A Self-Calibration Method of Varying Internal Camera Parameters Based on Binocular Stereo Vision

Zetao Jiang, Lianggang Jia, Shutao Guo

School of Information Engineering, Nanchang Hangkong University, Nanchang, China,
zetaojiang@126.com
School of Information Engineering, Nanchang Hangkong University, Nanchang, China,
jlg804@126.com; gshtao0617@126.com

Abstract

This paper presents a self-calibration method of varying internal camera parameters based on binocular stereo vision. In this method, using a binocular vision system to obtain a focal length of the image, and get the foundation matrix by using the match points, then build the double linear equations about focal square via utilizing the kruppa equation. Secondly, make relative orientation to get the all external parameters get through letting the baseline length equal the distance between the two cameras. Finally, make the area network adjustment with importing the camera’s nonlinear distortion parameters using the above initial value and the characteristic of camera’s relatively fixed position and posture to get the all precise camera parameters. The experimental results show that the back projection of the method can be controlled in about 5 pixels, and this method has certain anti-noise trait.

Keywords: Vary Inner Parameters, Kruppa Equation, Binocular Stereo Vision, Relative Orientation, Area Network Adjustment

1. Introduction

Binocular stereo vision calibration is a bridge in the world coordinate between two-dimensional images and three-dimensional coordinates. The calibration accuracy directly determines the accuracy of three-dimensional reconstruction, and is one of the hot issues in the field of computer vision research [1,2]. The traditional calibration methods often create the corresponding relationship between two-dimensional images and three-dimensional coordinates of the world coordinates to solve the image of the inner and outer parameters through two-dimensional or three-dimensional high-precision calibration object. The traditional calibration method based on two-dimensional calibration block was proposed in 1999 by Zhang Zhengyou based on camera calibration of 2-dimensional plane template. Calibration method based on three-dimensional block is the uses of three-dimensional DLT calibration transform to obtain Initial, and re-use bundle adjustment to the exact solution outside the parameters of the camera. To avoid this calibration method’s shortcomings based on calibration object, from the 1990s, a lot of scholars in vision research began to study camera self-calibration [4-6]; this is also the future trend of camera calibration. Currently, most self-calibration methods are based on the absolute conic to the absolute dual quadric [7]; a typical representative is the self-calibration method based on Kruppa equations proposed by Maybank and Faugeras [8], this method usually requires camera to do several a special movements, cannot meet the general self-calibration.

The self-calibration methods are used in case camera intrinsic parameters is fixed, less research on the changes of camera calibration intrinsic parameter, most of them need camera to do a special the translation or rotation. If Pollefeys, etc. control the camera to do a flat move in order to calculate the initial focus, and then re-use model constraints to achieve the calibration of the focal length changes [9]. However, in many cases, camera are arbitrary, and in order to quickly recover three-dimensional information, often using binocular vision, which requires calibrated changes the intrinsic parameters in binocular vision system. In this paper, the self-calibration method is proposed based on the changes of binocular vision of intrinsic parameters, through simulation and real test analysis shows that the error of back-projection using the method can be controlled at about 5 pixels, and has some of the anti-noise.
2. imaging model of binocular vision

O−X_WY_WZ_W, the world coordinate system, the external parameters of the left camera in the world
coordinate system are R_l, T_l, and of the right are R_r, T_r, (R, T are the diagonal elements and line
elements of the outside parameters), The relationship of the right camera relative to the left camera are
R and T, binocular vision imaging model shown in Fig.1:

![Fig.1. Binocular vision imaging model](image)

Camera imaging follow the pinhole imaging, such as squares (x, y) and the total line
(X, Y, Z) meet in object-point equation, as shown in equation (1).

\[
\begin{align*}
    x - x_0 - \Delta x &= -f \frac{a_1 (X - X_s) + b_1 (Y - Y_s) + c_1 (Z - Z_s)}{a_1 (X - X_s) + b_1 (Y - Y_s) + c_1 (Z - Z_s)} = -f \frac{X}{Z} \\
    y - y_0 - \Delta y &= -f \frac{a_2 (X - X_s) + b_2 (Y - Y_s) + c_2 (Z - Z_s)}{a_2 (X - X_s) + b_2 (Y - Y_s) + c_2 (Z - Z_s)} = -f \frac{Y}{Z}
\end{align*}
\]

(1)

Where: \((x_0, y_0), f\), are respectively the image main point and focal length of the camera parameters,
\(\Delta x, \Delta y\) are distortion error correction coefficient, as the following shown.

\[
\begin{align*}
    \Delta x &= (x - x_0) \left(K_1 r^2 + K_2 r^4 \right) + P_1 \left(r^2 + 2(x - x_0)^2 \right) + 2P_2 (x - x_0)(y - y_0) \\
    \Delta y &= (y - y_0) \left(K_1 r^2 + K_2 r^4 \right) + P_3 \left(r^2 + 2(y - y_0)^2 \right) + 2P_4 (x - x_0)(y - y_0)
\end{align*}
\]

\(K_1, K_2\) are the radial distortion, \(P_1, P_2\) are the eccentric distortion, the intrinsic parameters matrix is:

\[
K = \begin{bmatrix}
    f & 0 & x_0 \\
    0 & f & y_0 \\
    0 & 0 & 1
\end{bmatrix}
\]

\(a_1, a_2, b_1, b_2, b_3, c_1, c_2, c_3\) are the camera’s diagonal elements of the external parameters.
\(X_s, Y_s, Z_s\) are the camera’s line elements of the external parameters. Assume the external parameters
the camera of the binocular vision, as the follow.
\[
R_r = \begin{bmatrix}
    a_{1r} & a_{2r} & a_{3r} \\
    b_{1r} & b_{2r} & b_{3r} \\
    c_{1r} & c_{2r} & c_{3r}
\end{bmatrix}, \quad R_l = \begin{bmatrix}
    a_{1l} & a_{2l} & a_{3l} \\
    b_{1l} & b_{2l} & b_{3l} \\
    c_{1l} & c_{2l} & c_{3l}
\end{bmatrix}, \quad R_r^{-1}T_r = \begin{bmatrix}
    X_{sl} \\
    Y_{sl} \\
    Z_{sl}
\end{bmatrix}, \quad T_l = \begin{bmatrix}
    X_{sr} \\
    Y_{sr} \\
    Z_{sr}
\end{bmatrix}
\]

\[
T = T_r - R_rR_l^{-1}T_l\]

the camera left and right relative position of binocular vision, and

\[
R = R_lR_r^{-1}\]

Attitude, is to calibrate the external parameters of This binocular stereo vision. \((x_{sl}, y_{sl})\), the left camera intrinsic parameters of binocular vision, \(f_l\) and \((x_{sr}, y_{sr})\) which are the right camera intrinsic parameters, \(f_r\) to be the calibration intrinsic parameters (the different focal length between the left and the right camera).

3. The solution of the left and right focal length of binocular vision

To solve fundamental matrix \(F\) though binocular visual match point, and then re-use matrix t Kruppia equation about the intrinsic parameters is derived, such as the type (2) shown:

\[
\frac{r^2v^2_iCv_i}{u^2_iC'u_2} = \frac{sr^2v^2_iCv_i}{u^2_iC'u_1} = \frac{s^2v^2_iCv_i}{u^2_iC'u_1}
\]

In the formula: \(C = (KK^T)^{-1}\), \(C' = (KK'^T)^{-1}\), \(F = UDV^T\), \(u_i\), \(v_i\) are the column vector of \(U\) and \(V\), \(r\), \(s\) are the proportional coefficient. Assume only the focal length \(f_l\), \(f_r\) unknown in the left and right camera intrinsic parameters matrix \(K\), \(K'\). As the main image point of the camera \(X_0, Y_0\), \(X_0', Y_0'\) are 0. (Now the industry camera can ensure that the image main point close to the center of the image, focal length based on the assumption that solving as the block adjustment initial value is sufficient.) So that we can build on the bilinear equations \(f_l^2\), \(f_r^2\), as the follows:

\[
\begin{align*}
    A_1f_l^2 + B_1f_r^2 + C_1f_l^2f_r^2 + D_1 &= 0 \\
    A_2f_l^2 + B_2f_r^2 + C_2f_l^2f_r^2 + D_2 &= 0
\end{align*}
\]

In the formula:

\[
\begin{align*}
    A_1 &= r^2u_1u_3(v_1^2 + v_2^2) - sr^2v_3(v_1v_2 + v_1v_2) \\
    B_1 &= r^2v_1v_3(u_1u_2 + u_1u_2) - sr^2v_3v_2(u_1v_2 + u_2v_2) \\
    C_1 &= r^2(v_1^2 + v_2^2)(u_1u_2 + u_2u_2) - sr^2(v_1^2 + v_2^2)(v_1v_2 + v_2v_2) \\
    D_1 &= r^2u_1u_3v_3^2 - sr^2v_3u_3^2 \\
    A_2 &= s^2u_1u_3(v_1^2 + v_2^2) - sr^2v_3(v_1v_2 + v_1v_2) \\
    B_2 &= s^2v_1v_3(u_1u_2 + u_1u_2) - sr^2v_3v_2(u_1v_2 + u_2v_2) \\
    C_2 &= s^2(v_1^2 + v_2^2)(u_1u_2 + u_2u_2) - sr(u_1^2 + u_2^2)(v_1v_2 + v_2v_2) \\
    D_2 &= s^2u_1u_3v_3^2 - sr^2v_3u_3^2
\end{align*}
\]
By solving the bilinear equations, can solve the left and right camera’s focal length \( f_l \) and \( f_r \).

4. The relative orientation of Binocular vision

Relative orientation is the areas of photogrammetric, often used to determine the relative position and gesture of the left and right camera, imaging of binocular visual and geometry, as shown in Fig.2:

![Fig.2. Binocular visual imaging geometry diagram](image)

In the figure, \( m_1, m_2 \), express the model point M on the imaging in the left and right images, \( S_1 m_1, S_2 m_2 \) are a pair of the same name light (Match point of the light generated), They are coplanar with the spatial baseline \( S_1 S_2 \), the plane can be expressed by using three vectors \( R_1, R_2 \) and \( B \) mixed product, \( B \cdot (R_1 \times R_2) = 0 \). The formula can be expressed as the form of coordinates, as shown in equation (3):

\[
F = \begin{bmatrix} B_x & B_y & B_z \\ X_1 & Y_1 & Z_1 \\ X_2 & Y_2 & Z_2 \end{bmatrix} = 0
\]

(3)

In the formula:

\[
\begin{bmatrix} X_1 \\ Y_1 \\ Z_1 \end{bmatrix} = \begin{bmatrix} a_{1} & a_{2} & a_{3} \\ b_{1} & b_{2} & b_{3} \\ c_{1} & c_{2} & c_{3} \end{bmatrix} \begin{bmatrix} x_1 \\ y_1 \\ -f \end{bmatrix} = R \begin{bmatrix} x_1 \\ y_1 \\ -f \end{bmatrix}
\]

\[
\begin{bmatrix} x_2 \\ y_2 \\ z_2 \end{bmatrix} = \begin{bmatrix} a_{1} & a_{2} & a_{3} \\ b_{1} & b_{2} & b_{3} \\ c_{1} & c_{2} & c_{3} \end{bmatrix} \begin{bmatrix} x'_2 \\ y'_2 \\ -f' \end{bmatrix} = R \begin{bmatrix} x'_2 \\ y'_2 \\ -f' \end{bmatrix}
\]

The solution process is divided into two steps. First, using of eight or more than eight match points to make the relative orientation of the camera directly, and solved initial value about the relative position and posture, then an accurate position and attitude relative are solved by using continuous relative orientations around the camera.

Equation (3) is started as:

\[
L_1 y_1 x_2 + L_2 y_1 y_2 - L_3 y_1 f + L_4 f x_2 + L_5 f y_2 - L_6 f f + L_7 x_1 x_2 + L_8 x_1 y_2 - L_9 x_1 f = 0
\]

\[
L_i (i = 1, 2...9)
\]

is obtained by linear least squares solution, then though \( L_7 \) the relative position T
and attitude $R = \begin{bmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{bmatrix}$ of the camera are obtained, as (4) shown. The value of $Bx$ is assumed (1000 is generally), then through the camera's distance value $\sqrt{(B^2_x + B^2_y + B^2_z)}$, $Bx$ can be solved.

$B_x = \frac{- (L_x L_x + L_y L_y + L_z L_z)}{Bx}$

$B_y = \frac{(L_y L_x - L_x L_y - Bz L_y)}{(Bx^2 + By^2 + Bz^2)}$

$B_z = \frac{(L_z L_x - L_x L_z)}{(Bx^2 + By^2 + Bz^2)}$

Among:

$L_s^2 = 2Bx^2 / (L_0^2 + L_1^2 + L_3^2 + L_4^2 + L_5^2 + L_6^2 - L_7^2 - L_8^2 - L_9^2)$

$\phi = \frac{1}{L_i (i = 1, 2, \ldots, 9)}$

The coplanar equation (3), according to Taylor’s multi-function formula expands to one-time items:

$F = F_0 + \frac{\partial F}{\partial \phi} d\phi + \frac{\partial F}{\partial \omega} d\omega + \frac{\partial F}{\partial k} dk + \frac{\partial F}{\partial B_y} dB_y + \frac{\partial F}{\partial B_z} dB_z = 0$ (5)

The R and T of the initial values into (5), to do the iterative calculation, get a continuous relative orientation precise solution, which is about the precise position and orientation relative of the camera.

5. with the constraint condition of a regional network adjustment

After relative orientation is completed, we get the parameter which is observed by the binocular vision of the internal and external camera, so the main point of the camera parameters is (0,0) and the distortion K1, K2, P1, P2 are also 0. Actually, the main point of the camera is not strictly for (0, 0), but a small figure, and the distortion is existence objectively. Moreover, considering the relative position and carriage of the binocular vision camera are different, this paper uses the constrained condition of a regional network adjustment, introduces the camera parameter of non-linear distortion and makes an exact solution of the inside and outside parameters.

Expand the type (1) by using the multivariate function Taylor formula to get equation (6), where the initial of X, Y, Z can be obtained by two (or more) images intersection.

$\nu_i = \frac{\partial \nu_i}{\partial X} \Delta X_i + \frac{\partial \nu_i}{\partial Y} \Delta Y_i + \frac{\partial \nu_i}{\partial Z} \Delta Z_i + \frac{\partial \nu_i}{\partial \omega} \Delta \omega_i + \frac{\partial \nu_i}{\partial \phi} \Delta \phi_i + \frac{\partial \nu_i}{\partial k} \Delta k_i + \frac{\partial \nu_i}{\partial B_y} \Delta B_y_i + \frac{\partial \nu_i}{\partial B_z} \Delta B_z_i$.

$\nu_i = \frac{\partial \nu_i}{\partial X} \Delta X_i + \frac{\partial \nu_i}{\partial Y} \Delta Y_i + \frac{\partial \nu_i}{\partial Z} \Delta Z_i + \frac{\partial \nu_i}{\partial \omega} \Delta \omega_i + \frac{\partial \nu_i}{\partial \phi} \Delta \phi_i + \frac{\partial \nu_i}{\partial k} \Delta k_i + \frac{\partial \nu_i}{\partial B_y} \Delta B_y_i + \frac{\partial \nu_i}{\partial B_z} \Delta B_z_i$.

In which:

$l_i = (x - x_i - \Delta x) + f \frac{X}{Z}$

$l_i = (y - y_i - \Delta y) + f \frac{Y}{Z}$
all the internal and external accurate parameters of binocular vision are obtained by iterative calculation on the type (6).

6. Experiment and analysis

To compare the feasibility and accuracy of self-calibration method of the binocular vision, this paper compare with "three-dimensional DLT + bundle adjustment" method of the basic traditional template. Fig.3 and Fig.4 are two images taken by binocular visual, in which two different cameras’ focal length, significantly longer focal length of the right camera.

The two Canons 5DMark2 SLR cameras are prepared for the experiments, with a resolution of 4500 × 3000, the size of pixel are 7.7 microns. Using the "three-dimensional DLT + bundle adjustment" method calibration, the results are shown as follows, (Length in millimeters), the back-projection error is less than 1.5 pixels.

\[
R = \begin{bmatrix}
0.9964 & -0.0336 & 0.0771 \\
0.0299 & 0.9983 & 0.0489 \\
-0.0787 & -0.0464 & 0.9958
\end{bmatrix}, \quad T = \begin{bmatrix}
978.3958 \\
630.3764 \\
1703.8885
\end{bmatrix}
\]

Three rotation angles are solved though the R-matrix decomposition (unit: degrees):

\[
\begin{bmatrix}
-4.43 \\
-2.80 \\
1.71
\end{bmatrix}, \quad \text{where Y-axis as the main corner for the rotation system.}
\]
Table 1. Self-calibration intrinsic parameters the binocular vision (traditional method)

<table>
<thead>
<tr>
<th></th>
<th>f</th>
<th>X0</th>
<th>Y0</th>
<th>K1</th>
<th>K2</th>
<th>P1</th>
<th>P2</th>
</tr>
</thead>
<tbody>
<tr>
<td>The left</td>
<td>27.7419</td>
<td>0.2042</td>
<td>0.0021</td>
<td>-9.5759e-009</td>
<td>8.7049e-016</td>
<td>9.0445e-009</td>
<td>1.9033e-007</td>
</tr>
<tr>
<td>The right</td>
<td>38.3887</td>
<td>0.1446</td>
<td>0.0341</td>
<td>2.3327e-010</td>
<td>3.9565e-016</td>
<td>-1.6426e-008</td>
<td>2.6483e-007</td>
</tr>
</tbody>
</table>

Using this method to solve the fundamental matrix $F$:

$$F = \begin{bmatrix}
8.50e-007 & 1.35e-005 & 0.0271 \\
-1.39e-005 & 3.35e-008 & -0.0347 \\
-0.0212 & 0.0294 & 0.9983
\end{bmatrix}$$

According to the 147 basis-point in the left and right images in matrix matched, the basis matrix is solved, and establishes the bilinear equations of focal length squared by using basic matrix, get the left and right camera's focal length: $f_l = 28.0776$, $f_r = 38.0325$. By using the baseline length 1966.92 (measured) of the camera relative distance to sure the relative orientation of, the relative rotation angle of the camera are $[-4.07, -2.72, 1.57]$, the relative position $T = [978.3958, 632.8162, 1701.0201]$. Using the above parameters as the initial value, and consider the relative position of around the camera and the property with the same attitude, then conduct block adjustment with the constraints, get all internal and external orientation elements of binocular vision.

Table 2. Self-calibration parameters of binocular vision (the method of the paper)

<table>
<thead>
<tr>
<th></th>
<th>f</th>
<th>X0</th>
<th>Y0</th>
<th>K1</th>
<th>K2</th>
<th>P1</th>
<th>P2</th>
</tr>
</thead>
<tbody>
<tr>
<td>The left</td>
<td>27.5573</td>
<td>0.2135</td>
<td>0.0013</td>
<td>-9.3632e-009</td>
<td>8.6742e-016</td>
<td>9.1456e-009</td>
<td>1.8929e-007</td>
</tr>
<tr>
<td>The right</td>
<td>38.3034</td>
<td>0.1425</td>
<td>0.0321</td>
<td>2.5478e-010</td>
<td>3.8745e-016</td>
<td>-1.8245e-008</td>
<td>2.7635e-007</td>
</tr>
</tbody>
</table>

Relative rotation angle around the camera are $[-4.35, -2.72, 1.73]$. The Relative position is $T = [978.39, 631.43, 1703.02]$. Using the results of this self-calibration, the space point go back projection around the image, the errors are less than 5.2 pixels. In this paper, the match point with about the 1.5 pixels random error, Results from the self-calibration can be seen that its accuracy is better and the noise has certain immunity.

7. Conclusion

In this paper, a set of practical and high-precision self-calibration system is achieved. First, use a binocular vision system to obtain the image of a focal length, and then solve the fundamental matrix by the image matching points, using Krupp equation to establish the bilinear equations about the square of focal length to solve the left and right camera focal length; then take a relative orientation of the binocular vision system using the distance of the left and right camera focal length as baseline, to obtain the external parameters of two pairs. Finally, use the initial value and the relatively fixed position and orientation of the left and right camera focal length as a constraint, which are obtained from the two steps above, to make regional network adjustment. And introduce the no-liner distortion parameters of the camera in the adjustment to obtain the high-precision intrinsic parameters of each camera, variable parameters and the relatively position and orientation of the left and right camera focal length. The test analysis shows that the back-projection error of this method can be controlled in 5 pixels or so, and has a certain noise immunity.
8. Acknowledgment

This paper is sponsored by Nature Science Foundation of China (60973096) and Science Foundation of aviation in China (2010ZC56007).

9. References


