IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received February 26, 2019, accepted March 20, 2019, date of publication March 25, 2019, date of current version April 8, 2019.

Digital Object Identifier 10.1109/ACCESS.2019.2907222

An Accurate and Stable Pose Estimation Method Based on Geometry for Port Hoisting Machinery

QI WANG[1,2] AND ZHANGYAN ZHAO[1]
[1]School of Logistics Engineering, Wuhan University of Technology, Wuhan 430063, China
[2]Wuhan Railway Vocational College of Technology, Wuhan 430000, China

Corresponding author: Zhangyan Zhao (13871001982@163.com)

This work was supported by the People's Republic of China Ministry of Science and Technology under Grant 2017YFC0805703.

ABSTRACT As a visual measurement method, photogrammetry is widely used in engineering measurement fields. Photogrammetry must first complete camera pose estimation. However, due to limitations of the port condition, photographs often have problems, such as attitude angles that are too large, coplanar or noncoplanar control point processing, and fewer control points, which leads to the fact that the existing algorithm may not be able to satisfy high-accuracy pose estimation. To solve this problem, a new pose estimation method is proposed in this paper. First, estimate the first pose by iterating between the optimal position and the attitude of the camera. Second, use geometric relationships and physical meaning to obtain the closest value to the actual value of the camera's attitude. Third, use the approximate camera attitude as the initial value of the iterative algorithm to obtain an accurate camera pose. Finally, the proposed method is compared with the best currently employed optimization methods through simulated and real crane image experiments. The experimental results show that our method works accurately, stable, and with no severe pose jump occurrences and can be used for pose estimation of port hoisting machinery.

INDEX TERMS Pose estimation projection operator, camera pose ambiguity, port hoisting machinery, arbitrary initial value.

I. INTRODUCTION

Port hoisting machinery is specialized equipment for hoisting and handling goods in coastal and inland ports [1]. If the port hoisting machinery is broken during a hoisting operation, it may cause incalculable losses to property and people's lives. Therefore, research on safety problems has attracted considerable attention [2] because it is necessary to monitor the deformation of port hoisting machinery. Currently, to monitor steel structure equipment, there are feasible measurement techniques: layout sensors [3], [4], laser scanning [5], [6], use of total station [7] and photogrammetry [8], [9]. The layout sensors method needs to attach sensors on the surface of equipment and lay transmission wires to receive signals, which is a time-consuming and labor-intensive task. In addition, the results obtained by this method are the results of the test points, which cannot reflect the situation of the whole measurement object. Laser measurement has the advantages of fast measurement speed and high accuracy, but the initial data obtained by the laser scanner is a large number of three-dimensional point clouds, and these require a large number of professional personnel and professional software for data processing, which is a tedious task. Additionally, the total station method requires considerable on-site measurement time and the surveyor to aim its sight at the marking point each time, which requires a large amount of professional work. However, port hoisting machinery is usually very busy, and many measurement tasks need to be performed under high load conditions, so the measurement fieldwork must be completed quickly.

Based on these reasons, photogrammetry, as a measure that can be automated, is the most suitable measure for a real-time monitoring system for port hoisting machinery. However, the performance of photogrammetry is still not satisfactory when it is used in monitoring port hoisting machinery. As one of the pioneers, Dr. Li of Wuhan University of Technology (WHUT) attempted to use photogrammetry to detect a portal crane in 2016 and planned to use the detection results in his study of the crane safety assessment method [10]. Unfortunately, this plan failed because of the port's special condition limitation. Because of site operating conditions and safety considerations, neither the seaside nor the logistics

The associate editor coordinating the review of this manuscript and approving it for publication was Jing Liang.

corridor was available for setting the measuring station in port. In addition, the outline size of port hoisting machinery is usually very large. These lead to a large attitude angle when taking photos. In addition, the external surface of the port hoisting machinery is uniformly painted and lacks texture features, which makes it difficult to identify control points in port hoisting machinery. Moreover, the coordinates of the control points are usually obtained by auxiliary tools such as a total station. Ports that use port hoisting machinery are usually very busy, and many measurement tasks need to be performed under high-load conditions; therefore, sufficient control cannot be achieved. These issues resulted in having to deal with coplanar or noncoplanar control points and fewer control points. However, photogrammetry must complete the camera pose estimation first. Therefore, accurately and stably determining the calibrated camera pose whose six-degree-of-freedom are three positions and three attitude angles is becoming increasingly significant. This process is normally known as pose estimation or space resection in photogrammetry. Pose estimation algorithms for port hoisting machinery must have stable and highly accurate solutions in these cases.

Existing pose estimation methods are broadly categorized into two types: noniterative and iterative methods. The former has the advantage of a fast calculation speed due to the application of a linear algorithm. However, noniterative methods are sensitive to the noise of image points, especially when there are fewer control points. Iterative methods are usually more accurate than noniterative methods; however, when the initial value is not appropriate, iterative methods can be difficult to converge to the true value that is a global minimum. The classical noniterative algorithm is the direct linear transformation (DLT) algorithm proposed by Abdel-Aziz and Karara [11], which regards the nine variables of the projection matrix as independent variables for establishing linear constraints. However, due to the neglect of the orthogonal constraints of the rotation matrix, the DLT algorithm requires a large number of control points to obtain an accurate solution. Then, Přibyl *et al.* [12] proposed an improved method that is DLT-Combined-Lines based on the DLT algorithm to solve large sets of lines using linear formulation of Perspective-n-Line (PnL). But this method is not accurate enough to be used in engineering practice. Ansar and Daniilidis [13] developed linear solutions for special linear types of not only n points but also n lines. For pose estimation from line features, this algorithms often lacks sufficient accuracy. By applying the Gröbner basis technique, Banno [14] and Zheng *et al.* [15], [16] devised the pose estimation as a functional minimization problem and retrieved all stationary points. However, those methods are limited by the speed and numerical stability of a Gröbner basis solver. Lepetit *et al.* [17] proposed a high-efficiency and high-precision linear algorithm named EPnP, which is considered to be the best noniterative algorithm. The EPnP method uses 4 virtual control points to show the coordinates of the space points. However, when the control points are coplanar, and the depth changes drastically, the EPnP algorithm may not obtain good estimation results. Thus, when there are no redundant control points, the existing noniterative algorithms are sensitive to image noise because they lack a target function and neglect part of the constraint conditions in the calculation process.

Iterative algorithms can obtain accurate results when redundant control points are not available. Oberkampf *et al.* [18]–[20] proposed a POSIT algorithm to estimate the perspective projection model by using the initial value obtained from the scaling model. The classic iterative algorithms convert pose estimation into a nonlinear least-squares problem and solve it by nonlinear optimization methods. The most representative algorithms define an objective function with the minimum error in the image space or the object space. Garro *et al.* [21] presented an iterative algorithm to define the minimum square sum of residuals in an image space. In addition, Lu *et al.* [22] proposed an orthogonal iterative (OI) algorithm that is one of the best iterative solutions for pose estimation. It establishes the objective function with a minimum collinear error in the object space. Compared with other iterative methods, the OI algorithm has a fast convergence speed and high accuracy. However, it rely on initial reference parameters, and if the initialization is insufficient, it tends to be trapped in a local minimum. Subsequently, Schweighofer and Pinz [23] considered the local minimum and improved the OI algorithm. However, we note that the accuracy of these methods [17], [23] may be reduced when the attitude angle is large. Considering pose ambiguity and large angle cases, Wu *et al.* [24] proposed a method to find the correct solution by calculating all the ambiguous poses. Nonetheless, this algorithm is only suitable for coplanar control points and too inefficient to be suitable for engineering practice. As mentioned above, existing pose estimation methods for port hoisting machinery have some drawbacks: (1) Some algorithms usually require a proper initial value to converge to the true value, which may become stuck in an inappropriate local minimum, especially when the camera has a large attitude angle. (2) Most existing pose estimation solutions may not be able to handle coplanar and noncoplanar control points.

Based on the discussion above, we propose an accurate and stable pose estimation method based on geometry for port hoisting machinery. The central contribution of this paper is that our method does not rely on a proper initial value to converge to the true value and is suitable for large attitude angles, coplanar or noncoplanar control points and fewer control points. Its main ideas are as follows. Step one. According to the collinear error of the space control points in the image space coordinate, the objective function with the minimum error is established by the projection operator. Then, the first pose is estimated by iterating between the optimal position and the attitude of the camera. Step two. The collinear residual error is used to determine whether the first attitude of the camera is a local minimum. If it is, we can determine the closest value to the actual value of the camera's attitude through the constraint of the combination matrix with

the geometric relationship and the actual physical meaning. Step three. The approximate camera attitude is used as the initial value of step one to obtain an accurate camera pose. Finally, the residual comparison and physical meaning tests of the calculated camera pose are conducted.

This paper is organized as follows: In Section 2, the projection operator is utilized in the objective function to estimate optimal camera poses. Then, we illustrate the geometric ambiguity that may lead to multiple solutions and determine the relationship between the local minimum and the correct solution by mathematical analysis in Section 3. In Section 4, we propose a stable pose estimation algorithm, and Section 5 presents the experimental results using both simulated data and real crane image data. Section 6 presents the conclusion.

II. OPTIMAL POSE BASED ON PROJECTION OPERATOR

The aim of pose estimation is to estimate the position and attitude of a camera given intrinsic parameters through a sequence of space control points and their corresponding image points. Due to measurement errors, the image points and the space control points in the image space coordinate have errors. This paper uses the projection operator to establish the objective function based on the collinear error of the space control point in the image space coordinate system. Then, the optimal camera position and attitude can be obtained through the minimum objective function. In addition, in the process of establishing the error function, a unit quaternion is used to form the rotation matrix representing the attitude of the camera to avoid the Euler angles expressed by complex trigonometric functions.

A. ROTATION REPRESENTATION

In photogrammetry, not considering image distortion, the classic collinear equation is as follows:

$$\begin{aligned} x &= -f\frac{a_1(X-X_s)+b_1(Y-Y_s)+c_1(Z-Z_s)}{a_3(X-X_s)+b_3(Y-Y_s)+c_3(Z-Z_s)}, \\ y &= -f\frac{a_2(X-X_s)+b_2(Y-Y_s)+c_2(Z-Z_s)}{a_3(X-X_s)+b_3(Y-Y_s)+c_3(Z-Z_s)}. \end{aligned} \tag{1}$$

where (x, y) denotes the coordinates of an image point in image plane $(X, Y, Z)^T$ denotes the coordinates of the space control points in the object space, $(X_s, Y_s, Z_s)^T$ denotes the coordinates of the camera position in the object space f is the focal length of the camera lens; and $a_1 \sim c_3$ are elements of rotation matrix R, which can be transformed into the attitude of the camera.

According to expression (1), the collinear equation can be written as follows:

$$\begin{bmatrix} \bar{X} \\ \bar{Y} \\ \bar{Z} \end{bmatrix} = \begin{bmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{bmatrix} \begin{bmatrix} X-X_S \\ Y-Y_S \\ Z-Z_S \end{bmatrix} = R \begin{bmatrix} X-X_S \\ Y-Y_S \\ Z-Z_S \end{bmatrix}. \tag{2}$$

The vector form is expressed as:

$$p' = R(p - p_s). \tag{3}$$

where $p' = \left(\bar{X}, \bar{Y}, \bar{Z}\right)^T$, $p = (X, Y, Z)^T$, $p_s = (X_s, Y_s, Z_s)^T$, p' point is the coordinate of the space control point in the image space coordinate system, R is the 3∗3 rotation matrix.

The collinear equation can also be expressed as

$$t = -f\frac{p'}{\bar{Z}}. \tag{4}$$

where $t = (x, y, -f)^T$

According to the property of orthogonal projection, the orthogonal projection of the vector p' in the direction of t should be equal to p' Therefore, the expression is as follows:

$$p' = Vp' \tag{5}$$

$$R(p - p_s) = VR(p - p_s). \tag{6}$$

where $V = \frac{tt^T}{t^T t}$ is the projection operator

1) EULER ANGLE

The Euler angles are three angles that represent the orientation of the object presented in one reference frame. The rotation angles corresponding to the rotation order of the y-x-z-axis are α, β, γ, which can be represented in (7), as shown at the bottom of this page.

2) UNIT QUATERNION

In this paper, the unit quaternion is used in calculating the rotation matrix, which can avoid complicated trigonometric functions and improve the speed of computation. The quaternion can express the rotation around an arbitrary axis using complex algebra extended to 3 imaginary dimensions. Complex units i, j, k within the 3 imaginary dimensions can be defined as $k^2 = j^2 = i^2 = kji = -1$, $jk = -kj = iki = -ik = j$, $ij = -ji = k$. The unit quaternion can be defined with a vector form as follows: $\dot{q} = |q_0, q_1, q_2, q_3|^T$

$$R = R_\gamma R_\beta R_\alpha = \begin{bmatrix} \cos\alpha\cos\gamma - \sin\alpha\sin\beta\sin\gamma & \cos\beta\sin\gamma & \sin\alpha\cos\gamma + \cos\alpha\sin\beta\sin\gamma \\ -\cos\alpha\sin\gamma - \sin\alpha\sin\beta\cos\gamma & \cos\beta\cos\gamma & -\sin\alpha\sin\gamma + \cos\alpha\sin\beta\cos\gamma \\ -\sin\alpha\cos\beta & -\sin\beta & \cos\alpha\cos\beta \end{bmatrix} \tag{7}$$

$$R = \begin{bmatrix} q_0^2+q_1^2-q_2^2-q_3^2 & 2(q_1q_2-q_0q_3) & 2(q_1q_3+q_0q_2) \\ 2(q_1q_2+q_0q_3) & q_0^2-q_1^2+q_2^2-q_3^2 & 2(q_2q_3-q_0q_1) \\ 2(q_3q_1-q_0q_2) & 2(q_2q_3+q_0q_1) & q_0^2-q_1^2-q_2^2+q_3^2 \end{bmatrix} \tag{8}$$

and $\sqrt{q_0^2+q_1^2+q_2^2+q_3^2}=1$. The conjugate quaternion is expressed as: $\dot{q}^*=|q_0,-q_1,-q_2,-q_3|^T$.

The rotation matrix is expressed with the unit quaternion form as in (8), as shown at the bottom of the previous page [25], [26]:

The unit quaternion can be converted to Euler angles by the following:

$$\begin{aligned}\alpha &= \arctan\left(-R_{31}/R_{33}\right)\\ \beta &= \arcsin\left(-R_{32}\right)\\ \gamma &= \arctan\left(R_{12}/R_{22}\right)\end{aligned} \tag{9}$$

B. OBJECTIVE FUNCTION

Because of the measurement error, function (5) has a collinear error from which the error function can be derived. The error function can be expressed as:

$$e_i = \left\|p_i' - V_i p_i'\right\|. \tag{10}$$

where i is the number of control points

The error function based on the least squares principle can be expressed as:

$$\begin{aligned}E &= \sum_{i=1}^{n}\|e_i\|^2 = \sum_{i=1}^{n}\left\|p_i' - V_i p_i'\right\|^2\\ &= \sum_{i=1}^{n}\|(I-V_i)R(p_i-p_s)\|^2\end{aligned} \tag{11}$$

which is the objective function.

C. OPTIMAL POSE ESTIMATION

1) OPTIMAL POSITION OF THE CAMERA

Assuming that R is known, only P_s in the objective function is unknown. According to the mean value theorem, we can obtain the optimal solution of the camera position by deriving the objective function (11) which can be expressed as:

$$\frac{\partial E}{\partial p_s} = \sum_{i=1}^{n} -2R^T(I-V)R(p-p_s) = 0. \tag{12}$$

The $(I-V)^T(I-V)=(I-V)$ since the projection operator $(I-V)^T$ is symmetric.

The optimal position of the camera is

$$p_s(R) = \left[\sum_{i=1}^{n} R^T(I-V_i)R\right]^{-1}\cdot\left[\sum_{i=1}^{n} R^T(I-V_i)Rp_i\right] \tag{13}$$

2) OPTIMAL ATTITUDE OF THE CAMERA

The objective function only containing R as an unknown quantity can be expressed as:

$$\begin{aligned}E(R) &= \sum_{i=1}^{n}\left\|Rp_i - Rp_s(R) - V_i p_i'\right\|^2\\ &= \sum_{i=1}^{n}\left\|Rp_i - p_{v,i}' + T\right\|^2\end{aligned} \tag{14}$$

where $p_{v,i}' = V_i p_i'$, $T=-Rp_s(R)$.

The sets of measurements are centralized as follows:

$$\begin{aligned}\bar{p}_i &= p_i - \frac{1}{n}\sum_{i=1}^{n} p_i\\ \bar{p}_{v,i}' &= p_{v,i}' - \frac{1}{n}\sum_{i=1}^{n} p_{v,i}'\\ \bar{T} &= T + R\frac{1}{n}\sum_{i=1}^{n} p_i - \frac{1}{n}\sum_{i=1}^{n} p_{v,i}'\end{aligned} \tag{15}$$

Then, the objective function is expressed as:

$$\begin{aligned}E(R) &= \sum_{i=1}^{n}\left\|R\bar{p}_i - \bar{p}_{v,i}' + \bar{T}\right\|^2\\ &= \sum_{i=1}^{n}\left\|R\bar{p}_i - \bar{p}_{v,i}'\right\|^2 + 2\left(\sum_{i=1}^{n}\left|R\bar{p}_i - \bar{p}_{v,i}'\right|\right)\bar{T} + \sum_{i=1}^{n}\left\|\bar{T}\right\|^2\end{aligned} \tag{16}$$

where $\sum_{i=1}^{n}\bar{p}_i = \sum_{i=1}^{n}\left(p_i - \frac{1}{n}\sum_{i=1}^{n} p_i\right) = 0$ and

$$\sum_{i=1}^{n}\bar{p}_{v,i}' = \sum_{i=1}^{n}\left(p_{v,i}' - \frac{1}{n}\sum_{i=1}^{n} p_{v,i}'\right) = 0.$$

To minimize the expression (16), $\sum_{i=1}^{n}\left\|\bar{T}\right\|^2$ has to be equal to zero.

The above expression can be expressed as:

$$\begin{aligned}E(R) &= \sum_{i=1}^{n}\left\|R\bar{p}_i - \bar{p}_{v,i}'\right\|^2\\ &= \sum_{i=1}^{n}\|R\bar{p}_i\|^2 + \sum_{i=1}^{n}\left\|\bar{p}_{v,i}'\right\|^2 - 2\sum_{i=1}^{n} R\bar{p}_i \bullet \bar{p}_{v,i}'\end{aligned} \tag{17}$$

where $\bullet$ represents the dot products.

In each case, the remaining error is minimized when the last terms are as large as possible. That is,

$$C = \sum_{i=1}^{n} R\bar{p}_i \bullet \bar{p}_{v,i}' \tag{18}$$

should be as large as possible.

To use the unit quaternion to represent the rotation matrix, $\bar{p}_i, \bar{p}_{v,i}'$ represented as:

$$\begin{aligned}\dot{\bar{p}}_i &= \left|0, \bar{X}_i, \bar{Y}_i, \bar{Z}_i\right|^T,\\ \dot{\bar{p}}_{v,i}' &= \left|0, \bar{X}_{v,i}, \bar{Y}_{v,i}, \bar{Z}_{v,i}\right|^T.\end{aligned} \tag{19}$$

in the form of the unit quaternion should be included in expression (18)

Expression (18) uses the unit quaternion as in (20) and (21), as shown at the bottom of the next page, where $N_{XX} = \sum_{i=1}^{n}\tilde{X}_i\tilde{X}_{v,i}$, $N_{XY} = \sum_{i=1}^{n}\tilde{X}_i\tilde{Y}_{v,i}$, and so on.

Since N is a real symmetric matrix, its eigenvalues are real numbers. Without loss of generality, we assume that the eigenvalues corresponding to the eigenvector $(\varepsilon_1, \varepsilon_2, \varepsilon_3, \varepsilon_4)$

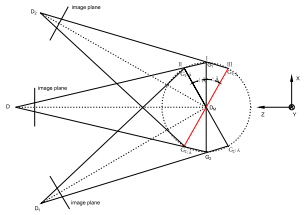


FIGURE 1. Pose ambiguities with a geometric interpretation.

are $\lambda_1 \geq \lambda_2 \geq \lambda_3 \geq \lambda_4$ Because the eigenvectors are linearly independent and orthogonal, the unit quaternions based on $\varepsilon_1, \varepsilon_2, \varepsilon_3, \varepsilon_4$ can be decomposed into

$$\dot{q} = d_1\varepsilon_1 + d_2\varepsilon_2 + d_3\varepsilon_3 + d_4\varepsilon_4, d_1^2 + d_2^2 + d_3^3 + d_4^4 = 1 \quad (22)$$

where d_1, d_2, d_3, d_4 are real numbers.

Expression (20) can be expressed as:

$$\begin{aligned} C &= \dot{q}^T N \dot{q} = \dot{q} \bullet (N\dot{q}) \\ &= d_1^2\lambda_1 + d_2^2\lambda_2 + d_3^2\lambda_3 + d_4^2\lambda_4 \leq \lambda_1 \end{aligned} \quad (23)$$

This equation describes the condition for the establishment of the upper equivalents, which are the maximum conditions for C such that the unit quaternion $\dot{q}$ is the eigenvector corresponding to the maximum eigenvalue of the matrix N. At this point, the expression is as follows:

$$\dot{q} = \arg\max_{\varepsilon} eig\,(N) \quad (24)$$

Then, we can obtain the rotation matrix R through expression (8) and obtain the optimal attitude of the camera.

III. MULTIPLE SOLUTIONS

A. GEOMETRIC INTERPRETATION

Camera pose estimation is defined as the position and attitude of the calibrated camera with respect to the object reference frame to find the minimum value of the corresponding objective function (11). We demonstrate pose ambiguity through a specific example in which the objective function has a local minima. Fig 1 shows three points D, D_1, D_2 that are the image plane, the model of coplanar control points (I, II, III) and the projection centers of different positions of the camera. Without loss of generality, we assume that the coplanar model coordinate axis is located in the model center D_M and is consistent with the origin of the object reference frame. When the camera moves from D point to D_1 point, the corresponding model plane I rotates counterclockwise α degrees around the Y axis of the object reference frame and then obtains plane II However, it can be seen from the graph that the D_2 point

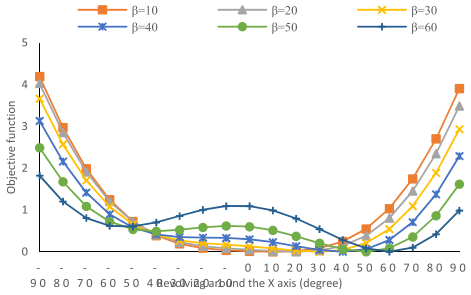


FIGURE 2. Objective function values for various rotation angles.

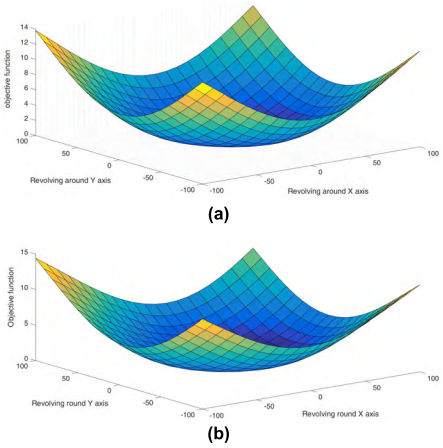


FIGURE 3. The change of the objective function in the case of a small rotating angle. (a) $\alpha = 10°, \beta = 15°$. (b) $\alpha = 20°, \beta = 30°$.

corresponds to the minimum of the objective function as well, corresponding to plane III (G_1, G_2 points rotate clockwise β around the Y axis). Therefore, the existence of multiple solutions for the minimum value of the objective function leads to camera pose ambiguity

Then, we assign values to some parameter configurations and only change one angle parameter. The setting of our model is $p_1 = [1, 1, 0]^T, p_2 = [1, -1, 0]^T, p_3 = [-1, 1, 0]^T$, $p_4 = [-1, -1, 0]^T$, $p_s = (0, 0, 3)^T$. Fig 2 shows how the objective function changes for varying rotation angles β and zero noise. For $\beta = 40°, 50°$, and $60°$, the local minimum of the objective function can be found at $-38.42°, -49.56°$, and $-58.26°$. For $\beta < 34.2°$, no second local minimum of the objective function exists.

With angle $\gamma = 15°$ fixed, when attitude angles α and β are small, the distribution of the objective function in two directions (x-axis and y-axis) is presented in Fig 3. The Fig 4 shows that the attitude angles α and β are large. We can

$$C = \sum_{i=1}^{n} R^T \dot{\bar{p}}_i \bullet \dot{\bar{p}}'_{v,i} = \sum_{i=1}^{n} \left(\dot{q}\dot{\bar{p}}_i\dot{q}^*\right) \bullet \dot{\bar{p}}'_{v,i} = \sum_{i=1}^{n} \left(\dot{q}\dot{\bar{p}}_i\right) \bullet \left(\dot{\bar{p}}'_{v,i}\dot{q}\right) = \dot{q}^T N \dot{q}. \quad (20)$$

$$N = \begin{bmatrix} N_{XX} + N_{YY} + N_{ZZ} & N_{YZ} - N_{ZY} & N_{ZX} - N_{XZ} & N_{XY} - N_{YX} \\ N_{YZ} - N_{ZY} & N_{XX} - N_{YY} - N_{ZZ} & N_{XY} + N_{YX} & N_{ZX} + N_{XZ} \\ N_{ZX} - N_{XZ} & N_{XY} + N_{YX} & -N_{XX} + N_{YY} - N_{ZZ} & N_{YZ} + N_{ZY} \\ N_{XY} - N_{YX} & N_{ZX} + N_{XZ} & N_{YZ} + N_{ZY} & -N_{XX} - N_{YY} + N_{ZZ} \end{bmatrix} \quad (21)$$

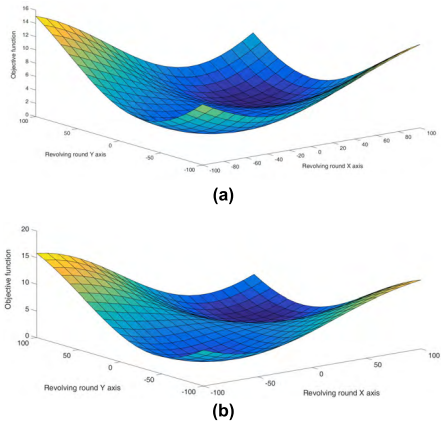


FIGURE 4. The change of the objective function in the case of a large rotating angle. (a) $\alpha = 40°, \beta = 50°$. (b) $\alpha = 50°, \beta = 60°$.

see that in small attitude angles ($\alpha = 10°, \beta = 15°$), ($\alpha = 20°, \beta = 30°$), the objective function only has one minimum value. However, when the attitude angles become large ($\alpha = 40°, \beta = 50°$) and ($\alpha = 50°, \beta = 60°$), the objective function has a local minimum at ($\alpha = -38.44°, \beta = -52.32°$), ($\alpha = -47.58°, \beta = -63.24°$ An Accurate and Stable Pose Estimation Method Based on Geometry for Port Hoisting Machinery), which results in camera pose ambiguity.

B. MATHEMATICAL ANALYSIS

There are at most 4 solutions for pose estimation problems [27], [28]. Previous studies have mainly constructed polynomials using the distance between the control points and the central point of the camera lens. These polynomials determine the number of solutions. However, it is difficult to use in engineering practice because of the complex judgment methods. In our study, a new method that can quickly find the correct pose estimation solution through the attitude angles with definite geometric features is proposed. In expression (14), the objective function only contains R as an unknown quantity. By expression (8), the objective function contains only four variables (q_0, q_1, q_2, q_3). Based on an interval based splitting approach, the stationary points to the objective function can be found [29]. Beginning from intervals $\alpha \in \left[-\frac{\pi}{2}, \frac{\pi}{2}\right], \beta \in \left[-\frac{\pi}{2}, \frac{\pi}{2}\right]$, and $\gamma \in \left[-\frac{\pi}{2}, \frac{\pi}{2}\right]$, we calculate the corresponding image point coordinate t, projection operator V, and then estimate the stationary points of the objective function using Jacobian and Bernstein expansions (readers can refer to [29] for detail). The minimum of the objective function can be obtained by the estimated Hessian matrix

$$H_{q0,q1,q2,q3} = \begin{bmatrix} E_{q0,q0} & E_{q0,q1} & E_{q0,q2} & E_{q0,q3} \\ E_{q1,q0} & E_{q1,q1} & E_{q1,q2} & E_{q1,q3} \\ E_{q2,q0} & E_{q2,q1} & E_{q2,q2} & E_{q2,q3} \\ E_{q3,q0} & E_{q3,q1} & E_{q3,q2} & E_{q3,q3} \end{bmatrix} \quad (25)$$

for each stationary point, which may be a saddle point, maximum, or minimum.

Obtaining the minimum requires that all the eigenvalues of the Hessian matrix are positive. To estimate these eigenvalues, the roots of the characteristic polynomial $\det\left(H_{q0,q1,q2,q3} - \lambda I\right) = 0$ have to be calculated. If all the eigenvalues of the Hessian matrix are not all positive, the stationary point is not a minimum. While the other eigenvalues are not all positive or negative, the stationary point is the local minima of the objective function.

The stationary point (minimum, local minima) expressed in quaternions are transformed into α, β, γ, which is a camera attitude with definite geometric meaning. Our new algorithm is based on these major observations:

(1) The correct solution to the pose estimation problem should have a lower error value.

(2) When α, β, γ are small, there is no local minimum. However, when α, β, γ are large, local minima exist, leading to multiple solutions.

(3) If a local minimum exists, this attitude angle $(\alpha', \beta', \gamma')$ is approximately equal or opposite to the correct solution (α, β, γ).

IV. STABLE POSE ESTIMATION ALGORITHM

The analytical results in Section 3 indicate that the solution tends to fall into a local minimum when the attitude angle of the camera is large. If the local minimum is taken as the camera attitude, pose estimation error will occur, and the port hoisting machinery measurement cannot be completed. To avoid this, it is necessary to find the minimum instead of the local minimum point. According to the analysis above, an accurate and stable pose estimation algorithm based on geometry for port hoisting machinery is presented. That is, estimate the first pose by iterating between the optimal position and the attitude of the camera. In this case, the solution may be the local minimum. To examine the correctness of the solution, the first pose is brought into the objective function (11) to check the residual error. First, we impose a physical meaning constraint $\sum_{i}^{n} V_i R(p - p_s) \geq 0$, which confirms that the observed control points are ahead of the estimated optimal pose in the realistic scene. Theoretically, the value of the function should be zero, but due to the measurement error, it cannot be zero. If the first pose is the correct value, the minimum corresponds to the correct solution. Then, the objective function value will be very small. Therefore, we can set a threshold value to determine whether the first pose is correct. Bringing the first estimate of data into the objective function, if the objective function value is less than the threshold, the camera pose is correct; otherwise, the local minimum solution was found. According to the law found in Section 3 through mathematical analysis, if the local minimum exists, the attitude angles are approximately equal or opposite to the correct solution. In addition, as R_γ only rotates around the optical axis of the camera, the geometric relation between the image plane and the space control points remains constant, and the rotation only influences image coordinates. Thus, we do not need to consider R_γ.

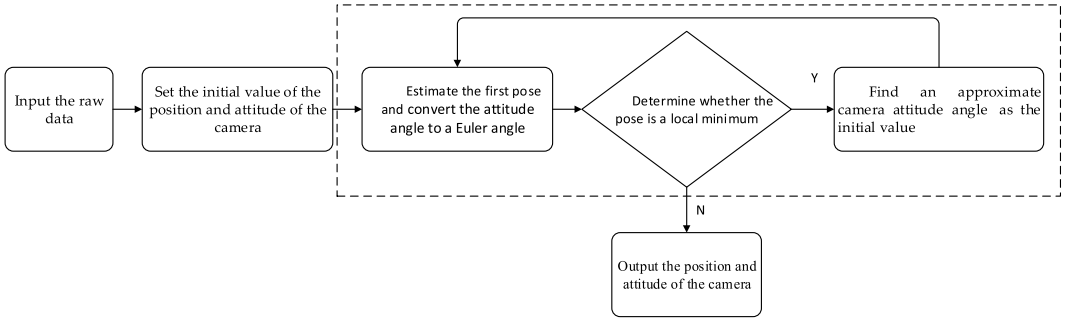


FIGURE 5. The flowchart of the proposed method.

The steps of the stable algorithm are as follows:

(1) Import the raw data, including the camera-intrinsic parameters and measured values of point coordinates t and the coordinates p of the corresponding control points in the object reference frame

(2) Set the initial value of the position and attitude of the camera. Since the proposed algorithm is independent of the initial value, without loss of generality, the initial position of the camera is set to $p_s = (0, 0, 0)^T$, and the unit quaternion corresponding to the attitude of the camera is set to $\dot{q} = (1, 0, 0, 0)^T$

(3) Estimate the first pose by iterating between the optimal position and the attitude of the camera. Then, convert the camera attitude expressed in the form of the unit quaternion to the Euler angle based on function (9).

(4) Determine whether the pose is a local minimum. If it is not, output the position and attitude of the camera. If it is, find a column vector that satisfies the physical meaning constraint, and the minimum of the objective function from a combination matrix $G = \begin{bmatrix} -\alpha_1 & -\alpha_1 & \alpha_1 \\ -\beta_1 & \beta_1 & -\beta_1 \\ \gamma & \gamma & \gamma \end{bmatrix}$ composing three rotation matrices. Then, use the found column vector as the initial value of the camera attitude angle.

(5) Repeat steps (3) to (4) to obtain the accurate camera pose.

The flow chart of the proposed method is shown in Fig. 5.

V. EXPERIMENTS AND RESULTS

To validate the accuracy and stability performance of the presented method, we compared the presented method with the DLT [11], EPNP [17] and OI algorithms [22] under the condition of noncoplanar control points and POSIT [18], RPNP [23] and Ansar and Daniilidis [13] under the condition of coplanar control points based on simulated data and real image data.

A. SIMULATED DATA TESTS AND RESULTS

In the simulated experiment, we used the following setup:

1. A virtual perspective camera with focal length 800 pixels and image size 640*480 pixels was given.

2. In the coplanar case, the control points were distributed uniformly in the range $[-2, 2] \times [-2, 2] \times [0, 0]$ in the object reference frame. In the noncoplanar case, the control points were distributed uniformly in the range $[-2, 2] \times [-2, 2] \times [0, 4]$.

3. For each test, a random attitude rotation of R_{true} was generated.

4. The selection of the camera position $p_{s,ture}$ was determined when the measured image points were located at the camera coordinate frame.

We measured the attitude error between the R_{true} and the calculated rotation matrix R, which was represented as $e_{attitude}(\deg rees) = \max_{k=1}^{3} a\cos(dot(r_{ture}^k, r^r)) \times 180/\pi$ where r_{ture}^k and r^r were the k-th column of R_{true} and R. $a\cos(.)$ and $dot(.,.)$ indicate the arccosine operation and dot product. The position error was defined as $e_{position}(\%) = \|p_{s,ture} - p_s\| / \|p_s\| \times 100$ which was used to measure the relative difference between $p_{s,ture}$ and p_s

To compare the accuracy of various algorithms with various numbers of control points, the number of control points varied from 4 to 15 with a step length of 1 and image point coordinates with Gaussian noise with a mean of 0 and a standard deviation of 3 pixels were randomly generated in the tests. The graphs in Figs 6 and 7 show the error of the position and attitude of the camera for 1000 independent random tests of each algorithm at each number of control point parameters in the noncoplanar and coplanar cases.

As illustrated in Fig 6, in noncoplanar cases, our algorithm is obviously better in accuracy and stability than other algorithms, especially in the case of fewer control points.

As shown in Fig 7, in coplanar cases, the performance of our algorithm is somewhat higher than RPNP [23] but obviously better than the other algorithms [13], [18]. The occurrence of this phenomenon can be interpreted as a local minimum leading to the existence of multiple solutions for the pose estimation problem. When the number of control points is small, it is difficult for other algorithms to distinguish the correct solution from the local minimum value. In addition, although the RPNP [23] method can obtain better results, it is only suitable for coplanar control points.

To compare the accuracy of various algorithms with different noise levels, the number of control points was n = 4, and Gaussian noise with a mean of 0 and the standard deviation of [0, 6] whose step length was 0.5 were added to the image point coordinates in the tests. The graphs in Figs 8 and 9 show the

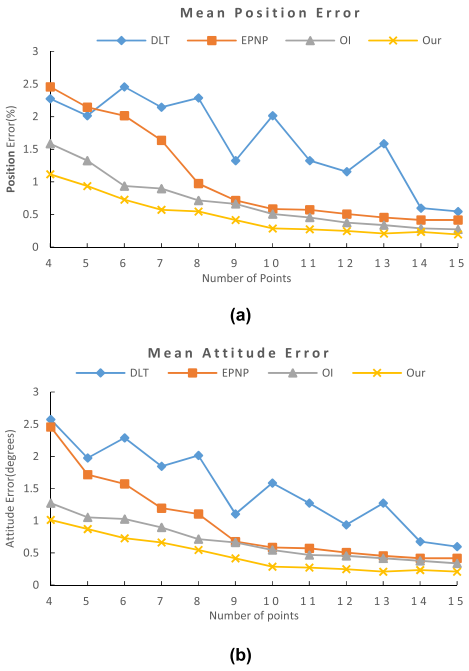


FIGURE 6. The relationship between the number of points and the mean error (noncoplanar). (a) Mean position error. (b) Mean attitude error.

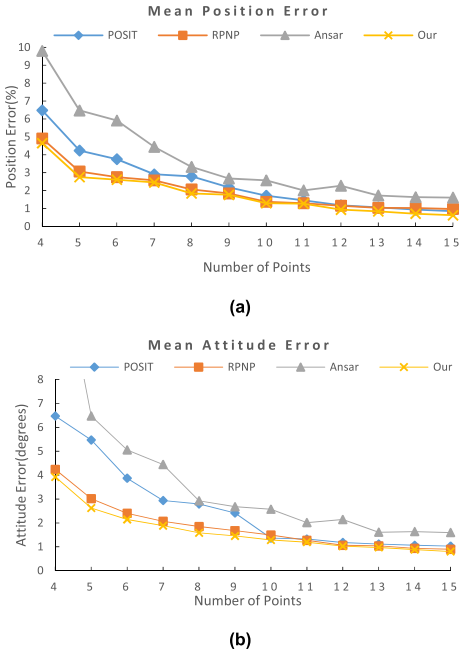


FIGURE 7. The relationship between the number of points and the mean error (coplanar). (a) Mean position error. (b) Mean attitude error.

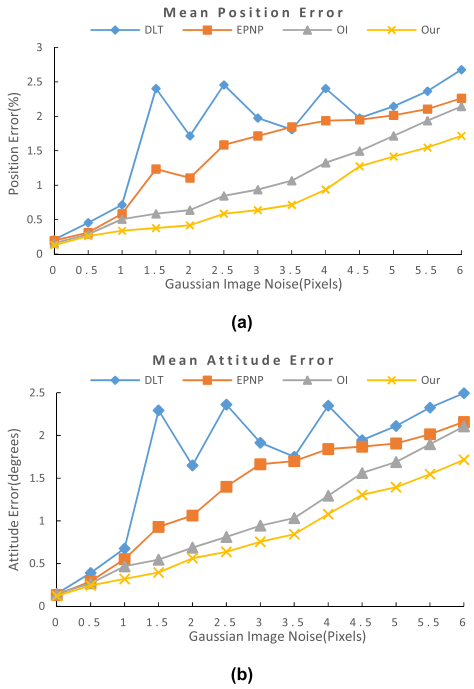


FIGURE 8. Mean error of different image pixel errors (noncoplanar). (a) Mean position error (b) Mean attitude error.

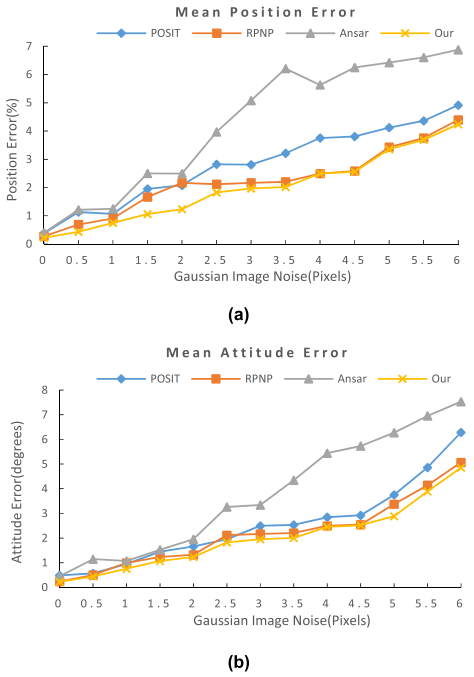


FIGURE 9. Mean error of different image pixel errors (coplanar). (a) Mean position error. (b) Mean attitude error.

error of the position and attitude of the camera for 1000 independent random tests of each algorithm at each noise level parameter in noncoplanar and coplanar cases, respectively.

As illustrated in Fig 8, in noncoplanar cases, the proposed method is consistently much more stable and accurate than the other methods, particularly when the noise level is higher than 1 pixel. Moreover, along with increasing noise disturbance, the errors increased quite slowly. As illustrated in Fig 9, in coplanar cases, the algorithm produces slightly better results than RPNP [23], but it is obviously better than the other two algorithms. This behavior can be interpreted as follows: unlike non-iterative methods, which are sensitive to noise, the proposed algorithm is an iterative

algorithm that has a strong immunity to noise. After obtaining the first pose by iterating between the optimal position and the attitude of the camera the algorithm can update the first pose as the initial value again by searching the minimum objective function and imposing an additional physical meaning constraint.

To test the rate of correct poses for the proposed algorithm in different attitude angles, we generated 1000 different models and poses for each attitude angle α from 0 to 80 degrees in the noncoplanar case. Different algorithms from [11], [17], [22] were used for comparison with the proposed method.

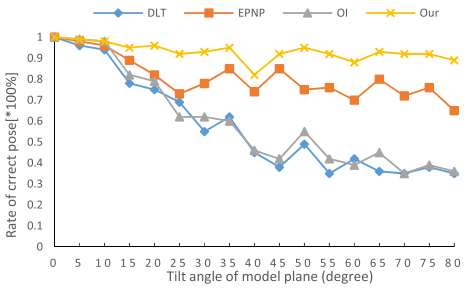


FIGURE 10. Correct probability of pose estimation at different tilt angles. (noise=2 pixels).

Fig. 10 shows the rates of correct poses for the four algorithms as one angle increases. In the process, we fixed Gaussian noise at 2 pixels. As we mentioned in Section 3, no local minimum existed in the case of small attitude angle; therefore, the rate of correct poses between 0 and 10 degrees for all four algorithms is above 94%. However, when the attitude angle continues to increase, other algorithms become much more sensitive to the angle. In particular, when the angle is greater than 40 degrees, the rate of correct poses of the DLT and OI algorithm decreases to less than 49%. The rate of our algorithm correctly calculating solutions stayed above 89% for almost attitude angles except for angles from 35 degrees to 45 degrees. Therefore, our algorithm has high stability at almost all angles, which are important for port shooting conditions.

B. REAL CRANE IMAGE TESTS AND RESULTS

1) TEST OBJECTS AND PREPARATIONS

To verify the feasibility of the pose estimation method based on geometry for port hoisting machinery, a real experiment was conducted.

As shown in Fig. 11, the object of the test was a gantry crane commonly used in a port's cargo storage yard in the port machinery testing laboratory of WHUT.

The test camera was a Canon EOS 5DS. The parameters of the test camera's charge-coupled device (CCD) are as shown in TABLE 1:

The camera calibration method in this experiment was provided by Heikkilä [30], which provided the MATLAB

FIGURE 11. Gantry crane in the port machinery testing laboratory of WHUT.

TABLE 1. Parameters of the camera.

Resolution	8688*5792
Pixel dimension	4.14μm

TABLE 2. The results of the calibration.

f	57.516
△x	0.2898
△y	-1.7181
k1	-0.1259
k2	-0.1309
q1	0
q2	0
s1	0
s2	0

Camera Calibrator toolbox. The results of the calibration are shown in Table 2:

f is the focal length of the camera, Δx and Δy are the coordinates of the principal point, k1 and k2 are the parameters of the radial distortion, q1 and q2 are parameters of the eccentric distortion, and s1 and s2 are parameters of the thin prism distortion.

To validate the estimated camera pose, we used the self-developed photogrammetric system for port hoisting machinery (PSPHM), which can record the corresponding camera pose of each photo. As shown in Figure 12, the PSPHM consists of a rotating platform, a bracket and a common camera. The rotating platform equipped with a position sensor, gyroscopes and dials obtains the camera pose when shooting.

2) TEST AND RESULT

The experiment was divided into two cases: noncoplanar and coplanar control points, and the number of control points was 4. The coordinates of control points were obtained by the total station before the test. In addition, the selection and

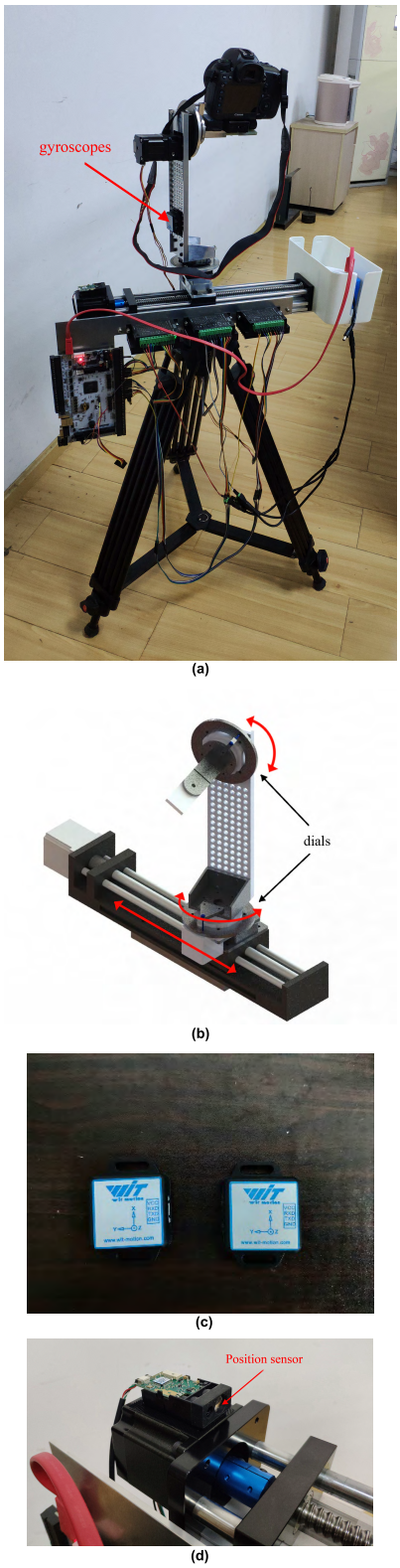


FIGURE 12. Photogrammetric system for port hoisting machinery. (a) Measurement system. (b) Rotating platform. (c) Gyroscope. (d) Position sensor.

numbering of control points are shown in Fig. 13. The number of coplanar selection points is 1,2,3,4, and the number of noncoplanar points is 1,3,5,6.

FIGURE 13. Numbering and selection of control points.

TABLE 3. The coordinates of control points.

Number	X(mm)	Y(mm)	Z(mm)
1	-1255	-530	20130
2	455	6970	20130
3	23015	-530	20130
4	24725	6970	20130
5	455	-1030	3050
6	455	8530	3050

TABLE 4. The true values of the camera's position and attitude.

X(mm)	Y(mm)	Z(mm)	α (degree)	β (degree)	γ (degree)
44630.3	-44985.8	1526.5	0.9	56.8	9.4

The origin of the measurement was the center of the left front leg of the crane on the ground. The coordinates of the control points 1, 2, 3, 4, 5, and 6 are shown in Table 3

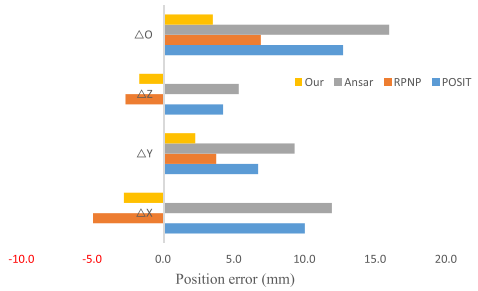
The true values of the camera's position and attitude were obtained through PSPHM, as shown in Table 4.

Let ΔX, ΔY and ΔZ be the difference between the measurements and actual values of the coordinates of camera's position in the object space and let $\Delta\alpha$, $\Delta\beta$ and $\Delta\gamma$ be the difference between the measurements and actual values of the coordinates of camera's attitude in the object space.

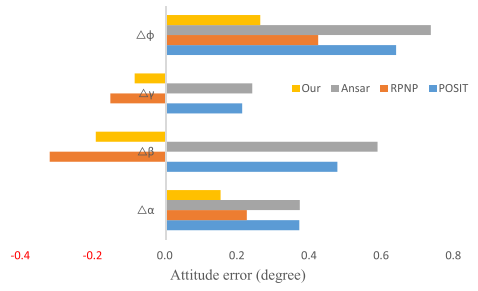
The relationship between ΔO and ΔX, ΔY and ΔZ and the relationship between $\Delta\varphi$ and $\Delta\alpha$, $\Delta\beta$ and $\Delta\gamma$ is as follows:

$$\Delta O = \sqrt{\Delta X^2 + \Delta Y^2 + \Delta Z^2},$$
$$\Delta \Phi = \sqrt{\Delta \alpha^2 + \Delta \beta^2 + \Delta \gamma^2}. \quad (26)$$

The experimental results of the coplanar and noncoplanar control points are shown in Fig. 14 and Fig. 15, respectively.

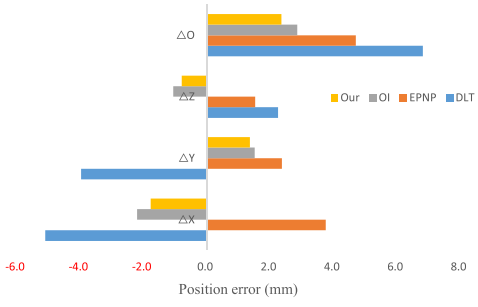


(a)



(b)

FIGURE 14. Test results under coplanar conditions. (a) Position error. (b) Attitude error.



(a)

(b)

FIGURE 15. Test results under noncoplanar conditions. (a) Position error. (b) Attitude error.

As illustrated in Fig. 14 and Fig. 15, our algorithm has a higher estimation accuracy than the other algorithms in the real image experiments. Therefore, the proposed algorithm can be used for camera pose estimation of port hoisting machinery in photogrammetry.

VI. CONCLUSIONS

In this paper, with consideration of multiple solutions, we described an accurate and stable pose estimation algorithm based on geometry for monitoring port hoisting machinery. According to the collinear error of the space control points in the image space coordinate, the error function with the minimum error was established by the projection operator. In the process of establishing the error function, a unit quaternion was used to form the rotation matrix representing the attitude of the camera to avoid the Euler angles expressed by complex trigonometric functions. In addition, the geometric characteristics of the attitude angle can help us to understand the relationship between the local minimum and the global optimal solution. We proposed a convenient and accurate solution for this problem. The central contribution of this paper is that our method does not rely on a proper initial value to converge to the true value and is suitable for large attitude angles, coplanar or noncoplanar control points and fewer control points, which are suitable for port shooting conditions.

The experimental results in Section 5 demonstrate our method's accuracy, stability and absence of severe pose jumps occurrence and its possession of stronger applicability and generality. Therefore, the paradox between the limitation of the port condition and the high-accuracy demand of camera pose estimation for port hoisting machinery has been addressed.

Although the algorithm mentioned in this paper is proposed for camera pose estimation for port hoisting machinery, it can still be applied in the measurement of large industrial equipment. Moreover, as a basic algorithm in photogrammetry and computational vision, the proposed algorithm can be applied to many relevant applications in AR [31] and UAV [32]. Thus, further research will be valuable.

REFERENCES

[1] L. B. Zhu, X. Wu, and H. Chen, "Study on special equipment safety risk assessment and control measures," (in Chinese), *China Saf. Sci. J.*, vol. 24, no. 1, pp. 149–155, Jan. 2014.
[2] S. Ruud and Å. Mikkelsen, "Risk-based rules for crane safety systems," *Reliab. Eng. Syst. Saf.*, vol. 93, no. 9, pp. 1369–1376, Sep. 2008.
[3] O. Ogundipe, G. W. Roberts, and C. J. Brown, "GPS monitoring of a steel box girder viaduct," *Struct. Infrastruct. Eng.*, vol. 10, no. 1, pp. 25–40, Jan. 2014.
[4] Y. Yu, Z. Zhao, and L. Chen, "Research and design of tower crane condition monitoring and fault diagnosis system," in *Proc. AICI*, Sanya, China, Oct. 2010, pp. 405–408.
[5] K. Kregar *et al.*, "Control measurements of crane rails performed by terrestrial laser scanning," *Sensors*, vol. 17, no. 7, p. 1671, Jul. 2017.
[6] D. González-Aguilera, J. Gómez-Lahoz, and J. Sánchez, "A new approach for structural monitoring of large dams with a three-dimensional laser scanner," *Sensors*, vol. 8, no. 9, pp. 5866–5883, Sep. 2008.
[7] S. D. Cheng, "Application of total station in deformation monitoring of crane tower," (in Chinese), *Eng. Technol.*, vol. 8, no. 2, pp. 217–218, 2016.
[8] C. Feng, V. R. Kamat, and H. Cai, "Camera marker networks for articulated machine pose estimation," *Autom. Construct.*, vol. 96, no. 12, pp. 148–160, Sep. 2018.
[9] L. Licev, J. Hendrych, J. Tomecek, R. Cajka, and M. Krejsa, "Monitoring of excessive deformation of steel structure extra-high voltage pylons," *Periodica Polytechn. Civil Eng.*, vol. 62, no. 2, pp. 323–329, Nov. 2018.
[10] A. Li and Z. Zhao, "Crane safety assessment method based on entropy and cumulative prospect theory," *Entropy*, vol. 19, no. 1, p. 44, Jan. 2017.

[11] Y. Abdel-Aziz, H. Karara, and M. Hauck, "Direct linear transformation from comparator coordinates into object space coordinates in close-range photogrammetry," *Photogramm. Eng. Remote Sens.*, vol. 81, no. 2, pp. 103–107, 2015.
[12] B. Přibyl, P. Zemčík, and M. Cadík, "Absolute pose estimation from line correspondences using direct linear transformation," *Comput. Vis. Image Understand.*, vol. 161, no. 11, pp. 130–144, May 2017.
[13] A. Ansar and K. Daniilidis, "Linear pose estimation from points or lines," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 25, no. 5, pp. 578–589, May 2003.
[14] A. Banno, "A P3P problem solver representing all parameters as a linear combination," *Image Vis. Comput.*, vol. 70, no. 2, pp. 55–62, Feb. 2018.
[15] Y. Q. Zheng, S. Sugimoto, and M. Okutomit, "ASPnP: An accurate and scalable solution to the perspective-n-point problem," *IEICE Trans. Inf. Syst.*, vol. E96D, no. 7, pp. 1525–1535, Jul. 2013.
[16] Y. Q. Zheng, Y. B. Kuang, S. Sugimoto, K. Astrom, and M. Okutomi, "Revisiting the PnP problem: A fast, general and optimal solution," in *Proc. ICCV*, Sydney, NSW, Australia, 2013, pp. 2345–2351.
[17] V. Lepetit, F. M. Noguer, and P. Fua, "EP*n*P: An accurate O(*n*) solution to the P*n*P problem," *Int. J. Comput. Vis.*, vol. 81, no. 2, pp. 155–166, Feb. 2009.
[18] D. Oberkampf, F. Daniel, D. Menthon, and L. S. Davis, "Iterative pose estimation using coplanar feature points," *Comput. Vis. Image Understand.*, vol. 63, no. 3, pp. 495–511, May 1996.
[19] D. F. Dementhon and L. S. Davis, "Model-based object pose in 25 lines of code," *Int. J. Comput. Vis.*, vol. 15, no. 1, pp. 123–141, Jun. 1995.
[20] T. Gramegna, L. Venturino, G. Cicirelli, G. Attolico, and A. Distante, "Optimization of the POSIT algorithm for indoor autonomous navigation," *Robot. Auton. Syst.*, vol. 48, nos. 2–3, pp. 145–162, Sep. 2004.
[21] V. Garro, F. Crosilla, and A. Fusiello, "Solving the PnP problem with anisotropic orthogonal procrustes analysis," in *Proc. 3DIMPVT*, Zurich, Switzerland, Oct. 2012, pp. 262–269.
[22] C.-P. Lu, G. D. Hager, and E. Mjolsness, "Fast and globally convergent pose estimation from video images," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 22, no. 6, pp. 610–622, Jun. 2000.
[23] G. Schweighofer and A. Pinz, "Robust pose estimation from a planar target," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 28, no. 12, pp. 2024–2030, Dec. 2006.
[24] P. C. Wu, H. Y. Tseng, M. H. Yang, and S. Y. Chien, "Direct pose estimation for planar objects," *Comput. Vis. Image Understand.*, vol. 172, no. 7, pp. 50–66, Mar. 2018.
[25] K. Fathian, J. Jin, S. G. Wee, D. H. Lee, Y. G. Kim, and N. R. Gans, "Camera relative pose estimation for visual servoing using quaternions," *Robot. Auton. Syst.*, vol. 107, no. 9, pp. 45–62, Jun. 2018.
[26] A. Janota, V. Šimák, D. Nemec, and J. Hrbček, "Improving the precision and speed of Euler angles computation from low-cost rotation sensor data," *Sensors*, vol. 15, no. 3, pp. 7016–7039, Mar. 2015.
[27] X.-S. Gao, X.-R. Hou, J. Tang, and H.-F. Cheng, "Complete solution classification for the perspective-three-point problem," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 25, no. 8, pp. 930–943, Aug. 2003.
[28] M. Vynnycky and K. Kanev, "Mathematical analysis of the multisolution phenomenon in the P3P problem," *J. Math. Imag. Vis.*, vol. 51, no. 2, pp. 326–337, Feb. 2015.
[29] M. Zettler and J. Garloff, "Robustness analysis of polynomials with polynomial parameter dependency using Bernstein expansion," *IEEE Trans. Autom. Control*, vol. 43, no. 3, pp. 425–431, Mar. 1998.
[30] J. Heikkila, "Geometric camera calibration using circular control points," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 22, no. 10, pp. 1066–1077, Oct. 2000.
[31] S. Q. Li and C. Xu, "Efficient lookup table based camera pose estimation for augmented reality," *Comput. Animation Virtual Worlds*, vol. 22, no. 1, pp. 47–58, Jan. 2011.
[32] M. Shahbazi, G. Sohn, J. Théau, and P. Menard, "Development and evaluation of a UAV-photogrammetry system for precise 3D environmental modeling," *Sensors*, vol. 15, no. 11, pp. 27493–27524, Oct. 2015.

QI WANG received the B.S. degree in industrial design from Jianghan University, Wuhan, China, and the M.S. degree in vehicle operation engineering from East China Jiaotong University, Nanchang, China. He is currently pursuing the Ph.D. degree with the Wuhan University of Technology, Wuhan. His research is based within the School of Logistics Engineering under the supervision of Prof. Zhao.

He is currently a Research Assistant with the School of Logistics Engineering and a Lecturer with the Wuhan Railway Vocational College of Technology. His research interests include computer vision measurement, and close range photogrammetry and measurement of large engineering equipment.

ZHANGYAN ZHAO received the B.S. degree in mechanical engineering from Southwest Jiao Tong University, Chengdu, China, and the M.S. degree and the Ph.D. degree in mechanical engineering from the Wuhan University of Technology, Wuhan, China, in 2001.

He is currently a Professor and a Doctoral Supervisor with the School of Logistics Engineering, Wuhan University of Technology. He has authored three books and published over 100 prestigious journal papers and conference papers. He holds six patents. His research interests include metal structure fault diagnosis and safety evaluation, computer vision measurement, and computer-aided engineering and manufacturing.