[4] C. R. Deepak, M. Pugazhaandi, S. Paul, M. A. Shajahan, G. Janakiraman, M. A. Atmanand, K. Annamalai, R. Jeyamani, M. Ravindran, E. Schulte, J. Panthel, H. Grebe, and W. Schwarz, "Underwater sand mining system for shallow waters," presented at the 3rd Ocean Mining Symp., Goa, India, Nov. 8–10, 1999.
[5] Z. Janosi and B. Hanamoto, "The analytical determination of drawbar pull as a function of slip for tracked vehicles in deformable soils," presented at the 1st Int. Conf. Soil Mechanics of Soil-Vehicle Systems, Torino, Italy, 1961.
[6] M. A. Atmanand, M. A. Shajahan, C. R. Deepak, R. Jeyamani, M. Ravindran, E. Schulte, J. Panthel, H. Grebe, and W. Schwarz, "Instrumentation for under water crawler for mining in shallow water," in *Proc. Int. Symp. Autonomous Robots and Agents*, Singapore, May 23–26, 2000.
[7] T. Qui and S. C. Jin, "Successive learning track-keeping control (SLTC) algorithm for seafloor vehicle," in *Proc. 9th Int. Offshore and Polar Engineering Conf.*, Brest, France, Jun. 4, 1999, pp. 441–446.
[8] Jong-su Choi, S. Hong, Hyung-Woo Kim, and H. L. Tae, "An experimental study on tractive performance of tracked vehicle on cohesive soft soil," in *Proc. 5th Ocean Mining Symp.*, Tsukuba, Japan, Sep. 15–19, 2003, pp. 139–143.
[9] G. W. Turnage and W. C. Seabergh, "Traficability and stability analysis for bottom crawling work vehicles in the nearshore region," in *Proc. Offshore Technology Conf.*, Houston, TX, pp. 1913–1926.
[10] H. Grebe and J. Panthel, "In-situ soil penetrometer," in *Berichte Aus Dem Zentrum für Meeres- und Klimaforschung (Reihe E: Hydrobiologie und Fischereiwesen, Nr. 11)* (in German), G. Schriever, A. Koschinsky, and H. Bluhm, Eds. Hamburg: Inst. Hydrobiolog. Fischereiwesen, 1996, pp. S. 47–55.
[11] O. Bode, Simulation von raupenfahrwerken tiefsee-böden (in German). Hannover, Germany, 1991.
[12] G. Dörfler, Untersuchungen der fahrwerk-boden-interaktion zur gestaltung von raupenfahrzeugen für die befahrung weicher tiefseeböden (in German). Karlsruhe, Germany, 1995.
[13] J. Y. Wong, *Theory of Ground Vehicles*, 3rd ed. Singapore: Wiley (Asia), 2001, pp. 560–560.
[14] I Rehorn, "Development of a deep-sea full-track vehicle (Caterpillar vehicle) and testing its internal tractional resistance," Ph.D. dissertation, IKS, Univ. Siegen, Siegen, Germany, 1994.
[15] K. Herzog, Entwicklung Einer Antriebs-Schlupf-Regelung Für ein Tiefseefahrzeug, (in German) Univ. Siegen, Siegen, Germany, 2000.

An Efficient Approach for the Calibration of Multiple PTZ Cameras

I-Hsien Chen and Sheng-Jyh Wang

***Abstract*—In this paper, we propose an efficient approach to infer the relative positioning and orientation among multiple pan–tilt–zoom (PTZ) cameras. In this approach, the tilt angle and altitude of each PTZ camera are estimated first based on the observation of some simple objects lying on a horizontal plane. With the estimated tilt angles and altitudes, the calibration of multiple cameras can be easily accomplished by comparing the back-projected world coordinates of some common vectors in 3-D space. This calibration method does not require particular system setup or specific calibration patterns. The sensitivity analysis with respect to parameter fluctuations and measurement errors is also discussed. Experiment results over real images have demonstrated the efficiency and feasibility of this approach.**

***Note to Practitioners*—This paper was motivated by the problem that pan–tilt–zoom (PTZ) cameras may change their poses from time to time to achieve better monitoring results. Whenever there is a change, we need to recalibrate the extrinsic parameters again. In this paper, we demonstrate a new and efficient approach to calibrate multiple PTZ cameras. The concept of our approach originated from the observation that people could usually make a rough estimate about the tilt angle of the camera simply based on some clues revealed in the captured images. Based on our approach, we can simply use some A4 papers on a table to calibrate multiple PTZ cameras. In our approach, there is no need to calculate the commonly used homography matrix. For real cases, once a set of PTZ cameras is settled, we can simply place a few simple patterns on a horizontal plane. These patterns can be A4 papers, books, boxes, etc.; and the horizontal plane can be a tabletop or the ground plane. The whole procedure does not need specially designed calibration patterns and requires only a light computational load. In the near future, we will work on the extension of the proposed approach so that we will be able to perform dynamic calibration when the PTZ cameras are under movement.**

***Index Terms*—Camera calibration, multiple cameras, pan-tilt-zoom (PTZ) camera.**

I. Introduction

For a surveillance system with multiple pan–tilt–zoom (PTZ) cameras, we may change the cameras' pan angle and tilt angle from time to time to achieve better monitoring results. However, whenever the pose of a PTZ camera is changed, we may need to recalibrate its extrinsic parameters again. Hence, how to quickly and efficiently calibrate the extrinsic parameters of multiple PTZ cameras has become an important issue. So far, various kinds of approaches have been developed to calibrate the static camera's intrinsic and extrinsic parameters, such as the techniques proposed in [1]–[9]. However, it would be impractical to repeatedly perform these elaborate calibration processes over a PTZ camera when that camera is under panning, tilting, or zooming all of the time.

Manuscript received October 23, 2005; revised March 15, 2006. This paper was recommended for publication by Associate Editor Y. F. Li and Editor M. Wang upon evaluation of the reviewers' comments. This work was supported by the Ministry of Economic Affairs of Taiwan, R.O.C., under Grant No. 95-EC-17-A-02-S1-032.

The authors are with the Department of Electronics Engineering, the Institute of Electronics, National Chiao Tung University, Hsinchu 300, Taiwan, R.O.C. (e-mail: kace.ee89g@nctu.edu.tw; shengjyh@cc.nctu.edu.tw).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TASE.2006.884040

On the other hand, [10]–[12] have proposed plane-based calibration methods especially designed for the calibration of multiple cameras. These approaches, proposed in [10] and [11], belong to factorization-based methods. In [12], the proposed process is similar to the integration of static camera calibration and moving camera calibration. It needs a multicamera rig to change the specific orientation to capture two or more views of a calibration grid. Another kind of method is proposed in [13] that uses the relationship between epipoles and trifocal tensors to calibrate multiple cameras. However, their method needs the assumption that at least two cameras are projected to each camera. For surveillance systems with multiple PTZ cameras, these elaborate processes and the constraints do not seem to be practical choices.

In this paper, we will first show the 3-D-to-2-D coordinate transformation in terms of the tilt angle of a PTZ camera. In [14], a similar scene model based on pan angle and tilt angle has also been established. In this paper, however, we will give a more complete formula that takes into account not only the translation effect but also the rotation effect when a PTZ camera is under tilt movement. We will then infer the tilt angle and the altitude of a PTZ camera based on the observation of some simple objects lying on a horizontal plane. The objects could be a corner with a known angle, a few corners with unknown but equal angles, a line segment with a known length, or a few line segments with unknown but equal lengths. In comparison, [15] proposed an approach that recovers the transformation between the image and the ground plane to find the look-down angle. In their method, however, the height information of the observed object has to be known in advance. In comparison, the height is not a necessity in our approach. After having calibrated the tilt angle and the altitude of each PTZ camera, we can easily achieve the calibration of multiple PTZ cameras based on some commonly observed objects. In some sense, our approach can be thought of to have decomposed the computation of homography matrix into two simple calibration processes so that the computational load becomes lighter for the calibration of multiple PTZ cameras.

This paper is organized as follows. First, in Section II, we will describe the mapping between the horizontal plane in the 3-D space and the 2-D image plane on a tilted camera. Based on the back projection formula, the pose (the tilt angle and the altitude) of the camera can be estimated by viewing some simple patterns on a horizontal plane. Next, in Section III, we will introduce how to utilize the estimation results to achieve the calibration of multiple PTZ cameras. The sensitivity analysis with respect to parameter fluctuations and measurement errors will be discussed in Section IV. Some experimental results on real data are demonstrated in Section V to illustrate the feasibility of this method. Finally, the conclusion is drawn.

II. Pose Estimation of a PTZ Camera

In this section, we show the mapping between the 3-D space and the image plane on a tilted PTZ camera, under the constraint that all observed points are located on a horizontal plane. We will then demonstrate that, with a few corners or a few line segments lying on a horizontal plane, we can easily estimate the tilt angle of the camera based on the back projection of the captured image.

A. Constrained Coordinate Mapping on a Tilted Camera

Fig. 1 illustrates the modeling of our camera setup. Here, we assume that the observed objects are located on a horizontal plane $\prod$, while the PTZ camera lies above $\prod$ with a height h. The PTZ camera may pan or tilt with respect to the rotation center O_R. Moreover, we assume the projection center of the camera, denoted as O_C, is away from O_R with distance r. We define the origin of the rectified world coordinates to be the projection center O_C of a PTZ camera with zero tilt angle. The Z-axis of the world coordinates is along the optical axis of the camera, while the X- and Y-axis of the world coordinates are parallel to the x- and y-axis of the projected image plane, respectively. When the camera tilts, the projection center moves to O_C' and the projected image plane is changed to a new 2-D plane.

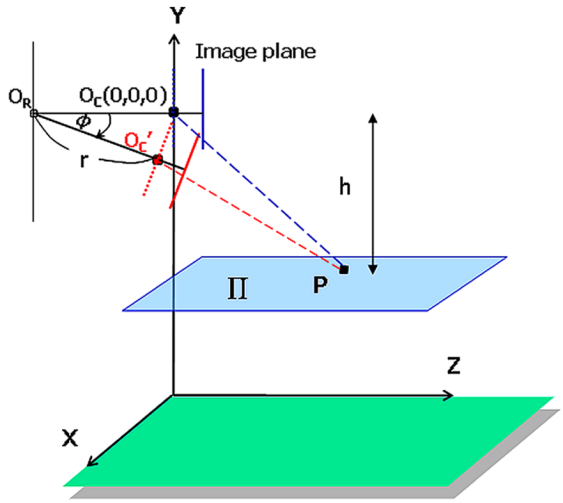

Fig. 1. Model of camera setup.

Assume $P = [X, Y, Z, 1]^T$ denotes the homogeneous coordinates of a 3-D point $\mathbf{p}$ in the world coordinates. For the case of a PTZ camera with zero tilt angle, we denote the perspective projection of $\mathbf{p}$ as $p = [x, y, 1]^T$. Under perspective projection, the relationship between P and p can be expressed as

$$p = \frac{1}{Z}\begin{bmatrix}\alpha & s & u_0\\ 0 & \beta & v_0\\ 0 & 0 & 1\end{bmatrix}[R \quad t]P \tag{1}$$

where α and β are the scale parameters expressed in pixel units for the x and y axes in the image plane, s is the skew factor due to some manufacturing error, and (u_0, v_0) is the principal point [16].

With respect to the rectified world coordinate system, the extrinsic term $[R \quad t]$ becomes $[\mathbf{I} \quad \mathbf{0}]$. To further simplify the mathematical deduction, we ignore the skew factor s and assume the image coordinates have been translated by a translation vector $(-u_0, -v_0)$. Hence, (1) can be simplified and reversed as

$$\begin{bmatrix}X\\ Y\\ Z\end{bmatrix} = Z\begin{bmatrix}\frac{1}{\alpha} & 0 & 0\\ 0 & \frac{1}{\beta} & 0\\ 0 & 0 & 1\end{bmatrix}\begin{bmatrix}x\\ y\\ 1\end{bmatrix} = \begin{bmatrix}\frac{xZ}{\alpha}\\ \frac{yZ}{\beta}\\ Z\end{bmatrix}. \tag{2}$$

When the PTZ camera tilts with an angle ϕ, the homogeneous coordinates of the projected image point will move to $[x', y', 1]^T$. In our previous work [17], we obtained (3). This equation indicates the back projection formula from the image coordinates of a tilted camera to the rectified world coordinates, under the constraint that all of the observed points are lying on a horizontal plane with $Y = -h$

$$\begin{bmatrix}X\\ Y\\ Z\end{bmatrix} = \begin{bmatrix}\frac{(x'-u_0)\beta(r\sin\phi - h)}{\alpha[(v_0 - y')\cos\phi - \beta\sin\phi)]}\\ -h\\ \frac{[(v_0-y')\sin\phi + \beta\cos\phi](r\sin\phi - h)}{(v_0-y')\cos\phi - \beta\sin\phi} - r + r\cos\phi\end{bmatrix}. \tag{3}$$

B. Estimation of Pose Based on Back Projections

Suppose we use a tilted camera to capture the image of a corner, which is located on a horizontal plane. Based on the captured image and a guessed tilt angle, we may use (3) to back project the captured image onto a horizontal plane on $Y = -h$. Assume three 3-D points P_A, P_B, and P_C, on a horizontal plane form a rectangular corner at P_A. The original image is captured by a camera with $\phi = 16°$, as shown in Fig. 2(a). In Fig. 2(b), we plot the back-projected images for various

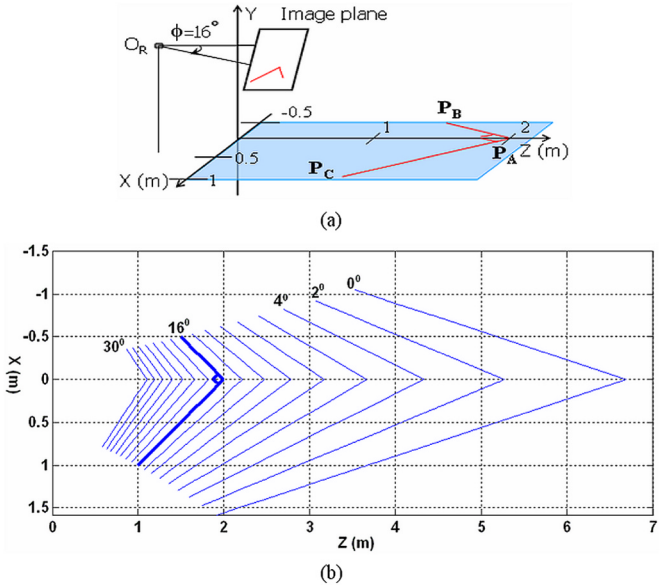

Fig. 2. (a) Rectangular corner captured by a tilted camera. (b) Illustration of back projection onto a horizontal plane for different choices of tilt angles.

choices of tilt angles. The guessed tilt angles range from 0 to 30°, with a 2° step. The back projection for the choice of 16° is plotted in bold, specifically. It can be seen that the back-projected corner becomes a rectangular corner only if the guessed tilt angle is correct. Besides, it is worth mentioning that a different choice of h only causes a scaling effect of the back-projected shape.

To formulate this example, we assume the corner at P_A has the angle ψ. Besides, P_A and P_B form a line segment with length L. In [17], we have built the relation between the back-projected angle and the guessed tilt angle as shown in (4) at the bottom of the next page. Note that in (4) we have ignored the offset terms, u_0 and v_0, to reduce the complexity of the formulation.

In Fig. 3, we show the back-projected angle ψ with respect to the guessed tilt angle, assuming α and β are known in advance. In this simulation, these two curves are generated by placing the rectangular corner on two different places of the horizontal plane. Again, the back-projected angle is equal to 90° only if we choose the tilt angle to be 16°. This simulation demonstrates that if we know in advance the angle of the captured corner, we can easily deduce the camera's tilt angle. Moreover, if we do not know in advance the actual angle of the corner, we can simply place that corner on more than two different places of the horizontal plane. In this way, based on the intersection of the deduced ψ-*v.s.*-ϕ curves, we may estimate not only the tilt angle of the camera but also the actual angle of the corner.

In [17], we have also derived a similar relationship between the back-projected length and the guessed tilt angle as

$$L = l(\phi) = \left\{ \left(\frac{(x'_B - u_0)\beta(r\sin\phi - h)}{\alpha[(v_0 - y'_B)\cos\phi - \beta\sin\phi]} - \frac{(x'_A - u_0)\beta(r\sin\phi - h)}{\alpha[(v_0 - y'_A)\cos\phi - \beta\sin\phi]} \right)^2 + \left(\frac{[(v_0 - y'_B)\sin\phi + \beta\cos\phi](r\sin\phi - h)}{(v_0 - y'_B)\cos\phi - \beta\sin\phi} - \frac{[(v_0 - y'_A)\sin\phi + \beta\cos\phi](r\sin\phi - h)}{(v_0 - y'_A)\cos\phi - \beta\sin\phi} \right)^2 \right\}^{1/2}. \quad (5)$$

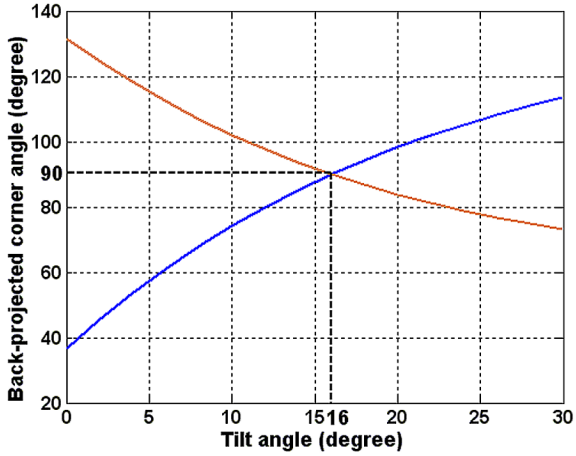

Fig. 3. Back-projected angle with respect to guessed tilt angles.

Similarly, if α, β, r, and h are known in advance, we can deduce the tilt angle directly based on the projected value of L. Note that in (5), the values of r and h only affect the scaling of L. Hence, even if the values of r and h are unknown, we may simply place more than two line segments of the same length on different places of a horizontal plane and seek to find the intersection of these corresponding L-*v.s.*-ϕ curves.

We may also seek to perform parameter estimation based on an optimization process. Here, we take (5) as an example. We assume several line segments with known lengths (not necessarily the same length) are placed on different positions of a horizontal plane and we use a tilted camera to capture the image. Assume the length of the ith segment is L_i, then we aim to find a set of parameters $\{\alpha, \beta, u_0, v_0, \phi, r, h\}$ that minimize

$$F(x'_1, y'_1, x'_2, y'_2, \ldots, x'_m, y'_m, \alpha, \beta, u_0, v_0, \phi, r, h) = \sum_{i_1}^{m} \left\| l_i\left(x'_i, y'_1, \alpha, \beta, u_0, v_0, \phi, r, h\right) - L_i \right\|^2. \quad (6)$$

In the optimization process, we adopt the Levenberg–Marquardt algorithm. Under our scene model, the tilt angle ϕ and altitude h can be roughly estimated simply based on visual observations. In our experiments, the error range for the guessed tilt angle ϕ is within $\pm 20°$ and the error range for the guessed altitude h is within ± 1.5 m. With these initial guesses, the optimization process is very stable and the estimation results are satisfactorily accurate.

III. Calibration of Multiple Cameras

A. Calibration of Multiple PTZ Cameras

In our camera model, each camera has its own world coordinate system. If a vector in the 3-D space, such as a line segment on a tabletop, is observed by several PTZ cameras at the same time, we can achieve the calibration of these cameras by mapping the individual back-projected world coordinates of this vector to a common reference world coordinates. We take two calibrated PTZ cameras as an example. Fig. 4(a) shows the scene model of these two cameras. Fig. 4(b) shows the vector locations in the world coordinates of these two cameras, respectively. Based on the estimated ϕ and h, together with the image projections of the vector points, we can get the world coordinates of points $\mathbf{A_{ref}}$, $\mathbf{B_{ref}}$, and $\mathbf{A}'$, $\mathbf{B}'$ from (3). The difference of the rotation angle ω between the two world coordinate systems can then be easily computed

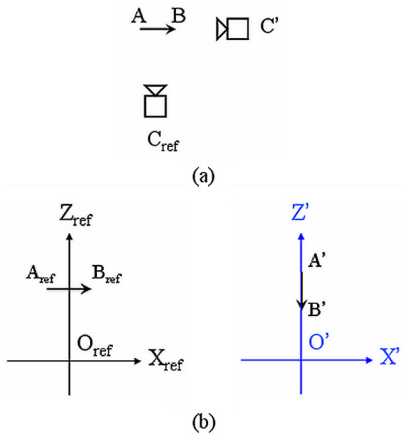

Fig. 4. (a) Top view of two cameras and a vector in the space. (b) The world coordinates of the vector with respect to these two cameras.

by

$$\cos w = \frac{\langle \overline{A'B'}, \overline{A_{\rm ref}B_{\rm ref}} \rangle}{\|A'B'\| \times \|A_{\rm ref}B_{\rm ref}\|}. \tag{7}$$

After applying the rotation to point A', the position translation $\mathbf{t}$ between these two cameras can be expressed as

$$\mathbf{t} = \mathbf{A_{ref}} - \begin{bmatrix} \cos w & 0 & -\sin w \\ 0 & 1 & 0 \\ \sin w & 0 & \cos w \end{bmatrix} \cdot \mathbf{A}'. \tag{8}$$

Hence, the 3-D relationship between these two cameras can be easily deduced.

B. Discussion of Pan Angle

Notice that, in the above deductions, we do not care about the pan angles of PTZ cameras. This is because, in the back-projection process, to guess a pan angle only implies to rotate the X and Z coordinates in our camera model. It does not change the spatial relationship between the PTZ camera and the back-projected objects. In Fig. 5, we show such an example. In Fig. 5(a), we show the image captured by a PTZ camera with three corner points being marked in circle. Fig. 5(b) shows the top view (i.e., X-Z plane) of the back-projected corner points with respect to four guessed pan angles, 0°, 30°, 60°, and 90°. The arrows indicate the optical axes of the camera with respect to these four pan angles. It can be seen that the spatial relationship between the optical axis and the back-projected corner points is almost the same when the camera pans. The little variation comes from the fact that the panning center is not the same as the projection center. However, since the rotation radius r is so small when compared with the distance between the camera and the object, this small variation can actually be ignored.

IV. Sensitivity Analysis

Due to errors in the estimation of camera parameters and in measurement, these ψ-*v.s.*-ϕ curves or L-*v.s.*-ϕ curves usually do not intersect at a single point. The impacts of parameter fluctuations have been analyzed via computer simulations in [18]. Here, we will deduce the formulae for the sensitivity analysis of the estimated ϕ and h.

In theory, for the estimation of ϕ and h, the optimization of (6) conforms to the following two equations:

$$\begin{cases} f_1 = \sum_{i=1}^{m} \frac{\partial l_i}{\partial \phi} - \sum_{i=1}^{m} L_i \frac{\partial l_i}{\partial \phi} = 0 \\ f_2 = \sum_{i=1}^{m} \frac{\partial l_i}{\partial h} - \sum_{i=1}^{m} L_i \frac{\partial l_i}{\partial h} = 0. \end{cases} \tag{9}$$

Hence, the estimated ϕ and h satisfy $f_1(\phi, h) = 0$ and $f_2(\phi, h) = 0$. We apply the implicit function theorem to (9) to find how ϕ and h deviate with respect to the measurement error in x_i'

$$\begin{bmatrix} \frac{\partial \phi}{\partial x_i'} \\ \frac{\partial h}{\partial x_i'} \end{bmatrix} = - \begin{bmatrix} \frac{\partial f_1}{\partial \phi} & \frac{\partial f_1}{\partial h} \\ \frac{\partial f_2}{\partial \phi} & \frac{\partial f_2}{\partial h} \end{bmatrix}^{-1} \begin{bmatrix} \frac{\partial f_1}{\partial x_i'} \\ \frac{\partial f_2}{\partial x_i'} \end{bmatrix}. \tag{10}$$

$$\begin{aligned} \psi = \cos^{-1} \Bigg\{ & \Bigg\{ \left(\frac{x'_B \beta}{\alpha (y'_B \cos\phi - \beta \sin\phi)} - \frac{x'_A \beta}{\alpha (y'_A \cos\phi - \beta \sin\phi)} \right) \\ & \times \left(\frac{x'_C \beta}{\alpha (y'_C \cos\phi - \beta \sin\phi)} - \frac{x'_A \beta}{\alpha (y'_A \cos\phi - \beta \sin\phi)} \right) \\ & + \left(\frac{(y'_B \sin\phi + \beta \cos\phi)}{y'_B \cos\phi - \beta \sin\phi} - \frac{(y'_A \sin\phi + \beta \cos\phi)}{y'_A \cos\phi - \beta \sin\phi} \right) \\ & \times \left(\frac{(y'_C \sin\phi + \beta \cos\phi)}{y'_C \cos\phi - \beta \sin\phi} - \frac{(y'_A \sin\phi + \beta \cos\phi)}{y'_A \cos\phi - \beta \sin\phi} \right) \Bigg\} \\ & \times \Bigg\{ \left(\frac{x'_B \beta}{\alpha (y'_B \cos\phi - \beta \sin\phi)} - \frac{x'_A \beta}{\alpha (y'_A \cos\phi - \beta \sin\phi)} \right)^2 \\ & + \left(\frac{(y'_B \sin\phi + \beta \cos\phi)}{y'_B \cos\phi - \beta \sin\phi} - \frac{(y'_A \sin\phi + \beta \cos\phi)}{y'_A \cos\phi - \beta \sin\phi} \right)^2 \Bigg\}^{-1/2} \\ & \times \Bigg\{ \left(\frac{x'_C \beta}{\alpha (y'_C \cos\phi - \beta \sin\phi)} - \frac{x'_A \beta}{\alpha (y'_A \cos\phi - \beta \sin\phi)} \right)^2 \\ & + \left(\frac{(y'_C \sin\phi + \beta \cos\phi)}{y'_C \cos\phi - \beta \sin\phi} - \frac{(y'_A \sin\phi + \beta \cos\phi)}{y'_A \cos\phi - \beta \sin\phi} \right)^2 \Bigg\}^{-1/2} \Bigg\}. \end{aligned} \tag{4}$$

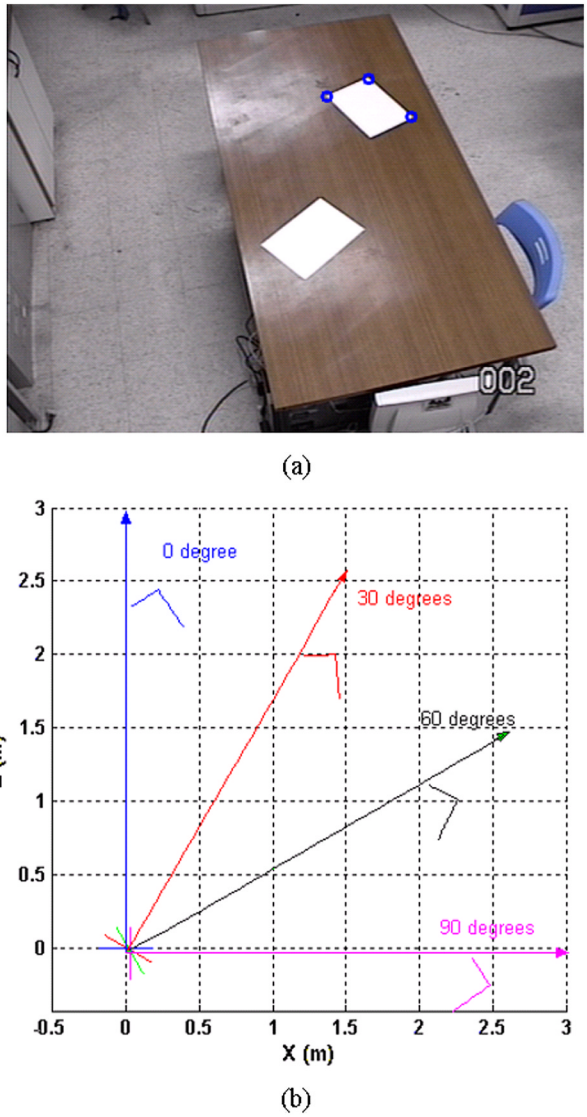

Fig. 5. (a) Three points marked in the image captured by a PTZ camera. (b) Top view of the back-projected corners and the optical axes with respect to different guessed pan angles.

Similarly, we can deduce the formulae for $\partial\phi/\partial\alpha$, $\partial\phi/\partial\beta$, and so on. If we assume the total variations of ϕ and h are caused by individual variations with respect to parameter fluctuations in $\{\alpha, \beta, u_0, v_0\}$ and measurement errors in $\{x_i', y_i'\}$, then we have

$$\Delta\phi \approx \sum_{i=1}^{m}\left(\frac{\partial\phi}{\partial x_i'}\Delta x_i'\right) + \sum_{i=1}^{m}\left(\frac{\partial\phi}{\partial y_i'}\Delta y_i'\right) + \frac{\partial\phi}{\partial\alpha}\Delta\alpha + \frac{\partial\phi}{\partial\beta}\Delta\beta + \frac{\partial\phi}{\partial u_0}\Delta u_0 + \frac{\partial\phi}{\partial v_0}\Delta v_0 \quad (11)$$

and

$$\Delta h \approx \sum_{i=1}^{m}\left(\frac{\partial h}{\partial x_i'}\Delta x_i'\right) + \sum_{i=1}^{m}\left(\frac{\partial h}{\partial y_i'}\Delta y_i'\right) + \frac{\partial h}{\partial\alpha}\Delta\alpha + \frac{\partial h}{\partial\beta}\Delta\beta + \frac{\partial h}{\partial u_0}\Delta u_0 + \frac{\partial h}{\partial v_0}\Delta v_0. \quad (12)$$

To verify the formulae deduced above, we perform the following simulations. Here, we assume two line segments of 0.283 m are placed on a horizontal plane and the altitude $h = 2$ m. The values of $\{u_0, v_0, \alpha, \beta\}$ are {348, 257, 770, 750}. The images of these two segments are assumed to be captured by a PTZ camera with the tilt angle $\phi = 60°$.

TABLE I
VARIATIONS OF TILT ANGLE AND ALTITUDE WITH RESPECT TO DIFFERENT PARAMETER FLUCTUATIONS AND MEASUREMENT ERRORS

Δu_0	-20	-10	+10	+20
simulated ($\Delta\phi$, Δh)	(0.00, 0.00)	(0.00, 0.00)	(0.00, 0.00)	(0.00, 0.00)
deduced ($\Delta\phi$, Δh)	(0.00, 0.00)	(0.00, 0.00)	(0.00, 0.00)	(0.00, 0.00)
Δv_0	-20	-10	+10	+20
simulated ($\Delta\phi$, Δh)	(-1.05, 0.00)	(-0.52, 0.00)	(0.51, 0.00)	(1.00, 0.00)
deduced ($\Delta\phi$, Δh)	(-1.02, 0.00)	(-0.51, 0.00)	(0.51, 0.00)	(1.02, 0.00)
$\Delta\alpha$	-20	-10	+10	+20
simulated ($\Delta\phi$, Δh)	(-1.44, -0.10)	(-0.72, -0.05)	(0.73, 0.05)	(1.47, 0.10)
deduced ($\Delta\phi$, Δh)	(-1.48, -0.11)	(-0.74, -0.05)	(0.72, 0.05)	(1.41, 0.10)
$\Delta\beta$	-20	-10	+10	+20
simulated ($\Delta\phi$, Δh)	(1.85, 0.05)	(0.91, 0.03)	(-0.88, -0.03)	(-1.72, -0.05)
deduced ($\Delta\phi$, Δh)	(1.72, 0.05)	(0.89, 0.03)	(-0.92, -0.03)	(-1.89, -0.06)
$\Delta x_1'$	-4	-2	+2	+4
simulated ($\Delta\phi$, Δh)	(-2.60, -0.18)	(-1.34, -0.09)	(1.45, 0.1)	(3.01, 0.21)
deduced ($\Delta\phi$, Δh)	(-2.84, -0.20)	(-1.41, -0.10)	(1.39, 0.1)	(2.73, 0.19)
$\Delta y_3'$	-4	-2	+2	+4
simulated ($\Delta\phi$, Δh)	(-3.62, -0.13)	(-1.94, -0.07)	(1.91, 0.07)	(3.90, 0.13)
deduced ($\Delta\phi$, Δh)	(-4.45, -0.17)	(-1.97, -0.07)	(1.78, 0.06)	(3.25, 0.11)

In the first simulation, the tilt angle $\phi = 60°$. We individually change the values of camera parameters and the measurement of (x_i', y_i') to see how the estimated values of ϕ and h vary. Here, the LM algorithm is applied to (6) for the estimation of ϕ and h. The variations of these estimation results, together with the variations deduced by (11) and (12) are listed in Table I. It can be seen that the deduced variations based on (11) and (12) well approximate the simulation results.

In the second simulation, we change the tilt angle ϕ from 15 to 75°, with a 15° step. Table II demonstrates that (11) and (12) conform to the variations of the simulation results for a wide rage of tilt angle ϕ. In practice, the fluctuations of camera parameters are likely to be less than 20 and the measurement errors of (x_i', y_i') are likely to be less than 4 pixels. Hence, the estimation errors of ϕ and h are expected to be acceptable in real cases.

V. EXPERIMENTS ON REAL IMAGES

The parameter estimation results of a single PTZ camera have been shown and discussed in [18]. In this section, some calibration results of multiple cameras on real data are demonstrated. In this simulation, the test images are captured by a PTZ camera mounted on the ceiling with an unknown tilt angle. The image resolution is 320 by 240 pixels. A few A4 papers are randomly placed on a horizontal table, as shown in Fig. 6. The corners of these A4 paper sheets could be easily identified either by hand or by a corner detection algorithm. So far, we identify these corners manually and we have developed a software package to facilitate the identification of corners and line segments in images. In the 3-D space, all of the corners are 90°, while the length and width of an A4 paper are 297 mm and 210 mm, respectively.

The upper part of Table III lists some experimental results for the calibration of multiple cameras. Here, the intrinsic parameters $\{\alpha, \beta, u_0, v_0\}$ of each camera are estimated in advance, based on Zhang's calibration method [3]. The parameter r can be estimated via direct measurement. Hence, (6) includes only two unknown variables ϕ and h. Each row of Table III lists the mean and standard deviation of the estimated parameters for a single camera. To calculate the

TABLE II
VARIATIONS OF TILT ANGLE AND ALTITUDE WITH RESPECT TO DIFFERENT CHOICES OF TILT ANGLE

		Tilt Angle				
	$(\Delta\phi, \Delta h)$	15°	30°	45°	60°	75°
$\Delta u_0 = 20$	simulated	(0.02, 0.00)	(0.01, 0.00)	(0.01, 0.00)	(0.00, 0.00)	(0.00, 0.00)
	deduced	(0.00, 0.00)	(0.00, 0.00)	(0.00, 0.00)	(0.00, 0.00)	(0.00, 0.00)
$\Delta v_0 = 20$	simulated	(1.82, 0.00)	(1.79, 0.00)	(1.50, 0.00)	(1.00, 0.00)	(0.46, 0.00)
	deduced	(1.81, 0.00)	(1.80, 0.00)	(1.51, 0.00)	(1.02, 0.00)	(0.47, 0.00)
$\Delta\alpha = 20$	simulated	(1.50, 0.11)	(1.49, 0.11)	(1.48, 0.10)	(1.47, 0.10)	(1.46, 0.10)
	deduced	(1.44, 0.10)	(1.43, 0.10)	(1.42, 0.10)	(1.41, 0.10)	(1.40, 0.10)
$\Delta\beta = 20$	simulated	(-0.47, -0.05)	(-1.00, -0.05)	(-1.46, -0.05)	(-1.72, -0.05)	(-1.73, -0.05)
	deduced	(-0.46, -0.05)	(-1.05, -0.05)	(-1.58, -0.06)	(-1.89, -0.06)	(-1.83, -0.06)
$\Delta x_1' = 4$	simulated	(2.75, 0.19)	(3.07, 0.22)	(3.16, 0.22)	(3.01, 0.21)	(2.65, 0.19)
	deduced	(2.52, 0.18)	(2.78, 0.20)	(2.85, 0.20)	(2.73, 0.19)	(2.43, 0.17)
$\Delta y_3' = 4$	simulated	(2.97, 0.10)	(3.80, 0.13)	(4.15, 0.14)	(3.90, 0.13)	(3.14, 0.11)
	deduced	(2.58, 0.09)	(3.18, 0.11)	(3.42, 0.11)	(3.25, 0.11)	(2.71, 0.09)

mean and standard deviation, five observations are made with each observation including eight selected line segments on the boundary of these A4 papers, as shown in Fig. 6(a). It can be seen that all of the estimated parameters have an acceptably small standard deviation.

In the lower part of Table III, each row corresponds to the estimations of the position and orientation of each camera with respect to Camera 2. The relative position and orientation are computed based on the mean value of ϕ and h and one common vector in Fig. 6(a). The top view of the relative positions in the 3-D space is illustrated in Fig. 6(b). The eight chosen corners of the A4 papers in Fig. 6(a) are also plotted in Fig. 6(b) to offer a clearer geometric sense.

To evaluate the calibration results, we randomly pick up a few test points in the image captured by Camera 2 and use the calibration result to find the corresponding points on the other three images. The results are shown in Fig. 7, with all corresponding points being represented by the same type of marker. It can be seen that all of the corresponding relationships are reasonably accurate.

In practice, a commonly used technique for camera pose estimation is to find the homography matrix between a reference plane in the 3-D space and the camera's image plane [19]. The rotation and translation matrices can then be extracted by applying the singular value decomposition (SVD) method over the homography matrix. In the homography approach, we need to define a reference world coordinate system and need to pick up a few spatial points with known reference world coordinates in advance. In other words, not only the distances but also the relative spatial information among the calibration points needs to be known. In comparison, our approach does not need to know the world coordinates of these calibration objects. We only need to measure the lengths or angles of the calibration objects. Hence, the preparation of calibration objects becomes much easier in our approach.

Besides, we may use fewer spatial points for the calibration of camera poses. This is because there is an implicit constraint in our approach. In a typical setup of PTZ cameras, the horizontal axis of the camera's image coordinate system is usually parallel to the ground plane. This parallelism is kept all of the time even though the camera is under the panning, tilting, and zooming operations from time to time. Moreover, in our approach, we do not actually care about the exact pan angle of the camera. These two conditions correspond to the constraints over the rotation about the Z axis and the rotation about the Y axis in our rectified world coordinate system. Hence, our method may lead to stable pose estimations even when we only use a few calibration points.

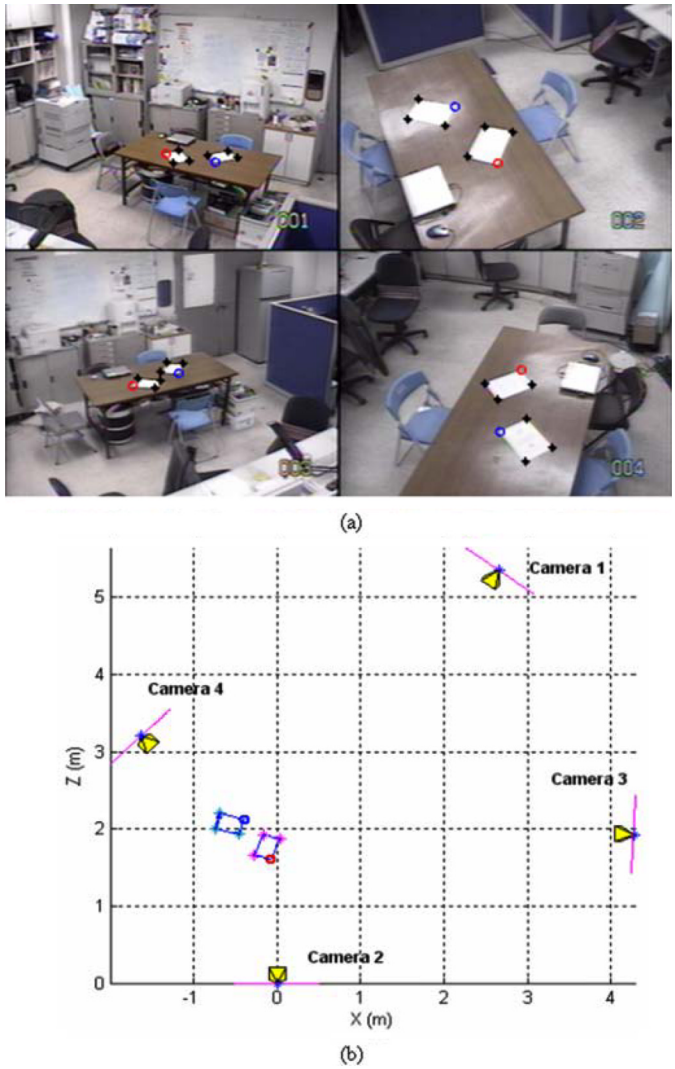

Fig. 6. (a) Test image captured by four PTZ cameras. (b) Top view of the relative positions between four PTZ cameras.

TABLE III
UPPER: ESTIMATED TILT ANGLE AND ALTITUDE. LOWER: SPATIAL RELATIONSHIP BETWEEN CAMERAS

	Mean		Deviation	
	ϕ(degree)	h(m)	ϕ(degree)	h(m)
Camera 1	20.4	1.96	1.5	0.12
Camera 2	51.1	2.32	0.6	0.03
Camera 3	24.8	1.95	1.2	0.17
Camera 4	44.2	2.01	0.3	0.14
	Relative position & orientation			
	X(m)	Y(m)	Z(m)	ϖ(degree)
Camera 1	2.7	-0.36	5.70	144.1
Camera 2	0	0	0	0
Camera 3	4.27	-0.37	1.93	87.1
Camera 4	-1.64	-0.31	3.2	-135.4

Fig. 7. Evaluation of calibration results.

Fig. 8. Test images with a rectangular calibration pattern.

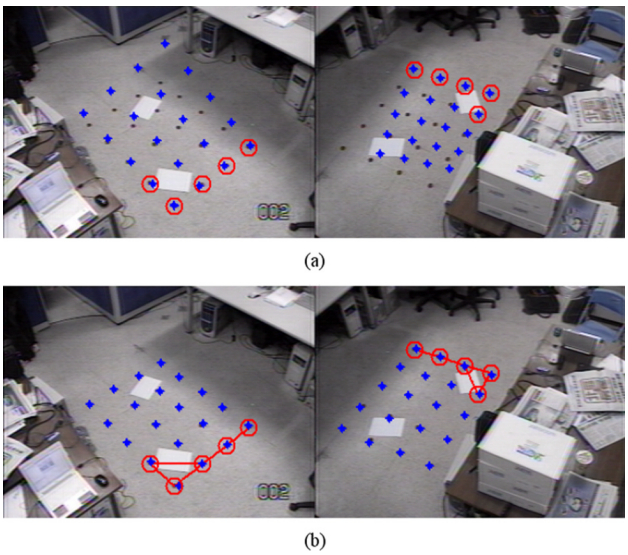

Fig. 9. Evaluation of calibration results by using five points. (a) Point correspondence based on the homography technique. (b) Point correspondence based on the proposed method.

To compare with the homography technique, we marked 20 points on the ground floor to form a rectangular pattern, as shown in Fig. 8. Five of these 20 points are chosen to be the calibration points, as marked by the circles in Fig. 9. The asterisk markers in Fig. 9(a) show the point correspondence based on the calibration result of the homography technique using the OpenCV library. On the other hand, the asterisk markers in Fig. 9(b) show the point correspondence based on our approach using five segments with known lengths, as shown in Fig. 9(b). Besides, we calibrated these two cameras 20 times by randomly choosing 5 of these 20 points as the calibration points. We then checked the point correspondence in the image captured by Camera 4 based on these 20 image points of Camera 2 and the calibration results. Table IV shows the mean absolute distance and standard deviation of the point-wise correspondence. It can be easily seen that our approach offers a more reliable and stable calibration result when we only use a few calibration points. When all 20 points are used, on the other hand, there would be no obvious difference between the performance of the homography technique and the performance of our approach. Nevertheless, for general surveillance environments, it could be difficult to place this kind of specific pattern for calibration. Hence, in general, simple calibration objects or a small amount of calibration points without known reference spatial coordinates will be preferred.

TABLE IV
MEAN ABSOLUTE DISTANCE AND STANDARD DEVIATION OF THE POINT-WISE CORRESPONDENCE BETWEEN CAMERA 2 AND CAMERA 4

	Mean Absolute Distance (pixels)	Deviation (pixels)
Homography technique	6.0	5.6
Proposed method	3.4	1.4

Another advantage of our method is the comprehensible sense of camera pose. The tilt angle, altitude, and orientation of the camera offer a more direct physical sense about the camera pose in the 3-D space, especially when the PTZ cameras are under panning, tilting, and zooming operations from time to time. In our approach, we derive some explicit formula to describe how the tilt angle and altitude of a PTZ camera affect the 3-D-to-2-D projection. This makes the calculation of 2-D-to-3-D backprojection much easier without the need of indirect depth computation. Besides, based on the comprehensible space sense, the relationships among multiple cameras can be easily obtained without complicated computations. In comparison, when using the conventional homography technique, the relative position and orientation between each pair of cameras offer less comprehensible sense about the setting of multiple cameras. Although these relative coordinate systems may still be transformed into an integrated coordinate system, the work on the calibration of multiple cameras will become more elaborate when the number of cameras increases.

VI. CONCLUSION

To summarize, in this paper, we adopt a system model that is general enough to fit for a large class of surveillance systems with multiple PTZ cameras. A calibration method is proposed that does not require a particular system setup or specific calibration patterns. When compared to some existing calibration approaches, which extract the homography matrix and the rotation matrix, our approach offers direct geometric sense and can simplify the calibration process. No coordinated calibration pattern is needed and the computational load is light. The experiment results over real images demonstrate the efficiency of this approach.

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 J. Robot. Autom.*, vol. RA-3, no. 4, pp. 323–344, Aug. 1987.

[2] M. Agrawal and L. S. Davis, "Camera calibration using spheres: A semi-definite programming approach," in *Proc. 9th IEEE Int. Conf. Computer Vision*, Oct. 2003, vol. 2, pp. 782–789.

[3] Z. Zhang, "Flexible camera calibration by viewing a plane from unknown orientations," in *Proc. 7th IEEE Int. Conf. Computer Vision*, Sep. 1999, vol. 1, pp. 666–673.

[4] G.-Q. Wei and S. D. Ma, "A complete two-plane camera calibration method and experimental comparisons," in *Proc. 4th Int. Conf. Computer Vision*, May 1993, pp. 439–446.

[5] P. F. Sturm and S. J. Maybank, "On plane-based camera calibration: A general algorithm, singularities, application," presented at the IEEE Comput. Soc. Conf. Computer Vision and Pattern Recognition, Jun. 1999.

[6] Z. Zhang, "Camera calibration with one-dimensional objects," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 26, no. 7, pp. 892–899, Jul. 2004.

[7] A. Basu and K. Ravi, "Active camera calibration using pan, tilt and roll," *IEEE Trans. Syst., Man Cybern. B, Cybern.*, vol. 27, no. 3, pp. 559–566, Jun. 1997.

[8] E. E. Hemayed, "A survey of camera self-calibration," in *Proc. IEEE Conf. Advanced Video and Signal Based Surveillance*, Jul. 2003, pp. 351–357.

[9] R. T. Collins and Y. Tsin, "Calibration of an outdoor active camera system," in *Proc. IEEE Comput. Soc. Conf. Computer Vision and Pattern Recognition*, Jun. 1999, vol. 1, pp. 534–534.

[10] P. Sturm, "Algorithms for plane-based pose estimation," in *Proc. IEEE Conf. Computer Vision and Pattern Recognition*, Jun. 2000, vol. 1, pp. 706–711.

[11] T. Ueshiba and F. Tomita, "Plane-based calibration algorithm for multi-camera systems via factorization of homography matrices," in *Proc. 9th IEEE Int. Conf. Computer Vision*, Oct. 2003, vol. 2, pp. 966–973.

[12] C. Wiles and A. Davison, "Calibrating a multi-camera system for 3D modelling," in *Proc. IEEE Workshop on Multi-View Modeling and Analysis of Visual Scenes*, Jun. 1999, pp. 29–36.

[13] Y. Sugimure and J. Sato, "Camera calibration and reconstruction from the chain connection of mutual camera projections," in *Proc. 17th Int. Conf. Pattern Recognition*, Aug. 2004, vol. 1, pp. 100–103.

[14] S.-N. Lim, L. S. Davis, and A. Elgammal, "A scalable image-based multi-camera visual surveillance system," in *Proc. IEEE Conf. Advanced Video and Signal Based Surveillance*, Jul. 2003, pp. 205–212.

[15] P. Remagnino and G. A. Jones, "Automated registration of surveillance data for multi-camera fusion," in *Proc. 5th Int. Conf. Information Fusion*, Jul. 2002, vol. 2, pp. 1190–1197.

[16] D. A. Forsyth and J. Ponce, *Computer Vision: A Modern Approach*. Upper Saddle River, NJ: Prentice-Hall, 2003, pp. 28–30.

[17] I.-H. Chen and S.-J. Wang, "Efficient vision-based calibration for visual surveillance systems with multiple PTZ cameras," in *Proc. IEEE Int. Conf. Computer Vision Systems*, Jan. 2006, pp. 24–31.

[18] I.-H. Chen and S.-J. Wang, "A vision-based approach to extracting the tilt angle and altitude of a PTZ camera," *Proc. SPIE*, vol. 6066, pp. 606609-1–9, Jan. 2006.

[19] R. Hartley and A. Zisserman, *Multiple View Geometry in Computer Vision*. Cambridge, U.K.: Cambridge University Press, 2003, ch. 7.