High accuracy linear method for axis measurement

Lijun Zhong
Zhang Li
Xiaohu Zhang
Yang Shang
Qifeng Yu

High accuracy linear method for axis measurement

Lijun Zhong,[a,b,]* **Zhang Li,[a,b]** **Xiaohu Zhang,[c]** **Yang Shang,[a,b]** **and Qifeng Yu[a,b]**

[a]National University of Defense Technology, College of Aerospace Science and Engineering, Changsha, China

[b]Hunan Provincial Key Laboratory of Image Measurement and Vision Navigation, Changsha, China

[c]Sun Yat-Sen University, School of Aeronautics and Astronautics, Guangzhou, China

Abstract. The attitude parameter is an important state parameter for a long axisymmetric target. The plane intersection (PI) method is a commonly used method for attitude estimation. However, this method only uses planes' information in the object space under a multiple camera system (more than two cameras simultaneously observing a target). We propose two methods to address the aforementioned issue. One method involves minimizing the square of the object-space angle residual (OAR) and the other method involves minimizing the square of the image-space angle residual (IAR). The linear optimization method is used for the above minimizing problems. The simulation results demonstrate that the IAR method has higher accuracy than the PI and OAR methods under multiple and dual camera systems because it incorporates information of a pair of corresponding image points. Furthermore, our experiments have shown that the linear method is generally faster, and it has an equivalent accuracy compared to the iterative method. © *The Authors. Published by SPIE under a Creative Commons Attribution 4.0 Unported License. Distribution or reproduction of this work in whole or in part requires full attribution of the original publication, including its DOI.* [DOI: 10.1117/1.OE.59.2.024101]

1 Introduction

The axis attitude is an important parameter for the motion state analysis, performance measurement, and aerodynamic parameter identification of rockets or missiles. Optical measurement is the main measurement method of a shooting range due to its noncontact, passive, high frequencies and high accuracy characteristics. Generally, a shooting range uses theodolites to record the target launching and running process, followed by image processing to obtain the target yaw angle and the pitch angle.

The pose measurement of the target mainly has the following types of methods: The first method is to use plane intersection (PI) to obtain the target's central axis attitude,[1] This method assumes that the target is a symmetrical object, and two cameras are used to simultaneously record the target. The central axis of the target can then be extracted from the image. Based on the principle of light propagation along a straight line, the target axis must be located on a plane consisting of the optical center and the target central axis in the image. The intersection line of the planes of the two cameras can be calculated to obtain the target axis representation. This method gives more accurate measurements than the method based on endpoints, and the form is intuitive. This method also shows that the longer the target axis in an image, the higher the pose measurement accuracy. Many researchers have extended this method to other applications.[2–4] The accuracy of target axis extraction in an image plays an important role in improving the accuracy of the PI method.[5] This kind of method can obtain only the target's central axis attitude, not the target's position or the roll angle. The second method is to use the information of the target's structure and then match its projection on the image with the contour extracted from the image.[6–16] Zhao et al.[7] proposed a pose estimation algorithm based on the inner angle and the triangle constraint. Liu and Hu[8] gave a pose

---

*Address all correspondence to Lijun Zhong, E-mail: trand1986@hotmail.com

estimation method based on a monocular image for a rotating solid shape spacecraft that utilized the inherent characteristics of a spacecraft. However, this method was not applicable for other types of targets. Peng et al.[9,10] proposed some methods for non-cooperative space targets that were based on only stereo cameras or fusion with laser radar. Becke and Schlegl[11] proposed a method for estimating a target axis based on the contours under monocular or multi-image conditions. Hanek et al.[15] used the linear extreme contours of points, lines, and cylinders to estimate the position and orientation of a camera in the world coordinate system. The third method is to use some cooperation signs (points or lines) on the target. First, we achieve the cooperation signs' coordinate in the target's coordinate system, and then, we solve the pose from the target to the camera's coordinate system.[17–26] Horn et al.[20] utilized the point cloud of the side surface, which achieved by a laser profile sensor to measure the pose of cylindrical components, but this method cannot work at high frequencies. Some researchers estimated camera's pose by the correspondence of lines or epipolar constraint.[21–25] Zhou et al.[26] proposed a method to track and estimate the pose of known rigid objects efficiently in a complex environment, which based on a 3D particle filter with M-estimation optimization. The fourth method is to directly achieve the point cloud by a laser profile sensor and then use the point cloud to estimate the target's attitude.[23]

For long cylindrical targets, the second method cannot achieve a high precision while the cameras are at a long distance of more than 100 m. For the third method, it is difficult to obtain the coordinates of the cooperation signs in the target coordinate system with high precision. There are two main methods for this. One method is based on model labeling, and the other method involves measurement with a total station after the mark is attached. With the first method, it is difficult to guarantee the accuracy of labeling when the target is large. In the total station measurement process, since it is necessary to move multiple positions to measure all of the feature points, it is easy to introduce conversion errors and affect the measurement accuracy. The fourth method cannot work at high frequencies. Later, in this paper, we will describe the accuracy advantage of the axis method relative to the feature points-based method in the measurement of the axis attitude for long symmetric targets. So, in this paper, we propose a method to improve the PI's accuracy.

For a multiple camera system (more than two cameras simultaneously observing a target), the planes given by the cameras from the optical center and the target axis in the image cannot intersect at the same straight line. This is caused by the inevitable errors of the camera parameter and the central axis. The solution for the PI method is to derive the direction vectors of the target using all of the plane equations. Then, the target attitude angle can be derived according to the definition of decomposition. However, the PI method lacks a geometric explanation when there are more than two cameras. In this paper, we propose two optimization methods to address this issue. The first method is based on minimizing the object-space angle residual (OAR), and the other method is based on minimizing the image-space angle residual (IAR). We also provide linear solutions and iterative solutions for the OAR and the IAR, called the OARL, IARL, OARI, and IARI, respectively. The simulation results showed the IAR obtained a higher accuracy than the PI and the OAR. Furthermore, the IARL was generally faster than the IARI.

## 2 Plane Intersection Method

2.1 Target Attitude Representation

For an axisymmetric long strip target, the pose of the target is usually determined by its symmetry axis direction, which is generally represented by the direction vector or the angles. Alternatively, the yaw and pitch angles could be used to represent the direction vector, as shown in Fig. 1. For the shooting range, X faces north, Y faces upward, and Z faces east. OA is the target axis, OA' is the projection of OA on the XOZ plane, the angle between the X axis and OA'(ω) is the yaw angle, and the angle between OA and OA'(φ) is the pitch angle.

2.2 Coordinate System Definition

The world coordinate system W is a right-handed coordinate system, as shown in Fig. 2. The origin of the camera coordinate system is the optical center. The Z axis is pointing in the positive
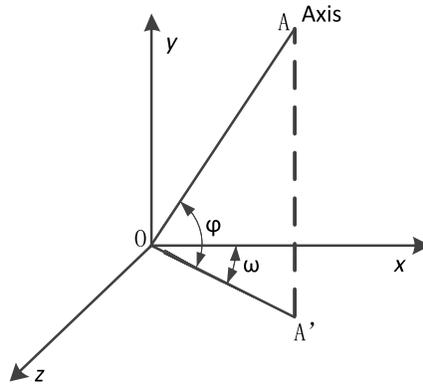
Fig. 1 Axis direction and attitude angle definition. For the shooting range, X faces north, Y faces upward, and Z faces east. OA is the target axis, OA' is the projection of OA on the XOZ plane, the angle between the X-axis and OA'(ω) is the yaw angle, and the angle between OA and OA'(φ) is the pitch angle.
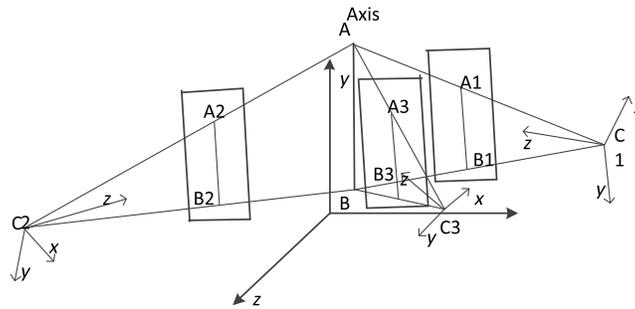
Fig. 2 Schematic diagram of PI. The three cameras C1, C2, and C3 observe the same target AB from different directions, and the projections on the image are A1B1, A2B2, and A3B3, respectively. Theoretically, the three planes C1A1B1, C2A2B2, and C3A3B3 intersect in a straight line.

direction of the optical axis, the X axis is perpendicular to the Z axis horizontally to the right, and the Y axis is perpendicular to the X axis and to the Z axis, whose direction is determined by the right-handed definition criterion. The origin of the image coordinate system is the upper left corner.

2.3 Plane Representation by a Single Camera

As shown in Fig. 2, AB is the target axis, C1 and C2 are the optical centers of the two cameras, A1B1 is the projection of the target axis on camera C1, and A2B2 is the projection of the target axis on camera C2. Assume that the equation of the target central axis in the i'th image coordinate system is

$$\mathbf{A}_i^T \mathbf{X} = 0, \tag{1}$$

where $\mathbf{A}_i = [\, a_i \quad b_i \quad c_i \,]^T$ and $\mathbf{X} = [\, x \quad y \quad 1 \,]^T$. The image plane can be seen as the $z = f$ plane in the camera coordinate system, where f is the physical focal length of the camera. The Z axis of the camera coordinate system passes through the principal point of the image plane (C_x, C_y), so the transformation from the image coordinate system to the camera coordinate system is determined as follows:

$$\begin{cases} x_c = (x - C_x)d_x \\ y_c = (y - C_y)d_y \end{cases}, \tag{2}$$

where (x_c, y_c) is in the camera coordinate system, (x, y) is in the image coordinate system, and (d_x, d_y) is the actual physical size of the camera pixel.

The central axis of the target in the image can be expressed in the camera coordinate system as follows:

$$\begin{cases} a_i(x_c/d_x + C_x) + b_i(y_c/d_y + C_y) + c = 0 \\ z - f = 0 \end{cases}.$$ (3)

The plane equation (composed of the axis and the optical center of the camera) can be defined as follows:

$$a_i x_c/d_x + b_i y_c/d_y + (aC_x + bC_y + c)z_c/f = 0.$$ (4)

The equivalent focal length of the camera is abbreviated as follows:

$$\begin{cases} F_x = f/d_x, \\ F_y = f/d_y. \end{cases}$$ (5)

The plane equation can be summarized as follows:

$$a_i F_x x_c + b_i F_y y_c + (a_i C_x + b_i C_y + c_i)z_c = 0.$$ (6)

Given the parameter matrix $\mathbf{n}_i = \begin{bmatrix} n_1 & n_2 & n_3 \end{bmatrix}^T = \begin{bmatrix} a_i F_x & b_i F_y & a_i C_x + b_i C_y + c_i \end{bmatrix}^T$, the plane equation is then rewritten as follows:

$$\mathbf{n}_i^T \mathbf{P}_c = 0,$$ (7)

where $\mathbf{P}_c = \begin{bmatrix} x_c & y_c & z_c \end{bmatrix}^T$ is a point from the camera coordinate system and n is the normal vector of the plane. It is assumed that the external parameters of the i'th camera in the world coordinate system are R_i, t_i, where R_i and t_i are the rotation matrix and the translation vector from the world coordinate system to the camera coordinate system. Therefore, the point in the world coordinate system can be transformed to the camera coordinate system using Eq. (8):

$$\mathbf{P}_c = \mathbf{R}_i \mathbf{P}_w + \mathbf{t}_i,$$ (8)

where $\mathbf{P}_w = \begin{bmatrix} x_w & y_w & z_w \end{bmatrix}^T$ is a point in the world coordinate system. The plane equation in the world coordinate system can then be derived as follows:

$$\mathbf{n}_i^T \mathbf{R}_i \mathbf{P}_w + \mathbf{n}_i^T \mathbf{t}_i = 0.$$ (9)

So far, the expression of the plane (composed of optical center of the camera and the central axis of the target in the image) has been obtained in the world coordinate system.

2.4 *Principle of Dual Camera Plane Intersection*

Supposing that the plane equations of the two cameras are

$$\begin{cases} \mathbf{n_1^T P_w} + d_1 = 0 \\ \mathbf{n_2^T P_w} + d_2 = 0 \end{cases},$$ (10)

where n_1 and n_2 are the normal vectors of the two planes. We do not consider the atmospheric refraction in this paper, so the direction vector of the intersection line $\mathbf{I} = (l, m, n)^T$ must be perpendicular to both normal vectors, it can be solved as

$$\mathbf{I} = \mathbf{n_1^T} \times \mathbf{n_2^T} = \begin{bmatrix} i & j & k \\ n_{11} & n_{12} & n_{13} \\ n_{21} & n_{22} & n_{23} \end{bmatrix}. \tag{11}$$

The yaw angle ω and pitch angle φ of the target can be determined as follows:

$$\begin{cases} \omega = \tan^{-1} n/l \\ \varphi = \tan^{-1} m/\sqrt[2]{l^2 + n^2} \end{cases}. \tag{12}$$

It can be seen that the representation of the direction vector obtained by the final intersection of this method is independent of d and only d for the plane parameter that is related to the translation vector of the camera. This means that the final direction vector is independent of the translation vector T. It can be concluded that the target position is related only to the attitude of the camera.

3 Our Method

3.1 Multicamera Intersection Scenario

As shown in Fig. 2, the three cameras C1, C2, and C3 observe the same target AB from different directions, and the projections on the image are A1B1, A2B2, and A3B3, respectively. Theoretically, the three planes C1A1B1, C2A2B2, and C3A3B3 should intersect in a straight line. However, due to the errors of the camera parameters and the axis, the direction vectors given by each set of two cameras are not completely consistent. Therefore, it is important to derive the optimal direction vector.

3.2 Optimization Conditions

This paper proposes two optimization methods for the axis measurement problem. These two methods are based on minimizing residuals:

1. In a similar manner to the object-space residual method for camera calibration, the target axis extracted from the image was used as the true value. Firstly, we calculated the angle between the target central axis and each plane that consisted of the optical center and the central axis. Secondly, we minimized the square sum of all of the angles. In this paper, this angle is called the OAR. As shown in Fig. 3, the C2A2B2 plane is the plane composed of the optical center and the target axis in the image; A′B′ is the target axis, and B′D is the projection of A′B′ on the plane C2A2B2. α is the OAR.

2. In a similar manner to the image-space residual method for camera calibration, the target axis was projected to the image plane according to the theoretical model, and the angle between the projection and the axis extracted on the image plane was considered the angle
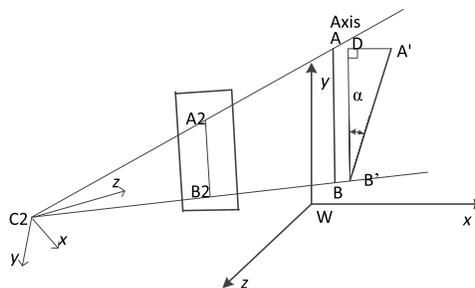
Fig. 3 Schematic diagram of the OAR. The C2A2B2 plane is the plane composed of the optical center and the target axis in the image; AB is the target axis; A′B′ is the measured result of the target axis, and B′D is the projection of A′B′ on the plane C2A2B2. α is the OAR.
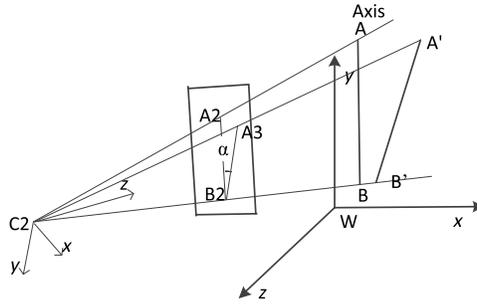
Fig. 4 Schematic diagram of the IAR. The C2A2B2 plane is the plane composed of the optical center and the axis of the camera; AB is the target axis; A′B′ is the measured result of the target axis, B2A3 is the projection of A′B′ on the image plane, and α is the IAR.

residual. We also minimized the square sum of all of the angles. In this paper, this term is called the IAR. As shown in Fig. 4, the C2A2B2 plane is the plane composed of the optical center and the axis of the camera; A′B′ is the target axis, B2A3 is the projection of A′B′ on the image plane, and α is the image angle residual.

3.3 Object-Space Angle Residual

3.3.1 Definition of the optimization problem

Supposing that multiple cameras simultaneously observe a target, one can extract the target axis in the image. The plane equations of the camera's optical center and the target axis in the image are expressed in the world coordinate system as follows:

$$\mathbf{n}_i^T \mathbf{P}_w + d_i = 0. \tag{13}$$

Supposing that the target axis direction is $\mathbf{I} = (l, m, n)^T$, then the angle between the central axis and the plane is

$$\alpha_i = \sin^{-1}\left(\frac{\mathbf{n}_i^T \mathbf{I}}{\|\mathbf{n}_i\| \|\mathbf{I}\|}\right). \tag{14}$$

For the convenience of calculation, the linear direction and the plane parameters are normalized, i.e.,

$$\begin{cases} \|\mathbf{n}_i\| = 1 \\ \|\mathbf{I}\| = 1 \end{cases}. \tag{15}$$

Then, α_i can be expressed as follows:

$$\alpha_i = \sin^{-1}(\mathbf{n}_i^T \mathbf{I}). \tag{16}$$

From the geometric understanding of the PI, one can use this angle as the residual of the obtained central axis relative to each camera plane. In a similar manner to the spatial residual for the camera calibration, minimizing this residual can provide the optimal solution:

$$\min \sum_{i=1}^{N} \alpha_i^2, \tag{17}$$

where N is the number of cameras. For a small angle, α_i; could be approximated by $\mathbf{n}_i^T I$, therefore, Eq. (17) could be rewritten as follows:

$$\min \sum_{i=1}^{N} (\mathbf{n}_i^T \mathbf{I})^2, \tag{18}$$

where $(\mathbf{n}_i^T \mathbf{I})$ is within the interval of $[-1, 1]$.

3.3.2 *Linear method*

The Levenberg–Marquardt[22] can be used to solve Eq. (17) and keep low computation efficiency. The following method provides a linear solution to the minimum value by deriving

$$y = \sum_{i=1}^{N} (\mathbf{n}_i^T \mathbf{I})^2. \tag{19}$$

Since y is a function of $(l, m, n)^T$, letting the partial derivative be 0:

$$\begin{cases} y_l' = \sum_{i=1}^{N} (\mathbf{n}_{i1} \mathbf{n}_i^T \mathbf{I}) = 0 \\ y_m' = \sum_{i=1}^{N} (\mathbf{n}_{i2} \mathbf{n}_i^T \mathbf{I}) = 0 \\ y_n' = \sum_{i=1}^{N} (\mathbf{n}_{i3} \mathbf{n}_i^T \mathbf{I}) = 0 \end{cases}. \tag{20}$$

We can solve the equations by the singular value decomposition (SVD)[27] of the coefficient matrix.

Theorem 1: It is assumed that n_i are the coefficients of the spatial plane $n_i^T P_w + d_i = 0$. The necessary and sufficient condition for rank $= 2$ of the matrix in Eq. (21) is that at least two of the planes are not parallel, and all of the planes are parallel to the same straight line.

The coefficient matrix of Eq. (20) can be expressed as the form of Eq. (21):

$$\begin{bmatrix} \mathbf{n}_1 & \mathbf{n}_2 & \cdots & \mathbf{n}_N \end{bmatrix} \begin{bmatrix} \mathbf{n}_1 & \mathbf{n}_2 & \cdots & \mathbf{n}_N \end{bmatrix}^T. \tag{21}$$

Proof of the theorem's sufficiency:

Supposing two planes p1 and p2 are not parallel, and $\frac{n_{11}}{n_{21}}, \frac{n_{12}}{n_{22}}, \frac{n_{13}}{n_{23}}$ are not all equal, then the matrix $\begin{bmatrix} \mathbf{n}_1 \\ \mathbf{n}_2 \end{bmatrix}$ has at least one second-order subdetermination, that is, not 0, i.e., the rank is equal to 2. One can then conclude that the rank of $\begin{bmatrix} \mathbf{n}_1 & \mathbf{n}_2 & \cdots & \mathbf{n}_N \end{bmatrix}$ is larger than or equal to 2.

Since all planes are parallel to the same straight line, all planes can be made across the same straight line by changing d. Then, all planes form a planar cluster. It can be known from the properties of planar clusters that other planes can be linearly represented by p1 and p2:

$$\omega(\mathbf{n}_1^T \mathbf{P}_w + d_1) + \mu(\mathbf{n}_2^T \mathbf{P}_w + d_2) = 0. \tag{22}$$

It can be seen that the rank of $\begin{bmatrix} \mathbf{n}_1 & \mathbf{n}_2 & \cdots & \mathbf{n}_N \end{bmatrix}$ is 2. From $R(A) = R(A^T * A)$,[28] the rank of the formula (21) is 2. The sufficiency certificate is completed.

Next, the necessity of the theorem is proven.

It is known from $R(A) = R(A^T * A)$ that the rank of $\begin{bmatrix} \mathbf{n}_1 & \mathbf{n}_2 & \cdots & \mathbf{n}_N \end{bmatrix}$ is 2. This means that at least two planes are not parallel. Without loss of generality, p1 and p2 are chosen for the following analysis. All other planes can be represented linearly by p1 and p2, i.e.,

$$\omega \mathbf{n}_1^T \mathbf{P}_w + \mu \mathbf{n}_2^T \mathbf{P}_w = 0. \tag{23}$$

Since D affects only the plane position, it does not affect the plane direction, so all planes are parallel to the straight line represented by Eq. (24):

$$\begin{cases} \mathbf{n}_1^T \mathbf{P}_w = 0 \\ \mathbf{n}_2^T \mathbf{P}_w = 0 \end{cases}. \tag{24}$$

The necessity certificate is completed.

According to the above Theorem 1, if there are only two cameras (not parallel), then the equation has a unique nonzero solution, and the geometric meaning of the solution is consistent with the PI method. If there are more than two cameras, unless the condition of Theorem 1 is satisfied, the system of equations does not have a strict nonzero solution, and only the least-squares solution $(l, m, n)^T$ can be obtained as the optimal solution.

3.3.3 *Efficiency analysis*

The linear method for minimizing the object angle residual is as follows.

1. Solve each plane parameter with time complexity of $O(n)$.
2. Solve the coefficient matrix with time complexity of $O(n)$.
3. Solve the final result in a fixed time.

It can be seen from the above analysis that the method can be solved only by calculating the coefficient matrix and solving equations with the least square method. Therefore, its time complexity is $O(n)$.

3.4 *Image-Space Angle Residual*

3.4.1 *Definition of the optimization problem*

It is assumed that the final central axis direction is $I = (l, m, n)^T$ and the axis passes through a point $P_0(x_0, y_0, z_0)^T$ in the world coordinate system. It is easy to see that the straight line passes through the point $P_1(x_0 + l, y_0 + m, z_0 + n)^T$. In the world coordinate system, P_0 can be obtained from the intersection of the pair of corresponding image point on the central axis. The plane consisting of the straight line and the optical center of the i'th camera can be expressed as follows:

$$\begin{vmatrix} x_w & y_w & z_w & 1 \\ Ox_i & Oy_i & Oz_i & 1 \\ x_0 & y_0 & z_0 & 1 \\ x_0 + l & y_0 + m & z_0 + n & 1 \end{vmatrix} = 0, \tag{25}$$

where (Ox_i, Oy_i, Oz_i) is the optical center of the i'th camera. The plane equation after simplification is

$$\mathbf{n}_{wi}^T \mathbf{P}_w + d_i = 0, \tag{26}$$

where n_{wi} is the normal vector of the plane and the subscript w means that it is from the world coordinate system. It can be seen from Eq. (25) that n_{wi} can be represented linearly by I. Letting the relationship matrix be M, i.e.,

$$\mathbf{n}_{wi}^T = \mathbf{I}^T \mathbf{M}_i. \tag{27}$$

It is then easy to prove that M is a full-rank matrix. The transformation matrix from the world coordinate system to the camera coordinate system is $\begin{bmatrix} \mathbf{R}_i & \mathbf{t}_i \\ 0 & 1 \end{bmatrix}$, which is also a full-rank matrix. The plane is expressed in the camera coordinate system as

$$\mathbf{n}_{ci}^T = \mathbf{n}_{wi}^T \begin{bmatrix} \mathbf{R}_i & \mathbf{t}_i \\ 0 & 1 \end{bmatrix}, \tag{28}$$

where n_{ci} is the normal vector of the plane and the subscript c means that it is from the camera coordinate system. From the image plane equation $z = f$, the direction of the intersection line is

$\mathbf{I}_i = (n_{ci1}, n_{ci2}, 0)$. Then, the angle between the axis of the target in the image plane and the intersection line is

$$\alpha_i = \cos^{-1}\left(\frac{\mathbf{I}_i[a_i \ b_i]^T}{\|\mathbf{I}_i\|\|a_i \ b_i\|}\right), \tag{29}$$

where a_i and b_i are coefficients of the linear equation $[a_i \ \ b_i \ \ c_i][x \ \ y \ \ 1]^T = 0$ in the image plane.

Therefore, the optimized objective function is

$$\min \sum_{i=1}^{N} \alpha_i^2, \tag{30}$$

where N is the number of cameras.

3.4.2 *Linear method*

A method of linearly solving for the minimum value is given by derivation below.

The original expression (30) is complicated, especially the partial derivative expression. The numerator and the denominator contain $(l, m, n)^T$, which cannot be solved linearly. Therefore, the expression is deformed.

If the two 2D lines $\mathbf{I}_1 = (a_1, b_1, c_1)^T$, $\mathbf{I}_2 = (a_2, b_2, c_2)^T$ are nearly parallel, the angle $\alpha_i = \tan^{-1} a_1/b_1 - \tan^{-1} a_2/b_2$ is a small angle and we have $|\alpha_i| \approx |\tan(\alpha_i)| = \left|\frac{\frac{a_1}{b_1} - \frac{a_2}{b_2}}{1 + \frac{a_1 a_2}{b_1 b_2}}\right| = \left|\frac{a_1 b_2 - a_2 b_1}{a_1 a_2 + b_1 b_2}\right|$.

For the sake of discussion, we divide a_1, b_1 and a_2, b_2 by the larger one, respectively, so that the maximum value is 1. Then, $1 \le |a_1 a_2 + b_1 b_2| \approx |a_1^2 + b_1^2| \le 2$, so $|\alpha_i|$ is positively correlated with $|a_1 b_2 - a_2 b_1|$. $|\alpha|$ is the absolute value of α. Then, Eq. (31) can be used instead of Eq. (28) as the optimization solution objective function:

$$\min\left[\sum_{i=1}^{N} (a_i n_{ci2} - b_i n_{ci1})^2\right]. \tag{31}$$

From the previous derivation, n_{ci1}, n_{ci2} is a linear function of $I = (l, m, n)^T$. Then, the equation can be recorded as $\sum_{i=1}^{N} (\mathbf{m}_i^T \mathbf{I})^2$ where m_i are the coefficients of I in Eq. (31) and we can think of it as representing a normal vector to a plane too.

This can be thought of as a function of $(l, m, n)^T$. Clearly, this function is a basic elementary function, so it is continuous and differentiable in the domain of the definition, and the minimum value must exist. At the point where the minimum value is obtained, the partial derivative of $(l, m, n)^T$ exists and it is 0, and the partial derivative is obtained separately:

$$\begin{cases} y_l' = \sum_{i=1}^{N}(\mathbf{m}_{i1}\mathbf{m}_i^T \mathbf{I}) = 0 \\ y_m' = \sum_{i=1}^{N}(\mathbf{m}_{i2}\mathbf{m}_i^T \mathbf{I}) = 0 \\ y_n' = \sum_{i=1}^{N}(\mathbf{m}_{i3}\mathbf{m}_i^T \mathbf{I}) = 0 \end{cases} \tag{32}$$

We can solve the equations by the SVD[27] decomposition of the coefficient matrix.

It can be understood from Theorem 1 that only two planes represented by m_i need to be nonparallel and all planes are parallel to the same line. The rank of this matrix is 2, which indicates that the system of equations has a unique nonzero solution.

Since all planes are parallel to the straight line (x, y, z), $(x + l, y + m, z + n)$, and only the optical center of the two cameras and the target line is not coplanar, then the rank of the matrix of Eq. (32) is 2, and the direction solved for by the image angle residual can be used to obtain the global minimum.

3.4.3 *Efficiency analysis*

The linear method for minimizing the image angle residual is as follows.

1. Solve the 3D coordinate of the corresponding image points with time complexity of $O(n)$.
2. Solve the straight-line parameters projected onto the image plane with time complexity of $O(n)$.
3. Solve the coefficient matrix with time complexity of $O(n)$.
4. Find the final result in a fixed time.

It can be seen from the above analysis that the image angle residual method can be solved for only by calculating the 3D point coordinate, the projection line parameter, the solution coefficient matrix, and the least squares solution, so the time complexity is $O(n)$, which is the same as that of the OARL method.

4 Experiments

Because the ground truth was not easy to obtain in the real experiment, we used only the real data to verify the validity of the algorithm, and then we used the simulation to verify the accuracy of the algorithm.
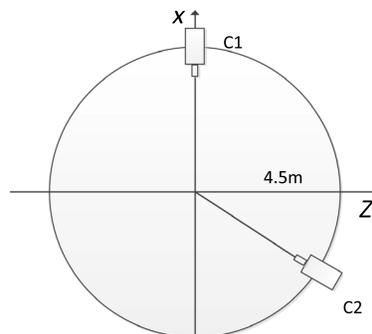
4.1 *Simulation*

4.1.1 *Simulation environment*

The simulation platform was Windows 7, and the processor was an Intel(R) Core TM i7-6820HQ 2.7 GHz.

Simulation conditions: The equivalent focal length was (2181.8, 2181.8) and the image main point was (1023.5, 1023.5). The cameras were arranged in a circle around the target, and the radius of the circle was 4.5 m. The origin of the world coordinate system was the center of the circle. The two endpoints of the target axis were $(-0.5, -0.5, -0.5)$ and $(1.5, 1.5, 1.5)$. The top view of simulation scenario is shown in Fig. 5.

4.1.2 *Simulation description*

1. The corresponding image points of the image-space residual method were obtained by adding the same error in the image to the endpoint (0.5, 0.5, 0.5). The corresponding image points were also affected by the error, ensuring that the image-space residual method was also solved under the same error conditions.
2. When there were only two cameras, the angle between the two cameras and the target center was 120 deg.

Fig. 5 The top view of the simulation scenario.

3. The PI, OARI, OARL, IARI, and IARL were used for each simulation. We used the angle between the result and the true value to evaluate the accuracy. All of the simulations were calculated 1000 times and the root mean square (RMS) of the angular error was obtained. The units were degrees except for simulation condition 7 (see below).

$$E = \cos^{-1} \frac{\mathbf{II}_t}{\|\mathbf{I}\|\|\mathbf{I_t}\|} . \tag{33}$$

Simulations were conducted for the following conditions, separately:

1. Number of cameras.
2. Axis extraction error.
3. Error of the camera angle.
4. Error of the camera optical center.
5. Errors of all camera external parameters.
6. All camera external parameters and extraction errors.
7. Solving for all external parameters and extraction errors and comparing the running time.

4.1.3 Simulation results

The different simulations and the related results are given below.

Simulation 1: A Gaussian error with a mean of 0 was added to the optical center of the camera, with a 10 mm RMS value. A Gaussian error with a mean of 0 was added to the camera angle, with a 0.5 deg RMS value. A Gaussian error with a mean of 0 was added to the x and y directions of the two endpoints of the axis with a 1-pixel RMS value. To verify the solution of this method in the case of only two cameras, the number of cameras ranged from two to nine. The cameras were arranged on a circle with a radius of 4.5 m, the center of the circle was the origin of the world coordinate system. The angle was calculated according to the average number of cameras. The corresponding image points of the IAR were obtained through the intersection of the head points with the error added. The simulation results for the different numbers of cameras are as follows.

1. If there were only two cameras, the geometric meaning of the OAR method was the same as that of the PI method, so, the result was the same, and the IAR used the corresponding image point to minimize the image residual, so, the geometric meaning was inconsistent and the accuracy was higher.
2. As the number of cameras increased, the accuracy of all of the methods increased gradually, indicating that these methods effectively utilized the constraints of the multi-camera.
3. It is observable from Table 1 that the accuracy of the OAR and PI was the same, and the IAR achieved the highest accuracy.

Table 1 Relationship between angle error and number of cameras.

Number	PI	OARI	OARL	IARI	IARL
2	1.16	—	1.16	—	**1.047**
3	0.61	0.61	0.61	**0.563**	**0.563**
4	0.594	0.594	0.594	**0.49**	**0.49**
5	0.473	0.473	0.473	**0.381**	**0.38**
6	0.435	0.435	0.435	**0.355**	**0.355**
7	0.402	0.402	0.402	**0.318**	**0.318**
8	0.371	0.371	0.371	**0.293**	**0.293**
9	0.351	0.351	0.351	**0.273**	**0.273**

Note: The bold means the best results.

4. The accuracy of the linear solution and iterative solution of the OAR and the IAR was the same, which shows that the linear method accuracy was equivalent to the iterative method.

In the subsequent simulation results, the linear and iterative results were essentially the same. For the convenience of comparison, only the OARL and IARL are discussed for the following experiment (except the running time comparison in simulation 7).

Simulation 2: A Gaussian error with a mean of 0 was added to the X and Y directions at the head and the tail of the axis with 0 to 4 pixels of RMS. We fixed the number of cameras at $N = 5$. The simulation results are shown in Fig. 6.

Simulation 3: A Gaussian error with a mean of 0 was added to the three angles of the external parameter with 0.1 to 1 deg of RMS. The step was 0.05 deg, and the number of cameras was 5. The simulation results are shown in Fig. 7.

Simulation 4: A Gaussian error with a mean of 0 was added to the three values of the camera's optical center, with 1 to 50 mm of RMS. The step size was 5 mm and the number of cameras was 5. The simulation results are shown in Fig. 8.

Simulation 5: A Gaussian error with a mean of 0 was added to the optical center of the camera, with 1 to 50 mm of RMS, and the step size was 1 mm. A Gaussian error with a mean of 0 was added to the camera angle, with 0.05 to 2.5 deg of RMS, and the number of cameras was 5. The simulation results are shown in Fig. 9.

Simulation 6: A Gaussian error with a mean of 0 was added to the optical center of the camera, with 1 to 50 mm of RMS, and the step size was 1 mm. A Gaussian error with a mean of 0 was added to the camera angle, with 0.05 to 2.5 deg of RMS. A Gaussian error with a mean
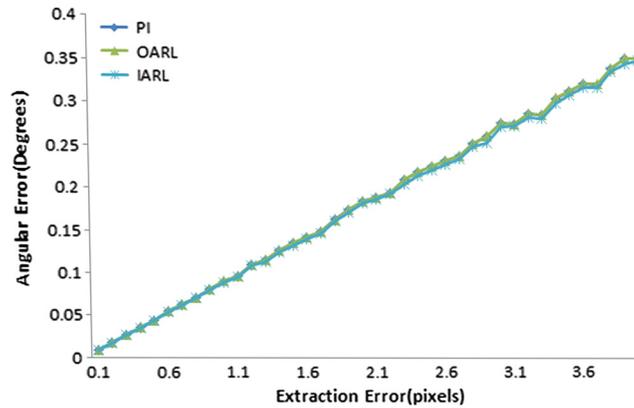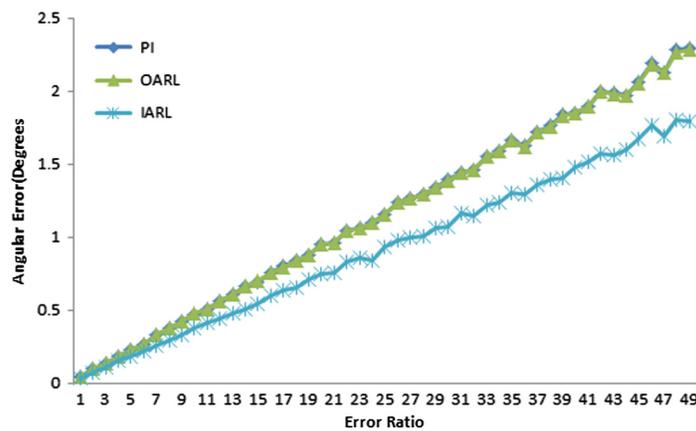

Fig. 6 Relationship between the angular error and the axis extraction error. It is clear that the accuracy of the three methods was similar. As the error increased, the errors of the three methods increased.
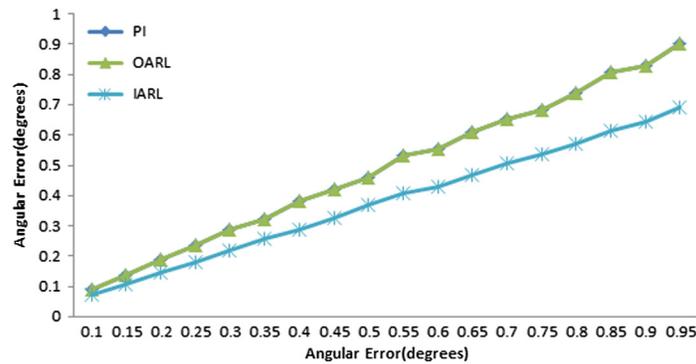
Fig. 7 Relationship between the angular error and the camera angle error (in degrees). It is apparent that the accuracy of the OAR and the PI was same, and the IAR achieved higher accuracy.
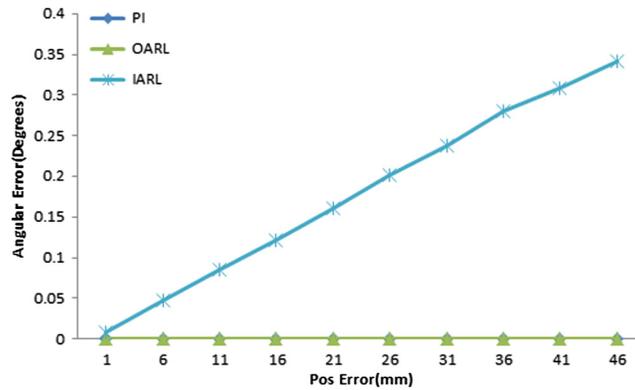
Fig. 8 Relationship between the angular error and the optical center error (in millimeters). It is evident for the case of only the optical center error, the error was 0 since the OAR and the PI did not use the optical center. However, the IAR method produced a certain error.

Fig. 9 Relationship between the angular error and the camera external parameter error. For the case of only an external parameter error, the accuracy of the OAR and PI was same, and the IAR had a higher accuracy.

of 0 was added to the *x* and *y* directions of the two endpoints of the axis, with 0.1 to 5 pixels of RMS, and the number of cameras was 5. The simulation results are shown in Fig. 10.

Simulation 7: The same error was added as in case 1. The number of cameras was increased from 3 to 100. The simulation was run 10,000 times. The results are shown in Fig. 11.

4.1.4 *Simulation conclusion*

We can conclude from the above simulations:

1. The accuracy of the IAR was higher than that of the OAR and PI.
2. The linear methods presented in this paper achieved consistent accuracy for the iterative method, and the computation burden was much lower than that of the iterative method.

4.2 *Experiment Verification*

Next, we will demonstrate the PI's advantage for the measurement of a long axisymmetric target with physical experiment 1 and demonstrate the correctness of the method proposed in this paper with physical experiment 2.
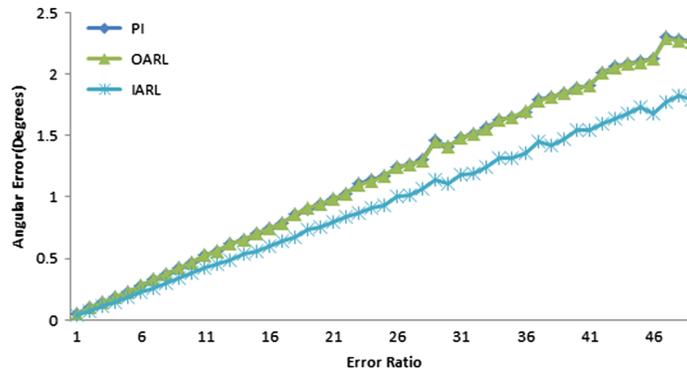
Fig. 10 Relationship between the angular error and the camera external parameters and the axis extraction error. The accuracy of the OAR and PI was same, and the IAR achieved a higher accuracy.
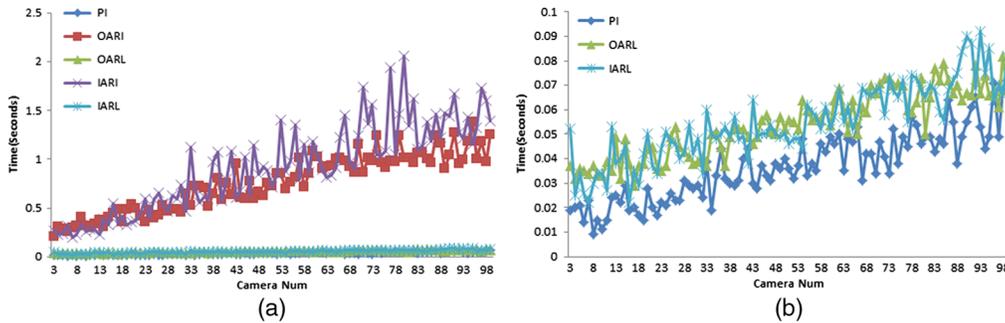

Fig. 11 (a) and (b) Relationship between the running time (in seconds) and the number of cameras. (1) The overall running time of the iterative method was higher than that of the other methods, especially the IARI, and its time-consuming instability was mainly caused by the inconsistent number of iterations for different random errors. (2) With the increase in the number of cameras, the cost times of the PI, OARL, and IARL were similar, because all of these methods were linear methods.

4.2.1 Physical experiment 1

Physical environment. To verify the accurate performance of the central axis intersection algorithm in the axial attitude measurement of the symmetric target, the design experiment was verified and the four commonly used methods were tested, namely

OPNP + OI: single camera position calculation method (OPNP[17] + OI[18]).

GOIAMCS: generalized orthogonal iterative Algorithm of multiple camera systems (GOIAMCS).[19]

DMPI+AO: dual-camera multiple points intersection [21] + absolutely orientation[20] (DMPI + AO).

PI: plane intersection.[1]

The test scenario was shown in Fig. 12. The target was a cylindrical barrel. The points on the barrel were used as marker points on the target for pose calculation. The points on the wall and on the pillar were calibration control points. The optical axes of the two cameras all faced the barrel, with the center of the barrel as the intersection point. The angle was ∼90 deg, and the optical center was about 1.5 m from the barrel. The target filled the width of the field of view as much as possible. A column of points was arranged in the direction of the camera and for 45 deg on both sides. The main purpose of this was to simultaneously improve the coverage area of the marker points and ensure the extraction accuracy of the points as much as possible. At the same time, some points were arranged in the common area of the two cameras for method 3.

(a)   (b)
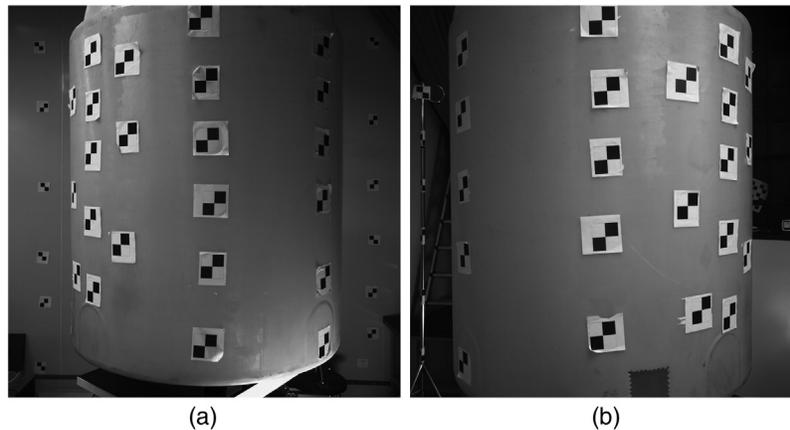
Fig. 12 (a), (b) Physical scene graphs are two images captured by two cameras from different positions, and the barrel is the target to be measured. The points on the barrel served as marker points on the target for pose calculation. The points on the wall and on the pillar are calibration control points.

Some cooperative signs were also arranged on the walls around the cameras to convert the control points and marking points to the same coordinate system when the total station frame was in two different positions.

The auxiliary measuring equipment consisted of a Leica TS30 and a Leica TM5100A. The Leica TS30 was used to obtain the coordinates of the control points for calibration and the coordinates of the target's points for posture calculation. Its accuracy was 0.6 mm. The Leica TM5100A was used to measure the change in the angle of the barrel. Its accuracy was 0.5 arc sec, and the accuracy of the vertical lifting platform to which the Leica TM5100A was attached was 2 arc sec.

Experimental procedure. The barrel was attached to a platform that could precisely control the direction of rotation, and a plane mirror was fixed on the barrel as the measuring reference of the Leica TM5100A.

The measurement steps were as follows:

1. The Leica TS30 was used to measure the markers on the target and the fixed point, and the fixed point was used to calibrate the camera.
2. The common points on the wall were used to unify the mark points on the target in the same coordinate system.
3. Four methods were used to measure the initial pose of the target at the initial position.
4. The platform was used to control the barrel to move in the same direction five times, with 1 deg of movement for each motion. The image was captured by two cameras and measured by the Leica TM5100A at the same time. The measured result of the Leica TM5100A was used as the true value of the actual motion. The results are presented in Table 2.
5. Four methods were used to solve for the target pose. The r-matrix obtained by the above three methods was decomposed to obtain its Y axis direction, and the Y axis direction angle obtained by the same method in the third step was calculated. For the angle obtained by the PI method, the Y axis angle obtained by the third step PI method was directly calculated. To measure the variation as a measure of the accuracy of the measured results, as presented in Table 2, the unit was degrees.

The RMS of error is $\sum_1^5 (I_c - I_t)^2$, where I_c is the value got by algorithm and I_t is the truth value got by the total station. It is evident from Table 2 that the RMS of attitude error measured in this experimental scenario from low to high was PI, GOIAMCS, DMPI + AO, and OPNP + OI. Therefore, it can be proven that the PI method had advantages in the axis attitude measurement of a long symmetric target.

Table 2 True values and algorithm results (unit: degrees).

| True value	OPNP + OI	GOIAMCS	DMPI + AO	PI
1.0135	1.0573	1.0164	1.0179	**1.0146**
2.0432	2.0517	**2.0361**	2.0317	2.0298
3.0672	**3.0687**	3.05351	3.0466	3.0555
4.0954	4.0718	4.07406	4.0657	**4.0846**
5.1154	5.0860	5.09529	5.0820	**5.0986**
RMS of error	0.026	0.015	0.023	**0.012**

Note: The bold means the best results.

4.2.2 *Physical experiment 2*

To verify the correctness of the algorithm, the method was also verified by physical experiments.

Physical environment. The physical environment for physical experiment 2 was shown in Fig. 13. The target was placed on the board in the center and the diagonal markers on the side were the calibration control points. The camera's pixel size was 4288×2848, and the equivalent focal distance was 5200.

Experimental procedure. The experiment was conducted as follows:

1. The total station was used to obtain the coordinates of the calibration control points in the total station coordinate system. The total station coordinate system was used as the world coordinate system. The coordinate system direction was Y vertically upward and XZ horizontal, and the system constituted a right-hand coordinate system.
2. The same camera was used to take photos from nine different locations. The calibration control points were extracted from the image and the 3D coordinates were used to calibrate all of the cameras for a unified world coordinate system.
3. The target central axis was extracted from the image.
4. The PI, OARI, OARL, IARI, and IARL were used separately to obtain the results. All results are shown in Table 3. It can be seen from the table that the results of the five methods were essentially the same, which proved the correctness of the method.
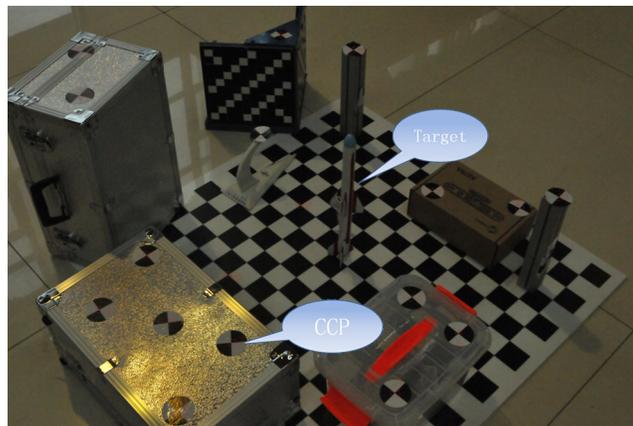
Fig. 13 Physical scene graph 2. The target was placed on the board in the center and the diagonal markers on the side were the calibration control point.

Table 3 Experimental results (unit: degrees).

Parameter	PI	OARI	OARL	IARI	IARL
Yaw	−0.97	−0.95	−0.97	−1.02	−1.02
Pitch	−89.31	−89.31	−89.31	−89.09	−89.09

5 Conclusion

This paper proposed two methods, which were based on minimizing the residuals of the object angle and image angle. These two methods can provide geometric explanations that are lacking in the PI method. The simulation demonstrated that in the case of only two cameras, the results of the OAR method were consistent with those of the PI method, and the IAR method achieved higher accuracy than the PI and OAR method using corresponding image point information. It was not difficult to extract a pair of corresponding image points on the target axis in the actual task, so the method had important practical significance for the case of only two cameras. For a multiple camera system, the accuracies of the OAR and the PI were the same, and the accuracy of the IAR was higher than the accuracies of the OAR and the PI. Furthermore, the linear method was compared with the iterative method. We proved that the IARL was generally faster than the IARI. The IARL had great significance in the actual task measurement scenario.

However, in the same manner as the PI, all of these proposed methods need the target's center axis for measurement, so they are suitable only for the measurement of an axisymmetric target.

Acknowledgments

References

1. Q. F. Yu, X. Y. Sun, and G. J. Chen, "A new method of measure the pitching and yaw of the axes symmetry object through the optical image," *J. Natl. Univ. Def. Technol.* **22**(2), 15–19 (2000).
2. Y. Zhang et al., "Attitude measurement method research for missile launch" *China Opt.* **8**(6), 997–1003 (2015).
3. K. Luo et al., "Measuring technology on elevation angle and yawing angle of space target based on optical measurement method," *J. Changchun Univ. Sci. Technol.* (2007).
4. C. Wang, *Study on Methods to Improve the Attitude Measurement Precision of the Rotary Object*, University of Chinese Academy of Sciences (Institute of Optics and Electronics) (2013).
5. S. W. Ding, *Research on Axis Extraction Methods for Target's Attitude Measurement*, National University of Defense Technology (2013).
6. L. Liu and Z. Zhao, "A new approach for measurement of pitch, roll and yaw angles based on a circular feature," *Trans. Inst. Meas. Control* **35**(3), 384–397 (2013).
7. R. Zhao et al., "Pose estimation based on the constraints of inner angles and areas of triangles," *Proc. SPIE* **7384**, 738411 (2009).
8. C. Liu and W. Hu, "Relative pose estimation for cylinder-shaped space-crafts using single image," *IEEE Trans. Aerosp. Electron. Syst.* **50**(4), 3036–3056 (2014).
9. J. Peng, W. Xu, and H. Yuan, "An efficient pose measurement method of a space non-cooperative target based on stereo vision," *IEEE Access* **5**, 22344–22362 (2017).
10. J. Peng et al., "Pose measurement and motion estimation of space non-cooperative targets based on laser radar and stereo-vision fusion," *IEEE Sens. J.* **19**(8), 3008–3019 (2018).
11. M. Becke and T. Schlegl, "Least squares pose estimation of cylinder axes from multiple views using contour line features," in *IECON 2015-Conf. IEEE Ind. Electron. Soc.*, IEEE, pp. 001855–001861 (2015).

12. J. B. Huang, Z. Chen, and T. L. Chia, *Pose Determination of a Cylinder Using Reprojection Transformation*, Elsevier Science Inc. (1996).
13. C. Doignon and M. De Mathelin, "A degenerate conic-based method for a direct fitting and 3-D pose of cylinders with a single perspective view," in *IEEE Int. Conf. Rob. Autom.*, IEEE, pp. 4220–4225 (2007).
14. A. Y. C. Shiu and H. Zhuang, "Pose determination of cylinders, cones, and spheres from perspective projections," *Proc. SPIE* **1708**, 771–782 (1992).
15. R. Hanek et al., "Yet another method for pose estimation: A probabilistic approach using points, lines, and cylinders," in *IEEE Comput. Soc. Conf. Comput. Vision and Pattern Recognit.*, Vol. 2, p. 550 (1999).
16. Z. Zhao and G. Wen, "Model-based estimation of axisymmetric target's pose and speed using line-scan camera," in *Int. Conf. Syst. Sci. Eng. Des. and Manuf. Inf.*, IEEE, pp. 31–34 (2011).
17. Y. Zheng et al., "Revisiting the PnP problem: a fast, general and optimal solution," in *ICCV 2013*, pp. 2344–2351 (2013).
18. C.-P. Lu, G. D. Hager, and E. Mjolsness, "Fast and globally convergent pose estimation from video images," *IEEE Trans. Pattern Anal. Mach. Intell.* **22**(6), 610–622 (2000).
19. Y. X. Xu, Y. L. Jiang, and C. Fang, "Generalized orthogonal iterative algorithm for pose estimation of multiple camera systems," *Acta Opt. Sin.* **29**(1), 72–77 (2009).
20. B. K. P. Horn, H. M. Hildren, and S. Negahdaripour, "Closed-form solution of absolute orientation using orthonormal matrices," *Int. J. Opt. Soc. Am. A* **5**, 1127–1135 (1988).
21. Y. Q. Zhang et al., "Pose optimization based on integral of the distance between line segments," *Sci. China Technol. Sci.* **59**(1), 135–148 (2016).
22. J. Heller, M. Havlena, and T. Pajdla, "Globally optimal hand-eye calibration using branch-and-bound," *IEEE Trans. Pattern Anal. Mach. Intell.* **38**(5), 1027–1033 (2015).
23. Z. Q. Zhang, L. Zhang, and G. Z. Yang, "A computationally efficient method for hand-eye calibration," *Int. J. Comput. Assist. Radiol. Surg.* **12**(10), 1775–1787 (2017).
24. J. J. Lee and M. H. Jeong, "Stereo camera head-eye calibration based on minimum variance approach using surface normal vectors," *Sensors* **18**(11), 3706 (2018).
25. Y. Zhang, Z. C. Qiu, and X. M. Zhang, "Calibration method for hand-eye system with rotation and translation couplings," *Appl. Opt.* **58**(20), 5375–5387 (2019).
26. L. M. Zhou et al., "3D rigid pose tracking based on new distance function of line segments," *Proc. SPIE* **11179**, 111794L (2019).
27. G. H. Golub, "Singular value decomposition and least squares solutions," *Numer. Math.* **14**(5), 403–420 (1970).
28. Q. Wang et al., *The Sixth Edition of Linear Algebra*, Department of Mathematics of Tongji University (2014).

Lijun Zhong received his BE degree in computer science and technology from Central South University, Changsha, China, in 2008 and his ME degree in software engineering from Central South University, Changsha, China, in 2008. He is currently pursuing his PhD with the College of Aerospace Science and Engineering, National University of Defense Technology, Changsha, China. His research interests include image processing, computer vision, deep learning, and photogrammetry.

Zhang Li is an assistant professor at the College of Aerospace Science and Engineering, National University of Defense Technology, China. He received his PhD in biomedical engineering from the Delft University of Technology in 2015. He authored more than 20 papers in high-ranking journals and conferences (e.g., TMI and TBME). His research interests include computer vision, particularly image registration. He is also interested in the applications of deep learning.

Xiaohu Zhang received his PhD in aerospace science and technology from the National University of Defense Technology, Changsha, China, in 2006. He is a professor with the School of Aeronautics and Astronautics, Sun Yat-Sen University. His research interests include image processing, computer vision, space situational awareness, and photogrammetry.

Yang Shang received his PhD in aerospace science and technology from the National University of Defense Technology, Changsha, China, in 2006. He is a professor with the College of Aerospace Science and Engineering, National University of Defense Technology, Changsha, China. His research interests include image processing, computer vision, vision navigation, and photogrammetry.

Qifeng Yu received his PhD in precision optical measurement from the University of Bremen, Germany, in 1996. He is a professor with the College of Aerospace Science and Engineering, National University of Defense Technology, Changsha, China. He is also an academician of the Chinese Academy of Sciences. His research interests include image processing, computer vision, vision navigation, and photogrammetry.