrightarrow{M_i''}$). The related equations are given below.

The matrix for the rotation by angle ω around the general normalized vector $\vec{a} = [x, y, z]^T$ is given by (8).

$$R_{\vec{a}}^{\omega} = \begin{bmatrix} (1-c)x^2 + (c) & (1-c)xy - (s)z & (1-c)xz + (s)y \\ (1-c)xy + (s)z & (1-c)y^2 + (c) & (1-c)yz - (s)x \\ (1-c)xz - (s)y & (1-c)yz + (s)x & (1-c)z^2 + (c) \end{bmatrix}, \quad (8)$$

$$c = \cos(\omega),\ s = \sin(\omega).$$

The rotation in the plane defined by the unique nonzero vectors $\overrightarrow{V_1}$ and $\overrightarrow{V_2}$ can be written as follows:

$$R_{V_1,V_2}^{\omega} = R_{\frac{[V_1,V_2]}{\|[V_1,V_2]\|}}^{\omega} \quad (9)$$

The *pan* and *tilt* difference angles mentioned above are given by the following:

$$\Delta\varphi = \varphi_j - \varphi_i,\ \Delta\theta = \theta_j - \theta_i. \quad (10)$$

The final error functional formula can be written using (8) through (10).

$$\overrightarrow{M_i'} = R_{\vec{N}}^{\Delta\varphi} \cdot \left(\overrightarrow{M_i} - \vec{T}\right), \qquad \overrightarrow{M_i''} = R_{\vec{N},\overrightarrow{M_i'}}^{\Delta\theta} \cdot \overrightarrow{M_i'}$$

$$Err_{ij} = \arccos\left(\frac{\left(\overrightarrow{M_i''}, \overrightarrow{M_j}\right)}{\left\|\overrightarrow{M_i''}\right\| \cdot \left\|\overrightarrow{M_j}\right\|}\right) \quad (11)$$

$$Err = \sum_{i=1}^{N} \sum_{j=i+1}^{N} Err_{ij}$$

According to the experimental results, the initial values for α and β can be chosen arbitrarily as soon as the error functional has a single global minimum in the ranges $\alpha \in [0, 2\pi]$ and $\beta \in [0, \pi/2]$.

The concept behind the calculation of $\vec{B}$ is that for each marker point M_i, we calculate $\overrightarrow{B_i}$ using (12) by projecting $\overrightarrow{M_i}$ onto the plane τ and rotating around $\vec{N}$ by the angle $-\varphi_i$. The median of $\overrightarrow{B_i}$ is assumed to be the estimation of $\vec{B}$.

$$\overrightarrow{M_i^P} = \overrightarrow{M_i} - \vec{T} - \vec{N} \cdot \left(\vec{N}, \overrightarrow{M_i} - \vec{T}\right),$$

$$\overrightarrow{B_i} = R_{\vec{N}}^{-\varphi_i} \cdot \left(\frac{\overrightarrow{M_i^P}}{\left\|\overrightarrow{M_i^P}\right\|}\right) \quad (12)$$

Finally, we obtain the Euler angles by observing how the vector $\overrightarrow{C_0} = [1,0,0]^T$ is transformed to $\vec{B}$ (Fig. 4). The intersection of the τ plane with the *X-Y* plane, shifted to the camera level, gives us the vector $\overrightarrow{C_1}$, which is the one we rotate round when tilting camera by θ. This means that:

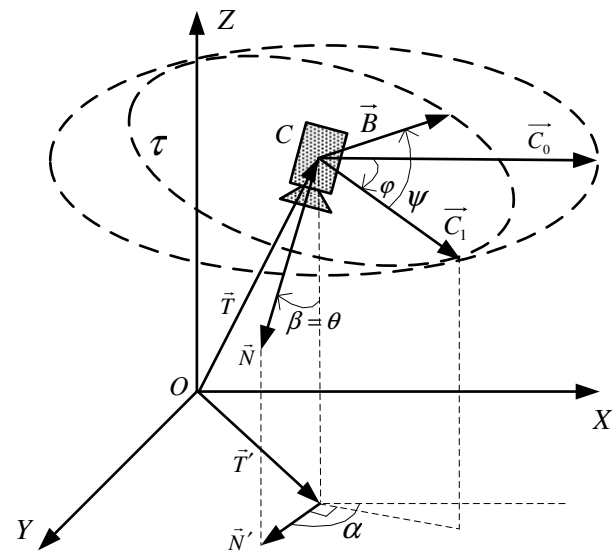

Figure 4: Scheme for extraction of Euler angles

$$\varphi = \alpha - \frac{\pi}{2},\ \theta = \beta,\ \psi = \angle\left(\overrightarrow{C_1}, \vec{B}\right), \quad (13)$$

where:

$$\overrightarrow{C_1} = R_{OZ}^{\frac{\pi}{2}-\alpha} \cdot \overrightarrow{C_0},$$
$$\angle\left(\overrightarrow{C_1}, \vec{B}\right) = \arccos\left(\overrightarrow{C_1}, \vec{B}\right) \cdot \operatorname{sgn}\left(\left[\overrightarrow{C_1}, \vec{B}\right], \vec{N}\right) \quad (14)$$

Thus, we have proven that the Euler angles for space conversion can be expressed in terms of the vectors $\vec{N}$ and $\vec{B}$, which were estimated earlier.

5. PROPOSED METHOD IMPLEMENTATION AND EVALUATION

To make the system as automatic as possible, we used visual markers that can be identified in a video stream by means of software. Recognition systems can be divided into two classes: those that simply register the existence of a marker in the video frame and those that also output additional information about the marker identity. The second class is more interesting from the perspective of system automation.

The calibration process is as follows.

1. An operator places markers (each having a unique identification number, or ID) in specific locations with respect to the camera's field of view, measures their 3D world coordinates and then inputs the ID numbers and coordinates into the system;
2. The system searches for markers by sequentially rotating the camera through all possible angles and by capturing frames and passing them to a marker detection routine. The scanning is performed with different zoom values. This allows locating both close and distant markers, which is especially useful in extensive environments;
3. After searching for the markers, the system assembles the set of direct measurements by decoding marker

coordinates and then passes the measurement set to the calibration routines.

The calibration process described above is suitable for simultaneous calibration of multiple PTZ cameras located in a single scope (all points of space from which at least three markers are visible and easily recognizable by means of marker detector), because stage 1 is done by the operator once, and other stages can be performed by a set of cameras independently.

There are certain requirements that the marker recognition system should meet: low false positive rate, high detection rate and insensitivity to scaling of marker images. Also, error correction is highly desirable to enable the system to still function correctly even in conditions with low lighting.

All these requirements are met by ARTag, a fiducial marker recognition system [8]. Each marker is represented by an image with a white background containing a black rectangular region (or vice-versa) and 36 bit fields. According to [8], there are 2002 unique markers. This number is much smaller than the 2^{36} possible bit sequences, because each ARTag marker contains redundant parity and recovery bits. This arrangement allows the algorithm to determine if the captured image contains a valid ARTag marker and to recover the ID in cases where some part of marker is overexposed or partially hidden by another object. The author claims that the false positive rate of the system is about 0.0039%, which is rather good result. A sample ARTag marker that acts as the calibration object is shown in Fig. 5.

In practice, the use of direct measurements introduces some difficulties related to the requirement for point M to be lying on the camera optical axis. Usually, the intersection of camera optical axis with the focal plane matches the center of the frame. A visual marker detection system may recognize marker at almost every point of frame. There are two possible ways to produce direct measurement in such a case:

- Rotate the camera in the direction of the marker until the marker projection onto the focal plane converges to the frame center and the direct measurement requirement is met;

Figure 5: ARTag visual marker acting as a calibration object

- Calculate *pan* and *tilt* values (PT_M), corresponding to world coordinates of the detected marker M and marker detection coordinates P using known camera intrinsic parameters μ and ν and internal pan-tilt coordinates PT_0 (Fig. 6) as follows:

$$\overrightarrow{OP_1} = R_{OZ}^{P_0} \cdot R_{OY}^{T_0} \cdot \overrightarrow{OP_0}$$

$$\overrightarrow{OP_0} = \left[\tan\left(\frac{\nu}{2}\right) P_0^Y,\ \tan\left(\frac{\mu}{2}\right) P_0^X,\ -1 \right]^T \tag{15}$$

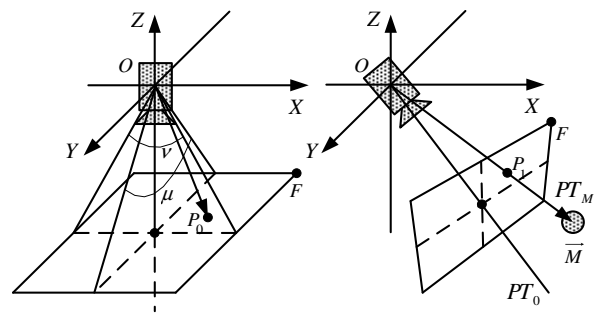

Figure 6: Calculation of *pan* and *tilt* values for any frame point

Here PT_M can be calculated by applying (4) to vector $\overrightarrow{OP_1}$. Point P_0 is the marker projection in Cartesian 2D coordinates with zero at the frame center; point F is the top-left frame corner and has coordinates $[-1,1]^T$.

The minimization routine used in both implementations is the Nelder-Mead simplex search, but the conventional Levenberg-Marquardt or even gradient descent searches are also suitable. Typical plots of error functionals for both position and rotation calibration methods are presented in Fig. 7. In this sample case, the input of calibration is six direct measurements. The real camera position is C=[0.90, 3.55, 2.68]^T, the real N-axis angles are (α,β)=(113°,18°). The ranges of α and β in rotation calibration are: $\alpha \in [0^\circ, 360^\circ]$, $\beta \in [0^\circ, 90^\circ]$. As can be seen from the plots, the functionals have single global minimums at specified ranges, and these minimums match real data.

The result of the calibration process depends on the number of measurements, precision of 3D coordinates specification and maximum distance between any pair of measurements. The last is defined by absolute angular differences between corresponding φ and ψ with respect to wraparound of φ.

The minimum number of measurements for position calibration is three. Geometrically, it is similar to the problem of finding the fourth vertex of the triangular pyramid with three vertices and some vertex angles given. It can be proven, that this problem is unsolvable if less than three vertices and angles are given. As for rotation calibration, two measurements (one for M_i and one for M_j) are clearly enough to perform it as in (11).

The comparisons of diverse approaches to PTZ camera calibration are non-traditional in this problem domain, as far as there is no set of conventional reference techniques. The prototype of the system has been tested for stability. Fig. 8 shows the response of calibration results to the normally distributed error in input 3D marker coordinates. The variance of error varies over the range [0,1] along the x-axis. The offset of the camera position from the real position in meters is shown at the left plot. The norm of the angular error of rotation calibration, in degrees, can be seen at the

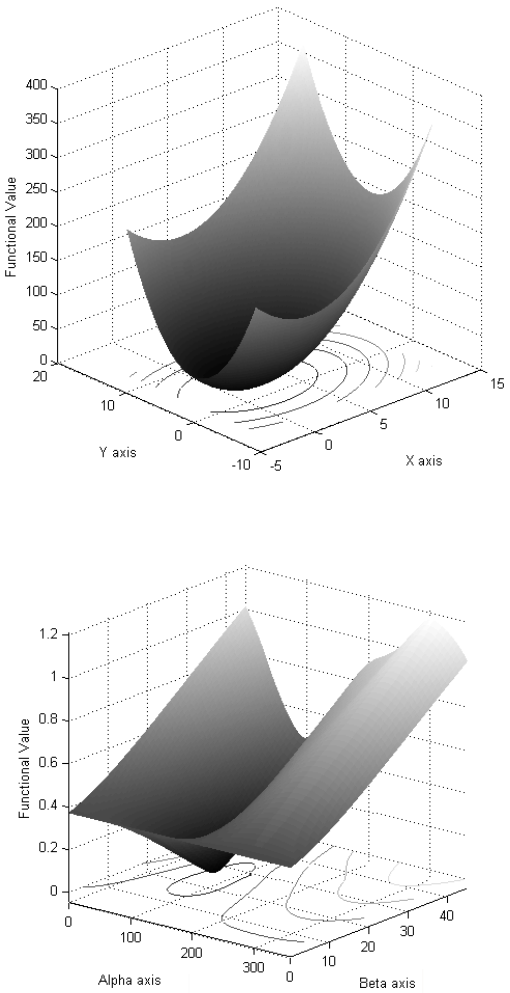

Figure 7: Position functional *Err(X,Y)* for fixed *Z=2.0* (Top), Rotation functional *Err(α,β)* (Bottom)

right plot. Tests on real data have proven that small errors in calibration input yield acceptable precision of the calibration output.

The calibration subsystem was implemented in MATLAB and C++. Both implementations show acceptable speed of operation (1-2 seconds) on a Pentium 4 1.7 GHz computer when calculating all of the parameters using the set of 50 synthetic measurements. This calculation is only the third step of the calibration process.

6. CONCLUSION

In this paper we have proposed a new, fully automatic PTZ camera calibration algorithm. The algorithm performs calibration of extrinsic camera parameters by applying a nonlinear search. The concept of direct measurements was introduced, and it was shown that their use yields calibration separability and total invariance from any intrinsic camera parameters. The implementation of the calibration system included a special visual marker detection and recognition technique that demonstrates good detection rates and very low error rates.

Areas of further work include research and development of combined methods for fully automatic multiple PTZ cameras calibration.

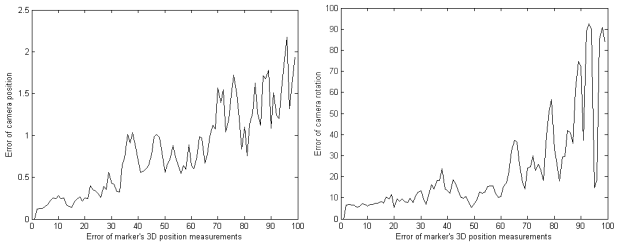

Figure 8: Error response (left – position calibration, right – rotation calibration)

7. REFERENCES

[1] R. Y. Tsai, "A versatile camera calibration technique for high-accuracy 3D machine vision metrology using off-the-shelf TV cameras and lenses", *IEEE Journal of Robotics and Automation*, vol. RA-3, no. 4, Aug. 1987, pp. 323-344.

[2] I. Everts, N. Sebe, G. A. Jones, "Cooperative Object Tracking with Multiple PTZ Cameras", *IEEE Proceedings of the 14th International Conference on Image Analysis and Processing*, 2007, pp. 323-330.

[3] M. Agrawal, L.S. Davis, "Camera calibration using spheres: a semi-definite programming approach", *Proceedings Ninth IEEE International Conference on Computer Vision*, vol. 2, Oct. 2003, pp. 782-789.

[4] I.-H. Chen, S.-J. Wang, "Efficient Vision-Based Calibration for Visual Surveillance Systems with Multiple PTZ Cameras", *IEEE International Conference on Computer Vision Systems*, Jan. 2006, pp. 24-24.

[5] A. Basu, K. Ravi, "Active camera calibration using pan, tilt and roll", *IEEE Transactions on Systems, Man, and Cybernetics*, Part B 27(3), pp. 559-566.

[6] Y. Seo, K. S. Hong, "Theory and practice on the self-calibration of a rotating and zooming camera from two views", *IEE Proceedings Vision Image and Signal Processing*, vol. 148, June 2001, pp. 166-172.

[7] H. Kim, K. S. Hong, "A Practical Self-Calibration Method of Rotating and Zooming Cameras", *15th International Conference on Pattern Recognition*, vol. 1, 2000, pp. 354-357.

[8] Mark Fiala, "ARTag, a fiducial marker system using digital techniques", *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 2005, vol. 2, pp. 590-596.

About the authors

Anton Obukhov graduated with honors from Graphics and Media Laboratory, Department of Computational Mathematics and Cybernetics, Moscow State University in 2008. aobukhov@graphics.cs.msu.ru

Konstantin Strelnikov is a Ph.D. student in Keldysh Institute of Applied Mathematics of Russian Academy of Sciences. kstrelnikov@graphics.cs.msu.ru

Dmitriy Vatolin received his Ph.D. in 2000 from Moscow State University. He is an expert in image, video and data compression, IEEE Member, a senior researcher and the head of the Video Group at Graphics and Media Laboratory, Department of Computational Mathematics and Cybernetics, Moscow State University. dmitriy@graphics.cs.msu.ru