

Planar Conic Based Camera Calibration

Changjiang Yang[†] Fengmei Sun^{†‡} Zhanyi Hu[‡]

[†]Department of Comp. Sci. and Eng., Michigan State University, East Lansing, MI 48824

[‡]National Lab. of Pattern Recognition, Chinese Academy of Sciences, Beijing 100080, China

[‡]College of Fundamental Research, North China University of Technology, Beijing 100041, China
yangcha1@cse.msu.edu

Abstract

Inspired by the technique proposed by Zhang in [13], we proposed a new camera calibration technique, which only requires observing three or more planar con-centric conics at a few (at least two) different orientations. All computations involved are linear matrix manipulations. Compared with the classical techniques where an expensive calibration pattern is commonly used, our technique is easy to implement and more flexible. Using Conics also simplifies the problem of correspondence. Both computer simulation and real data are used to test the proposed technique.

1. Introduction

Camera calibration is an essential step in 3D computer vision to obtain a metric reconstruction from 2D images [1, 2]. With the rapid development in computer technology, network and multi-media, cameras are becoming more and more cheap and popular. A typical computer user who is not expert in computer vision will not be willing to invest money for expensive calibration equipment. Therefore, flexibility, robustness and low cost are critical factors for camera calibration [3, 6, 11].

Recently Zhang [13] presented a flexible camera calibration method which utilize the planar pattern. The technique only requires the camera to observe a planar pattern shown at a few (at least two) different orientations. Either the camera or the planar pattern can be freely moved. The motion need not be known. The pattern can be printed on a laser printer and attached to a planar surface. Sturm and Maybank [9] proposed a similar plane-base calibration method and the singularity of the method was discussed. Triggs [10] developed a more flexible self-calibration technique which needs 5 views of a planar scene. But the computational complexity is high.

In this paper, we use the correspondences of the conics rather than points to calibrate camera. Because conics

are compact and global primitive, easy to match correspondences, our method can gain considerable robustness and flexibility [5, 8, 7]. At the mean time, a closed-form solution can be obtained and the computations required are only the matrix manipulations.

2. Camera Model and Homography

An image point is denoted by $\mathbf{m} = [u, v]^T$. A 3D point is denoted by $\mathbf{M} = [X, Y, Z]^T$. The homogeneous point coordinates of them are denoted by $\tilde{\mathbf{m}} = [u, v, 1]^T$ and $\tilde{\mathbf{M}} = [X, Y, Z, 1]^T$ respectively. The camera is pin-hole model: the relationship between 3D point \mathbf{M} and image projection \mathbf{m} is given by

$$s\tilde{\mathbf{m}} = K[R \ \mathbf{t}]\tilde{\mathbf{M}} \quad (1)$$

where s is a non-zero scale factor. R , \mathbf{t} are rotation and translation respectively. K , called the camera intrinsic matrix, is given by

$$K = \begin{bmatrix} \alpha & \gamma & u_0 \\ 0 & \beta & v_0 \\ 0 & 0 & 1 \end{bmatrix}$$

Without loss of generality, we assume that the model plane is on $Z = 0$ of the world coordinate system. Denoting the i^{th} column of the rotation matrix R by \mathbf{r}_i , we have

$$s \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = K \begin{bmatrix} \mathbf{r}_1 & \mathbf{r}_2 & \mathbf{t} \end{bmatrix} \begin{bmatrix} X \\ Y \\ 1 \end{bmatrix}$$

Therefore, a point on the model plane \mathbf{M} and its image \mathbf{m} is related by a homography H

$$s\tilde{\mathbf{m}} = H\tilde{\mathbf{M}} \quad (2)$$

where $\tilde{\mathbf{M}} = [X, Y, 1]^T$ is reduced notation of point \mathbf{M} , $H = \lambda K \begin{bmatrix} \mathbf{r}_1 & \mathbf{r}_2 & \mathbf{t} \end{bmatrix}$ is a 3×3 matrix, λ is a scale factor. Denoting $H = \begin{bmatrix} \mathbf{h}_1 & \mathbf{h}_2 & \mathbf{h}_3 \end{bmatrix}$, we have

$$\begin{bmatrix} \mathbf{h}_1 & \mathbf{h}_2 & \mathbf{h}_3 \end{bmatrix} = \lambda K \begin{bmatrix} \mathbf{r}_1 & \mathbf{r}_2 & \mathbf{t} \end{bmatrix}$$

Clearly \mathbf{t} does not lie on the plane $\mathbf{r}_1, \mathbf{r}_2$, so $\det(H) \neq 0$.

3. Camera Calibration from Planar Conics

Given a planar conic C and its image Q , we have

$$\widetilde{\mathbf{M}}^T C \widetilde{\mathbf{M}} = 0 \quad (3)$$

$$\widetilde{\mathbf{m}}^T Q \widetilde{\mathbf{m}} = 0 \quad (4)$$

Substitute Eq.(2) into Eq.(4), yields

$$\widetilde{\mathbf{M}}^T H^T Q H \widetilde{\mathbf{M}} = 0 \quad (5)$$

Comparing Eq.(3) and Eq.(5), we obtain

$$s'C = H^T Q H \quad (6)$$

If we assume $\det(C) = \det(Q)$, and $\det(H) = 1$, then from Eq.(6) we can get the following equation:

$$C = H^T Q H \quad (7)$$

For the sake of convenience, we can easily print out the concentric conics on a paper using laser printer. Here we assume the calibrating conics are concentric. Without loss of generality, we assume the center of the conics locate at the origin of the coordinates system, then any of the conics can be expressed as

$$C = \begin{bmatrix} \frac{1}{a^2} & 0 & 0 \\ 0 & \frac{1}{b^2} & 0 \\ 0 & 0 & -1 \end{bmatrix}$$

From Eq.(7) we have

$$Q^{-1} = H C^{-1} H^T \quad (8)$$

Substitute $H = [\mathbf{h}_1, \mathbf{h}_2, \mathbf{h}_3]$ into Eq.(8), yields

$$Q^{-1} = a^2 H_{x1} + b^2 H_{x2} - H_{x3} \quad (9)$$

where $H_{x1} = \mathbf{h}_1 \mathbf{h}_1^T$, $H_{x2} = \mathbf{h}_2 \mathbf{h}_2^T$, $H_{x3} = \mathbf{h}_3 \mathbf{h}_3^T$. To solve the three unknown matrices H_{x1} , H_{x2} , H_{x3} , we need at least three concentric conics.

From H_{x1} , H_{x2} , H_{x3} , we can easily get the absolute values of vectors $\mathbf{h}_1, \mathbf{h}_2, \mathbf{h}_3$ by computing the square root of the diagonal elements of the corresponding three matrices. The signs can be determined from the upper triangular submatrix of H_{x1} , H_{x2} , H_{x3} . But this method is proved sensitive to the noise. Instead, we employ the matrix factorization technique [4] to solve $\mathbf{h}_1, \mathbf{h}_2, \mathbf{h}_3$.

Here we take \mathbf{h}_1 as an example to display how to solve $\mathbf{h}_1, \mathbf{h}_2, \mathbf{h}_3$. Because $H_{x1} = \mathbf{h}_1 \mathbf{h}_1^T$ is a rank 1 symmetric matrix, it has the following singular value decomposition:

$$H_{x1} = U \Sigma V$$

where $\Sigma = \text{diag}(\sigma_1, 0, 0)$ is the singular value matrix, and U is orthogonal. Because of errors, the computed singular values $\hat{\sigma}_i$, ($i > 1$) are nonzero, but far smaller than $\hat{\sigma}_1$. We can set them to zero, so

$$H_{x1} = \mathbf{h}_1 \mathbf{h}_1^T = \begin{bmatrix} \sqrt{\hat{\sigma}_1} & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} \sqrt{\hat{\sigma}_1} & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

Let

$$G = U \begin{bmatrix} \sqrt{\hat{\sigma}_1} & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} = \begin{bmatrix} g_{11} & 0 & 0 \\ g_{21} & 0 & 0 \\ g_{31} & 0 & 0 \end{bmatrix}$$

then we can obtain $\mathbf{h}_1 = [g_{11}, g_{21}, g_{31}]$. Similarly we can solve \mathbf{h}_2 and \mathbf{h}_3 .

Once H is solved, from the fact that \mathbf{r}_1 and \mathbf{r}_2 are orthonormal, we have

$$\begin{aligned} \mathbf{h}_1^T K^{-T} K \mathbf{h}_2 &= 0 \\ \mathbf{h}_1^T K^{-T} K \mathbf{h}_1 &= \mathbf{h}_2^T K^{-T} K \mathbf{h}_2 \end{aligned}$$

From above two equations, we can get two constraints on K . If $n(n \geq 3)$ images of the model plane are observed, we will have in general a unique solution up to a scale factor. More details can be found in [13].

4. Experimental Results

The proposed technique has been tested on both simulated data and real images. For the simulated experiments, the parameters of the simulated camera are: $\alpha = 1250$, $\beta = 1250$, $\gamma = 1.0908$, $u_0 = v_0 = 500$, the image resolution is 1000×1000 . The orientation of the model plane is represented by a 3D vector \mathbf{r} which is parallel to the rotation axis and whose magnitude is equal to the rotation angle. Its position is represented by \mathbf{t} (unit in cm). The parameters of the first set of experiments are: $\mathbf{r}_1 = [20, 0, 0]^T$, $\mathbf{t}_1 = [-9, -12.5, 100]^T$, $\mathbf{r}_2 = [0, 20, 0]^T$, $\mathbf{t}_2 = [-9, -12.5, 110]^T$, $\mathbf{r}_3 = [-30, -30, 15]^T$, $\mathbf{t}_3 = [-10.5, -12.5, 120]^T$. There are three concentric ellipses on the model plane. One is a circle with radius $30cm$, the half long axis and short axis of the other two ellipses are $30cm$, and $20cm$, the intersection angle of them are 90° . Each of them consists of 1000 data points uniformly distributed along the arc. Gaussian noise with mean 0 and standard deviation σ is added to the projected image points. The computed camera parameters are then compared with the ground truth. We measure the absolute errors for u_0 and v_0 , and relative errors for the ratios of α and β (equivalent to effective focal length on x axis and y axis). We vary the noise level from 0.1 pixels to 3.0 pixels. For each noise level, 100 independent trials are conducted and the results shown in Fig.1.

The second set of simulating experiments are to investigate the performance with the respect to the number of the images. The setup of the first three images is same as the first experiment. From the fourth images, we randomly choose a rotation axis from a unit sphere, and rotation angle is 30° . We vary the number of the image from 2 to 20. For each number of images, 100 trials of independent plane orientations and noise with mean 0 and standard deviation 0.5 pixels are conducted. The results are shown in Fig.2.

For the real image experiments, the camera to be calibrated is an off-the-shelf CCD digital camera. The image resolution is 1280×960 . We printed three concentric circles with radiuses 9.5cm , 5cm , 3cm , two concentric conics with half long axis and half short axis 12.5cm , 7.5cm , and 9cm , 6cm , as shown in Fig.3. Five images were taken from five distinct orientations. The conic of the image was detected by Randomized Hough Transform and fitted by Orthogonal Distance Regression (shown in Fig.4). We applied our calibration algorithm to the images. Firstly, we used the first 2, 3, 4, or all 5 images to calibrate the camera. The results are list in Table 1. We also used the four of the five images to calibrate the camera. The results are shown in Table 2. Both results are very consistent with each other.

In order to evaluate the performance of the proposed technique, we verified the calibrated camera parameters with a stereo vision experiment. Two images of a calibration pattern (see Fig.5) were taken by the same camera as used above for the calibration. Mainly two sides of the calibration pattern are visible in images. We manually picked 9 point matches on each side (indicated by cross), and use the structure-from-motion method proposed by Zhang [12] to reconstruct the calibration pattern with the above calibrated camera intrinsic parameters. Two views of the reconstructed calibration pattern are shown in Fig.6, from which we can see the reconstructed points are indeed coplanar. The computed angle between the two reconstructed planes is 91.1148031866572° , which is very close to the ground truth (90°).

5. Conclusions

We proposed a planar conic based calibration technique, which only requires the camera to observe a model plane with 3 or more than 3 concentric conics from a few orientations. The motion does not need to be known. Only matrix computations are needed once the conic parameters are obtained, and there is no need to refine the results by nonlinear optimizations. Using conics as geometric primitives, the proposed technique gains considerable robustness, and greatly reduces the difficulty of the primitive correspondences. Both computer simulation and real data have been used to test the validity of our new calibration technique.

References

- [1] D.C.Brown, Close-range camera calibration, *Photogrammetric Engineering*, 37(8):855-866, 1971.
- [2] W.Faig, Calibration of close-range photogrammetry system: Mathematical formulation, *Photogrammetric Engineering and Remote Sensing*, 41(12):1479-1486, 1975.
- [3] O.D.Faugeras, Q.T.Luong and S.J.Maybank, Camera Self-calibration: Theory and Experiments, In *Proc. European Conference on Computer Vision*, pp.321-334, Santa-Margherita, Italy, 1992.
- [4] G. Golub and C. van Loan, *Matrix Computations*, The John Hopkins University Press, Baltimore, Maryland, 3 edition, 1996.
- [5] S.Ma, Conics-based stereo, motion estimation, and pose determination, *International Journal of Computer Vision*, 10(1):7-25, 1993.
- [6] S.J.Maybank and O.D.Faugeras, A Theory of Self-calibration of a Moving Camera, *International Journal of Computer Vision*, 8(2):123-151, 1992.
- [7] J.L.Mundy and A.Zisserman, Eds., *Geometric Invariance in Computer Vision*, MIT Press, Cambridge, MA.
- [8] L.Quan, Conic reconstruction and correspondence from two views, *IEEE Trans. Pattern Analysis and Machine Intelligence*, 18(2):151-160, 1996.
- [9] P.Sturm and S.Maybank, On plane-based camera calibration: A general algorithm, singularities, applications, In *Proc. CVPR*, pp.432-437, Fort Collins, Colorado, 1999.
- [10] B.Triggs, Autocalibration from planar scenes, In *Proc. ECCV'98*, pp.89-105, Freiburg, Germany, June 1998.
- [11] 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 Robot. and Auto.*, 3(4):323-344, Aug. 1987.
- [12] Z.Zhang, Motion and structure from two perspective views: From essential parameters to Euclidean motion via fundamental matrix, *Journal of the Optical Society of America A*, 14(11):2938-2950, 1997.
- [13] Z.Zhang, A flexible new technique for camera calibration, *Technical Report MSR-TR-98-71*, Microsoft Corporation, 1998.

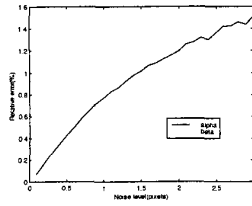


Fig.1 Error vs. the noise level of the image points

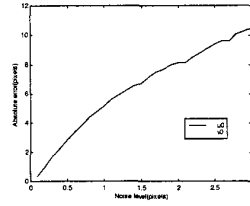


Fig. 2 Error vs. the number of the images

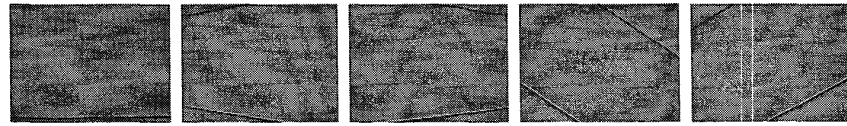
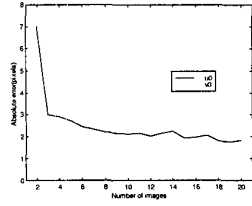
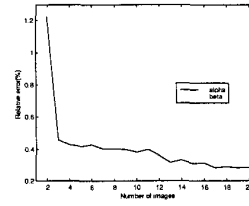


Fig. 3 Five images of the model plane

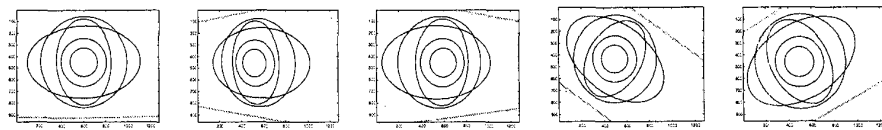


Fig.4 The detected conics of the five images

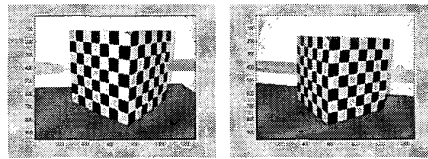


Fig. 5 Two images of the calibration pattern

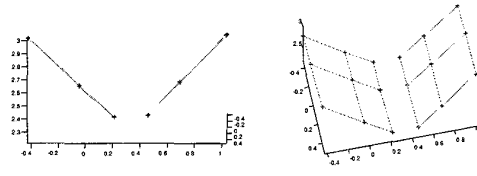


Fig. 6 Two views of the reconstructed calibration pattern

Table 1 Results with real data of 2 through 5 images

No. of images	2 images	3 images	4 images	5 images
α	1410.23419860267	1379.74660762767	1383.15984621025	1385.42091609659
β	1415.06194927586	1390.9277406454	1393.55533132077	1394.27758306234
γ	0	1.3954381033933	1.2110201911172	1.2872398332345
u_0	532.598317791235	540.329411003444	538.416307482848	538.683419549089
v_0	434.082610510191	430.780137322556	432.236975787966	432.934509446601

Table 2 Variation of the calibration results among all quadruples of images

Quadruple	α	β	γ	u_0	v_0
(1234)	1383.15984621025	1393.55533132077	1.2110201911172	538.416307482848	432.236975787966
(1235)	1379.98724247956	1390.75663401891	1.3203001422300	540.447318509619	429.366751940351
(1245)	1381.00720033379	1391.93045605753	1.3323443008141	541.162705608795	433.403689937113
(1345)	1380.31810999085	1393.12487911776	1.2936816569949	539.491335565401	431.912312046381
(2345)	1385.65586342016	1396.48894428389	1.3148142393008	540.723326178109	430.126720785189
Mean	1382.02565248692	1393.17124895977	1.2944321060914	540.048198668954	431.409290099401
Deviation	2.37554669203011	2.15213289696762	0.04868272766093	1.09890682944158	1.63803781761885