WUJNS
WHUAN UNIVERSITY JOURNAL OF NATURAL SCIENCES

Vol. 8 No. 4 2003 1090-1096

Article ID: 1007-1202(2003)04-1090-07

Flexible Planar-Scene Camera Calibration Technique

Zhang Yong-jun[1], Zhang Zu-xun[2], Zhang Jian-qing[2]

Abstract: A flexible camera calibration technique using 2D-DLT and bundle adjustment with planar scenes is proposed. The equation of principal line under image coordinate system represented with 2D-DLT parameters is educed using the correspondence between collinearity equations and 2D-DLT. A novel algorithm to obtain the initial value of principal point is put forward. Proof of Critical Motion Sequences for calibration is given in detail. The practical decomposition algorithm of exterior parameters using initial values of principal point, focal length and 2D-DLT parameters is discussed elaborately. Planar \ | scene camera calibration algorithm with bundle adjustment is addressed. Very good results have been obtained with both computer simulations and real data calibration. The calibration result can be used in some high precision applications, such as reverse engineering and industrial inspection.

Key words: camera calibration; 2D-DLT; bundle adjustment; planar grid; critical motion sequences; lens distortion

CLC number: TP 391, P 23

Received date: 2002-10-22

Foundation item: Supported by the Research Foundation of Doctoral Position Speciality of Universities (20010486011)

Biography: Zhang Yong-jun (1975-), male, Ph. D., research direction: digital photogrammetry and computer vision. E-mail: yjzhang@supresoft.com.cn

0 Introduction

Calibration of cameras is a prerequisite for the extraction of precise three-dimensional information from imagery in photogrammetry, computer vision, robotics and other areas. Much work has been done in the photogrammetry community[1-3], and more recently in computer vision[4,5]. A number of auto-calibration approaches based on active-vision technique have been addressed by computer vision communities, but many of them need non-linear optimization which takes more time than linear ones or camera motions need to be partly known[3\ 〗. In some cases, the result of auto-calibration can not be determined uniquely, which differs from the true value remarkably even with low noise level[6].

Faugeras proposed a camera self-calibration technique based on epipolar transformation and Kruppa equations[7]. Bill Triggs developed a self-calibration technique from at least 5 views of a planar scene[4], but this technique has difficulty to initialize. Zhang put forward a camera calibration technique for planar scenes based on the orthonormal property of the rotation matrix, and precision of about 0.35 pixel is obtained[8].

Direct Linear Transformation (DLT) is developed by Abdel-Aziz and Karara. 3D-DLT is widely used for camera calibration[3] in the literature, but no 2D-DLT-based calibration paper has been published, mainly because camera calibration is not an active area in Photogrammetry communities. Another reason is computer vision researchers are not familar with DLT. This paper mainly focuses on camera calibration technique using 2D-DLT and collinearity equations with planar scenes. The equation of principal line under image coordinate system represented by 2D-DLT parameters is worked out using the corres pondence between collinearity equations and 2D-DLT. Initial value of principal point can be obtained with at least two equations of principal lines. The practical decomposition algorithm of exterior parameters using initial values of principal point, focal length and 2D-DLT parameters is discussed. Planar-scene camera calibration algorithm with bundle adjustment is addressed.

1 2D-DLT and Initial Values

2D-DLT can be written as[9]

$$\begin{cases} x = \dfrac{h_1 X + h_2 Y + h_3}{h_7 X + h_8 Y + 1} \\ y = \dfrac{h_4 X + h_5 Y + h_6}{h_7 X + h_8 Y + 1} \end{cases} \tag{1}$$

where $H = (h_1, h_2, h_3, h_4, h_5, h_6, h_7, h_8)^T$ is the transformation parameters equivalent to homography matrix up to a scale factor when taken 1 as the ninth element, X, Y the space point under world coordinate system ($Z = 0$), x, y the corresponding image point.

Given an image of model plane, the values of transformation parameters can be estimated by $AH = 0$. The solution is the right singular vector of A associated with the smallest singular value. In order to eliminate the influence of outliers which may be introduced by miss-match of image points and the corresponding model points, the parameters can be refined with an iterative least-squares method after linearise Eq. (1).

The mostly used collinearity equations in photogrammetry can be written as[1]

$$\begin{cases} x - x_0 = -f \dfrac{a_1(X - X_s) + b_1(Y - Y_s) + c_1(Z - Z_s)}{a_3(X - X_s) + b_3(Y - Y_s) + c_3(Z - Z_s)} \\ y - y_0 = -f \dfrac{a_2(X - X_s) + b_2(Y - Y_s) + c_2(Z - Z_s)}{a_3(X - X_s) + b_3(Y - Y_s) + c_3(Z - Z_s)} \end{cases} \tag{2}$$

where x_0, y_0, f are the interior parameters, X_s, Y_s, Z_s the position of camera, X, Y, Z the space point under world coordinate system ($Z = 0$ for planar grid), x, y the corresponding image point and $R = \{a_i, b_i, c_i, i = 1, 2, 3\}$ the rotation matrix composed of three rotation angles ϕ, ω, κ.

When the same coordinate system is choosed, Eq. (2) can be written as

$$\begin{cases} x = \dfrac{\left(f\dfrac{a_1}{\lambda} - \dfrac{a_3}{\lambda}x_0\right)X + \left(f\dfrac{b_1}{\lambda} - \dfrac{b_3}{\lambda}x_0\right)Y + \left(x_0 - \dfrac{f}{\lambda}(a_1 X_s + b_1 Y_s + c_1 Z_s)\right)}{-\dfrac{a_3}{\lambda}X - \dfrac{b_3}{\lambda}Y + 1} \\ y = \dfrac{\left(f\dfrac{a_2}{\lambda} - \dfrac{a_3}{\lambda}y_0\right)X + \left(f\dfrac{b_2}{\lambda} - \dfrac{b_3}{\lambda}y_0\right)Y + \left(y_0 - \dfrac{f}{\lambda}(a_2 X_s + b_2 Y_s + c_2 Z_s)\right)}{-\dfrac{a_3}{\lambda}X - \dfrac{b_3}{\lambda}Y + 1} \end{cases} \tag{3}$$

where $\lambda = (a_3 X_s + b_3 Y_s + c_3 Z_s)$.

Comparing Eq. (1) with Eq. (3), we have

$$\begin{cases} h_1 = f\dfrac{a_1}{\lambda} - \dfrac{a_3}{\lambda}x_0 \\ h_2 = f\dfrac{b_1}{\lambda} - \dfrac{b_3}{\lambda}x_0 \end{cases} \tag{4}$$

$$\begin{cases} h_4 = f\dfrac{a_2}{\lambda} - \dfrac{a_3}{\lambda}y_0 \\ h_5 = f\dfrac{b_2}{\lambda} - \dfrac{b_3}{\lambda}y_0 \end{cases} \tag{5}$$

$$\begin{cases} h_3 = x_0 - \dfrac{f}{\lambda}(a_1 X_s + b_1 Y_s + c_1 Z_s) \\ h_6 = y_0 - \dfrac{f}{\lambda}(a_2 X_s + b_2 Y_s + c_2 Z_s) \end{cases} \tag{6}$$

$$\begin{cases} h_7 = -\dfrac{a_3}{\lambda} \\ h_8 = -\dfrac{b_3}{\lambda} \end{cases} \tag{7}$$

From Eq. (4), Eq. (5) and Eq. (7), we obtain the following equations

$$\begin{cases} \dfrac{(h_1 - h_7 x_0)}{f} = \dfrac{a_1}{\lambda} \\ \dfrac{(h_2 - h_8 x_0)}{f} = \dfrac{b_1}{\lambda} \end{cases} \tag{8}$$

$$\begin{cases} \dfrac{(h_4 - h_7 y_0)}{f} = \dfrac{a_2}{\lambda} \\ \dfrac{(h_5 - h_8 y_0)}{f} = \dfrac{b_2}{\lambda} \end{cases} \tag{9}$$

$$\begin{cases} -h_7 = \dfrac{a_3}{\lambda} \\ -h_8 = \dfrac{b_3}{\lambda} \end{cases} \quad (10)$$

Multiplying the upper and lower parts of Eq. (8), Eq. (9) and Eq. (10) respectively, considering $a_1 b_1 + a_2 b_2 + a_3 b_3 = 0$, if the principal point (x_0, y_0) is known or obtained with certain approaches, the focal length can be obtained as follows

$$f = [-(h_1 - h_7 x_0) \cdot (h_2 - h_8 x_0) - (h_4 - h_7 y_0) \cdot (h_5 - h_8 y_0)]^{\frac{1}{2}} / (h_7 h_8) \quad (11)$$

Self-multiplying each items of Eq. (8), Eq. (9) and Eq. (10), taking $a_1^2 + a_2^2 + a_3^2 = 1$, $b_1^2 + b_2^2 + b_3^2 = 1$ into account, and canceling out λ, we obtain

$$\begin{aligned} &[(h_1 - h_7 x_0)^2 - (h_2 - h_8 x_0)^2 \\ &\quad + (h_4 - h_7 y_0)^2 - (h_5 - h_8 y_0)^2] f^{-2} \\ &\quad + (h_7^2 - h_8^2) = 0 \end{aligned} \quad (12)$$

Focal length f can be canceled out using Eq. (11) and Eq. (12), then we have

$$\begin{aligned} F_h &= (h_1 h_8 - h_2 h_7)(h_1 h_7 - h_7^2 x_0 + h_2 h_8 - h_8^2 x_0) \\ &\quad + (h_4 h_8 - h_5 h_7)(h_4 h_7 - h_7^2 y_0 + h_5 h_8 - h_8^2 y_0) \\ &= 0 \end{aligned} \quad (13)$$

In most cases, as we know, the principal point is different from image center. There are 9 interior and exterior parameters $(f, x_0, y_0, \phi, \omega, \kappa, X_s, Y_s, Z_s)$ of a camera when lens distortion, skew and aspect ratio are ignored. Obviously, these 9 parameters can not be decomposed uniquely from the 8 parameters of 2D-DLT. Theoretically, the principal point (x_0, y_0) can move freely on the principal line of image, proof is given in appendix A.

Eq. (13) can also be written in the form of $(L_x, L_y) \cdot (x_0, y_0)^T = L_c$. As we know, the principal point always lies on the principal line of image[2], so if we have at least two nonparallel principal lines, the principal point (x_0, y_0) can be obtained by solve the over-definite linear equation $LX = C$.

Note that we should avoid the so-called Critical Motion Sequences (CMS)[10]. 2D-DLT parameters among images are linearly correlated in the case of images taken with a fixed camera while the planar grid is rotating around its Z-axis. All the principal lines actually overlap each other in such cases, so the principal point can not be obtained from these lines. In practice, we only need to change the orientation of camera from one snapshot to another when the table turns around its Z-axis. Proof of CMS is given in appendix B.

After 2D-DLT parameters, focal length and principal point are determined, the initial values of camera exterior parameters can be decomposed as follows.

Replace λ in Eq. (8) and Eq. (9) with which derived from Eq. (10), we have

$$\begin{cases} \dfrac{a_1}{a_3} = -\dfrac{(h_1 - h_7 x_0)}{f h_7} \\ \dfrac{b_1}{b_3} = -\dfrac{(h_2 - h_8 x_0)}{f h_8} \\ \dfrac{a_2}{a_3} = -\dfrac{(h_4 - h_7 y_0)}{f h_7} \\ \dfrac{b_2}{b_3} = -\dfrac{(h_5 - h_8 y_0)}{f h_8} \end{cases} \quad (14)$$

Considering $b_1^2 + b_2^2 + b_3^2 = 1$, we can obtain $b_3^2 = \dfrac{1}{1 + \dfrac{(h_2 - h_8 x_0)^2}{f^2 h_8^2} + \dfrac{(h_5 - h_8 y_0)^2}{f^2 h_8^2}}$ from Eq. (14).

Under the ϕ, ω, κ system where Y-axis is taken as the primary axis, $\tan \kappa = b_1 / b_2$, and from Eq. (14) we have $\tan \kappa = \dfrac{b_1}{b_2} = \dfrac{h_2 - h_8 x_0}{h_5 - h_8 y_0}$, so κ can be determined uniquely.

The value of b_3 can initially take the positive value of the square root. Compare κ which is already determined with κ' calculated from b_1 and b_2 corresponding to positive b_3 from Eq. (14). If $\kappa ! = \kappa'$, b_3 should take the negative value of the square root, then b_1, b_2 and ω can be determined from Eq. (14) and $\sin \omega = -b_3$ respectively.

Using the knowledge that the row vectors of rotation matrix are orthonormal, we have

$$\begin{pmatrix} c_1 \\ c_2 \\ c_3 \end{pmatrix} = \begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix} \times \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix} = \begin{pmatrix} a_2 b_3 - a_3 b_2 \\ a_3 b_1 - a_1 b_3 \\ a_1 b_2 - a_2 b_1 \end{pmatrix} \quad (15)$$

As we know,

$\tan \phi = \dfrac{a_3}{c_3} = \dfrac{a_3}{a_1 b_2 - a_2 b_1} = \dfrac{1}{\dfrac{a_1}{a_3} b_2 - \dfrac{a_2}{a_3} b_1}$ where b_1 and b_2 have been determined along with ω, a_1 / a_3, and a_2 / a_3 are already determined in Eq. (14), so ϕ can also be determined uniquely.

The values of λ determined from Eqs. (8), (9) and (10) can be averaged to get the mean value, then from Eq. (6) and the definition of λ, we have

$$\begin{cases} h_3 = x_0 - \dfrac{f}{\lambda}(a_1 X_s + b_1 Y_s + c_1 Z_s) \\ h_6 = y_0 - \dfrac{f}{\lambda}(a_2 X_s + b_2 Y_s + c_2 Z_s) \\ \lambda = (a_3 X_s + b_3 Y_s + c_3 Z_s) \end{cases} \quad (16)$$

This linear equation can be easily solved to obtain the initial values of X_s, Y_s and Z_s. Note, as shown in above, not all the 9 elements of rotation matrix are calculated when determine ϕ, ω and κ. So elements of the rotation matrix should be recalculated through ϕ, ω and κ to derive the initial values of X_s, Y_s and Z_s.

2 Camera Calibration with Bundle Adjustment

Bundle adjustment is the problem of refining a visual reconstruction to produce jointly optimal 3D structure and viewing parameter (camera pose and/or calibration) estimates[11]. It is widely used by photogrammetry and computer vision communities. Generally, lens distortion is larger in non-metric cameras than in metric ones, so it must be determined in calibration along with the interior and exterior parameters. Skew of the two image axes will be ignored in the camera model since it is very close to zero in most current cameras. To take a full calibtation, the left part of collinearity Eq. (2) should consider lens distortion:

$$\begin{cases} \Delta x = (x - x_0)(K_1 r^2 + K_2 r^4) \\ \quad + P_1(r^2 + 2(x - x_0)^2) \\ \quad + 2P_2(x - x_0)(y - y_0) \\ \Delta y = (y - y_0)(K_1 r^2 + K_2 r^4) \\ \quad + P_2(r^2 + 2(y - y_0)^2) \\ \quad + 2P_1(x - x_0)(y - y_0) \end{cases} \tag{17}$$

where $r^2 = (x - x_0)^2 + (y - y_0)^2$, K_1 and K_2 are the first two orders of radial distortion, P_1 and P_2 are decentering distortion. Other distortion parameters such as prism distortion are not considered here since they are almost equal to zero.

Linearise Eq. (2) and (17) with Taylor series, the error equations for calibration can be writtenas as

$$\begin{aligned} v_x = {} & \frac{\partial x}{\partial X_s}\Delta X_s + \frac{\partial x}{\partial Y_s}\Delta Y_s + \frac{\partial x}{\partial Z_s}\Delta Z_s + \frac{\partial x}{\partial \phi}\Delta \phi \\ & + \frac{\partial x}{\partial \omega}\Delta \omega + \frac{\partial x}{\partial \kappa}\Delta \kappa + \frac{\partial x}{\partial X}\Delta X + \frac{\partial x}{\partial Y}\Delta Y + \frac{\partial x}{\partial Z}\Delta Z \\ & + \frac{\partial x}{\partial f_x}\Delta f_x + \frac{\partial x}{\partial f_y}\Delta f_y + \frac{\partial x}{\partial x_0}\Delta x_0 + \frac{\partial x}{\partial y_0}\Delta y_0 \\ & + \frac{\partial x}{\partial K_1}\Delta K_1 + \frac{\partial x}{\partial K_2}\Delta K_2 + \frac{\partial x}{\partial P_1}\Delta P_1 + \frac{\partial x}{\partial P_2}\Delta P_2 - l_x \\ v_y = {} & \frac{\partial y}{\partial X_s}\Delta X_s + \frac{\partial y}{\partial Y_s}\Delta Y_s + \frac{\partial y}{\partial Z_s}\Delta Z_s + \frac{\partial y}{\partial \phi}\Delta \phi \\ & + \frac{\partial y}{\partial \omega}\Delta \omega + \frac{\partial y}{\partial \kappa}\Delta \kappa + \frac{\partial y}{\partial X}\Delta X + \frac{\partial y}{\partial Y}\Delta Y + \frac{\partial y}{\partial Z}\Delta Z \\ & + \frac{\partial y}{\partial f_x}\Delta f_x + \frac{\partial y}{\partial f_y}\Delta f_y + \frac{\partial y}{\partial x_0}\Delta x_0 + \frac{\partial y}{\partial y_0}\Delta y_0 \\ & + \frac{\partial y}{\partial K_1}\Delta K_1 + \frac{\partial y}{\partial K_2}\Delta K_2 + \frac{\partial y}{\partial P_1}\Delta P_1 + \frac{\partial y}{\partial P_2}\Delta P_2 - l_y \end{aligned} \tag{18}$$

After the initial values of camera parameters are determined, they can be refined with the forenamed bundle adjustment. Due to the non-linear characteristics of the problem, iterations need to be performed. The corresponding items of ΔX, ΔY and ΔZ can be removed if the control points are considered to be of no errors. The status of normal equation is ill conditioned generally, parameters should be weighted properly to ensure the stability of calibration results.

3 Results of Experiments

3.1 Computer Simulations

Several tests have been done with the proposed technique. Firstly, in order to assess the influence of noise level, the proposed technique has been tested with computer simulated data. Five images of a planar scene with 500 points are used. The space points under world coordinate system are given artificially. Gaussian noise with 0 mean and δ standard deviation is added to the projected image points. The estimated camera parameters are then compared with the true values.

The simulated camera has the following property: $f = 4\,457.00$, $x_0 = 650.00$, $y_0 = 515.00$, with image resolution of 1 300 ×1 030. We can easily get focal length represented by millimeter given the physical size of image plane. Noise levels are varied from $\delta = 0.1$ pixels to $\delta = 1.5$ pixels. 100 independent trials are performed for each noise level. The results shown in Fig. 1 are the root-mean-squares (RMSs) of differences among the estimated values and the true values. It is obvious in Fig. 1 that the errors of focal length and principal point increase quasi-linearly along with the increase of noise level.

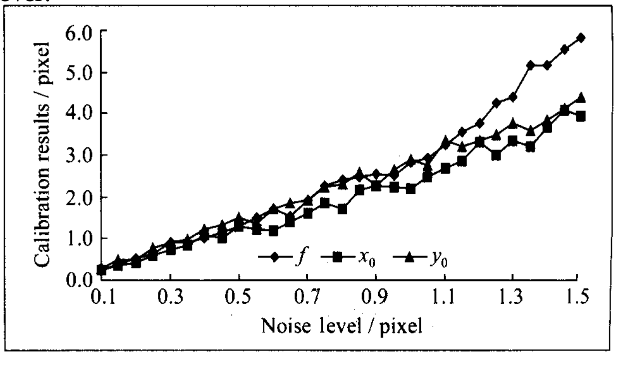

Fig. 1 Calibration results vs. noise levels

Since the principal point is different from the image center in most cases, another experiment is taken to test the effectivity of the proposed principal point calculation technique. In this test, x_0 varies from the image center to no larger than 30 pixels, y_0 is calculated from the equation of principal line obtained from the true values. Noises of $\delta = 0.3$ pixels are added to the projected image points of the five images used in the test. Camera exterior parameters are decomposed after 2D-DLT parameters, principal point and focal length are determined, and then the estimated parameters of bundle adjustment are compared with the true values. 100 independent trials are taken for each principal point deviation from the image center. RMSs of these results are shown in Fig. 2. It is obvious that the errors of estimated principal point do not increase along with the increase of deviation and the errors are all below 1 pixels, which shows the effectivity of the proposed principal point calculation technique.

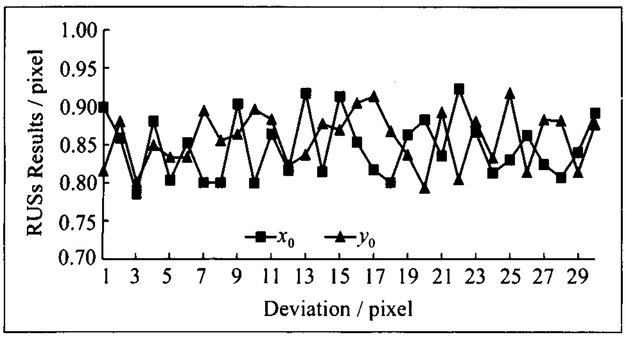

Fig. 2 Results vs. principal point deviations

2-15 images are used for calibration to test the influence of image numbers. Projected image points are also added by noise of 0.5 pixels. 100 independent trials are performed with each image number, and the estimated values are compared with the true values. Results are shown in Table 1.

Obviously, the errors decrease when more images are used for calibration. The precision of estimated values improves significantly while images vary from 2 to 3. To ensure the stability of calibration result, 5-8 images should be used in practice.

3.2 Real Data Experiment

The proposed technique is also tested with real image data. In this test, the planar grid (about 45 cm ×45 cm) rotates along with a table which turns around its vertical axis. 8 images are taken, one of them is shown in Fig. 3, the crosses are the matched grid points. There are 900 designed points in the planar grid, and the precision of each designed coordinate is about 0.1 mm. Actually, it is enough with a precision of about 0.5 mm for the grid point. It is quite easy for mechanical engineering researchers or industries to make such a grid. The image resolution of the CCD camera to be calibrated is 1 300 pixels ×1 030 pixels. Image grid points are detected as the intersection of straight lines fitted to each square with precision of higher than 0.1 pixel.

Table 1 RMSs of focal length and principal point

Images	f	x_0	y_0
2	2.360	1.840	1.690
3	1.400	1.090	0.920
4	1.390	1.090	0.880
5	1.230	0.890	0.760
6	1.170	0.710	0.563
7	0.991	0.658	0.770
8	1.069	0.574	0.741
9	0.970	0.496	0.651
10	0.967	0.554	0.715
11	0.900	0.568	0.620
12	0.600	0.485	0.439
13	0.475	0.426	0.536
14	0.464	0.303	0.390
15	0.418	0.220	0.360

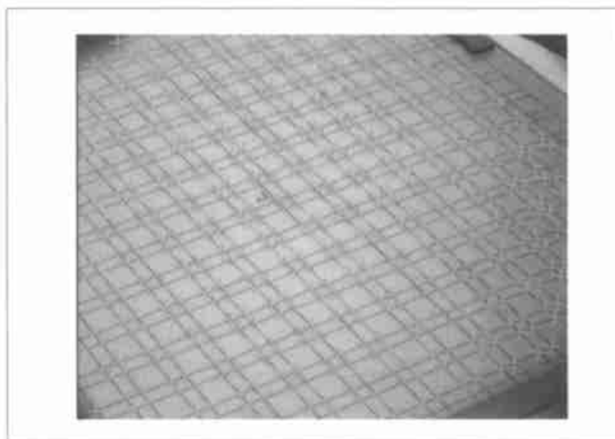

Fig. 3 One of the images used for calibration

There are about 500 grid points visible in each image. As we know, planar grid coordinates should be taken as weighted-unknowns in the rigorous bundle adjustment in order to eliminate the influence of imprecision of the designed grid coordinates and to get more precise camera parameters. The unit weight RMS of the rigorous bundle adjustment is 0.08 pixel. Values and errors of interior parameters are shown in Table 2.

It can be seen from Table 2 that the principal point is very close to the image center, and RMS errors are below 0.1 pixels. The aspect ration is 0.998 2, i.e. the pixels are nearly square. The error of focal length is about 0.2 pixels. As can be calculated from the values in Table 2 and Eq. (17), the maximum lens distortion is about 3 pixels. The deviations between detected image points and projected ones with the calibrated camera parameters and grid coordinates are also calculated. The RMSs of the deviations are below 0.10 pixels for all the 8 images, equivalent to about 0.03 mm in the grid coordinate system.

Table 2 Estimated result and precision with real image

Items	f_x	f_y	x_0	y_0
Estimated	4 426.135	4 418.137	652.120	514.730
RMS error	0.201	0.228	0.080	0.081

Items	$K_1/10^{-9}$	$K_2/10^{-15}$	$P_1/10^{-7}$	$P_2/10^{-7}$
Estimated	-7.416 0	-4.522 0	6.489 0	6.684 0
RMS error	0.116 2	0.182 8	0.465 4	0.125 1

The proposed algorithm has been used in our deformation inspection system of industrial sheetmetal parts successfully.

4 Conclusions

In this paper, we propose a new technique for camera calibration. The proposed technique only requires the camera to observe a planar pattern at a few (at least two) different orientations. Either the camera or the pattern can be moved freely, and the motion need not be known. Compared with classical techniques that use expensive equipment such as 3D calibration field, the proposed technique is flexible considerably.

The proposed technique consists of two steps. In the first step, initial values are decomposed from 2D-DLT parameters. The initial values are refined in the second step with an iterative linear bundle adjustment using collinearity equations based on least-squares criterion.

Both computer simulation and real data have been used to test the proposed technique, and very good results have been obtained, which verifies the feasibility of the proposed planar-scene camera calibration technique.

Appendix A: Proof of Ambiguities with Single Image

Consider the rotation matrix composed of A, ν, κ where Z-axis is taken as the primary axis[1], $\tan\kappa = c_1/c_2$, where κ is the angle between the principal line and y-axis. Substituting the corresponding items of Eq. (15) in $\tan\kappa = c_1/c_2$ results in

$$\tan\kappa = \frac{c_1}{c_2} = \frac{a_2 b_3 - a_3 b_2}{a_3 b_1 - a_1 b_3} = \frac{a_2/a_3 - b_2/b_3}{b_1/b_3 - a_1/a_3} \tag{19}$$

Substituting the corresponding items of Eq. (14) in the above equation results in

$$\tan\kappa = -\frac{h_4 h_8 - h_5 h_7}{h_1 h_8 - h_2 h_7} \tag{20}$$

Eq. (13) can be represented in the form of a line $y_0 = Ax_0 + C$, where $A = \tan\alpha = -\frac{h_1 h_8 - h_2 h_7}{h_4 h_8 - h_5 h_7}$ is the slope of the line and α is the angle between the line and x-axis. Obviously, $\tan\kappa = 1/\tan\alpha$, which means $\kappa = 90° - \alpha$, i.e. Eq. (13) is actually the equation of principal line in image. Clearly, as long as the principal point locates on the principal line, each group of decomposition is valid for the perspective relationship, i.e. the camera parameters can not be decomposed uniquely from single image. Mathematically, it is impossible to calibrate a camera completely from single image of a planar pattern without any other information.

Appendix B: Proof of Critical Motion Sequences

Under the pinhole model, projection relationship between image and planar pattern can be written as[8]:

$$s\begin{pmatrix} x \\ y \\ 1 \end{pmatrix} = A(r_1 \quad r_2 \quad r_3 \quad t)\begin{pmatrix} X \\ Y \\ 0 \\ 1 \end{pmatrix} = A(r_1 \quad r_2 \quad t)\begin{pmatrix} X \\ Y \\ 1 \end{pmatrix} \tag{21}$$

where A is called the camera interior matrix, r_1, r_2 and t are the first two columns of rotation matrix and the camera translation respectively. $H = A(r_1 \quad r_2 \quad t)$ is called Homography between the model plane and the image. It is obvious that Eq. (21) is actually the equation of 2D-DLT when scale s is canceled out with the third row. So 2D-DLT parameters are equivalent to homography matrix while 1 is taken as the ninth element. Now we take the form of homography for 2D-DLT parameters, then two rotation matrix are related by

$$R_2 = R_1\begin{pmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{pmatrix} \tag{22}$$

where θ is the angle of the relative rotation. We will use superscript $^{(1)}$ and $^{(2)}$ to denote vectors related to image 1 and 2, respectively. Substitute Eq. (22) into $H = A(r_1 \quad r_2 \quad t)$ results in

$$\begin{pmatrix} h_1^{(2)} \\ h_4^{(2)} \\ h_7^{(2)} \end{pmatrix} = \begin{pmatrix} h_1^{(1)} \\ h_4^{(1)} \\ h_7^{(1)} \end{pmatrix}\cos\theta + \begin{pmatrix} h_2^{(1)} \\ h_5^{(1)} \\ h_8^{(1)} \end{pmatrix}\sin\theta$$

$$\downarrow\begin{pmatrix} h_2^{(2)} \\ h_5^{(2)} \\ h_8^{(2)} \end{pmatrix} = -\begin{pmatrix} h_1^{(1)} \\ h_4^{(1)} \\ h_7^{(1)} \end{pmatrix}\sin\theta + \begin{pmatrix} h_2^{(1)} \\ h_5^{(1)} \\ h_8^{(1)} \end{pmatrix}\cos\theta \qquad (23)$$

As we know, the principal line in image can be represented in the form of $y_0 = Ax_0 + C$, and the slope of this line can be written as follows

$$\tan\alpha = -\frac{h_1 h_8 - h_2 h_7}{h_4 h_8 - h_5 h_7} \qquad (24)$$

where α is the angle between the principal line and x-axis of image.

Substitute corresponding items of Eq. (23) in Eq. (24) results in

$$\begin{aligned} & h_1^{(2)} h_8^{(2)} - h_2^{(2)} h_7^{(2)} \\ & = (h_1^{(1)}\cos\theta + h_2^{(1)}\sin\theta)(-h_7^{(1)}\sin\theta + h_8^{(1)}\cos\theta) \\ & \quad - (-h_1^{(1)}\sin\theta + h_2^{(1)}\cos\theta)(h_7^{(1)}\cos\theta + h_8^{(1)}\sin\theta) \\ & = h_1^{(1)} h_8^{(1)} - h_2^{(1)} h_7^{(1)} \qquad (25) \end{aligned}$$

We have $h_4^{(2)} h_8^{(2)} - h_5^{(2)} h_7^{(2)} = h_4^{(1)} h_8^{(1)} - h_5^{(1)} h_7^{(1)}$ similarly. Clearly, two principal lines are mutually parallel to each other. Further more, C can be written as follows from Eq. (13)

$$\begin{aligned} C = & [(h_1 h_8 - h_2 h_7)(h_1 h_7 + h_2 h_8) \\ & + (h_4 h_8 - h_5 h_7)(h_4 h_7 + h_5 h_8)] / (h_7^2 + h_8^2) \end{aligned} \qquad (26)$$

From Eq. (23) we can obtain $C^{(2)} = C^{(1)}$ without difficulty. So it is obvious that two principal lines are actually overlapped lines under image coordinate system, which indicates that principal point can not be obtained from these configurations, namely critical motion sequences. In practice, we only need to change the orientation of the camera or the model plane from one snapshot to another to avoid these degenerances.

References

[1] Wang Zhi-zhuo. *Principles of Photogrammetry*. Beijing: Publishing House of Surveying and Mapping, 1990(Ch).

[2] Zhu Zhao-guang, Sun Hu, Cui Bing-guang. *Photogrammetry*. Beijing: Publishing House of Surveying and Mapping, 1995 (Ch).

[3] Chen Fang-Jing. Application of Least Squares Adjustment Technique to Geometric Camera Calibration and Photogrammetric Flow Visualization. *ISA 43rd International Instrumentation Symposium*, Orlando: Florida. NASA, 1997, 659-668.

[4] Triggs Bill. Autocalibration from Planar Scenes. 1998 *European Conference on Computer Vision*. Freiburg: Germany. Springer-Verlag, 89-105, 1998.

[5] Sturm Peter, Maybank Stephen. On Plane-Based Camera Calibration: A General Algorithm, Singularities, Applications. *CVPR*, Freiburg: Springer-Verlag, 1999. 432-437.

[6] Qiu Mao-lin. Overview of Camera Calibration for Computer Vision, *ACTA Automatica SINICA*, 2000, **26**(1): 43-55 (Ch).

[7] Faugeras Olivier, Luong Quang-Tuan, Maybank Stephen. Camera self calibration: Theory and experiments. In G. Sandini, editor, *European Conf. Computer Vision*, Freiburg, Germany: Springer-Verlag, 1992, **588**: 321-334.

[8] Zhang Zheng-you. A Flexible New Technique for Camera Calibration. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2000, **22**(11): 1330-1334.

[9] Feng Wen-hao. *Close Range Photogrammetry*. Wuhan: Publishing house of Wuhan University, 2002 (Ch).

[10] Sturm Peter. Critical motion sequences for monocular self-calibration and uncalibrated Euclidean reconstruction, *CVPR*, Freiburg, Germany: Springer-Verlag, 1997. 1100-1105.

[11] Triggs Bill, Mclauchlan Peter, hartley Richard. Bundle Adjustment -- A Modern Synthesis. *Proceedings of the Int. Workshop on Vision Algorithms: Theory and Practice*, Freiburg, Germany: Springer-Verlag, 1999, 298-372.