A Monocular Vision Relative Displacement Measurement Method Based on Bundle Adjustment Optimization and Quadratic Function Correction

Lei Ren, Xiucheng Dong*, Dongbo Liu and Fan Zhang

School of Electrical Engineering and Electronic Information, Xihua University, Chengdu, Sichuan 610039, China

*E-mail: Levi5182009@163.com; dxc136@163.com

Abstract. In order to meet the requirements of high-precision target displacement measurement, this paper proposed a measurement method based on monocular vision measurement system by which the advantages of simple structure, convenient debugging, easy to install and calibrate is recognized. This method first uses the bundle adjustment algorithm with conditional constraints to optimize the solution results of the orthogonal iterative algorithm, then calculates the relative poses of the optimization results, finally establishes an error correction model to perform secondary optimization on the relative poses. The experimental results show that the method proposed in this paper can reduce more than 54% average error of displacement measurement results, and the error between the final displacement measurement results and the actual displacement values is controlled at 0.4mm level, which proves that the method has good practical value.

1. Introduction

In recent years, accurately measuring the pose parameters of the target is significance in many research fields [1]. With the advantages of non-contact and high accuracy, visual measurement has been widely used. By extracting the pixel coordinates and corresponding actual coordinates of the cooperative target on the measured target, the pose parameters between the target coordinate system and the camera coordinate system can be solved.

How to measure the pose parameters more accurately has been the focus of scholars. The classic direct linear transformation (DLT) algorithm proposed by Abdel-Aziz and Karara [2] regards the nine variables of the projection matrix as independent variables that establish linear constraints. However, because the orthogonal constraints of the rotation matrix are ignored, the DLT algorithm requires a large number of control points to obtain an accurate solution. Přibyl [3] proposed the DLT-Combined-Lines algorithm based on the DLT algorithm, which uses Perspective-n-Line (PnL) linear formula to solve a large number of lines, but it is not accurate enough to use in engineering. Lepetit [4] proposed a high-efficiency and high-precision linear algorithm called EPNP, which uses 4 virtual control points to represent points in object coordinates. However, the algorithm may get bad results when the control points are coplanar and the depth changing dramatically. Chen [5] proposed a measurement method using three marker points, but the three-point method does not have enough points to optimize, resulting in large error and susceptibility to interference. In order to avoid the influence of interference, Oberkampf [6], [7], [8] proposed a POSIT algorithm, which iteratively estimates the perspective projection model by using the initial value obtained from the scaled model.
Lu [9] proposed an orthogonal iterative algorithm (OI). Compared with other iterative methods, the OI algorithm has faster convergence speed and higher accuracy [10]. The solution results are still good when a good initial value cannot be provided [11]. The completeness of the iterative algorithm makes it widely used, but using only the orthogonal iterative algorithm still can not meet the requirements of high-precision measurement.

So in order to meet the requirements of high-precision measurement, this paper proposed a method of monocular vision measurement based on the OI algorithm, the measurement results are optimized by bundle adjustment algorithm with constraints. After calculating the relative poses of the optimization results, establishing an error correction model based on the quadratic function for further optimization.

2. The Monocular Vision Pose Measurement Method

As shown in figure 1, pasting the cooperation target flat on the surface of the object and collecting images with a calibrated single camera. After image processing, contour extraction, center positioning, the center coordinates of the cooperative target can be found.

![Figure 1. The schematic of monocular vision measurement.](image)

Placing the camera and adjusting the field of view, using Zhang's [12] calibration method to get the camera’s internal parameters. As shown in figure 2, an object coordinate system is established on the cooperative target. The actual coordinates of each point can be calculated from the actual distance between the points.

Establishing the projection matrix (1) between the object coordinates and the corresponding pixel coordinates. The model (2) for minimizing collinear errors of objects in object space based on the collinear projection relationship between object space points and corresponding image points in object space.

\[
V_i = \frac{PP^T}{P^T}P_i
\]

\[
E(R,T) = \min_{R,T} \sum_{i=1}^{n}\|e_i\|^2 = \sum_{i=1}^{n}\|(I - V_i)(RP_i + T)\|^2
\]

Formula (2) can be regarded as a quadratic function about the translation matrix $T$. When the optimal rotation matrix $R$ can be given, the optimal solution of $T$ with $R$ can be obtained by using partial derivatives.

\[
T(R) = \frac{1}{n} \left( I - \frac{1}{n} \sum_{i=1}^{n} V_i \right) \left( \sum_{i=1}^{n} (V_i - I)RP_i \right)^{-1}
\]

After the new rotation matrix $R$ and the translation matrix $T$ are obtained, the new projection point...
\( q' \) of point \( P_i \) on the line of sight \( V_i \) can be expressed by combining \( R \) and \( T \), \( q'_i \) can be expressed as formula (4).

\[
q'_i(R) = V_i \left( R P_i + T \left( R \right) \right), \quad R^{n+1} = \min_{R} \sum_{i=1}^{n} \left\| \left( R P_i + T \left( R \right) \right) - q'_i \left( R^n \right) \right\|^2
\]  

(4)

It can be seen from the above formulas that the problem becomes the absolute orientation problem. The SVD singular value decomposition method can be used to obtain the new rotation matrix \( R \). Until the error is less than the set value or the number of iteration reaches the set value, the obtained \( R^{n+1} \) and \( T^{n+1} \) are considered as the optimal rotation and translation matrix.

3. The Bundle Adjustment Algorithm with Conditional Constraints

In order to obtain more accurate pose parameters, this paper uses the bundle adjustment algorithm with constraints for the first optimization.

The external parameters of the camera and the object coordinates of the points are both regarded as observation values. And establishing a formula with the constraints such as the distance between the 4 control points and the object coordinates of all points should be in the same plane. The constraint formula and the bundle adjustment formula are combined to obtain formula (5).

\[
\begin{align*}
V &= AR + BW - L, \\
V' &= BW - L,
\end{align*}
\]  

(5)

3.1. The establishment of bundle adjustment error formula

According to the imaging rule of the camera, the calculation formula (6) of the pixel coordinates involves the distortion correction formula and the collinear formula. Combining these formulas, using Taylor expansion and take the first order term.

\[
\begin{align*}
\begin{cases}
\theta = f x + f y + u_0, \\
y = f x + v_0
\end{cases}, \\
\begin{pmatrix}
x = x' + x' \left( k_1 r^2 + k_2 r^4 + k_3 r^6 \right) + 2p_1 x y' + p_2 \left( r' + 2x' \right) \\
y = y' + y' \left( k_1 r^2 + k_2 r^4 + k_3 r^6 \right), + p_1 \left( r'^2 + 2y'^2 \right) + 2p_2 x y'
\end{pmatrix}
\end{align*}
\]

\[
\begin{align*}
x' &= a_1 X + a_2 Y + a_3 Z + T, \\
y' &= b_1 X + b_2 Y + b_3 Z + T
\end{align*}
\]  

(6)

The obtained formula is shown in formula (7). Formula (8) can be obtained from observation value + observation correction value = approximate value + approximate correction value.

\[
\begin{align*}
\begin{cases}
\theta = u + \frac{\partial u}{\partial \theta_1} d\theta_1 + \frac{\partial u}{\partial \theta_2} d\theta_2 + \frac{\partial u}{\partial \theta_3} d\theta_3 + \frac{\partial u}{\partial T_1} dT_1 + \frac{\partial u}{\partial T_2} dT_2 + \frac{\partial u}{\partial T_3} dT_3, \\
y = v + \frac{\partial v}{\partial \theta_1} d\theta_1 + \frac{\partial v}{\partial \theta_2} d\theta_2 + \frac{\partial v}{\partial \theta_3} d\theta_3 + \frac{\partial v}{\partial T_1} dT_1 + \frac{\partial v}{\partial T_2} dT_2 + \frac{\partial v}{\partial T_3} dT_3 + \frac{\partial v}{\partial X_u} dX_u + \frac{\partial v}{\partial Y_u} dY_u + \frac{\partial v}{\partial Z_u} dZ_u.
\end{cases}
\end{align*}
\]  

(7)
\[
\begin{align*}
&u + V_u = (u) + du, \\
v + V_v = (v) + dv.
\end{align*}
\]

Combining formulas (5) (6) (7) (8), the matrix form of error formula is obtained:

\[
V_{ij} = \begin{bmatrix} V_u^i \\ V_v^i \end{bmatrix}, \quad L_{ij} = \begin{bmatrix} u_{ij} - (u)_j \\ v_{ij} - (v)_j \end{bmatrix}
\]

\[
R_j = \left[ \partial \theta^i_x \partial \theta^i_y \partial \theta^i_z \partial T^i_x \partial T^i_y \right]^T
\]

\[
W_i = \left[ \partial X^i \partial Y^i \partial Z^i \right]^T
\]

\[
A_j = \begin{bmatrix} \frac{\partial u_{ij}}{\partial \theta^i_x} d\theta^i_x & \frac{\partial u_{ij}}{\partial \theta^i_y} d\theta^i_y \\ \frac{\partial v_{ij}}{\partial \theta^i_x} d\theta^i_x & \frac{\partial v_{ij}}{\partial \theta^i_y} d\theta^i_y \\ \frac{\partial u_{ij}}{\partial T^i_x} dT^i_x & \frac{\partial u_{ij}}{\partial T^i_y} dT^i_y \\ \frac{\partial v_{ij}}{\partial T^i_x} dT^i_x & \frac{\partial v_{ij}}{\partial T^i_y} dT^i_y \end{bmatrix}, \quad B_j = \begin{bmatrix} \frac{\partial u_{ij}}{\partial X^w} dX^w & \frac{\partial u_{ij}}{\partial Y^w} dY^w & \frac{\partial u_{ij}}{\partial Z^w} dZ^w \\ \frac{\partial v_{ij}}{\partial X^w} dX^w & \frac{\partial v_{ij}}{\partial Y^w} dY^w & \frac{\partial v_{ij}}{\partial Z^w} dZ^w \end{bmatrix}
\]

In the above formulas, \( i = 1, 2, \ldots, n \) refers to the number of object points, and \( j = 1, 2, \ldots, m \) refers to the number of photos.

### 3.2. The Establishment of Constraint Formula

Calculating the distance between two points by formula (13).

\[
S^2_j = \left( X^j_w - X^i_w \right)^2 + \left( Y^j_w - Y^i_w \right)^2 + \left( Z^j_w - Z^i_w \right)^2
\]

And all points should be in the same \( Z \) plane, so the plane constraint formula (14) is established.

\[
Z^j_w - Z^i_w = 0
\]

In formulas (13) (14), \( i \) and \( j \) refer to the serial numbers of two points, and \( S \) is the actual distance between the two points. Combining (5) (13) and (14), the matrix form of the constraint formulas (15) is established:

\[
V_r = \begin{bmatrix} X^j_w - X^i_w \\ Y^j_w - Y^i_w \\ Z^j_w - Z^i_w \end{bmatrix}, \quad V_s = \begin{bmatrix} \Delta X_i \\ \Delta Y_i \\ \Delta Z_i \end{bmatrix}, \quad V_r = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}^T - \frac{s_i^2 - s_j^2}{2}, \quad V_s = \begin{bmatrix} \Delta X_j \\ \Delta Y_j \\ \Delta Z_j \end{bmatrix}
\]

\[
- \Delta X_j \Delta Y_j \Delta Z_j
\]
In these formulas, $S_K$ represents the true distance, and $S_M$ represents the calculated distance. The compensation values of external parameters and object coordinates can be obtained.

4. Correcting Systematic Error of Displacement Measurements

Formula (16) represents the transformation relationship between the camera coordinate system, the starting object coordinate system $P_s$, and the target position object coordinate system $P_E$.

$$Q_S = R_s P_s + T_1$$
$$Q_E = R_s P_E + T_2$$  \hspace{1cm} (16)

Both $Q_S$ and $Q_E$ belong to the same camera coordinate system during the measurement process, so the camera coordinate system can be used as a reference to establish the relationship between the target position object coordinate system and the starting position object coordinate system.

$$P_{E(S)} = R_{i^1} \left( R_{s} P_E + T_2 - T_1 \right)$$ \hspace{1cm} (17)

Because of the limitations of a single camera, there will be a systematic error in measuring relative displacement. The error is caused by the magnification of the uncorrected distortion due to the large measurement field of view, and the insensitivity of the single camera to depth direction information.

As shown in figure 3, when the target is moving in non-depth of field, the errors in the X and Y directions are linear, while the errors in the Z direction is quadratic. In order to facilitate the establishment of the modified model, a quadratic function model is used.

As shown in figure 4, selecting some positions which the relative distances is known, as reference positions. Establishing an optimization model based on the pose parameters, the calculated relative displacements and the real values of these reference positions. Then inputting the pose parameters of other positions in the model, the optimized relative displacement can be obtained.

![Figure 3](image)

**Figure 3.** The changing trend of systematic error.

![Figure 4](image)

**Figure 4.** Setting up error-correction Model.

From formula (17), the indirect displacement between the two coordinate systems can be represented by $T_s = R_{i^1} \left( T_2 - T_1 \right)$. First, calculating the relative displacement $T_C$ of the reference position and the starting position, and then calculating the error between the true value $T_K$ and the calculated value $T_C$.

$$\Delta X = X_R - X_C$$
$$\Delta Y = Y_R - Y_C$$
$$\Delta Z = Z_R - Z_C$$  \hspace{1cm} (18)

According to formula (17), the relative displacement is only related to the external parameters, so a modified model is established.
\[ \begin{align*}
\Delta X &= a_1R + a_2T + a_3R^2 + a_4T^2 + a_5 \\
\Delta Y &= b_1R + b_2T + b_3R^2 + b_4T^2 + b_5 \\
\Delta Z &= c_1R + c_2T + c_3R^2 + c_4T^2 + c_5
\end{align*} \]

The calculated value is modified by the model to make it closer to the true relative displacement.

5. Experiment and Analysis

In this paper, the experiment used Nikon D7200 SLR camera photos, the lens model of the camera is AF-S DX NIKKOR 18-140mm f/3.5-5.6 ED VR, and the focal length \( f \) is fixed at 35mm when taking photos. After placing and calibrating the camera, orthogonal iteration algorithm (OI), orthogonal iteration + bundle adjustment algorithm with constraints (OI+BA), orthogonal iteration + bundle adjustment algorithm with constraints + systematic error correction algorithm (OI+BA+EC) are used to solve, and analyzing the results. Three experiments were conducted, in each experiment, the moving distance along the X-axis of the model was 1000mm.

As shown in figure 5, the reference positions are taken every 200mm to establish the model. The model is placed on a high-precision captive ballistic system (CTS) to ensure the accuracy of each movement.

Figure 5. Part of the experimental pictures.

| Method | direction | 100 | 300 | 500 | 700 | 900 | 1000 |
|--------|-----------|-----|-----|-----|-----|-----|------|
| OI     | X         | 0.2195 | -0.317 | -0.2432 | -0.8777 | -1.1915 | -1.0356 |
|        | Y         | 0.2917 | 0.9613 | 1.1842 | 1.2742 | 1.3631 | 1.3538 |
|        | Z         | -1.7784 | -2.9485 | -1.6463 | 2.0949 | 3.0667 | 4.4749 |
| OI+BA  | X         | 0.3059 | -0.1652 | -0.1149 | -0.7655 | -1.0347 | -0.8537 |
|        | Y         | 0.2145 | 0.8475 | 1.0831 | 1.1329 | 1.1869 | 1.1716 |
|        | Z         | -0.6625 | -1.8383 | -1.7942 | -0.7685 | -1.0851 | -0.4952 |
| OI+BA+EC | X      | 0.3086 | -0.2461 | -0.3824 | 0.2103 | 0.0727 | 0.067 |
|        | Y         | 0.0775 | 0.072 | -0.1014 | 0.1105 | -0.0759 | -0.088 |
|        | Z         | -0.2038 | 0.0192 | -0.2041 | 0.1846 | 0.1326 | 0.1906 |

As shown in figure 6 and table 1, compared with the original algorithm, after using the bundle adjustment optimization, the standard deviation of the X direction error is reduced from 0.349 to 0.324, while the standard deviation of the Y direction error is reduced from 0.312 to 0.287, and the standard deviation of the Z direction error is reduced from 0.852 to 0.507. After adding the error correction method, the standard deviation of the X direction error is further reduced to 0.162, the standard deviation of the Y direction error is further reduced to 0.0398, and the standard deviation of the Z direction error is further reduced to 0.0705.
Compared with the errors of the initial result, the average error of the final results is reduced by more than 54%, and the standard deviation of the errors is reduced by more than 53%. The maximum value of the final errors are controlled at 0.4mm level.

6. Conclusions
This paper proposed a method to measure the relative displacement of objects, which adopts the method of monocular vision measurement based on the OI algorithm, and the measurement results are optimized by bundle adjustment algorithm with constraints. Calculating the relative poses of the optimization results and selecting some positions in the measurement space as reference positions. Analyzing the error between the calculated pose and the actual value and establishing an error correction model based on the quadratic function to improve measurement accuracy. It can be seen from the experimental data that under the 1000mm field of view, comparing with the relative displacements obtained from original method, the average error of the final result obtained from the method proposed by this paper is reduced by more than 54%, and the standard deviation of the error is reduced by more than 53%. The error between the final calculation result and the actual value is controlled at 0.4mm level, proving the effectiveness of the method. The accuracy can be further improved by improving the accuracy of image coordinate extraction.

Acknowledgements
This paper is supported by National Natural Science Foundation of China (11872069), Key Project of Sichuan Science and Technology Department (2018JY0463), National Ministry of Education “Chunhui Plan” Scientific Research Project (Z2017076), Sichuan University Research Innovation Team – Machine Vision and Intelligent Control (18TD0024), Joint Laboratory of Industrial, Academic and Research, Siwei High Technology- Xihua University (2016-YF04-00044-JH).

References
[1] Wen G, Ming Z, Bogen Z, and Xiao Tian W, 2013 Overview of target tracking technology Chinese Optics 7(3) 365-75 (In Chinese)
[2] Abdel-Aziz Y, Karara HM, and Hauck M 2015 Direct linear transformation from comparator coordinates into object space coordinates in close-range photogrammetry Photogrammetric Engineering & Remote Sensing 81(2) 103-7
[3] Přibyl B, Zemčík P, and Čadík M 2017 Absolute pose estimation from line correspondences using direct linear transformation Computer Vision and Image Understanding 161 130-44
[4] Lepetit V, Moreno-Noguer F, and Fua P 2009 Epnp: An accurate o (n) solution to the pnp problem International Journal of Computer Vision 81(2) 155
[5] Jiechun C, Min G, Zhengliang D and Feng Y 2010 Stereo vision based attitude measurement of aircraft model in wind tunnel Journal of Applied Optics 31(3) 437-41 (In Chinese)
[6] Oberkampf D, DeMenthon D F and Davis L S 1996 Iterative pose estimation using coplanar feature points Computer Vision and Image Understanding 63(3) 495-511
[7] Dementhon D F, and Davis L S 1995 Model-based object pose in 25 lines of code International Journal of Computer Vision 15(1-2) 123-41
[8] Gramenega T, Venturino L, Cicirelli G, Attolico G and Distanza A 2004 Optimization of the POSIT algorithm for indoor autonomous navigation Robotics and Autonomous Systems 48(2-3) 145-62
[9] C.-P. L, Hager G D and Mjolsness E 2000 Fast and globally convergent pose estimation from video images IEEE Transactions on Pattern Analysis And Machine Intelligence 22(6) 610-22
[10] Qi W and Zhangyan Z 2019 An Accurate and Stable Pose Estimation Method Based on Geometry for Port Hoisting Machinery IEEE Access 7 39117-28
[11] Run Z, Zhengyu Z and Xuhui H 2018 Weighted Orthogonal Iteration Algorithm for Camera Pose Estimation Acta Optica Sinica 38(5) 0515002 (In Chinese)
[12] Zhengyu Z, 2000, A flexible new technique for camera calibration IEEE Transactions on Pattern Analysis and Machine Intelligence 22(11) 1330-4
