Camera Calibration from Profile of Revolution

ZHENG Shunyi  YANG Ling  SUN Mingwei

ABSTRACT  This paper focuses on the problem of calibrating a pinhole camera from images of profile of a revolution. In this paper, the symmetry of images of profiles of revolution has been extensively exploited and a practical and accurate technique of camera calibration from profiles alone has been developed. Compared with traditional techniques for camera calibration, for instance, it may involve taking images of some precisely machined calibration pattern (such as a calibration grid), or edge detection for determining vanish points which are often far from images center or even do not physically exist, or calculation of fundamental matrix and Krupp equations which can be numerically unstable, the method presented here uses just profiles of revolution, which are commonly found in daily life (e.g. bowls and vases), to make the process easier as a result of the reduced cost and increased accessibility of the calibration objects. This paper firstly analyzes the relationship between the symmetry property of profile of revolution and the intrinsic parameters of a camera, and then shows how to use images of profile of revolution to provide enough information for determining intrinsic parameters. During the process, high-accurate profile extraction algorithm has also been used. Finally, results from real data are presented, demonstrating the efficiency and accuracy of the proposed methods.

KEYWORDS  camera calibration; profile of revolution; harmonic homology; symmetry

CLC NUMBER  P246.1

Introduction

Camera Calibration is the process determining the intrinsic parameters (i.e. focal length and principal point) of a camera. It is an essential step for both motion estimation and 3D reconstruction, which are two important tasks in computer vision. It is also an important part of photogrammetry. Consequently, camera calibration has been extensively studied in computer vision and photogrammetry. Classical methods usually involve special complex test fields which must be set on special site and be constructed and maintained with high cost. So classical methods are only used for special tasks. Besides classical method, many new methods have arisen, which often taking images of some precisely machined calibration pattern (such as a calibration grid). These methods do not require direct mechanical measurements on the cameras, and often produce very good results. Nevertheless, they need to design and use highly accurate tailor-made calibration pattern, which are both difficult and expensive to be manufactured.

In this paper, we will introduce a novel technique for camera calibration. The method presented here uses just profiles of revolution, which are commonly found in daily life (e.g. bowls and vases), making the process easier as a result of the reduced cost and increased accessibility of the calibration objects.

1 Existing camera calibration techniques

From the starting in the photogrammetry community to recent years in computer vision, much
work has been done for camera calibration. These techniques can be seen in detail in References [1, 2] (in photogrammetry) and References [3-7] (in computer vision).

Classical calibration techniques [1, 2] in photogrammetry involve full-scale nonlinear optimizations with large number of parameters. These techniques are able to cope with complex camera models and produce accurate results, but require a good initialization and are computationally expensive.

One of most commonly used camera calibration techniques in computer vision is direct linear transformation (DLT) technique, which was introduced in Reference [9]. DLT ignored lens distortion and treated the coefficients of 3 X 4 projection matrix as unknown, so it only involves solving a system of linear equations, which can be done by a linear-squares method. But the accuracy is often not high.

Recently planar scene based calibration methods became more popular and attract more attention [8-14]. In Reference [8], a calibration method with 2D-DLT and bundle adjustment was developed. By using the correspondence between collinearity equations and 2D-DLT it worked out the equation of principal vertical line. Then the initial value of principle point can be obtained with at least two equations of principal vertical lines, and finally bundle adjustment had been used to calculate the final results, with this method good results can be obtained, but high-accurate calibration pattern is necessary and the accuracy of such pattern is directly related with final calibration results.

The methods mentioned above often require a certain number of reference points and both their 3D coordinates and corresponding image coordinates. Self-calibration is another commonly used camera calibration techniques in computer vision. It does not use any calibration object and only required image point correspondences, but it has a large number of parameters to be estimated, resulting in a much harder mathematical problem [15]. More important is that such method is not yet mature [14], because there are many parameters to be estimated, we cannot always obtain reliable results.

Other techniques exist; calibration from vanishing point [16] involves the calculation of vanishing points which are often far from images center or even do not physically exist so that it is very difficult to produce high accurate results; some calibration methods use fundamental matrix and Kruppa equations [11] whose calculation may be numerically unstable.

2 Theoretical background

2.1 Camera model

In this paper, the camera model that we consider is the simple pinhole (perspective) camera. The focal length, aspect ratio, skew and principal point are referred to as the intrinsic parameters [12].

Consider a pinhole (perspective) camera, whose camera calibration matrix $K$ is a 3 X 3 matrix given by [12]

$$
K = \begin{bmatrix}
  f_x & \xi & u_0 \\
  0 & f_y & v_0 \\
  0 & 0 & 1
\end{bmatrix} \begin{bmatrix}
f_x \\
\xi \\
u_0
\end{bmatrix} \begin{bmatrix}
f \\
v_0
\end{bmatrix}
$$

where $f_x$ and $f_y$ are the focal lengths; $\xi$ is the skew which depends on the angle between the image axes; $f$ is the focal length; $a = f_x / f_y$ is the aspect ratio; $(u_0, v_0)$ is the principle point, and it is the point at which the optical axis intersects the image plane (see Fig. 1). $K$ is called intrinsic matrix which is just related to camera’s intrinsic parameters.

![Extrinsic parameters and intrinsic parameters of a camera](image)

Here we assume that the image axes are orthogonal to each other, so $\xi = 0$, and $f_x = f_y$, then $a = 1$, which is known as unit aspect ratio. For most cameras, these assumptions can be met and therefore, are reasonable. At this time, the camera matrix can be expressed with the following form.
Then, we have the projection matrix $P$, which models the pinhole camera. It is given by

$$P = K[R \ t]$$

where $K$ is camera calibration matrix; $t$ is a $3 \times 1$ translation vector and $R$ is $3 \times 3$ rotation matrix. $R$ and $t$ are the extrinsic parameters of camera.

Considering a point $x$, whose 3D coordinates and image coordinates are $(X, Y, Z)$ and $(u, v)$ respectively, and then we have the relationship between 3D coordinates and image coordinates:

$$\begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \begin{bmatrix} X \\ Y \\ Z \end{bmatrix}$$

where $a$ is an arbitrary scale factor.

### 2.2 Properties of surface of revolution

The profile of a surface of revolution is symmetric with respect to the image of its rotation axis. There are three possibilities for the transformation that maps the sides of the profile to each other: if the camera is pointing towards the axis of rotation and has zero skew and aspect ratio one, the transformation is a bilateral symmetry; if the intrinsic parameters are arbitrary but the camera is still pointing towards the rotation axis, the transformation is a skew symmetry; if the camera is pointing away from the axis of rotation, the transformation is a harmonic homology. The images (Fig. 2) below show examples of such transformations.

![Bilateral symmetry](image1.png)  ![Skew symmetry](image2.png)  ![Harmonic homology](image3.png)

**Fig. 2** Symmetric of profile of revolution

In fact, the bilateral symmetry is a particular case of the skew symmetry when the symmetry lines are orthogonal to rotate axis $l_e$. Furthermore, the skew symmetry is a particular case of the harmonic homology, occurring when vanish point $v_e$ is at infinity. The harmonic homology provides important information for camera calibration.

### 2.3 Harmonic homology

Consider a surface of revolution $S$, the image of $S$ taken by a pinhole camera $P$ is a curve $\varepsilon$, as Fig. 2 shows. Let $l_e$ be the image of the axis of rotation of $S$, in the camera $P$. The optical center of $P$ and the axis of revolution define a plane $\Pi$, whose normal direction is $n_e$. The image of the point at infinity in the direction $n_e$ is the vanishing point $v_e$ (see Fig. 3).

If $v_e$ and $l_e$ are represented in homogeneous coordinates as $v_e = \begin{bmatrix} u & v & 1 \end{bmatrix}$ and $l_e = \begin{bmatrix} \cos \theta & \sin \theta & -d \end{bmatrix}^T$, the harmonic homology $W$ is given by:

$$W = I - 2 \frac{v_e l_e^T}{v_e^T l_e}$$

The profile $\varepsilon$ will be invariant to this transformation, which simply maps one side of the profile (with respect to the image of the axis of rotation) to the other.
Now consider any two vectors \( \mathbf{n}_1 \) and \( \mathbf{n}_2 \), parallel to \( \Pi, \mathbf{n}_3 \) and \( \mathbf{n}_4 \), are orthogonal to each other. Then \( \mathbf{n}_3, \mathbf{n}_4, \mathbf{n}_5 \) form a set of three mutually orthogonal directions. By construction, the vanishing points corresponding to the directions of \( \mathbf{n}_1 \) and \( \mathbf{n}_2 \) will lie on \( \ell \). Let \( v_s, v_t, v_u \) be three vanishing points associated with three mutually orthogonal directions \( \mathbf{n}_3, \mathbf{n}_4, \mathbf{n}_5 \) respectively. As shown in Reference [10], under the assumption of zero skew and unit aspect ratio, the vanishing points associated with these three directions can be used to determine the principal point and the focal length of \( P \); it is proved that the principal point will coincide with the orthocenter of the triangle with vertices given by the vanishing points \( v_s, v_t, v_u \), and it follows that the square root of the product of the distances from the orthocenter to any vertex and to the opposite side will give the focal length. The principle point \( u_0 \) of the camera coincides with the orthocenter of the triangle with vertices given by the vanishing points \( v_s, v_t, v_u \), and the focal length is equal to square root of \( d_1 \) and \( d_2 \) (see Fig. 4(a)). As a result, given a harmonic homology \( W \) defined by the vanishing point \( v_s \) and the axis of revolution \( \ell \), the principle point of camera \( P \) will lie on a line \( \ell \), passing through \( v_s \) and perpendicular to \( \ell \) (see Fig. 4(b)).

The harmonic homology \( W \) that maps each side of the silhouette \( \varepsilon \) to its symmetric counterpart is then estimated by minimizing the distances between the original silhouette \( \varepsilon \) and its transformed version \( \varepsilon' = W\varepsilon \). This can be done by sampling \( N \) evenly spaced points \( x_i \) along the silhouette \( \varepsilon \) and optimizing the cost function:

\[
\text{cost}_W(v_s, \ell) = \sum_{i=1}^{N} \text{dist}(W(v_s, \ell)x_i, \varepsilon)^2 \tag{6}
\]

where \( \text{dist}(W(v_s, \ell)x_i, \varepsilon) \) is the orthogonal distance from the transformed sample point \( x'_i = W(v_s, \ell)x_i \) to the original silhouette \( \varepsilon \).

In order to give a good initialization to avoid convergence to local minima, we use bitangents of the silhouette\[^{[12]}\]. Two bitangents are selected manually and make sure that the four bitangent points are homologous in the homology coordinate.

Given two bitangents \( l(p_1, p_2) \) and \( l(q_1, q_2) \) on two sides of the profile \( \varepsilon \) with bitangent points \( p_1, p_2 \) and \( q_1, q_2 \) respectively, the intersection of two bitangents \( l(p_1, p_2) \) and \( l(q_1, q_2) \) and the intersection of the diagonals \( \{l(p_1, q_1), l(p_2, q_2)\} \) give two points \( A \) and \( B \). Points \( A \) and \( B \) define a line for an estimate of \( \ell \). And the point of intersection of the lines \( l(p_1, q_1) \) and \( l(p_2, q_2) \) gives an estimate for the vanishing point \( v_s \) (see Fig. 6). The initialization of \( \ell \) and \( v_s \) from bitangents often provides an excellent initial guess for the optimization problem. This
is good enough to avoid any local minimum and allows convergence to the global minimum in a small number of iterations.

Fig. 6 Initialization of the optimization parameters $l_0$ and $v_0$ from the bitangents lines formed by bitangent points

3.2 Estimation of intrinsic parameters

3.2.1 Vanishing point based method (Method 1)

In this paper, we assume that the camera has zero skew and unit aspect ratio. Then the line $l_0$ passing through the principal point $(u_0, v_0)$ and the vanishing point $v_0$ will be orthogonal to the image of the revolution axis $l_0$. Let $v_0 = [v_1, v_2, v_3]^T$ and $l_0 = [l_1, l_2, l_3]^T$. The line $l_0$ can be given by:

$$l_0 = \frac{1}{\sqrt{(l_2v_3)^2 + (l_1v_3)^2}} \begin{bmatrix} l_2v_3 \\ l_1v_3 \\ l_1v_2 - l_2v_1 \end{bmatrix}$$ (7)

When more than two lines are available, the principal point $(u_0, v_0)$ can be estimated by a linear least-squares method:

$$\begin{bmatrix} l_{11}^T \\ l_{21}^T \\ \vdots \\ l_{M1}^T \end{bmatrix} \begin{bmatrix} u_0 \\ v_0 \end{bmatrix} = \begin{bmatrix} f^2 + u_0^2 \\ u_0v_0 \\ v_0 \end{bmatrix} l_0^T$$ (8)

where $M \geq 2$ is the total number of lines and $\alpha$ is a scale factor. The estimated principle point $(u_0, v_0)$ is then projected onto each line $l_{ni}$ orthogonally as $x_{ni}$, and the focal length $f$ will be given by:

$$f = \frac{1}{M} \sum_{i=1}^{M} \sqrt{\text{dist}(x_{ni}, v_0) \times \text{dist}(x_{ni}, l_0)}$$ (9)

where dist$(x_{ni}, v_0)$ is the distance between $x_{ni}$ and $v_0$, and dist$(x_{ni}, l_0)$ is the orthogonal distance from $x_{ni}$ to the image of the revolution axis $l_0$.

3.2.2 Homology based method (Method 2)

As described above, $l_0$ and $v_0$ are respectively the image axis of revolution and the vanishing point of the normal direction of the plane passing through $l_0$ and the camera center. And they are strictly related to the calibration matrix $K$, which embeds the information about the internal camera parameters. In particular it holds $v_0 = w_0$ where $w_0 = KK^T$ is the image of absolute conic $\Gamma$. So we can get:

$$v_0 = \begin{bmatrix} f^2 + u_0^2 \\ u_0v_0 \\ f^2 + v_0^2 \\ v_0 \end{bmatrix} l_0$$ (10)

From known $l_0$ and $v_0$, we can calculate the focal length $f$ and principle point $(u_0, v_0)$ by solving Eq. (10), which is an ordinary non-linear optimization problem.

4 Experiments

The results will be given in this section. The experiments were done under the assumption of zero skew and unit aspect ratio, and on two real cameras.

4.1 Ground truth

Two cameras have been used in this experiment. One camera is Rollei metric camera D70 and the camera resolution is $1280 \times 1024$. The ground truth for this camera’s intrinsic parameter was fixed when produced. The other camera is a CCD camera, which image size is $1300 \times 1030$. And the ground truth for this camera’s intrinsic parameter is obtained by a method based on 2D-DLT and bundle adjustment. Because the accuracy of such 2D-DLT method is very high, it is reasonable to deem its results to be the ground truth. The parameters of Rollei D70 and CCD camera are shown in Table 1.

Table 1 Ground truth of two cameras/pixel

| Parameters | $f$ | $u_0$ | $v_0$ |
|------------|-----|-------|-------|
| Ground truth Rollei D70 | 1080 | 617 | 478 |
| Ground truth CCD | 2434 | 648 | 499 |
4.2 Results on real data

Two sets of real images have been used for calibration of the Rollei metric camera D70 and CCD camera respectively. One set consists of three images of a vase taken by CCD camera, and the other consists of three images of a jar taken by Rollei metric camera D70 (see Fig. 7). The jar is bigger than the vase. This is because the focus length of two cameras greatly different, and different objects used can make the image of revolution compatible with the whole image size. Fig. 8 shows the estimated principle points of the two cameras by use of Method 1. In which, extracted profiles and estimated axes of revolution represented by grey lines, and principle point estimated by Method 1 represented by black cross. Fig. 8 (a) is calibration of Rollei D70. Fig. 8(b) is calibration of CCD camera. Fig. 9 compares the principle points of two cameras estimated by Method 1 and Method 2 respectively. In which black cross is the principle point calculated by Method 1 and grey cross is the principle point calculated by Method 2. Fig. 9 (a) is principle points of Rollei D70. Fig. 9(b) is principle points of CCD camera.

The results of calibration from the images using the algorithm presented above are shown in Table 2. In Table 3 the errors of the two methods are listed. The errors shown are the percentage error of each parameter relative to the ground truth.

![Fig. 7 Images of object taken by cameras](image1)

![Fig. 8 Camera calibration results by Method 1](image2)

![Fig. 9 Comparison of Method 1 with Method 2](image3)

Table 2 Results of calibration from the images using the algorithm presented in this paper/pixel

| Parameters        | Ground truth | Method 1 | Method 2 |
|-------------------|--------------|----------|----------|
| Rollei metric camera D70 |              |          |          |
| $f$               | 1 080        | 1 092    | 1 090    |
| $u_0$             | 617          | 628      | 603      |
| $v_0$             | 478          | 480      | 479      |
| $f$               | 2 434        | 2 410    | 2 420    |
| CCD camera        |              |          |          |
| $u_0$             | 648          | 674      | 613      |
| $v_0$             | 499          | 465      | 485      |

Table 3 Errors of the two methods

| Camera            | Method | $f$      | $u_0$    | $v_0$    |
|-------------------|--------|----------|----------|----------|
| Rollei metric camera D70 | Method 1 | 1.11     | 1.78     | 0.42     |
|                   | Method 2 | 0.93     | 2.27     | 0.21     |
| CCD camera        | Method 1 | 0.99     | 4.01     | 6.81     |
|                   | Method 2 | 0.58     | 5.40     | 2.81     |
5 Conclusions

By using symmetry property of profile of revolution, a practical technique for camera calibration has been developed. The use of a surface of revolution makes the calibration process easier, for it does not require the use of any precisely adjusted device such as calibration grid. Besides, a surface of revolution can be easily found in daily life, this makes camera calibration cheaper.

The methods introduced here are promising as demonstrated by the experiment results on real data. The focal lengths were estimated in high accuracy, having errors of only around 1% with respect to the ground truth. \( u_0 \) and \( v_0 \) were also estimated in high accuracy. Generally, Method 2 works better than Method 1.

Now we are considering a statistically optimal way, which integrates all the information provided by the silhouettes to estimate the intrinsic parameters and the harmonic homologies simultaneously. Better results are expected and will be reported in near future.

REFERENCES

[1] Brown D C (1971) Close-range camera calibration.[J]. Photogrammetric Engineering, 37(8), 855-855
[2] Faig W (1975) Calibration of close-range photogrammetry systems: mathematical formulation.[J]. Photogrammetric Engineering and Remote Sensing, 41 (12): 1 479-1 486
[3] Gennery D (1979) Stereo-camera calibration.[C]. The 10th Image Understanding Workshop, Los Angeles, CA
[4] Ganapathy S (1984) Decomposition of transformation matrices for robot vision.[J]. Pattern Recognition Letters, 2, 401-412
[5] Faugeras O, Toscani G (1985) The calibration problem for stereo.[C]. IEEE Conference Computer Vision and Pattern Recognition, Miami Beach. Florida
[6] Wei Guoqing, Ma Songde (1993) A complete two-plane camera calibration method and experimental comparisons.[C]. The 4th International Conference

Computer Vision, Berlin
[7] Zhang Zhengyou (2004) Camera calibration with one-dimensional objects[J]. IEEE Trans on Pattern Analysis and Machine Intelligence, 26(7):892-899
[8] Zhang Yongjun, Zhang Zuxun, Zhang Jianqing (2002) Camera calibration using 2D-DLT and bundle adjustment with planar scenes[J]. Geomatics and Information Science of Wuhan University, 27 (6); 566-571 (in Chinese)
[9] Abdel-Aziz Y I, Karara H M (1971) Direct linear transformation from comparator coordinates into object space coordinates in close-range photogrammetry [C]. ASP/UI Symp Close-Range Photogrammetry, Urbana, IL
[10] Caprile B, Torre V (1990) Using vanishing points for camera calibration[J]. Int Journal of Computer Vision, 4(2):127-140
[11] Luong Q T, Faugeras Q (1997) Camera calibration, scene motion and structure recovery from point correspondences and Fundamental matrices[J]. The International Journal of Computer Vision, 22(3); 261-289
[12] Faugeras O D (1993) Three-dimensional computer vision; a geometric viewpoint[M]. Cambridge, MA; MIT Press
[13] Zisserman A, Forsyth D, Mundy J L, et al. (1992) Recognizing general curved objects efficiently[M]/ Mundy J L, Zisserman A. Geometric Invariance in Computer Vision. Cambridge, MA;MIT Press, 228-251
[14] Zhang Zhengyou (1998) A flexible new technique for camera calibration. Technical Report MSR-TR-98-71[R]. Redmond, USA: Microsoft Research
[15] Comanicu D, Meer P (2002) Mean shift, a robust approach toward feature space analysis[J]. IEEE Trans Pattern Anal Machine Intell, 24;603-619
[16] Meer P, Georgescu B (2001) Edge detection with embedded confidence[J]. IEEE Trans Pattern Anal Machine Intell, 23; 1 351-1 365
[17] Wong K Y (2001) Structure and motion from silhouettes[D], Cambridge; Department of Engineering, University of Cambridge
[18] Hartley R L, Zisserman A (2000) Multiple view geometry in computer vision[M]. Cambridge: Cambridge University Press