A Novel Stratified Self-calibration Method of Camera Based on Rotation Movement

Canlin Li*,
School of Computer and Communication Engineering, Zhengzhou University of Light Industry, Zhengzhou, China
Email: lcl_zju@aliyun.com

Rijian Su
School of Computer and Communication Engineering, Zhengzhou University of Light Industry, Zhengzhou, China
Email: srj@zzuli.edu.cn

Abstract—This paper proposes a novel stratified self-calibration method of camera based on rotation movement. The proposed method firstly captures more than three images of the same scene with small relative rotation angles among the captured images. After feature extraction and matching of captured images, the pixel coordinates of feature point are normalized. Then the stratified self-calibration is performed, following projective, affine and metric calibration. Projective calibration determines the camera projective matrix of every image in the projective reconstruct space. Affine calibration calculates the parameters of infinity reference plane in this space and the homography according to approximately equal relationship among the main diagonal elements of homography, which is inferred by virtue of small relative rotation angles among the captured images and the property of internal reference matrix corresponding to normalized pixel coordinates. Lastly metric calibration acquires internal reference matrix from the calculated homography. The proposed method can be applied to quickly, accurately and stably obtain internal parameters of camera. Real data has been used to test the proposed approach, and very good results have been achieved.

Index Terms—camera calibration, self-calibration, rotation movement, stratified calibration

I. INTRODUCTION

The process of acquiring the parameters of camera geometric model is called camera calibration[1]. It is an essential step to extract three-dimensional space information from two-dimensional images in the applications of image processing and computer vision, and is widely used in three-dimensional reconstruction, navigation, visual surveillance and other fields[2-3]. Under certain camera model, camera calibration strikes camera parameters of the model with a series of mathematical transformations and calculations by virtue of the image processing. At present, there have been a lot of camera calibration methods, which contain traditional calibration and self-calibration.

Traditional calibration is relatively mature and stable but based on specific experimental conditions, such as calibration references with known shape or size[4-6]. The traditional methods of using calibration references obtain a wide range of applications, which typically involve Tsai two-step method[4]. These traditional methods require using the calibration references in the process of shooting calibration images, which would bring great inconvenience into the camera calibration, since the use and position adjustment of calibration reference may result in the online task interrupted.

Camera self-calibration does not require the use of calibration reference, and estimates the camera internal parameters based solely on the corresponding relationship between feature pixels of images, which makes online and real-time calibrating camera parameters possible. However, many existing self-calibration methods still cannot get stable results and be mature enough for practical application[7-11], and therefore it is necessary to improve the existing self-calibration techniques through some ways.

This paper presents a novel stratified self-calibration method of camera based on rotation movement. The proposed method firstly shoots more than three image of the same scene following the requirements of small angle rotation under the condition of internal parameters unchanged. After feature extraction and matching on the captured images, the projection, affine and metric calibration[12] of three levels are performed by virtue of feature correspondence among images. The proposed method can be applied to quickly, accurately and stably obtain internal parameters of camera. This makes online and real-time calibrating the internal parameters of camera possible in the case of no dependent calibration references, thus improving the existing camera self-calibration.
camera can also be rotated, but the relative rotation angles among the captured images should be as small as possible, for example, the relative rotation angles along three coordinate axes should be less than 15°. Suppose that the series of captured images are \(i_1, i_2, \ldots, i_M\). Figure 1 shows four self-calibration images with different angles in the same scene.

\(K\) is an upper triangular matrix representing camera internal parameters.

\[
K = \begin{bmatrix} k_u & s & p_v \\ 0 & k_v & p_u \\ 0 & 0 & 1 \end{bmatrix},
\]

(1)

The goal of this self-calibration method is to determine the internal reference matrix \(K\) of camera, where \(k_u\) is the magnification in \(u\) direction (horizontal) of image, in pixels,

\(k_v\) is the magnification in \(v\) direction (vertical) of image, in pixels,

\(s\) is the distortion factor corresponding to the distortion of camera axis,

\(p_u\) and \(p_v\) are the principal point coordinates, in pixels.

The parameters \(k_u\) and \(k_v\) are closely linked to focal length \(f\) of camera. If \(k_u = k_v\) and \(s = 0\), \(k_u\) and \(k_v\) are just focal length \(f\) in pixels. \(k_u = k_v\) means that the photosensitive array of camera contains square pixels. If \(k_u \neq k_v\), \(k_u\) is the ratio of the focal length \(f_u\) and the pixel size in \(u\) direction, and \(k_v\) is the ratio of the focal length \(f_v\) and the pixel size in \(v\) direction. Thus, there are the following relationships \(f_u = k_u \cdot dpx\), and \(f_v = k_v \cdot dpy\), where \(f_u\) and \(f_v\) are effective focal length, or distance from image plane to projective center, unit in millimeters, and \(dpx\) and \(dpy\) are the pixel size of image plane, respectively in \(u\) and \(v\) directions. \(k_u \neq k_v\) means that the photosensitive array of camera contains non-square pixels, such as CCD camera situation.

**B. Feature Extraction and Matching of Captured Images**

The following step is to extract the feature points for every captured image. The feature extraction could be based on the general extraction method or SIFT extraction method[13]. Then feature matching between the images is performed to acquire the pixel coordinate of every matched feature point. The matched feature point in an image can be matched to one feature point in each of other images, and they correspond to the same 3D feature in the world. For the proposed stratified self-calibration in this paper, at least eight matched feature points in every captured image should be determined and found out. The more correctly matched feature points, the more constraints provided, the more accurate the constraint correspondence among the images, thus the more desirable results come from the proposed self-calibration.

**C. Normalizing the Pixel Coordinates of Feature**

To make approximately equal relationship established in affine calibration process, the original pixel coordinates of feature points are required to transform into normalized pixel coordinates. The details are as follows. The original pixel coordinate \(u = (u, v, 1)\)' of every feature point is transformed into the normalized pixel coordinate through \(u' = Tu\). In all following steps of stratified self-calibration, we use \(u'\), instead of \(u\). Here,

\[
T = \begin{bmatrix} 1 & -2c_v & -1 \\ 0 & 2c_u & -1 \\ 0 & 0 & 1 \end{bmatrix},
\]

(2)

where \(c_u\) and \(c_v\) are the coordinates of the image center, Through this coordinate normalization process, the image center of every image is aligned to the coordinate origin of the image plane, and two items of normalized pixel coordinate are all in the range of (-1,1).

According to camera perspective projection model, the following relationship is established.

\[
u \sim PX \sim K[R] [-Rt] X,
\]

(3)

where \(u\) is the pixel coordinate of feature point in the homogeneous form, and \(X = (x, y, z, 1)\)' is the coordinate of the corresponding 3D feature point in the homogeneous form, and \(P\) is the corresponding projection matrix, and \(R\) and \(t\) are respectively rotation matrix and translational vector in camera coordinate system relative to world coordinate system.

After normalization transformation about pixel coordinates of the image, the following relationship is established.

\[
u' \sim Tu' \sim TK[R] [-Rt] X = K'[R] [-Rt] X,
\]

(4)

where

\[
K' = TK
\]

(5)

In the following steps, on the basis of the perspective projection equation (4), \(K'\) is firstly determined by virtue of stratified self-calibration, then the original internal parameter matrix \(K\) is acquired through (5).

**D. Projective Calibration**

The projective calibration aims to obtain camera projective matrix \(P_i\) (\(i = 1, 2 \cdots M\)) of every image in the projective reconstruct space. To simplify the solving problem, we let world coordinate system align to the camera coordinate system of \(i\), so \(i\) is considered as the reference image. Also, there is the following formula.

\[
P_{i0} = [I_{3 \times 3} 0_3]
\]

(6)

The other projective matrices \(P_{io}\) (\(i>1\)) are solved as follows.

1) Calculating the fundamental matrix \(F_{io}\) between \(i\) and \(i_1\).
According to the corresponding relationship of matched feature points between $I_1$ and $I_i$, the fundamental matrix $F_i$ is calculated by virtue of the eight-point algorithm [14].

2) Forcing matrix $F_i$, rank reduced

According to the nature of the fundamental matrix $F_i$, we let $F_i$ meet the reduced rank constraint $\text{rank}(F_i)=2$. The details are as follows.

After SVD decomposition of the matrix $F_i$, we can get

$$ F_i = U_i D_i V_i^T $$

(7)

where $D_i$ consists of the singular values of $F_i$. Set the smallest singular value of $D_i$ to 0, and construct a new matrix $F'_i$. Let

$$ F'_i = U_i D'_i V_i^T $$

(8)

where $F'_i$ is the solved fundamental matrix between $I_1$ and $I_i$.

3) Computing the coordinate of epipole $e_i$ about the image $I_i$

The epipole $e_i$ of the image $I_i$ is the projection of camera optical center in the image $I_i$. According to the knowledge of epipole geometry, there is the following equation.

$$ F'_i e_i = 0 $$

(9)

Since the matrix $F'_i$ has reduced rank, the solution of (9) or the value of $e_i$ is unique, and proportional to the last column of $U_i$.

4) Calculating the projective matrices $P_i$ ($i>1$)

According to the selected world coordinate system and the knowledge of projection transformation, the projective matrices $P_i$ ($i>1$) of camera is determined as follows.

$$ P_i = [H_i | e_i] \quad (i>1), $$

(10)

$$ H_i = [e_i] f_i e_i^T, $$

(11)

where $[e_i]$ is an anti-symmetric matrix from $e_i$. Let $e_i = (e_x, e_y, e_z)^T$, there is

$$ [e_i] \pi = \begin{bmatrix} 0 & -e_z & e_y \\ e_z & 0 & -e_x \\ -e_y & e_x & 0 \end{bmatrix} $$

(12)

where $\pi$ is a non-zero column vector, for example $\pi=[1,1,1]^T$.

E. Affine Calibration

Affine calibration aims to determine the parameters of infinity reference plane in the space of projection reconstruction, i.e. $\pi = (\pi_x, \pi_y, \pi_z)^T$, and thus determine the homography $H_a$ associated with the affine calibration. This parameter uses the position of the image $I_1$ as a reference position.

According to the relationships among projective, affine and metric calibration in the stratified method, there are the relationships as follows.

$$ P_a = P_{aT} - [K'_{aT} K'_{a}] $$

(13)

where $P_a$ is affine transformation matrix, and $P_{aT}$ is metric transformation matrix, and $K_{a}$ is the transition matrix from projective transformation matrix to affine transformation matrix, and $T_{a}$ is the transition matrix from affine transformation matrix to metric transformation matrix, and $K'_{aT}$ and $K'_{a}$ are respectively the translation vector and the rotation matrix in the position and orientation of the image $I_i$ relative to the reference $I_1$.

$$ T_{PD} \sim \begin{bmatrix} x_3 \\ y_3 \\ 1 \end{bmatrix} $$

(15)

$$ T_{DM} \sim \begin{bmatrix} K'_{aT} x_3 \\ 0 \\ 1 \end{bmatrix} $$

(16)

$$ R_{MD} \sim \begin{bmatrix} r_1 \\ r_2 \\ r_3 \\ r_4 \\ r_5 \\ r_6 \end{bmatrix} $$

(17)

Thus, according to the first three columns in the formulas (13) and (14), we get the following formulas.

$$ H_a = H_{PD} e_a \pi_k  \sim K'_{aT} K'_{a} $$

(18)

$H_a$ is the homography which transforms the projection in the reference image $I_1$ of a point in the infinity plane into the projection in the image $I_i$. With the effect of approximating from the small-angle rotation, we can infer that [15]

$$ \text{Diag}(H_a) \sim [1,1,1]^T. $$

(19)

$\text{Diag}(H_a)$ represents the vector from main diagonal elements of $H_a$. That is to say, the main diagonal elements of $H_a$ can be approximately equal. Thus, according to formulas (10) and (18), there is the following equation (20) containing unknown $\pi_a$.
where \( P_n(m,n) \) is the item in the \( m \)th row and \( n \)th column of matrix \( P_n \). \( \pi_n \) can be determined using the method of least squares\[16\]. Furthermore, we refine the unknown \( \pi_n \) by applying Levenberg-Marquardt optimization technique to the optimization problem as formula (21). After this, we can get \( H_n(i>1) \) from \( \pi_n \) and formula (18).

Then we normalize the matrix \( H_n \). For every \( H_n \), let \( \hat{H}_n = \alpha \pi_n \) so as to \( \det(H_n) = 1 \). It is easy to know that \( \alpha = \sqrt[3]{1/\det(H_n)} \), thus \( H_n \) can be transformed into \( \hat{H}_n \) with unit determinant as follows.

\[
\hat{H}_n = \sqrt[3]{1/\det(H_n)} \cdot H_n
\]

(22)

Moreover, the formula (18) is transformed into as follows.

\[
H_n = K R_n K^{-T}
\]

(23)

**F. Metric Calibration**

This stage consists of two steps. The first is to calculate internal reference matrix \( K \) of camera. We can get \( K^{-T} H_n K = R_n \) from (23). According to the property \( R_n = R_n \) of the rotation matrix, we get \( K^{-T} H_n K = K^{-T} H_n K^{-T} \). Thus we have the following important equation.

\[
K K^{-T} = H_n (K K^{-T}) H_n^T
\]

(24)

Let

\[
C = K K^{-T} = \begin{pmatrix} a & b & c \\ b & d & e \\ c & e & f \end{pmatrix}
\]

(25)

Since \( C \) is a symmetric matrix, equation (24) is become as follows.

\[
H_n C H_n^T = C
\]

(26)

For every \( H_n \), after expanding matrix multiplication in (26), and eliminating redundant equations, we can get six homogeneous linear equations involving six unknowns \( a, b, c, d, e \) and \( f \).

From M-1 \( H_n \) (\( M \geq 3 \)), we establish a homogeneous over-determined system of linear equations as follows.

\[
X C^T = 0
\]

(27)

where \( C = (a,b,c,d,e,f)^T \) is a vector consisted of the independent items of \( C \), and \( X \) is a \( 6(\text{M}-1) \times 6 \) matrix. The least-square solution of \( C^T \) is the eigenvector corresponding to the minimum eigenvalue of \( X^T X \)[16].

By conducting SVD decomposition of \( X \) or using Jacobi method, we can find the smallest eigenvalue of symmetric matrix \( X^T X \). According to \( C = K K^{-T} \), solve \( K \) by virtue of Cholesky decomposition. If the diagonal terms of \( K \) are required to be positive, the Cholesky decomposition is unique.

The second step in this stage is to determine original internal reference matrix \( K \) of camera. According to the formula (5), we have the following formula.

\[
K = T^{-1} K'
\]

(28)

In the case that \( T \) and \( K' \) are known, \( K \) can be solved.

**III. EXPERIMENTAL RESULTS**

We took real images to confront our proposed framework with the real world. The camera to be calibrated is an off-the-shelf Panasonic AW-E300 CCD camera. The image resolution is 720×576. Only three groups of real images which come from three difference scenes are shown in the paper due to space limitation, as can be seen in Figures 1-3.

We adopted the proposed self-calibration to calibrate each image group, and presented the calibrated results of intrinsic parameters in Table 1, where the second, third and last rows are for Group 1 of image in Figure 1, Group 2 of image in Figure 2 and Group 3 of image in Figure 3, respectively. In Table 1, \( f \) is effective focal length, or distance from image plane to projective center, unit in millimeters. Here we let \( f = f_s \cdot q \) is the image size factor about internal parameters and determined by \( q = f_s / f \).

As we can see from the calibration results of intrinsic parameters, the calibrated coordinate \( (p_u, p_v) \) is close to the ideal computer image coordinate (360, 288) of the origin in the image plane, and the calibrated distortion factor \( s \) is close to 0. We further evaluated the self-calibration accuracy by how well it can sense the 3D world. For each group of image, we applied the calibrated camera parameters to back project every 3D feature point into the image plane and computed the Euclidean distance between back-projected 2D feature point and real feature point. Then we averaged the Euclidean distances about all back-projected 2D feature points to obtain the mean back-projected error. In addition, we computed the standard deviation of the back-projected error. The mean back-projected error and the standard deviation are presented in the second and third columns in Table 2, respectively. As we can see from Table 2, the mean back-projected error and the standard deviation for the self-calibration are all lower than 1 pixel, which is satisfactory in many real computer vision and photogrammetry tasks, especially in the case of no
calibration reference with known 3D information. So our proposed self-calibration method actually yields good calibration results.

![Figure 1. Group 1 of real images for camera self-calibration](image1)

![Figure 2. Group 2 of real images for camera self-calibration](image2)

![Figure 3. Group 3 of real images for camera self-calibration](image3)

Table 1.

<table>
<thead>
<tr>
<th>Images</th>
<th>(k_u)</th>
<th>(k_v)</th>
<th>(f)</th>
<th>(p_u)</th>
<th>(p_v)</th>
<th>(s)</th>
<th>(q)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Group 1</td>
<td>968.42</td>
<td>1020.73</td>
<td>9.72</td>
<td>378.40</td>
<td>259.07</td>
<td>4.20</td>
<td>1.08</td>
</tr>
<tr>
<td>Group 2</td>
<td>902.24</td>
<td>1043.54</td>
<td>9.05</td>
<td>371.88</td>
<td>302.63</td>
<td>12.48</td>
<td>0.98</td>
</tr>
<tr>
<td>Group 3</td>
<td>909.03</td>
<td>1034.80</td>
<td>9.12</td>
<td>392.38</td>
<td>314.09</td>
<td>-1.36</td>
<td>1.00</td>
</tr>
</tbody>
</table>

Table 2.

<table>
<thead>
<tr>
<th>Images</th>
<th>Mean back-projected error</th>
<th>Standard deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Group 1</td>
<td>0.403</td>
<td>0.514</td>
</tr>
<tr>
<td>Group 2</td>
<td>0.285</td>
<td>0.372</td>
</tr>
<tr>
<td>Group 3</td>
<td>0.572</td>
<td>0.683</td>
</tr>
</tbody>
</table>

For our experiments, we used a PC platform with 2.5GHz Intel Core i5-2450M processor, 4G RAM and Windows 7 to run the presented self-calibration method implemented in C++. We recorded the computation time of obtaining the calibrated results as shown in Table 3.

The computation time for the calibration method involves the time \(t_1\) for detecting and matching SIFT feature points, the time \(t_2\) for self-calibration and the time \(t_3\) for reconstructing and back projecting 3D feature point. As we can see from Table 3, since LM optimization method is used in the process of rebuilding the 3D coordinates of the feature points, the time \(t_3\) for reconstructing and back projecting 3D feature point is even more than three second, but actually \(t_3\) is not due to the proposed self-calibration method and should not be included into the time of self-calibration. The time \(t_1\) of detecting and matching feature points is less than two second, and the computation time \(t_2\) for self-calibration is far less than one second, which has been very pleasing and encouraging without interrupting the execution of computer vision and photogrammetry tasks.

Table 3.

<table>
<thead>
<tr>
<th>Images</th>
<th>(t_1) (s)</th>
<th>(t_2) (s)</th>
<th>(t_3) (s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Group 1</td>
<td>1.513</td>
<td>0.0004</td>
<td>1.560</td>
</tr>
<tr>
<td>Group 2</td>
<td>1.388</td>
<td>0.016</td>
<td>3.182</td>
</tr>
<tr>
<td>Group 3</td>
<td>1.420</td>
<td>0.015</td>
<td>3.947</td>
</tr>
</tbody>
</table>

© 2014 ACADEMY PUBLISHER
We illustrated the software implementation on the proposed self-calibration method in Figures 4-7. Figure 4 showed the program interface, where the self-calibration images in Group 1 were loaded into the program by “Input Images” button and we can select an image to display in the left zone of interface through drop-down combobox “Select Image”. Also, we set CCD size and effective picture elements of camera through the program interface. Then the self-calibration was performed by using the method proposed in this paper when the “Self Calibration” button was pressed, as shown in Figure 5. Figure 5 also displayed the computation time of obtaining the calibrated results and the mean back-projected error as well as the standard deviation, where the red and white marks represented matched 2D feature points. If the “Show Results” button is pressed, the self-calibration results of the camera internal parameters will be displayed as a popped dialog like in Figure 6. Figure 7 showed the back-projected 2D feature point marked by blue and white after we pressed the “Feature/Reproject” button.

IV. CONCLUSION

This paper proposed a novel stratified self-calibration method of camera based on rotation movement. This method requires inputting more than three images from the same scene with constant internal reference. During capturing the images, the camera can be performed translational motion and also rotated, but the relative rotation angles among the captured images should be as small as possible. The proposed self-calibration method firstly extracts and matches feature points for all images. Then the stratified calibration involving projective, affine and metric calibration of three levels is carried out. Projective calibration determines the camera projective matrix of every image in the projective reconstruct space. Affine calibration calculates the parameters of infinity reference plane in this space and the homography according to approximately equal relationship among the main diagonal elements of homography. Finally metric calibration acquires internal reference matrix by virtue of the calculated homography. The proposed method provides a fast self-calibration solution with stable and accurate results. This makes online and real-time calibrating the internal parameters of camera possible in the case of no dependent calibration references, thus improving the existing camera self-calibration.

ACKNOWLEDGMENT

The work is supported by the National Nature Science Foundation of China (No.61374014), the Natural Science Foundation of Henan Educational Committee (No.12B520068) and the PHD Foundation of Zhengzhou University of Light Industry (No.2011BSJJ002).

REFERENCES


Canlin Li received PhD degree in computer science from Shanghai Jiaotong University in 2010. Since 2010 he has been in the School of Computer and Communication Engineering at Zhengzhou University of Light Industry. His research interests include image processing, graphics, multimedia and digital entertainment. Dr. Li is a member of IEEE as well as ACM.

Rijian Su received PhD degree in control science and engineering from Huazhong University of Science & Technology in 2010. He is an associate professor of School of Computer and Communication Engineering at Zhengzhou University of Light Industry. His research interests include control engineering, image and signal processing and computer application.