Trinocular Calibration Method Based on Binocular Calibration

XU Zhen-Ying*, LI Wen-Bin, WANG Yun, LUO Chun, GAO Shu-Yuan, CAO Dan-Dan
School of Mechanical Engineering, Jiangsu University, Zhenjiang, 212013, China
*corresponding author, e-mail: zhyxu77@163.com

Abstract
In order to solve the self-occlusion problem in plane-based multi-camera calibration system and expand the measurement range, a tri-camera vision system based on binocular calibration is proposed. The three cameras are grouped into two pairs, while the public camera is taken as the reference to build the global coordinate. By calibration of the measured absolute distance and the true absolute distance, global calibration is realized. The MRE (mean relative error) of the global calibration of the two camera pairs in the experiments can be as low as 0.277% and 0.328% respectively. Experiment results show that this method is feasible, simple and effective, and has high precision.

Keywords: computer vision, calibration, trinocular calibration, multi-camera vision system, global calibration

1. Introduction
The calibration of tri-camera vision system for large-scale measurement includes local and global calibration, that is, the calibration of camera intrinsic parameters and the calibration of the whole system. When we calibrate multi-cameras (three or more) with a planar pattern, self-occlusion occurs because of the pattern’s limited field-of-view (FOV) [1,2]. To solve this problem, Wang Liang et al. proposed a kind of one-dimensional (1D) calibration object, which is comprised of several collinear points (usually spheres) whose mutual distances had already been known [3]. However, errors may exist because of sphere eccentricity and errors in the location detection of the spheres. Research work has also been carried out by Zhang Guangjun et al. [4] to study the calibration based on double calibration targets, while the added position constraint requires high-precision calibration targets and increases matrix conversion errors, also, the targets are difficult to prepare when the cameras are at a distance. Li bin [5] has successfully calibrated three cameras by fixing them together, and moving the planar pattern. Frank Pagel [6] installed the cameras on the vehicle, rotated the vehicle so that the camera can observe the objects with known-structure, thus to calibrate the cameras. However, the above-mentioned two methods both require obtaining the camera group’s motion information qualitatively and quantitatively. The self-calibration approach presented by Chen X. [7] is to use a laser as a calibrator in a low-light room, but this approach can not be applied in a high-light situation, also it is prone to inaccuracies.

In order to overcome such drawbacks above, here we propose a tri-camera calibration system based on binocular calibration. Cameras are grouped into two binocular pairs which have common field-of-view (FOV) so that the problem of self-occlusion can be solved. The method can be used for global calibration of the cameras more than three.

2. Extrinsic Calibration of Binocular Vision System
In the tri-camera calibration system, each camera-pair forms a binocular vision system. Angular variation between the two cameras is allowed as long as common (FOV) exists, as shown in Figure 1.

In the system shown in figure 1, the relation between the left camera coordinate system $O_1 - x_1 y_1 z_1$ and the global coordinate system $O_w - X_w Y_w Z_w$ can be expressed as a space conversion matrix $M_1$.
\[
\begin{bmatrix}
x_r \\
y_r \\
z_r \\
1
\end{bmatrix}
= M_y
\begin{bmatrix}
x_w \\
y_w \\
z_w \\
1
\end{bmatrix}
= \begin{bmatrix}
r_{11} & r_{12} & r_{13} & t_{ls} \\
r_{21} & r_{22} & r_{23} & t_{ls} \\
r_{31} & r_{32} & r_{33} & t_{ls} \\
0 & 0 & 0 & 1
\end{bmatrix}
\begin{bmatrix}
x_w \\
y_w \\
z_w \\
1
\end{bmatrix}
\]

\[M_y = [R_y \ T_y]\] 

(1)

(2)

Where, \(R_y\) and \(T_y\) respectively represent the Rotation matrix between the camera coordinate and global coordinate system, and the vector of translation transformation between the origins.

Figure 1. Model of binocular vision with an angle

Similarly, the relation between the right camera coordinate system \(O_r - x_r y_r z_r\) and the global coordinate system \(O_w - X_w Y_w Z_w\) can be expressed as a space conversion matrix \(M_r\):

\[
\begin{bmatrix}
x_r \\
y_r \\
z_r \\
1
\end{bmatrix}
= M_r
\begin{bmatrix}
x_w \\
y_w \\
z_w \\
1
\end{bmatrix}
= \begin{bmatrix}
r_{r1} & r_{r2} & r_{r3} \\
r_{r4} & r_{r5} & r_{r6} \\
r_{r7} & r_{r8} & r_{r9} \\
0 & 0 & 0 & 1
\end{bmatrix}
\begin{bmatrix}
x_w \\
y_w \\
z_w \\
1
\end{bmatrix}
\]

\[M_r = [R_r \ T_r]\] 

(3)

(4)

where, \(R_r\) represents the rotation matrix between the camera coordinate and global coordinate system, \(T_r = \begin{bmatrix} t_{rx} & t_{ry} & t_{rz} \end{bmatrix}^T\) is the vector of translation transformation between the origins.

From Equation (1) and (3) we can see that, to any point \(P\) in the coordinate system \(O_w - X_w Y_w Z_w\), the corresponding relationship between the two cameras is

\[
\begin{bmatrix}
x_r \\
y_r \\
z_r
\end{bmatrix}
= R_y R_y' \begin{bmatrix} x_w \\
y_w \\
z_w
\end{bmatrix}
+ T_y R_y' T_y
\]

(5)
Therefore, to achieve the coordinate transformation between the two cameras, the rotation matrix and translation matrix can be expressed as [8]:

\[
\begin{align*}
R_i &= R_iR_{i+1}^{-1} \\
T_i &= T_i - R_iR_{i+1}^{-1}T_{i+1}
\end{align*}
\] (6)

Where: \(R_i\), \(T_i\) is the rotation matrix and translation matrix from the right camera to the left, respectively.

In binocular vision system, intrinsic and extrinsic parameters of the camera-pair are calibrated by one of the cameras, and then relative position of the two cameras can be worked out according to equation (6).

3. Global Calibration of Tri-camera

The purpose of the global calibration is to determine the relative position and orientation of the various visual sensor coordinate systems to the global coordinate, i.e. rotation matrix and a translation vector [9]. Such position and orientation can rarely be obtained by sensors. Instead, global calibration on the spot is essential, which is also the key to determine the accuracy of the entire system [10].

To reach a more precise calibration, one of the three cameras (usually the middle one) is regarded as the public camera, and respectively forms the binocular vision system with the other two cameras, as is shown in Figure 2. In this way, positional relationship of the two camera-pairs can be obtained with the calibration of the space translation matrix as shown in equation (6).

3.1 The Positional Relations Solving

In Figure 2, we take \(C_1\) as the public camera, take \(C_1-C_2\) and \(C_1-C_3\) as two sets of binocular stereo vision. After the binocular calibration, their position relations can be obtained according to equation (6):

\[
\begin{align*}
R_{21} &= R_2R_1^{-1} \\
T_{21} &= T_2 - R_2R_1^{-1}T_1
\end{align*}
\] (7)

Where: \(R_{21}\), \(T_{21}\) are the rotation matrix and translation matrix from \(C_2\) coordinate system to \(C_1\), respectively.

\[
\begin{align*}
R_{31} &= R_3R_1^{-1} \\
T_{31} &= T_3 - R_3R_1^{-1}T_1
\end{align*}
\] (8)

Where: \(R_{31}\), \(T_{31}\) are the rotation matrix and translation matrix from \(C_1\) coordinate system to the \(C_2\), respectively.

In tri-camera system, we integrate the measured data by global calibration which integrates three camera coordinate systems into one [11]. As is shown in figure 2, coordinate system of the camera \(C_1\) is the unified coordinate system. In this coordinate, position and orientation of the other two cameras can be respectively obtained through equation (7) and equation (8).

3.2 Nonlinear Optimization

Locate the planar target in position \(i\), \(p_{i,j}\) is the \(j\)th feature point, \(p_{w,1,j}^c\) and \(p_{w,2,j}^c\) respectively represent the homogeneous coordinates of undistorted image and re-projection image in the coordinate system of \(C_1\). Summarily, \(p_{w,1,j}^c\) and \(p_{w,2,j}^c\), \(p_{w,3,j}^c\) and \(p_{w,4,j}^c\) are the coordinates in the coordinate system of \(C_2\), respectively. \(p_{w,5,j}^c\) is the coordinates in the coordinate system of \(C_1\) according to the camera perspective projection model, the following equations can be obtained:

\[
\begin{align*}
\rho_1p_{w,1,j}^c &= A_1(I 0)p_{w,5,j}^c \\
\rho_2p_{w,2,j}^c &= A_2(I 0)M_2^c p_{w,5,j}^c \\
\rho_3p_{w,3,j}^c &= A_3(I 0)M_3^c p_{w,5,j}^c
\end{align*}
\] (9)
Where $\rho_1$, $\rho_2$, $\rho_3$ are arbitrary non-zero scale factor. $A_{c1}$, $A_{c2}$, $A_{c3}$ are the matrices of camera intrinsic parameters, $M_{c2}$, $M_{c3}$ are the conversion matrixes from C1 to C2, C1 to C3, respectively.

Assume the noise subject to Gaussian distribution and independent distribution, In order to get the optimal solution of $M_{c2}$, $M_{c3}$ under the maximum likelihood criterion, optimization equations can be established as followed:

$$f(a) = \sum_{i=1}^{m} \sum_{j=1}^{n} d(p_{im,j}^{c1}, \hat{p}_{im,j}^{c1})^2 + \sum_{i=1}^{m} \sum_{j=1}^{n} d(p_{im,j}^{c2}, \hat{p}_{im,j}^{c2})^2 + \sum_{i=1}^{m} \sum_{j=1}^{n} d(p_{im,j}^{c3}, \hat{p}_{im,j}^{c3})^2$$

(10)

where, $m$, $n$ respectively represent the times of placing of the target and the number of feature points, $a = (M_{c2}^T, M_{c3}^T)$.

Then, with the nonlinear optimization method (Levenberg-Marquardt), the optimal solution under the maximum likelihood criterion can finally be obtained.

4. Experiments and Results Analysis

4.1 Calibration of Tri-camera and Measurement Experiments

In the experiment, we use three same industrial cameras (AVT GuppyPro the F-201B) with the resolution 1024 * 768; Lens: Computar M3Z1228C-MP. Focal length: 12-36mm. The system is shown in Figure 3. A high-precision calibration board from Shanghai Luhua Technology Co, Ltd. is used, which has the dimensional accuracy and position accuracy ± 3µm, as is shown in Figure 4.

The left binocular stereo vision is comprised by C1 and C2, the right binocular stereo vision is comprised by C1 and C3. After binocular calibration, all the camera intrinsic parameters are obtained, as shown in Table 1 and Table 2.

According to equation (8) and nonlinear optimization method, the rotation matrix and translation matrix from C2 to C1 are obtained as followed:

$$R_{21} = \begin{bmatrix} 0.9838 & -0.0362 & 0.1757 \\ -0.0321 & 0.9992 & 0.0264 \\ -0.1765 & -0.0203 & 0.9841 \end{bmatrix}$$

$$T_{21} = \begin{bmatrix} -252.5587 \\ -0.8062 \\ -70.1793 \end{bmatrix}$$

According to equation (8) and nonlinear optimization method, the rotation matrix and translation matrix from C3 to C1 are obtained as followed:
\[
R_{31} = \begin{bmatrix}
0.9811 & -0.0060 & 0.1934 \\
0.0040 & 0.9999 & 0.0109 \\
-0.1934 & -0.0099 & 0.9810 \\
\end{bmatrix}
\]
\[
T_{31} = \begin{bmatrix}
-280.645 \\
-1.4581 \\
32.5050 \\
\end{bmatrix}
\]

Table 1 Camera intrinsic parameters of the left binocular vision module

<table>
<thead>
<tr>
<th>Intrinsic parameters</th>
<th>(C_1)</th>
<th>(C_2)</th>
</tr>
</thead>
<tbody>
<tr>
<td>(f_x)</td>
<td>5559.0651</td>
<td>8226.5977</td>
</tr>
<tr>
<td>(f_y)</td>
<td>5558.9282</td>
<td>8225.8982</td>
</tr>
<tr>
<td>(u_0)</td>
<td>513.1371</td>
<td>503.4972</td>
</tr>
<tr>
<td>(v_0)</td>
<td>380.9580</td>
<td>390.8671</td>
</tr>
<tr>
<td>(k_1) (radial distortion once)</td>
<td>0.1092</td>
<td>0.41547</td>
</tr>
<tr>
<td>(k_2) (radial distortion twice)</td>
<td>-0.3936</td>
<td>0.4876</td>
</tr>
<tr>
<td>(p_1) (tangential distortion once)</td>
<td>0.0056</td>
<td>0.0031</td>
</tr>
<tr>
<td>(p_2) (tangential distortion twice)</td>
<td>0.00008</td>
<td>0.0079</td>
</tr>
</tbody>
</table>

Table 2 Camera intrinsic reference value of the right binocular vision module

<table>
<thead>
<tr>
<th>Intrinsic parameters</th>
<th>(C_1)</th>
<th>(C_2)</th>
</tr>
</thead>
<tbody>
<tr>
<td>(f_x)</td>
<td>5558.7716</td>
<td>5641.0477</td>
</tr>
<tr>
<td>(f_y)</td>
<td>5559.3893</td>
<td>5640.6993</td>
</tr>
<tr>
<td>(u_0)</td>
<td>528.6553</td>
<td>506.8564</td>
</tr>
<tr>
<td>(v_0)</td>
<td>390.7990</td>
<td>388.6202</td>
</tr>
<tr>
<td>(k_1) (first radial distortion)</td>
<td>0.1090</td>
<td>0.0174</td>
</tr>
<tr>
<td>(k_2) (second radial distortion)</td>
<td>-0.3994</td>
<td>0.1338</td>
</tr>
<tr>
<td>(p_1) (first tangential distortion)</td>
<td>0.00553</td>
<td>0.0034</td>
</tr>
<tr>
<td>(p_2) (second tangential distortion)</td>
<td>0.0001</td>
<td>-0.0042</td>
</tr>
</tbody>
</table>

Figure 5 Corner points \(p\) and \(q\) in the calibration board

The absolute distance between the two corner-points of the high-precision calibration board is considered to verify the results of calibration, as is shown in figure 5. Coordinate value of the points are \(p(0, 100, 0), q(100, 0, 0)\) in the panel, so the absolute distance between them is \(|pq| = 141.421\text{mm}\).

Take 6 photos for the board. Take coordinate of \(C_1\) as the space unified system, consider the above-mentioned global calibration results and corner-points detection method, then through the photos, the coordinate values of space points \(p\) and \(q\) of the two camera-pairs can be respectively obtained, they are \(p, p, q, q, |pq|\). The absolute distance between \(p\) and \(q\) can be solved, as is shown in table 3. Compare \(|pq|\) with the truth distance \(|pq|_0\) [12], then the relative error can be worked out. Table 3 shows the corresponding values and the average error and the standard deviation of the values.

As Table 3 shows, measurement accuracy of the binocular stereo vision systems are about 0.09% and 0.145% respectively. RMS can reach 0.014 mm and 0.084 mm respectively. Compare with the above-mentioned methods with one-dimensional (1D) calibration object or double calibration targets [6] (in which, 1280 * 1024 resolution CCD had been used, 380 points had been measured), trinocular calibration method has advantages.

Trinocular Calibration Method Based on Binocular Calibration (XU Zhen-Ying)
4.2 Accuracy Analysis

The experiment results show that the proposed global calibration based on binocular stereo calibration method has the advantages of high precision and simple operation. However, errors still exist due to some objective and man-made factors. The factors are as follows:
1) Laboratory equipment error: the errors caused by the camera lens distortion.
2) The standard amount error: the errors in the distance of $p$ and $q$ in checkerboard.
3) Errors caused by illumination: gray scale of the camera imaging changes with the change of outside light, thus lead to the corner extraction errors.
4) Noise and corner detection error.
5) Human improper operation or negligence.

To avoid or reduce the errors mentioned above, the global calibration accuracy can still be improved in some extent.

5. Conclusions

Global calibration of multi-camera is not only the hotspot, but also the research difficulties in large-size measurement techniques based on machine vision, which can greatly affect the measurement accuracy of the system. To solve the self-occlusion problems exist in the calibration of the multi-camera (3 or more), a global calibration method based on binocular calibration has been proposed. Experiments showed a good result of the measurement, proving it a good solution for the self-occlusion problem in multi-camera calibration. Furthermore, as it doesn’t need a common field-of-view (FOV) between each camera pairs, this method is also a good way to extend the measurement range to a degree, which shows the practical value in the industrial measurements.

Acknowledgement

This paper is from the project of National Natural Science Foundation (No. 50975126), the key project of Natural Science Foundation of Jiangsu province (No.BK2010042), and the project of Talent Foundation of Jiangsu Province (No. BRA2011180).

References: