Practical Pure Pan and Pure Tilt Camera Calibration

Imran N. Junejo*
IRISA/INRIA Rennes, France
imran.junejo@inria.fr

Hassan Foroosh
University of Central Florida, Orlando, U.S.A.
foroosh@cs.ucf.edu

Abstract

Often the deployed pan-tilt-zoom (PTZ) cameras undergo a pure pan or pure tilt rotation. This is a degenerate case for most of the PTZ camera calibration methods. That is, under this motion, the estimated camera parameters are not unique. In this regard, we present a novel camera calibration method to estimate five camera parameters from these pure pan/tilt cameras by using only two images. Our solution is based on using the infinite homography and performing its eigendecomposition. Our solutions and analyses are thoroughly validated and tested on both synthetic and real data, whereby the proposed method is shown to be accurate and noise resilient.

1 Introduction

Pure rotating cameras are now very commonly used everywhere, for e.g. in video surveillance. The traditional off-line calibration methods [10, 11] are not practical due to the dynamic changes in internal and external parameters of these camera. As a result it is important that one can auto-calibrate the camera online, when required. In this regard, the first auto-calibration method was due to Faugeras et al. [2] but the earlier work on cameras of special motion was done by [8] for pure translation and by [4] for pure rotation. [3] use known rotations to perform camera calibration of the rotating cameras. [9] perform the calibration by analyzing the inter-image homographies, requiring at least 4 homographies. However, the state of the art auto-calibration method for rotating and zooming cameras is that of Agapito et al. [1], who used the mapping of the image of the absolute conic (IAC) via the infinite homography to impose linear constraints on camera internal parameters using five or more images.

Motivated by practical consideration regarding PTZ cameras, we focus primarily on the degenerate rotations i.e. pure pan and pure tilt, and investigate the limits and the stability of auto-calibration under these degenerate cases. Assuming that the aspect ratio is known (e.g. $\lambda = 1$), and that except for the focal length all other intrinsic parameters remain invariant, our constraints lead to low order polynomials, which can be readily solved. Our solution provides four unknown intrinsic parameters (other than λ) and the rotation angle from only two images for pure panning under variable focal length and zero skew, a considerable improvement over [1], where pure pan or tilt motion are unstable cases for their method. Therefore, using zero skew constraint is not sufficient and they have to assume known λ. Although assuming a non-zero skew introduces instability in our method as well, we are able to solve for 4 intrinsic parameters and the rotation angle using only an image pair. Experimental results demonstrate the superiority of the proposed method in terms of accuracy and noise resilience.

The remainder of this paper consists of a brief description of background and notations, followed by the main section that carefully outline the above mentioned contributions of this paper. We then present thorough experimental results on both synthetic and real data to validate our analysis and test our solution, followed by concluding remarks.

2 Background and Notations

For a pinhole camera model used in this paper, a 3D point $\mathbf{M} = [X\ Y\ Z\ 1]^T$ and its corresponding image projection $\mathbf{m} = [x\ y\ 1]^T$ are related via a 3×4 matrix $\mathbf{P}$ by

$$\mathbf{m} \sim \underbrace{\mathbf{K}[\mathbf{r}_1\ \mathbf{r}_2\ \mathbf{r}_3\ \mathbf{t}]}_{\mathbf{P}}\mathbf{M}, \quad \mathbf{K} = \begin{bmatrix} \lambda f & \gamma & u_0 \\ 0 & f & v_0 \\ 0 & 0 & 1 \end{bmatrix}, \quad (1)$$

where $\sim$ indicates equality up to multiplication by a non-zero scale factor, $\mathbf{r}_i$ are the columns of the rotation matrix $\mathbf{R}$, $\mathbf{t}$ is the translation vector, and $\mathbf{K}$ is a nonsingular 3×3 upper triangular matrix known as the camera calibration matrix including five parameters, i.e. the focal length f, the skew γ, the aspect ratio λ and the

*This is part of the author's work at the University of Central Florida, Orlando, U.S.A.

978-1-4244-2175-6/08/$25.00 ©2008 IEEE

principal point at $(u_0,\ v_0)$. In this work, we assume known λ and $\gamma = 0$, a widely used assumption [1, 9]. Let $\mathbf{K}_1$ and $\mathbf{K}_2$ be the camera calibration matrices for a pair of images obtained by a fixed rotating and zooming camera and let also $\mathbf{R}_{12}$ denote the relative rotation between the two orientations of the camera, then the two images are related by the infinite homography given by

$$\mathbf{H}_{21} \sim \mathbf{K}_1 \mathbf{R}_{21} \mathbf{K}_2^{-1} \qquad (2)$$

The IAC, denoted by $\boldsymbol{\omega}$, is an imaginary point conic directly related to the camera internal matrix $\mathbf{K}$, via $\boldsymbol{\omega} \sim \mathbf{K}^{-T}\mathbf{K}{-1}$. Once $\boldsymbol{\omega}$ is obtained, $\mathbf{K}$ is recovered by Cholesky decomposition [5].

3 Pure Pan & Pure Tilt: Degenerate Cases

When the camera rotation is reduced to either pure pan or pure tilt, many existing solutions [4, 1] in the literature degenerate. As a result they cannot provide all the unknown parameters from only two images [1]. Below, we describe a new approach that allows to solve for 4 intrinsic parameters and the unknown rotation angle from two images in both pure pan and pure tilt.

Pure Pan: We show that the case of pure pan can be solved by direct construction of a set of homogeneous equations. For pure pan, we obtain 5 independent equations from two images in terms of the unknown intrinsic parameters using eigendecomposition of the infinite homography and direct use of equation (2).

As pointed out in [6] the eigendecomposition of the infinite homography $\mathbf{H}_{21}$ provides three fixed points under the homography given by the eigenvectors: one real eigenvector $\mathbf{v}$, which corresponds to the vanishing point of the rotation axis, and two complex ones $\mathbf{I}$ and $\mathbf{J}$ that correspond to the imaged circular points of any plane orthogonal to the rotation axis. When the camera intrinsic parameters are fixed, these points provide four independent constraints on the image of the absolute conic $\boldsymbol{\omega}$ [6]:

$$\mathbf{I}^T \boldsymbol{\omega} \mathbf{I} = 0, \quad \mathbf{J}^T \boldsymbol{\omega} \mathbf{J} = 0, \quad \mathbf{l_v} \sim \mathbf{I} \times \mathbf{J} \sim \boldsymbol{\omega} \mathbf{v} \quad (3)$$

where the first two impose the constraints that the circular points of a plane must lie on the IAC and the third one impose the constraint that the vanishing point of the rotation axis direction has pole-polar relationship with the vanishing line of any plane orthogonal to the axis of rotation. We *propose* to also look at eigendecomposition of the line homography i.e. $\mathbf{H}_{21}^T$. It also has one real eigenvector corresponding to a real eigenvalue, and two complex ones corresponding to a pair of complex conjugate eigenvalues ($\mathbf{l_I}$ and $\mathbf{l_J}$), as shown in Fig. 1. Let $\mathbf{a}_y \sim [0\ 1\ 0]^T$ be the axis of rotation for a panning camera. By definition this axis must be invariant to panning, i.e. $\mathbf{R}_{21}^T \mathbf{a}_y = \mathbf{R}_{12}\mathbf{a}_y = \mathbf{a}_y$. Since the infinite homography $\mathbf{H}_{21}$ is a conjugate rotation matrix,

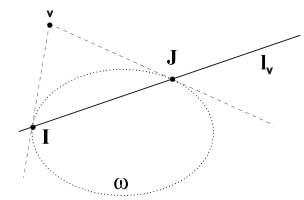


Figure 1. Constraints on IAC induced by the infinite homography.

we have

$$\mathbf{H}_{21}^T \mathbf{K}_1^{-T} \mathbf{a}_y \sim \mathbf{K}_2^{-T} \mathbf{R}_{21}^T \mathbf{a}_y \qquad (4)$$
$$\sim \mathbf{K}_2^{-T} \mathbf{a}_y \qquad (5)$$

Therefore, the vanishing line of the pencil of planes perpendicular to the axis of rotation is also given by $\mathbf{K}_2^{-T}\mathbf{a}_y$. Thus $\mathbf{l_I}$ and $\mathbf{l_J}$ may be viewed as the imaged vanishing lines of some imaginary planes that intersect the absolute conic at the circular points. As a result, the four constraints imposed by the infinite homography on the IAC are encoded in the following three homogeneous equations:

$$\mathbf{l_v} \sim \mathbf{K}_1^{-T} \mathbf{a}_y \sim \boldsymbol{\omega}\mathbf{v}, \quad \mathbf{l_I} \sim \boldsymbol{\omega}\mathbf{I}, \quad \mathbf{l_J} \sim \boldsymbol{\omega}\mathbf{J} \quad (6)$$

To see what happens when the rotation degenerates note that these equations are linear in $\boldsymbol{\omega}$, and upon taking cross-products of both sides as usual [5], they can reduce to a homogeneous equation of the form

$$\mathbf{A}\mathbf{c}_{\boldsymbol{\omega}} = 0 \qquad (7)$$

where $\mathbf{c}_{\boldsymbol{\omega}}$ is the vector of unknown components of IAC arranged in some order. When the rotation is general it can be shown that $\mathbf{A}$ has a one dimensional null space representing the solution to the four unknowns of $\boldsymbol{\omega}$. However, when the rotation *degenerates* to pure pan, or pure tilt the null space becomes 2-dimensional [5], and only two independent constraints can be imposed on the IAC from the set of equations in (6). In particular, one of the constraints applies directly to the principal point:

Proposition 1 *In a zero-skew camera, for pure pan the principal point lies on the vanishing line of the pencil of planes that are perpendicular to the axis of rotation.*

To demonstrate this, denote the principal point by $\mathbf{p} \sim [u_0\ v_0\ 1]^T$. It follows that

$$\mathbf{a}_y^T \mathbf{K}_2^{-1} \mathbf{p} = \mathbf{a}_y^T \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} = 0 \qquad (8)$$

which proves the result being sought.

Remark: The above propositions hold for pure tilt if we simply exchange the role of u_0 with v_0.

In summary, under degenerate rotation the eigenvector $\mathbf{l_v}$ corresponding to the real eigenvalue of $\mathbf{H}_{21}^T$ provides one constraint on the location of the principal point in the form

$$\mathbf{p}^T \mathbf{l_v} = 0 \qquad (9)$$

It is important to note that (9) does not hold under general rotation.

In order to solve for a camera model under pure pan and zoom from a minimum set of two images, we resort to a solution based on direct construction of a set of homogeneous equations. For this purpose, we first verify that under pure pan and zoom the imaged circular points of the plane perpendicular to the axis of rotation will become of the form

$$\begin{bmatrix} a \pm ib \\ v_0 \\ 1 \end{bmatrix} \tag{10}$$

where a and b can be written in terms of the unknown intrinsic parameters and the panning angle. Therefore the real and imaginary parts of the circular points may be used directly to impose constraints on the intrinsic parameters and the rotation angle. On the other hand, we can also construct additional homogeneous equations directly from (2) as follows:

Let $\mathbf{H}_{21} = [\mathbf{h}_1^T, \mathbf{h}_2^T, \mathbf{h}_3^T]^T$, $\mathbf{K}_1 = [\mathbf{k}_{11}^T, \mathbf{k}_{12}^T, \mathbf{k}_{13}^T]^T$, $\mathbf{R}_{21} = [\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3]^T$, and $\mathbf{K}_2 = [\mathbf{k}_{21}, \mathbf{k}_{22}, \mathbf{k}_{23}]$, where $\mathbf{H}_{21}$ and $\mathbf{K}_1$ are expressed in terms of their rows, and $\mathbf{R}_{21}$ and $\mathbf{K}_2$ are expressed in terms of their columns. We can then write the following set of homogeneous equations

$$\mathbf{h}_i^T \mathbf{k}_{2j} \sim \mathbf{k}_{1i}^T \mathbf{r}_j, \;\; i, j = 1, ..., 9 \tag{11}$$

The above equations together with the two constraints derived from the circular points provide only 5 independent constraints on the unknown rotation angle and the intrinsic parameters. Unfortunately, unlike the general case described earlier, for pure panning and zooming it is not possible to establish a constraint on the aspect ratio λ. Therefore, assuming that the aspect ratio is known (e.g. $\lambda = 1$), and that except for the focal length all other intrinsic parameters remain invariant, our constraints lead to low order polynomials, which can be readily solved. Therefore, our solution provides four unknown intrinsic parameters (other than λ and γ) and the rotation angle from only two images for pure panning under variable focal length and zero skew.

Pure Tilt: The case for pure tilt is quite similar to pure pan, with minor differences. In particular, as in pure pan, it can be proved that for pure tilt and zooming the principal point must lie on the vanishing line of the pencil of planes that are perpendicular to the axis of rotation. This provides a constraint similar to (9) on the principal point of the camera. Also, the real and the imaginary parts of the imaged circular points depend on the intrinsic parameters and the rotation angle as before, and can be used to impose constraints on the unknown parameters. However, the construction in (11) is somewhat different for the case of pure tilt, because the infinite homography in the case of pure tilt is of the form

$$\mathbf{H}_{2,1} \sim \begin{bmatrix} 1 & h_{12} & h_{13} \\ 0 & h_{22} & h_{23} \\ 0 & h_{32} & h_{33} \end{bmatrix} \tag{12}$$

providing only 5 equations. Again, it can be shown that in the case of pure tilt, none of the above constraints depends on the camera aspect ratio λ. As a result, it is not possible to recover λ for a purely tilting and zooming camera. Therefore, our solution provides again four unknown intrinsic parameters (i.e. the two focal lengths, and the principal point) plus the rotation angle from only two images for pure tilting under zero skew and variable focal length.

Cascading degenerate cases: One interesting and practical solution for the degenerate case occurs when the camera first pans and then tilts (or vice versa), leading to a minimum case of three images, with the corresponding infinite homographies $\mathbf{H}_{21}$ and $\mathbf{H}_{32}$. In such case, the principal point can be recovered immediately using

$$\mathbf{p} \sim \mathbf{l}_{\mathbf{v}}^{21} \times \mathbf{l}_{\mathbf{v}}^{32} \tag{13}$$

where $\mathbf{l}_{\mathbf{v}}^{21}$ and $\mathbf{l}_{\mathbf{v}}^{32}$ are the eigenvectors corresponding to the real eigenvalues of $\mathbf{H}_{21}^T$ and $\mathbf{H}_{32}^T$. Therefore, the problem would immediately reduce to the simple case of known principal point, which in most auto-calibration methods, including ours, simplifies the remaining set of equations. This scenario can be, for instance, used in a network of PTZ cameras at the cold start, for determining the principal point once and use it throughout the operation of the network, assuming that it remains invariant. Note also that in this case our method recovers all camera parameters including the aspect ratio, since the first and the third image have general rotation, although the other two pairs of combinations are degenerate.

4 Experimental Results

Due to space considerations, we show results on synthetic and real data for pure pan motion of the cameras only, the results for the tilt case should be similar.

Synthetic Data: We performed detailed experimentation on the effect of noise on camera parameter estimation over 1000 independent trials. For this purpose, a point cloud of 1000 random points [1] was produced inside a unit cube to generate image point correspondences while arbitrarily selecting the rotation angles. Simulated camera has a focal length of 1000, aspect ratio of $\lambda = 1.5$, skew $\gamma = 0$, and the principal point at $(u_0, v_0) = (512, 384)$, for image size of 1024×768.

(a)

Combo	θ_y	u_0	v_0	f_1	f_2
c_1	0.97	196.185	206.07	682.07	678.01
c_2	0.91	120.5	187.91	637.81	635.98
c_3	1.02	195.35	210.89	785.97	774.81
c_4	0.88	124.24	147.58	662.13	659.43
c_5	1.26	187.33	198.13	759.02	756.57
c_6	0.8	128.92	179.62	768.4	765.42
c_7	1.08	130.96	188.92	662.13	659.43
c_8	0.93	135.49	155.49	786.22	778.48
c_9	1.16	165.89	182.02	786.22	778.48
c_{10}	0.89	141.03	194.5	673.24	668.83
c_{11}	1.19	118.52	198.93	666.35	660.04
c_{12}	0.94	196.9	198.13	756.17	755.07
c_{13}	1.05	184.08	181.48	770.77	775.22
c_{14}	0.87	175.58	200.17	662.59	660.79
c_{15}	0.97	153.52	162.65	797.4	790.03
Mean	0.99	156.97	186.17	723.77	719.77
M. Dev.	0.08	29.28	9.29	41.23	34.95

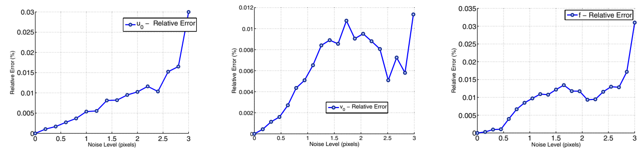
(b)

Figure 2. Sample images from pan sequence. Estimated parameters and their statistics.

The results are shown in Figure 3, where we gradually add a Gaussian noise of up to 3 pixels. Figure 3(a) shows the relative error in u_0, which is found to be less than 0.2% for a noise of up to 3 pixels. Similarly, noise for the v_0 and f is also very low. Thus, the error in the proposed estimated method is very low for all the estimated parameters.

Real Data: Several experiments are performed on real data. The data was obtained by a SONY® SNC-RZ30N PTZ camera with an image resolution of 320×240. Hence the ground truth rotation angles are known. Image features and correspondences are obtained by using the SIFT algorithm [7]. In order to evaluate our results, we use an approach similar to [11], i.e. use the uncertainty associated with the estimated intrinsic parameters characterized by their median deviation over many images, while taking into account the ground truth rotation angles. We deliberately keep f_1 and f_2 same so that we can estimate the accuracy of parameters estimations in the absence of ground truth for intrinsic camera parameters.

Around 15 images were captured while *panning* the camera. The rotation between the successive frames is 1°. In order to further investigate the stability of the proposed method, we apply it to all the successive image pairs (14 combinations out of the 15 images). The results are shown in Figure 2(b). A few of the images are shown in Figure 2(a). The second column depicts the estimated rotation angles to be $.99^\circ$, which is almost equal to the ground truth rotation angle. Camera zoom remained constant in the sequences; hence column 5 and 6 i.e. f_1 and f_2 are very close to each other. The results also demonstrate low median deviation for the estimated parameters.



(a) u_0 (b) v_0 (c) f

Figure 3. Performance vs. Noise Level: Relative Error in estimated parameters averaged over 1000 independent trials.

5 Concluding Remarks

We propose a novel PTZ camera calibration method that is well suited for the degenerate cases of pure pan and pure tilt. The proposed method has some advantages over the current state of the art PTZ calibration algorithm proposed by Agapito et al.[1]. Pure pan or tilt motion are unstable cases for their method. Therefore, using $\gamma = 0$ and assuming a known λ, they are able to solve for fewer parameters. We, on the other hand, are able to solve for 4 intrinsic parameters (i.e. f_1, f_2, u_o, v_o) and the rotation angle (θ_x or θ_y) using only an image pair. We perform experiments on real and synthetic data, and demonstrate the usefulness of the proposed method.

References

[1] L. D. Agapito, E. Hayman, and I. Reid. Self-calibration of rotating and zooming cameras. *Int. J. Comput. Vision*, 45(2):107–127, 2001.

[2] O. Faugeras, T. Luong, and S. Maybank. Camera self-calibration: theory and experiments. *Proc. of ECCV*, pages 321–334, 1992.

[3] J. Frahm and R. Koch. Camera calibration with known rotation. *Proc. IEEE ICCV*, pages 1418–1425, 2003.

[4] R. I. Hartley. Self-calibration of stationary cameras. *Int. J. Comput. Vision*, 22(1):5–23, 1997.

[5] R. I. Hartley and A. Zisserman. *Multiple View Geometry in Computer Vision*. Cambridge University Press, ISBN: 0521540518, second edition, 2004.

[6] D. Liebowitz and A. Zisserman. Combining scene and auto-calibration constraints. *Proc. IEEE ICCV*, pages 293–300, 1999.

[7] D. G. Lowe. Distinctive image features from scale-invariant keypoints. *International Journal of Computer Vision*, 6(2):91–110, 2004.

[8] T. Moons, L. Gool, M. Proesmans, and E. Pauwels. Affine reconstruction from perspective image pairs with a relative object-camera translation in between. *IEEE Trans. Pattern Anal. Mach. Intell.*, 18(1):77–83, 1996.

[9] Y. Seo and K. Hong. About the self-calibration of a rotating and zooming camera: Theory and practice. *Proc. IEEE ICCV*, pages 183–189, 1999.

[10] R. Tsai. A versatile camera calibration technique for high-accuracy 3D machine vision metrology using off-the-shelf tv cameras and lenses. *IEEE J. of Robotics and Automation*, 3(4):323–344, 1987.

[11] Z. Zhang. A flexible new technique for camera calibration. *IEEE Trans. Pattern Anal. Mach. Intell.*, 22(11):1330–1334, 2000.