Camera calibration approach based on iteration

Min Han, Jiangming Kan*, and Yutan Wang

Abstract—Tsai’s camera calibration method is widely used, but the accuracy of calibration must be improved. To further improve the calibration accuracy, this paper first introduces Tsai’s calibration principles and analyzes the factors that affect the accuracy of Tsai’s calibration method. Then, an improved method of camera calibration based on iteration is proposed. The method solves the problem of approximating the principal point to the image center in Tsai’s method, and then the objective function and the initial value of search are designed considering the speed of the optimizing search and calculated amount. Finally, the experimental results show that the accuracy of the calibration method is significantly improved compared with Tsai’s.

Keywords—calibration accuracy, iteration, principal point, speed

I. INTRODUCTION

In computer vision and visual measurements, camera calibration is the primary task [1-2]. Camera calibration is used to solve parameters of camera, and the solution method is the camera calibration method. The accuracy of the calibrated parameters will have an important effect on the subsequent measurement [6]. Many camera calibration methods have been proposed with the development of camera production processes, but the number of iterations is often fixed, and thus the improvement of the calibration accuracy is limited [11].

Tsai’s method is divided into two major steps at present: calibration of the external parameters (the rotation matrix R and the translation matrix t) and calibration of the internal parameters (the principal point, focal length and the distortion coefficient) in the case of lens distortion [8]. Tsai’s method is mainly achieved by Radial Alignment Constraint (RAC), which is simple, efficient and suitable for most camera calibration [4]. With the extensive application of Tsai’s method, the requirements for calibration accuracy are increasing. Many improvement measures have been proposed such as increasing the complexity of its distortion model [5], but the degree of improvement is still inadequate:

i. Although the accuracy is improved, the factors that affect the speed of calibration are not considered, such as the form of objective function, the unknown variable to be searched for, and the initial value of the unknown variable (the starting point of the search).

ii. Some improved methods do not address the inability of Tsai’s method to precisely calculate the scale factor s, for coplanar points [12], [16].

iii. Some of the improved methods involve iterative processes, but the number of iterations is often fixed, and thus the improvement of the calibration accuracy is limited [11].

iv. The problem of principal point initial value approximation in the Tsai method is not solved, that is, when solving the rotation matrix R and the translation matrix t, it is assumed that the principal point is approximately the image center [6], [14-15].

The rest of this paper is organized as follows. In Section II, the principle of Tsai’s Calibration is introduced, and then the problems of the Tsai method are analyzed in detail. In Section III, an improved method based on iterative operation is proposed to solve these problems. In section IV, Tsai’s method and the improved method are compared by three evaluation criteria, and then the experimental results verify the superiority of the proposed method. Section V summarizes the main contributions of this paper.

II. PRINCIPLE OF TSAI’S CALIBRATION METHOD AND ANALYSIS OF FACTORS INFLUENCING ACCURACY

Under the condition that the origin of the world coordinate system is not in the field of view and is not projected near the Y coordinate axis of the image coordinate system, the world coordinate system and the camera coordinate system exist in a relationship as shown in (1) for the i-th point [7].

\[
\begin{bmatrix}
  x_{ci} \\
  y_{ci} \\
  z_{ci}
\end{bmatrix} = \begin{bmatrix}
  R & t \\
  0 & 1
\end{bmatrix} \begin{bmatrix}
  x_{ci} \\
  y_{ci} \\
  z_{ci}
\end{bmatrix} = \begin{bmatrix}
  R_{1} & R_{2} & R_{3} & T_{x} \\
  I & I & I & 0 \\
  0 & 0 & 0 & 1
\end{bmatrix} \begin{bmatrix}
  x_{ci} \\
  y_{ci} \\
  z_{ci}
\end{bmatrix}
\]

\[(x_{ci}, y_{ci}, z_{ci}) \text{ and } (x_{ci}, y_{ci}, z_{ci})\text{ represent the coordinates of the world coordinate system and the camera coordinate system.}\]
system, respectively. $R$ is the rotation matrix. $t$ is the translation matrix.

A. Tsai’s Camera Calibration Method

Roger Tsai described a two-stage calibration method based on RAC. The model analyzes the ideal case of perspective projection imaging and then also considers the radial distortion. The calibration is divided into two steps.

In the first step, there exists RAC of $x_{ci}/y_{ci} = x_{di}/y_{di}$ when $z_{ci} \neq 0$. Then, (2) can be obtained by considering RAC and (1).

$$
\begin{bmatrix}
x_{gi}y_{di} & y_{gi}y_{di} & z_{gi}y_{di} & y_{di} - x_{gi}x_{di} & -y_{gi}x_{di} & -z_{gi}x_{di}
\end{bmatrix}
\begin{bmatrix}
s_{i}t_{i}/T_{y} \\
s_{i}r_{i}/T_{y} \\
s_{i}T_{i}/T_{y} \\
r_{i}/T_{i} \\
r_{i}/T_{i} \\
1
\end{bmatrix} = x_{di}
$$

(2)

$x_{di}, y_{di}$ is the distorted projection coordinate. It can be obtained by the following equation.

$$
\begin{bmatrix}
u_{i} \\
v_{i}
\end{bmatrix} = \begin{bmatrix}
s_{i}^{-1}d_{x}^{-1}x_{di} + c_{x} \\
dy^{-1}y_{di} + c_{y}
\end{bmatrix}
$$

(3)

($u_{i}, v_{i}$) is the image coordinate. ($c_{x}, c_{y}$) is the coordinate of the principal point. $d_{x}, d_{y}$ is the scale factor in the X axis and Y axis respectively.

Equation (2) is processed by the least squares method, and $R$, $T_{i}$ and $T_{r}$ are be obtained by the constraint that $R$ is the orthogonal unit matrix. At this point, the first step of Tsai’s method is complete.

In the second step, there exists the projective theorem

$$
\begin{bmatrix}
x_{ci} \\
y_{ci}
\end{bmatrix} = \begin{bmatrix}
f \cdot x_{ci} \\
f \cdot y_{ci}
\end{bmatrix}
$$

(4)

($x_{ci}, y_{ci}$) is the projective projection coordinate, which is the same coordinate point as ($x_{di}, y_{di}$) when distortion is not considered. $f$ is the focal length. (5) is given by considering (1) and the equation in the y direction of (4).

$$
\begin{bmatrix}
y_{ci} & -y_{di}
\end{bmatrix} \cdot \begin{bmatrix}
f \\
T_{r}
\end{bmatrix} = \left( r_{x}x_{ci} + r_{y}y_{ci} + r_{z}z_{ci} \right) y_{di}
$$

(5)

Then, the initial value of the focal length $f$ and $T_{r}$ can be obtained by the least squares method. $x_{ci}, y_{ci}$ can be obtained from (1) and (3).

When the radial distortion coefficient $k$ is considered, a first-order radial distortion relation is obtained as shown in (6).

$$
\begin{bmatrix}
x_{di} = x_{di} \left( 1 + k \left( x_{di}^{2} + y_{di}^{2} \right) \right) \\
y_{di} = y_{di} \left( 1 + k \left( x_{di}^{2} + y_{di}^{2} \right) \right)
\end{bmatrix}
$$

(6)

Equation (5) can be deformed as (7).

$$
y_{di} = f \cdot \left( r_{x}x_{ci} + r_{y}y_{ci} + r_{z}z_{ci} + T_{r} \right) \left( r_{x}x_{ci} + r_{y}y_{ci} + r_{z}z_{ci} + T_{r} \right)
$$

(7)

$(x_{di}, y_{di})$ in (6) is represented by the principal point coordinate $(c_{x}, c_{y})$. Then, $y_{di}$ is substituted into 7. The final values of $f$, $k$, $T_{r}$ and $(c_{x}, c_{y})$ are searched as the unknown variables.

B. Factors that Affect the Performance of the Tsai’s Method

An analysis of the process of Tsai’s calibration method reveals the following factors that affect the performance of camera:

a. In the first step of Tsai’s method, when $(x_{di}, y_{di})$ is calculated by (3), initial principal point coordinate $(c_{x}, c_{y})$ is approximately the coordinate of the image center point (half of the image size). Consequently, $R$, $T_{r}$ and $T_{i}$ obtained by the least squares method are also approximations, which will produce errors for the solution of other parameters.

b. The objective function (7) in Tsai only considers the constraint in the y-axis direction and does not take into account the x-axis coordinate constraint. The geometric information is not fully utilized, which will affect the calibration accuracy. In addition, some algorithms consider $f = \min(n \times \Delta e)$ (see also (10)) as the objective function, but the nonlinear quadratic equation (6) must be solved for $x_{di}$ and $y_{di}$. The discussion of the root case will increase the additional calculations and complexity of the algorithm. Thus, the objective function of this form is not desirable.

c. When corners are coplanar or $z_{ci} = 0$, $s_{i}$ is set to 1 directly in Tsai’s method, and the exact solution of $s_{i}$ is not given.

d. More iterations must be performed to achieve higher calibration accuracy.

Tsai’s calibration process is also a process of optimizing the search. The purpose is to find the value of the parameters of the vector by the time the objective function reaches a minimum. The initial value of the optimization search and the number of unknown variables in the objective function should be redesigned to improve the calibration speed and reduce the amount of calculations based on iteration:

e. The initial value of the optimization search will affect the starting point of the search and, consequently, speed of searches. Thus, an initial value close to the true value of the camera parameters is preferable.

f. Searching for too many unknown variables in the objective function will further slow the search. Therefore, the number of unknown variables in the objective function should be appropriately reduced to improve the speed of the search.

III. A Camera Calibration Method Based on Iteration

Based on the analysis in section B, the solutions are as follows:

a. Ideally, the principal point coordinate $(c_{x}, c_{y})$ is very precise when it is search for with all other parameters. However, such a search is impossible because solving too many variables simultaneously is time-consuming. In addition, the nonlinear solution adopted in the search is prone to error. An idea is proposed in this paper that is, the variables should be divided
into two searches, in which \((c_x, c_y)\) is set as the global variables, which are searched for twice, while the other parameters are searched for only once.

b. The new objective function must be redesigned taking into account the constraints of the x-axis and y-axis direction to solve the problem that the objective function (7) only considers the constraint of the y-axis direction. The intermediate variables of the objective function are the perspective projection coordinate \((x_{d_i}, y_{d_i})\) to solve the problem of the quadratic nonlinear solution.

c. \(s_i\) will be solved precisely in the second search of every iteration.
d. The iterative method is added for higher accuracy of calibration.
e. The result calculated by Tsai’s method is seen as the initial value of the improved method, which is more close to true value of parameters.
f. In every iteration, the variables are divided between two searches. The total number of variables to be solved is the same, but the number of variables in every objective function is reduced. Consequently, the speed of search increases.

The objective function of the improved method in this article is shown in (8).

\[
\begin{align*}
  f' &= \min \sum_{i=1}^{n} \sqrt{(x_{ui} - x_{ui}')^2 + (y_{ui} - y_{ui}')^2} \\
  (x_{ui}, y_{ui}) &\text{ is calculated by (6). (x_{ui}', y_{ui}') is the reconstructed perspective projection coordinate, which can be calculated by (9).}
\end{align*}
\]

\[
\begin{align*}
  x_{ui}' &= f' \left[ \left( r_1 x_{ei} + r_2 y_{ei} + r_3 z_{ei} + T_x \right) / \left( r_1 x_{ei} + r_2 y_{ei} + r_3 z_{ei} + T_z \right) \right] \\
  y_{ui}' &= f' \left[ \left( r_1 x_{ei} + r_2 y_{ei} + r_3 z_{ei} + T_y \right) / \left( r_1 x_{ei} + r_2 y_{ei} + r_3 z_{ei} + T_z \right) \right]
\end{align*}
\]

The reconstruction error is calculated as shown in (10).

\[
\Delta e = \frac{1}{n} \sum_{i=1}^{n} \sqrt{(u_i - \tilde{u}_i)^2 + (v_i - \tilde{v}_i)^2}
\]

(\(u_i, v_i\)) is image coordinate extracted by the Harris algorithm, and \(n\) is the number of corners. (1) can be solved to calculate \((x_{ei}, y_{ei}, z_{ei})\). (4) is solved to calculate \((x_{ui}, y_{ui})\) by the Projective Theorem. \((x_{d_i}, y_{d_i})\) can be obtained by solving (6) nonlinearly way and then substituting \((x_{d_i}, y_{d_i})\) into (3) to obtain the reconstructed image coordinate \((\tilde{u}_i, \tilde{v}_i)\).

If \(\Delta e\) converges or the variation of \(\Delta e\) is less than a given positive \(\varepsilon\), the iteration ends. Otherwise the search is used twice to solve (8) until the reconstruction error converges or the variation is less than the given positive \(\varepsilon\). The algorithm flow is shown as follows.

**Step 1.** Initialize

The traditional Tsai’s method is used to obtain the parameters that will be the initial value of the optimizing search in the first iteration.

**Step 2.** Obtain the principal point coordinate \((c_x, c_y)\), \(R\), \(T_x\) and \(T_y\).

The scale factor \(s_i\), the focal length \(f\), \(T_z\) and the distortion coefficient \(k\) are fixed, and then (8) is solved to obtain \((c_x, c_y)\), \(R\), \(T_x\) and \(T_y\) by the L-M method.

**Step 3.** Obtain the principal point coordinate \((c_x, c_y)\), \(k\), \(f\), \(T_x\) and \(T_y\).

\(R\), \(T_x\) and \(T_y\) are fixed, and then (8) is solved to obtain \((c_x, c_y)\), \(k\), \(f\), \(T_x\) and \(s_i\) by the L-M method.

**Step 4.** Calculate the reconstruction error \(\Delta e\).

The reconstruction error \(\Delta e\) is calculated by (10). If \(\Delta e\) converges or the change in \(\Delta e\) is less than the given positive \(\varepsilon\), the iteration will end. Otherwise proceed to **Step 2**.

**IV. Experiment**

There are many algorithms for search optimization. The Levenberg-Marquardt (L-M) algorithm combines the advantages of the gradient method and Newton method [13] and is used in this experiment. The given value \(\varepsilon\) is set to 0.001. Because 2D targets are easy to produce, the experiment uses a 2D planar target with 8 corners on each row and 11 corners on each column, in which the distance between adjacent points is 20 mm. A Sony Cyber-shot DSC-P120 camera was used with
the following specific parameters: image size of 640 pixels × 480 pixels, image center coordinate of (320,240), maximum image resolution of 2592 pixels × 1944 pixels, number of pixels in the horizontal and vertical unit lengths $N_x = N_y = 358.4229$ (pixels/mm) so $d_x = d_y = 0.0113$ mm.

Six images shot by this camera were used in an experiment in Matlab2014a. First, the image was grayed up, and the Harris corner detection algorithm was used to extract the coordinates of the corners. The world coordinate system, the camera coordinate system and the image coordinate system were established, and then the corresponding coordinate of each corner was calculated. Finally, the internal and external parameters of the camera were calibrated by Tsai’s method or the improved method, and the calibration accuracies of two methods were compared.

### A. Evaluation Criteria

Common evaluation indicators of the calibration algorithm are as follows.

- **MAE (Mean Absolute Error):**
  
  $$M^A E = \frac{1}{MN} \sum_{i=1}^{M} \sum_{j=1}^{N} \sqrt{(u_{ij} - \overline{u}_{ij})^2 + (v_{ij} - \overline{v}_{ij})^2}$$  

- **RMSE (Root Mean Square Error):**
  
  $$R M S E = \sqrt{\frac{1}{MN} \sum_{i=1}^{M} \sum_{j=1}^{N} \left((u_{ij} - \overline{u}_{ij})^2 + (v_{ij} - \overline{v}_{ij})^2\right)}$$

- **σ (Standard Deviation):**

  $$\sigma = \sqrt{\frac{1}{MN} \sum_{i=1}^{M} \sum_{j=1}^{N} error(i,j) - \mu}$$  

$M$ and $N$ represent the $M$-th row and the $N$-th column of the image, respectively; $error(i,j)$ represents the error of the $i$-th row and the $j$-th column in the image; and $\mu$ represents the mean of the error. The first two indicators are used to describe the calibration accuracy, and the third is used to describe the discrete degree of the error of an image.

### B. Experimental Results and Analysis

The image was grayed out, and then the Harris algorithm was used to extract image coordinate $(u_i, v_i)$. The reconstructed image coordinate $(\overline{u}_i, \overline{v}_i)$ was obtained by the improved method. One of the six images (Image ID: Image 1) that was processed is shown in Fig. 2. The positions of $(\overline{u}_i, \overline{v}_i)$ (blue star mark) and $(u_i, v_i)$ (red cross mark) almost overlap, which indicates the error of the improved method is very small.

![Fig. 2 Corner point extraction and positions of two kind points](image)

Six images were tested and the calibration results of the two methods are shown in Table 1 and Table 2.

<table>
<thead>
<tr>
<th>Image ID</th>
<th>$c_x$</th>
<th>$c_y$</th>
<th>$s_x$</th>
<th>MAE (pixels)</th>
<th>Iterations</th>
</tr>
</thead>
<tbody>
<tr>
<td>Image 1</td>
<td>320.24</td>
<td>240.11</td>
<td>1</td>
<td>0.6742</td>
<td>1</td>
</tr>
<tr>
<td>Image 2</td>
<td>320.17</td>
<td>240.09</td>
<td>1</td>
<td>0.7385</td>
<td>1</td>
</tr>
<tr>
<td>Image 3</td>
<td>319.76</td>
<td>239.89</td>
<td>1</td>
<td>0.6409</td>
<td>1</td>
</tr>
<tr>
<td>Image 4</td>
<td>319.85</td>
<td>239.69</td>
<td>1</td>
<td>0.7005</td>
<td>1</td>
</tr>
<tr>
<td>Image 5</td>
<td>319.85</td>
<td>239.86</td>
<td>1</td>
<td>0.8065</td>
<td>1</td>
</tr>
<tr>
<td>Image 6</td>
<td>319.83</td>
<td>239.52</td>
<td>1</td>
<td>0.7850</td>
<td>1</td>
</tr>
<tr>
<td>Average</td>
<td>319.95</td>
<td>239.86</td>
<td>1</td>
<td>0.7243</td>
<td>1</td>
</tr>
</tbody>
</table>
Table 2. Calibration results based on the iterative method

<table>
<thead>
<tr>
<th>Image ID</th>
<th>$c_x$</th>
<th>$c_y$</th>
<th>$s_x$</th>
<th>$MAE$(pixels)</th>
<th>Iterations</th>
</tr>
</thead>
<tbody>
<tr>
<td>Image 1</td>
<td>320.27</td>
<td>240.13</td>
<td>1.0003</td>
<td>0.5816</td>
<td>3</td>
</tr>
<tr>
<td>Image 2</td>
<td>319.51</td>
<td>240.13</td>
<td>1.0001</td>
<td>0.6619</td>
<td>2</td>
</tr>
<tr>
<td>Image 3</td>
<td>319.77</td>
<td>239.92</td>
<td>0.9996</td>
<td>0.5209</td>
<td>2</td>
</tr>
<tr>
<td>Image 4</td>
<td>319.82</td>
<td>239.71</td>
<td>1</td>
<td>0.5922</td>
<td>3</td>
</tr>
<tr>
<td>Image 5</td>
<td>319.80</td>
<td>240.03</td>
<td>1.0008</td>
<td>0.6477</td>
<td>5</td>
</tr>
<tr>
<td>Image 6</td>
<td>319.84</td>
<td>239.53</td>
<td>1.0005</td>
<td>0.6032</td>
<td>2</td>
</tr>
<tr>
<td>Average</td>
<td>319.84</td>
<td>239.91</td>
<td>1.0002</td>
<td>0.6013</td>
<td>2.83</td>
</tr>
</tbody>
</table>

Table 1 shows that only one iteration is performed in Tsai’s method. The search of $s_i$ is ignored because $s_i$ is assumed to be 1. The average of $(c_i, c_j)$ is $(319.95, 239.86)$, and the average of $MAE(\Delta e$ of last iteration) is 0.7243 pixels.

As shown in Table 2, $(c_i, c_j)$ in the improved algorithm is close to the center of the image but is not the center of the image. The average of $(c_i, c_j)$ is $(319.84, 239.91)$; the average of $s_i$ is 1.0002; the average number of iterations is 2.83. The criterion for ending iteration is satisfied with fewer iterations. The average $MAE$ is 0.6013 pixels. The combination of several iterations, accurate solution of the principal point & $s_i$, and other improved measures reduced the reconstruction error by approximately 20.46%. The improved method significantly improved the accuracy of calibration and achieved the desired results.

In addition, when the speed of search is compared, the average time of the improved method was 0.2385 seconds, compared to an average running time of 0.2440 seconds for Tsai. Designing the form of the objective function and setting an initial value in the optimization search reduces computation and increases the speed of the search. Thus, a single iteration takes less time in the improved method. Because there are multiple iterations, the improved algorithm requires a greater total time than Tsai. This is an unavoidable disadvantage of iterative computations.

One of the six images (Image ID: Image 1) was selected to analyze the error distribution of 88 points. In Fig. 3, the ordinate represents the error of each point, and the abscissa represents the sequence of the corners set by the ranks. The improved method proposed in this paper effectively reduced the error of most corners and improved the error greatly. The method is designed to improve the overall error; consequently, the errors of partial points may be larger than those in Tsai’s method, as shown in Fig. 3.
Table 3 shows that the RMSE of the improved algorithm is lower than Tsai’s for every calibration. The RMSE of the improved algorithm is reduced by approximately 30%, and the calibration accuracy is significantly improved, consistent with the data in Table 1 and Table 2. As shown in Fig. 3 and Table 3, the standard deviation $\sigma$ of the error distribution of the improved algorithm is smaller, which indicates that the error fluctuation of the improved algorithm is smaller, and thus the calibration result is more stable and reliable. This is also an advantage of the improved algorithm.

V. CONCLUSION

Tsai’s method is widely used but its accuracy urgently requires further improvement. In this paper, the factors affecting the accuracy of Tsai’s method were analyzed, and an improved method based on iterations was proposed. By redesigning the form of the objective function, the nonlinear solution of the quadratic equation and the discussion of the root are omitted, thus reducing the computational complexity. Redesigning the initial value of the optimizing search increases the speed of the optimizing search. The improved algorithm sets the coordinate of the principal point as a global variable. The RMSE of the improved calibration method is approximately 30% smaller than Tsai’s. The improved algorithm has advantages in speed and accuracy of calibration and thus can be applied in engineering practice.

ACKNOWLEDGMENT

This work was partly supported by the National Natural Science Foundation of China (Grant No. 31660239) and Beijing municipal construction project special fund. The authors have no conflict of interest to declare.

REFERENCES


