Metrol. Meas. Syst., Vol. 30 (2023) No. 2, pp. 235–258
DOI: 10.24425/mms.2023.144871

PAN POLSKA AKADEMIA NAUK

METROLOGY AND MEASUREMENT SYSTEMS
Index 330930, ISSN 0860-8229
www.metrology.wat.edu.pl

M&MS

AN ACCURATE AND STABLE POSE ESTIMATION METHOD FOR PLANAR CASES CONSIDERING THE LINE CONSTRAINTS BETWEEN EVERY TWO POINTS

Zhang Zimiao[1)], Zhang Hao[1)], Zhang Fumin[2)], Zhang Shihai[1)]

1) School of Mechanical Engineering, Tianjin University of Technology and Education, Tianjin, China
2) State Key Laboratory of Precision Measurement Technology and Instruments, Tianjin University, Tianjin, China
(✉ *zhangzimiao1985@foxmail.com, zzm19850126@aliyun.com, zhanghaoaa@163.com, zhangfumin@tju.edu.cn, zshky77@163.com.edu.cn*)

Abstract

The current solutions for pose estimation problems using coplanar feature points (PnP problems) can be divided into non-iterative and iterative solutions. The accuracy, stability, and efficiency of iterative methods are unsatisfactory. Therefore, non-iterative methods have become more popular. However, the non-iterative methods only consider the correspondence of the feature points with their 2D projections. They ignore the constraints formed between feature points. This results in lower pose estimation accuracy and stability. In this work, we proposed an accurate and stable pose estimation method considering the line constraints between every two feature points. Our method has two steps. In the first step, we solved the pose non-iteratively, considering the correspondence of the 3D feature points with their 2D projections and the line constraints formed by every two feature points. In the second step, the pose was refined by minimizing the re-projection errors with one iteration, further improving accuracy and stability. Simulation and actual experiment results show that our method's accuracy, stability, and computational efficiency are better than the other existing pose estimation methods. In the −45° to +45° measuring range, the maximum angle measurement error is no more than 0.039°, and the average angle measurement error is no more than 0.016°. In the 0 mm to 30 mm measuring range, the maximum displacement measurement error is no more than 0.049 mm, and the average displacement measurement error is no more than 0.012 mm. Compared to other current pose estimation methods, our method is the most efficient based on guaranteeing measurement accuracy and stability.

Keywords: pose estimation, line constraints, coplanar, non-iterative.

1. Introduction

The absolute pose estimation based on a set of feature points is a classical vision measurement problem, also known as the Perspective-n-Point (PnP) problem [1]. Given n feature points and their 2D projections, a target object's pose (position and orientation) relative to a calibrated camera

Article history: received October 29, 2022; revised February 4, 2023; accepted February 6, 2023; available online June 18, 2023.

is solved. The quality of the camera is crucial to pose estimation accuracy [2, 3], and the camera's intrinsic parameters should be calibrated using the Zhang method [4] with a sufficient number of target images in various poses. The pose estimation has numerous applications, including the hand-eye calibration of industrial robots [5], Lidar-camera systems [6], medical surgery [7], visual *Simultaneous Localization and Mapping* (SLAM) [8], etc.

Estimating pose through non-coplanar feature points is mature [9, 10]. However, pose estimation using coplanar feature points is still a challenge. The current solutions for pose estimation using coplanar feature points can be divided into non-iterative and iterative solutions. The iterative methods usually establish a nonlinear objective function to minimize the image-space error (reprojection error) or object-space error. The *Pose from Orthography and Scaling with Iterations* (POSIT) method is the typical one to minimize the image-space error [11, 12]. The improved POSIT method for planar cases was proposed by Oberkampf *et al.* [13]. The object-space error is the error of the feature point and its orthogonal projection on the corresponding perspective projection line. The *Lu, Hager and Mjolsness* (LHM) method is the typical one to minimize the object-space error [14, 15]. Sun *et al.* made some improvements to the LHM method. They used the LHM method based on a perspective-ray-based camera model [16, 17].

The choice of the initial value impacts the iterative process. If the initial value selection is inappropriate, the number of iterations will increase. The entire iterative process will take a long time. In addition, the iterative method does not get a closed-form solution. Therefore, there may be some local minima for the objective function, and the ambiguity problem of the pose will occur [18]. Besides, the choice of the two possible error functions is also a problem.

The accuracy, stability, and efficiency of iterative methods are unsatisfactory. Therefore, non-iterative methods have become more prevalent in recent years. The non-iterative methods were initially derived from Zhang's camera calibration method [4], the *Direct Linear Transformation* (DLT) method. It is based on the homography between the coplanar feature points and corresponding image points. Lepetit *et al*. [19] proposed the *Efficient Perspective-n-Point* (EPnP) method. In the solving process, a linearization strategy was adopted. Hesch *et al.* [20] proposed the *Direct Least-Squares* (DLS) method. The pose of the target object included a rotation matrix and a translation vector. The Cayley parameters were used to express the rotation matrix, which was then used to represent the translation vector. The entire process of the DLS method requires lots of matrix decomposition and transformation. Li *et al.* [21] proposed the *Robust Perspective-n-Point* (RPnP) method. A new coordinate frame was constructed through the two feature points with the longest 2D projection. Every three feature points and their 2D projections could form a fourth-order polynomial according to the *Perspective-3-Point* (P3P) algorithm. The pose was finally obtained by solving these polynomial equations. Wang *et al.* [22] made some improvements to this method. Constructing the new coordinate frame using the two feature points with the longest 2D projection is not necessary because they add one iteration. Zheng *et al.* proposed the *Accurate and Scalable Perspective-n-Point* (ASPnP) [23] and *O(n) Solution Perspective-n-Point* (OPnP) [24] methods. They used the quaternions to represent the rotation matrix and then parameterized the translation vector by the rotation matrix to construct the polynomial equations. At last, the polynomial equations were solved with the Gröbner basis method. Kneip *et al.* [25] proposed the *Unified Perspective-n-Point* (UPnP) method which also solved the polynomial equations with the Gröbner basis method. The number n of features should be greater than or equal to 4. The stability of the non-iterative methods can be enhanced by introducing redundant points as additional information [26].

Although the non-iterative pose estimation methods have good performance, they are based on the imaging constraint provided by each single feature point, that is, the correspondence of each feature point with its 2D projection. Due to various noises in the measurement, the constraint

Metrol. Meas. Syst., Vol. 30 (2023), No. 2, pp. 235–258
DOI: 10.24425/mms.2023.144871

provided by each single feature point is insufficient. Thus, the constraints formed between the feature points should be considered to improve the measurement accuracy and stability of non-iterative pose estimation methods. This has been achieved in some iterative methods [27, 28]. As to the non-iterative pose estimation methods, it still needs to be solved.

In addition, the following conditions should also be considered. The pose estimation accuracy depends on the shape of the array itself [29]. Huo *et al.* [30, 31] proposed that trapezoidal arrays should be applied as they are better than square arrays. Pose estimation accuracy also depends on the distance of the camera from the array. Qu *et al.* [32] suggested that the pose estimation error is proportional to the ratio of the distance from the camera to the array and the target size. They also suggested that the circular feature points should be as small as possible [33].

Based on the discussion above, we proposed an accurate and stable pose estimation method for the planar cases. We constructed the pose estimation model considering the correspondences of the 3D feature points with their 2D projections and the line constraints formed by every two feature points. The pose could be obtained non-iteratively. Because all the constraints are fully considered when getting the pose, only one iteration is needed to refine the pose by minimizing the re-projection errors. We compared and analysed our method with some typical pose estimation methods through the simulation experiments with synthetic data and measurement experiments with the real images. In the experiments with real images, we compared and analysed the influence of the shape of the target array and the distance from the camera to the target on our pose estimation method.

The motivation of our study is to provide an accurate and stable pose estimation method using coplanar feature points. The main contributions are listed as follows:

- A novel non-iterative pose estimation model is established considering the imaging constraint provided by each feature point and the line constraints formed by every two feature points.
- Only one iteration is needed to refine the solved pose because we fully consider all the constraints in the non-iterative step.
- The measurement accuracy and stability are improved compared to other non-iterative methods. Based on the non-iterative step's efficiency, adding one iteration (unlike other iterative methods, we reduce the number of iterations to only one) still guarantees the solution's efficiency.

The remainder of our paper is arranged as follows. We presented non-iterative and one-iteration steps of our method in the corresponding sections. In Section 4, we compared and analysed the experimental results of our method and other representative methods. Some conclusions were made in Section 5.

2. Non-iterative step of the pose estimation method

Fig. 1 shows a set of 3D-to-2D correspondences between n coplanar feature points M_i and their 2D projections m_i. The world coordinates of M_i are $\mathbf{M}_i^w(x_i^w, y_i^w, z_i^w)^{\mathrm{T}}$. The pixel coordinates of each 2D projection are $\mathbf{m}_i(u_i, v_i)^{\mathrm{T}}$, which can be obtained at the image processing stage. The pose estimation problem is to solve the transformation matrix $[\mathbf{R}\,\mathbf{T}]$ from the world coordinate frame to the camera coordinate frame. $\mathbf{R}$ is a 3×3 matrix. $\mathbf{T}$ is a 3×1 vector.

Because the feature points are coplanar, z_i^w is equal to 0. The perspective projection equation can be expressed as Eq. (1), where $\mathbf{A}$ is the camera intrinsic parameters matrix and d_i is the depth

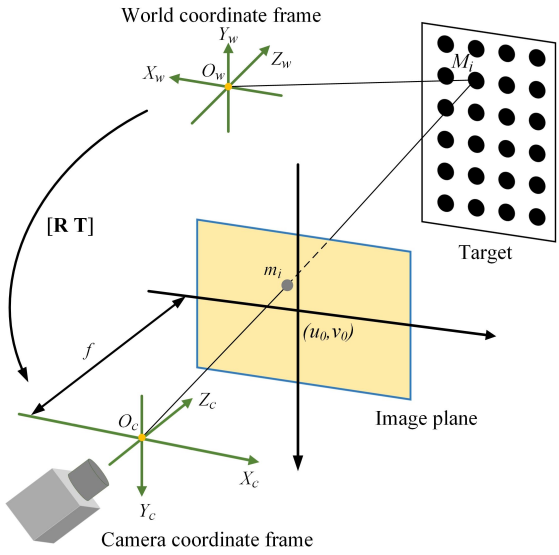


Fig. 1. Correspondence between feature point M_i and its 2D projection m_i.

of each feature point.

$$d_i \cdot \begin{bmatrix} u_i \\ v_i \\ 1 \end{bmatrix} = \begin{bmatrix} f_x & 0 & u_0 \\ 0 & f_y & v_0 \\ 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} \mathbf{R} & \mathbf{T} \end{bmatrix} \begin{bmatrix} x_i^w \\ y_i^w \\ 0 \\ 1 \end{bmatrix} = \mathbf{A} \cdot \begin{bmatrix} \mathbf{R} & \mathbf{T} \end{bmatrix} \begin{bmatrix} x_i^w \\ y_i^w \\ 0 \\ 1 \end{bmatrix},$$
$$\mathbf{R} = \begin{bmatrix} r_{11} & r_{12} & r_{13} \\ r_{21} & r_{22} & r_{23} \\ r_{31} & r_{32} & r_{33} \end{bmatrix}, \qquad \mathbf{T} = \begin{bmatrix} t_1 \\ t_2 \\ t_3 \end{bmatrix}. \tag{1}$$

When we multiply both sides of Eq. (1) by $\mathbf{A}^{-1}$, divide both sides of Eq. (1) by t_3, and eliminate the depth factor d_i, we obtain Eq. (2), where $\mathbf{H} = [f_x \cdot x_i^w, 0]^{\mathrm{T}}$, $\mathbf{I} = [f_x \cdot y_i^w, 0]^{\mathrm{T}}$, $\mathbf{J} = [0, f_y \cdot x_i^w]^{\mathrm{T}}$, $\mathbf{K} = [0, f_x \cdot y_i^w]^{\mathrm{T}}$, $\mathbf{L} = [(u_0 - u_i) \cdot x_i^w, (v_0 - v_i) \cdot x_i^w]^{\mathrm{T}}$, $\mathbf{M} = [(u_0 - u_i) \cdot y_i^w, (v_0 - v_i) \cdot y_i^w]^{\mathrm{T}}$, $\mathbf{N} = [f_x, 0]^T$, $\mathbf{O} = [0, f_y]^{\mathrm{T}}$, $\mathbf{P} = [u_0 - u_i, v_0 - v_i]^T$.

$$\begin{bmatrix} \mathbf{H} \, \mathbf{I} \, \mathbf{J} \, \mathbf{K} \, \mathbf{L} \, \mathbf{M} \, \mathbf{N} \, \mathbf{O} \, \mathbf{P} \end{bmatrix} [r_{11}/t_3, r_{12}/t_3, r_{21}/t_3, r_{22}/t_3, r_{31}/t_3, r_{32}/t_3, t_1/t_3, t_2/t_3, 1]^{\mathrm{T}} = 0 \tag{2}$$

By solving Eq. (2) and based on the unit orthogonality of $\mathbf{R}$, the pose $[\mathbf{R}\ \mathbf{T}]$ can be solved. However, more than the correspondence of each feature point with its 2D projection is required for the pose estimation. In some iterative methods for calibrating the parameters of the binocular vision system, the length or angle constraints formed between the calibration points are added to the iterative function to ensure calibration accuracy and robustness [27, 28]. The non-iterative methods are better than iterative ones for pose estimation, as illustrated above. However, for non-iterative methods, the constraints formed between the feature points should also be considered to strengthen the anti-noise ability.

As shown in Fig. 2, M_i and M_j are any two feature points that form a line M_iM_j. m_i and m_j are the corresponding image points. Line M_iM_j and optical centre O_c can form a plane $M_iO_cM_j$. The expression of the normal vector of plane $M_iO_cM_j$ is $\mathbf{N}_{ij}^w$ in the world coordinate frame. Line

Metrol. Meas. Syst., Vol. 30 (2023), No. 2, pp. 235–258
DOI: 10.24425/mms.2023.144871

$m_i m_j$ and O_c can form a plane $m_i O_c m_j$. The expression of the normal vector of plane $m_i O_c m_j$ is $\mathbf{n}_{ij}^c$ in the camera coordinate frame.

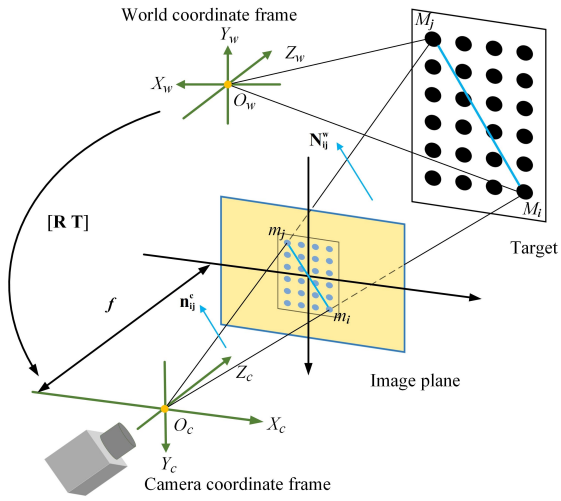


Fig. 2. Line constraint formed by any two feature points M_i and M_j.

The relationship of $\mathbf{N}_{ij}^w$ and $\mathbf{n}_{ij}^c$ is expressed by Eq. (3), where $[\mathbf{R}'\ \mathbf{T}']$ is the transformation matrix from the camera coordinate frame to the world coordinate frame, and μ_{ij} is a scale factor.

$$\mu_{ij} \cdot \mathbf{R}' \cdot \mathbf{n}_{ij}^c = \mathbf{N}_{ij}^w \tag{3}$$

$\mathbf{n}_{ij}^c$ can be expressed by $\mathbf{O}_c^c\mathbf{m}_i^c \times \mathbf{O}_c^c\mathbf{m}_j^c$. $\mathbf{N}_{ij}^w$ can be expressed by $\mathbf{O}_c^w\mathbf{M}_i^w \times \mathbf{O}_c^w\mathbf{M}_j^w$. $\mathbf{O}_c^w$ is the expression of O_c in the world coordinate frame. $\mathbf{O}_c^c$ is the expression of O_c in the camera coordinate frame. Then we can get Eq. (4).

$$\mu_{ij} \cdot \mathbf{R}' \cdot ((\mathbf{m}_i^c - \mathbf{O}_c^c) \times (\mathbf{m}_j^c - \mathbf{O}_c^c)) = (\mathbf{M}_i^w - \mathbf{O}_c^w) \times (\mathbf{M}_j^w - \mathbf{O}_c^w). \tag{4}$$

$\mathbf{O}_c^w$ is equal to $\mathbf{T}'$ and replace $\mathbf{O}_c^w$ with $\mathbf{T}'$. Eq. (5) can be obtained.

$$\mu_{ij} \cdot \mathbf{R}' \cdot (\mathbf{m}_i^c \times \mathbf{m}_j^c) = (\mathbf{M}_i^w - \mathbf{T}') \times (\mathbf{M}_j^w - \mathbf{T}'). \tag{5}$$

According to the cross-product rule, Eq. (5) is expanded, and similar items are combined. Then Eq. (6) can be deduced.

$$\mu_{ij} \cdot \mathbf{R}' \cdot (\mathbf{m}_i^c \times \mathbf{m}_j^c) = \mathbf{M}_i^w \times \mathbf{M}_j^w - \mathbf{T}' \times (\mathbf{M}_j^w - \mathbf{M}_i^w). \tag{6}$$

$\mathbf{T}'$ can be denoted with $-\mathbf{R}' \cdot \mathbf{T}$. Then Eq. (7) can be obtained.

$$\mu_{ij} \cdot \mathbf{R}' \cdot (\mathbf{m}_i^c \times \mathbf{m}_j^c) = \mathbf{M}_i^w \times \mathbf{M}_j^w - (-\mathbf{R}' \cdot \mathbf{T}) \times (\mathbf{M}_j^w - \mathbf{M}_i^w). \tag{7}$$

$\mathbf{R}$ and $\mathbf{R}'$ are mutually inverse matrices. When we multiply both sides of Eq. (7) by $\mathbf{R}$, Eq. (8) can be obtained, where $\mathbf{m}_i^c \times \mathbf{m}_j^c$, $\mathbf{M}_i^w \times \mathbf{M}_j^w$, and $\mathbf{M}_j^w - \mathbf{M}_i^w$ are known quantities. $\mathbf{m}_i^c$ and $\mathbf{m}_j^c$ can be obtained through pixel image coordinates, and the intrinsic parameters matrix $\mathbf{A}$. Eq. (8) is the expression of the line constraint formed by every two feature points, M_i and M_j.

$$\mu_{ij} \cdot (\mathbf{m}_i^c \times \mathbf{m}_j^c) = \mathbf{R} \cdot (\mathbf{M}_i^w \times \mathbf{M}_j^w) - (\mathbf{R} \cdot (\mathbf{M}_j^w - \mathbf{M}_i^w)) \times \mathbf{T}. \tag{8}$$

Next. we eliminate the scale factor μ_{ij} and divide both sides of Eq. (8) by t_3. There are 20 unknown quantities in total, according to Eq. (2) and Eq. (8). Given n feature points, according to Eq. (2), we can obtain $2n$ equations. By n feature points, $n(n-1)/2$ lines can be obtained. According to Eq. (8), we can get $n(n-1)$ equations. Therefore, for 20 unknown quantities, n should be greater than or equal to 4. These equations are solved simultaneously. According to the unit orthogonality of $\mathbf{R}$, the pose $[\mathbf{R}\ \mathbf{T}]$ can finally be obtained using the *Single Value Decomposition* (SVD) method. In the above-solving process, we fully consider all the constraints, including the 3D-to-2D correspondence of each feature point and the line constraints formed by every two feature points.

3. One iteration step of the pose estimation method

One iteration step can further improve the solution's accuracy and stability. We construct an objective function to minimize the re-projection errors of the n feature points, which is shown in Eq. (9).

$$re = \sum_{i=1}^{n} \left\| \mathbf{m}_i - \mathbf{m}_i(\mathbf{R}, \mathbf{T}) \right\|^2. \tag{9}$$

The rotation matrix $\mathbf{R}$ can be parameterized with the Cayley parameters (s_1, s_2, s_3) according to Eq. (10), where $\xi = 1 + s_1^2 + s_2^2 + s_3^2$.

$$\mathbf{R} = \frac{1}{\xi} \begin{bmatrix} 1 + s_1^2 - s_2^2 - s_3^2 & 2s_1s_2 - 2s_3 & 2s_1s_3 - 2s_2 \\ 2s_1s_2 + 2s_3 & 1 - s_1^2 + s_2^2 - s_3^2 & 2s_2s_3 - 2s_1 \\ 2s_1s_3 + 2s_2 & 2s_2s_3 + 2s_1 & 1 - s_1^2 - s_2^2 + s_3^2 \end{bmatrix}. \tag{10}$$

We define the matrix $\mathbf{D}$ and $\mathbf{E}$, as shown in Eq. (11).

$$\begin{aligned} \mathbf{D} &= \begin{bmatrix} -f_x & 0 & u_i - u_0 \\ 0 & -f_y & v_i - v_0 \end{bmatrix}, \\ \mathbf{E} &= \begin{bmatrix} f_x x_i^w & 2y_i^w(u_0 - u_i) & 2x_i^w(u_0 - u_i) & -2f_x y_i^w & f_x x_i^w \\ f_y y_i^w & 2y_i^w(v_0 - v_i) & 2x_i^w(v_0 - v_i) & 2f_y x_i^w & -f_y y_i^w \\ 2f_x y_i^w & 2x_i^w(u_0 - u_i) & -f_x x_i^w & 2y_i^w(u_0 - u_i) & -f_x x_i^w \\ 2f_x x_i^w & 2x_i^w(v_0 - v_i) & f_y y_i^w & 2y_i^w(v_0 - v_i) & -f_y y_i^w \end{bmatrix}. \end{aligned} \tag{11}$$

According to Eq. (1) and Eq. (11), we can deduce Eq. (12).

$$\begin{aligned} \xi \cdot \mathbf{D} \cdot \mathbf{T} &= \mathbf{E} \cdot \mathbf{CA}, \\ \mathbf{CA} &= \left[1, s_1, s_2, s_3, s_1^2, s_1s_2, s_1s_3, s_2^2, s_2s_3, s_3^2\right]^{\mathrm{T}}. \end{aligned} \tag{12}$$

From Eq. (12), we can easily express the vector $\mathbf{T}$ with vector $\mathbf{CA}$ and ξ, which are composed of the Caley parameters, as shown in Eq. (13).

$$\mathbf{T} = \frac{1}{\xi} \left(\mathbf{D}^{\mathrm{T}} \mathbf{D} \right)^{-1} \mathbf{D}^{\mathrm{T}} \mathbf{E} \cdot \mathbf{CA}. \tag{13}$$

The objective function is finally shown in Eq. (14) and solved with the Levenberg–Marquardt method.

$$re = \min \sum_{i=1}^{n} \left\| \mathbf{m}_i - \mathbf{m}_i(\mathbf{R}(s_1, s_2, s_3),\ \mathbf{T}(s_1, s_2, s_3)) \right\|^2. \tag{14}$$

Metrol. Meas. Syst., Vol. 30 (2023), No. 2, pp. 235–258
DOI: 10.24425/mms.2023.144871

We fully consider the constraints in the non-iterative step. Then a good enough pose can be obtained. Therefore, the number of iterations is only one. Fig. 3 depicts the complete procedure for estimating the pose [**R T**]. The line constraints are formed by every two feature points, including all the lines formed between the feature points. Taking 4 or 6 feature points as an example, as shown in Fig. 4, when the number of feature points is 4, the lines constraints are provided by M_1M_2, M_1M_3, M_1M_4, M_2M_3, M_2M_4, and M_3M_4. When the number of feature points is 6, the line constraints are provided by M_1M_2, M_1M_3, M_1M_4, M_1M_5, M_1M_6, M_2M_3, M_2M_4, $M_2M_5M_2M_6$, M_3M_4, M_3M_5, M_3M_6, M_4M_5, M_4M_6, and M_5M_6. Considering the line constraints formed between any two feature points in the non-iterative step of the method, the anti-noise ability of our method is strengthened. With one-step iteration, the anti-noise ability is further strengthened.

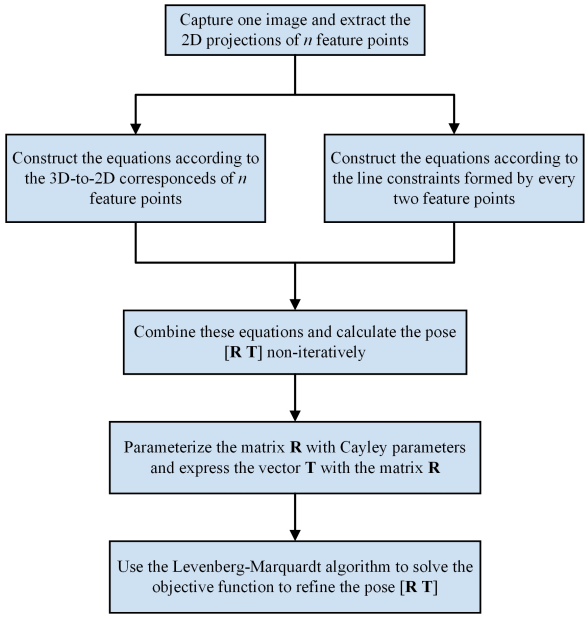


Fig. 3. Full procedure for estimating the pose [**R T**].

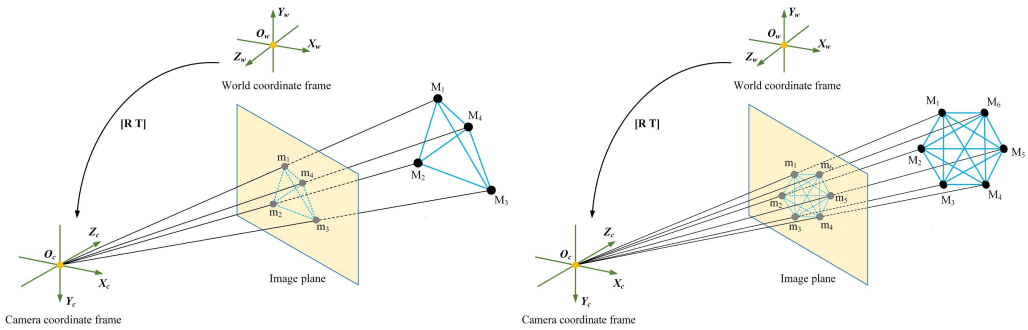


Fig. 4. Line constraints formed between feature points: (a) $n = 4$. (b) $n = 6$.

4. Results and discussion

The setup used in our experiments is shown in Fig. 5. The camera's resolution with a 12 mm lens is 1440 × 1080 pixels. The target with circle feature points is 27 × 27. The diameter of each circle is 5 mm. The distance between every two circle centres is 10 mm. A one-dimensional turntable with a positioning precision of ±0.005° and a one-dimensional linear translation stage with a positioning precision of ±0.003 mm are combined to alter the target pose. The CPU of the experimental computer is an i5–7200U. All the programs run on the MATLAB platform.

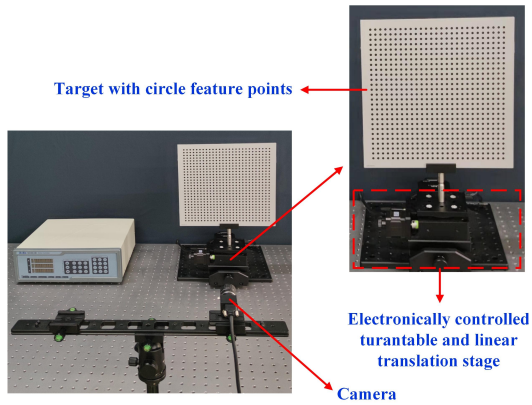


Fig. 5. Experimental setup.

We calibrated the camera's intrinsic parameters using the Zhang method [4] with 16 images of different poses. They are shown in Table 1. The calibration re-projection error is 0.05 pixels. All the pose estimation experiments are conducted with the same camera to ensure the consistency of verification of the measurement method.

Table 1. Calibrated intrinsic parameters

size	f_x	f_y	c_x	c_y	k_1	k_2	k_3	p_1	p_2
1440 × 1080	3417.741	3417.007	733.604	548.936	–0.082	0.151	6.026	0.000799	0.000540

To accurately evaluate the performance of our method, it was compared with some typical pose estimation methods during the experiments. These typical pose estimation methods are shown below:

- only use the non-iterative step of our method to solve the pose.
- the DLT method [4]: The traditional linear pose estimation method. It only considers the correspondence of each feature point with its 2D projection.
- the *Perspective-Ray-Based* (PRB)+LHM method [16, 17]: It is the LHM method [14] with the perspective-ray-based camera model. It is one of the iterative methods with the best accuracy.
- the OPnP method [24]: A robust and accurate pose estimation method. It is a typical method for solving the pose estimation problem using the Gröbner basis method.
- the Improved RPnP method [22]: A fast, robust, and non-iterative pose estimation method. Compared with the RPnP method [21], it constructs the target coordinate frame with any two feature points. It works well in all cases.

Metrol. Meas. Syst., Vol. 30 (2023), No. 2, pp. 235–258
DOI: 10.24425/mms.2023.144871

This study includes two categories of experiments: simulation experiments with synthetic data and measurement experiments with real images.

4.1. Simulation experiments with synthetic data

This part evaluates our pose estimation method through simulation experiments. First, we constructed the camera coordinate frame. Next, we synthesized a virtual perspective camera with a resolution of 1440×1080 pixels and an effective focal length of 1800 pixels. Then, n feature points M_i were randomly generated in the world coordinate frame. Finally, the feature points were projected to the image plane according to the virtual perspective camera and Gaussian noise could be added to the image plane. The pose $[\mathbf{RT}]$ was calculated with our and the other 5 methods. The measurement error of $\mathbf{R}$ was $\max(|\alpha - \alpha_{\text{true}}|, |\beta - \beta_{\text{true}}|, |\gamma - \gamma_{\text{true}}|)$. α, β, and γ were the Euler angles representing the rotation. The measurement error of $\mathbf{T}$ was $\|\mathbf{T} - \mathbf{T}_{\text{true}}\|/\|\mathbf{T}\|$.

We set the Gaussian noise level to 1 pixel and 2.5 pixels, respectively. Then, we changed the number of feature points. The measurement errors of $\mathbf{R}$ and $\mathbf{T}$ are shown in Fig. 6. It can be seen that with the increasing number of feature points, the measurement errors decrease for all the methods. This indicates that redundant feature points (n>4) can improve pose estimation accuracy.

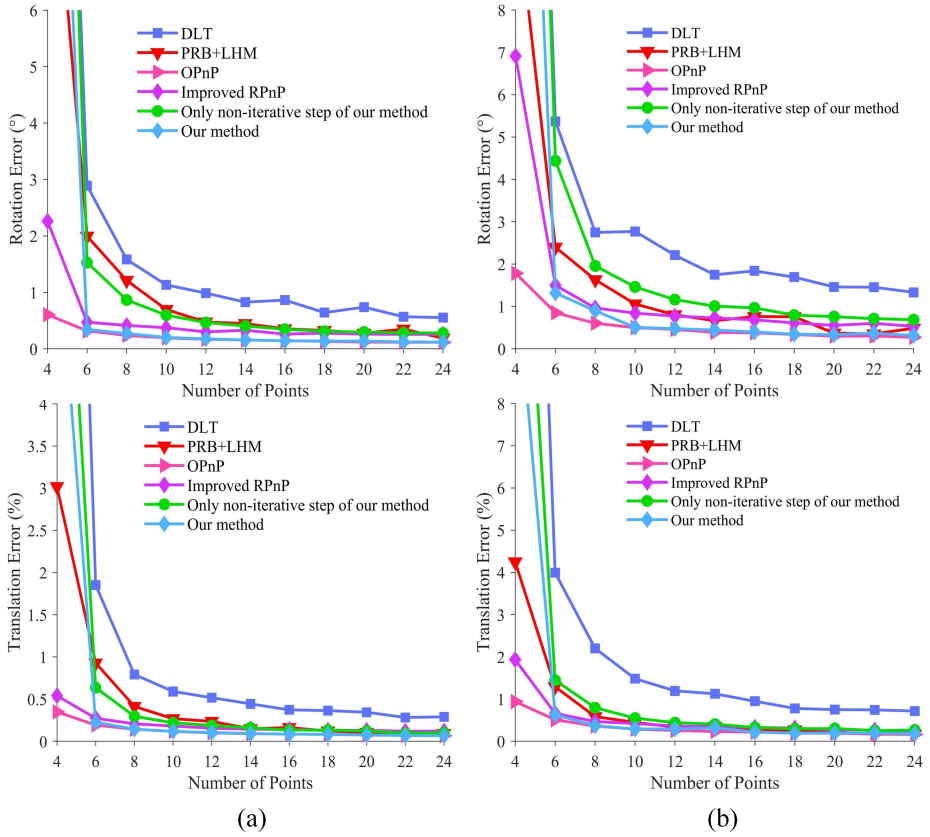


Fig. 6. Measurement errors of pose estimation for different numbers of feature points: (a) Noise level of 1 pixel, (b) Noise level of 2.5 pixels.

Then, we set the number of feature points to 12, 16, and 20, respectively, and observed the change in the measurement error as the noise level changed. Gaussian noise of 0 to 5 pixels with an interval of 0.5 pixels was added to the image plane. The measurement errors of **R** and **T** are shown in Fig. 7. It can be seen that with the increase in noise level, the solution accuracy of each method decreases. An approximately linear relationship exists between the measurement errors and the noise level. The PRB+LHM method is iterative. The OPnP method transforms the pose estimation problem into a nonlinear optimization problem. Therefore, their anti-noise ability is better. However, their solving process is complicated. Under the condition of redundant feature points, the disadvantage of solving efficiency is more prominent. Compared to the RPnP method, the improved RPnP method can establish the target coordinate system with any two feature points while still ensuring the pose estimation accuracy. The anti-noise ability of the DLT method is worse because it is a non-iterative method and considers less constraint information (only considering the correspondences of each feature point with its 2D projection). The non-iterative part of our method considers the line constraints formed by every two feature points. The anti-noise ability is strengthened. Adding only one iteration, the anti-noise ability of our method is further strengthened and almost equivalent to that of PRB+LHM and OPnP methods.

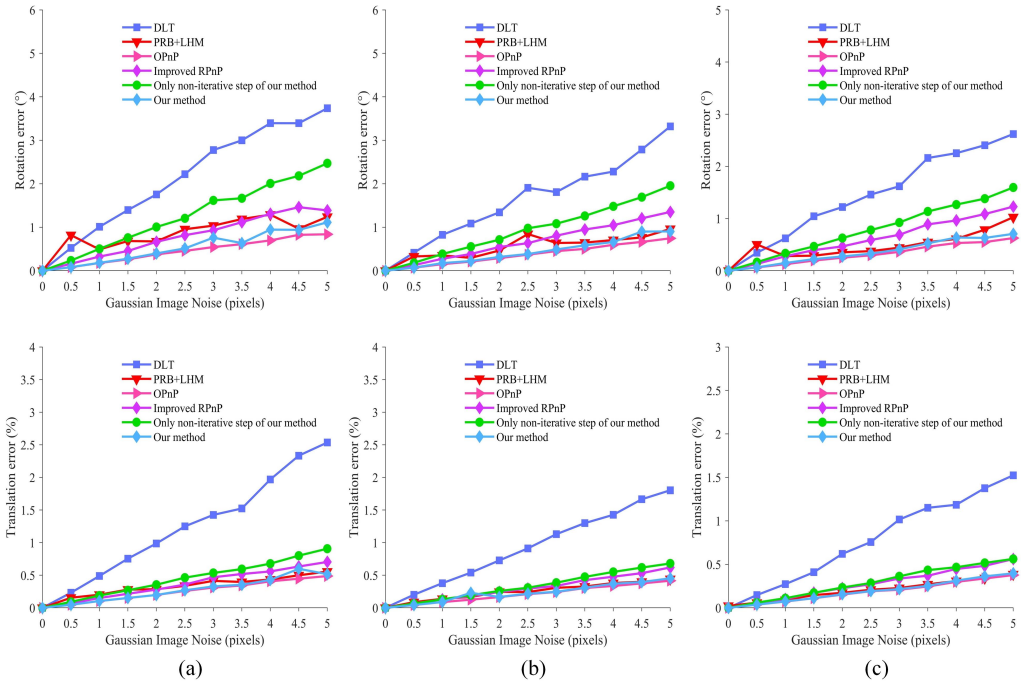


Fig. 7. Measurement errors of pose estimation under different noise levels: (a) $n = 12$, (b) $n = 16$, and (b) $n = 20$.

4.2. Pose estimation experiments with the real images

Pose estimation experiments – influence of the shape of the target array

We placed the target at 24 positions arbitrarily. The target's pose [**R** **T**] was estimated using our and the other 5 methods with different shapes of target arrays like a rectangle, an isosceles trapezoid, and a right-angled trapezoid which were formed by 4 feature points, as shown in Fig. 8.

Metrol. Meas. Syst., Vol. 30 (2023), No. 2, pp. 235–258
DOI: 10.24425/mms.2023.144871

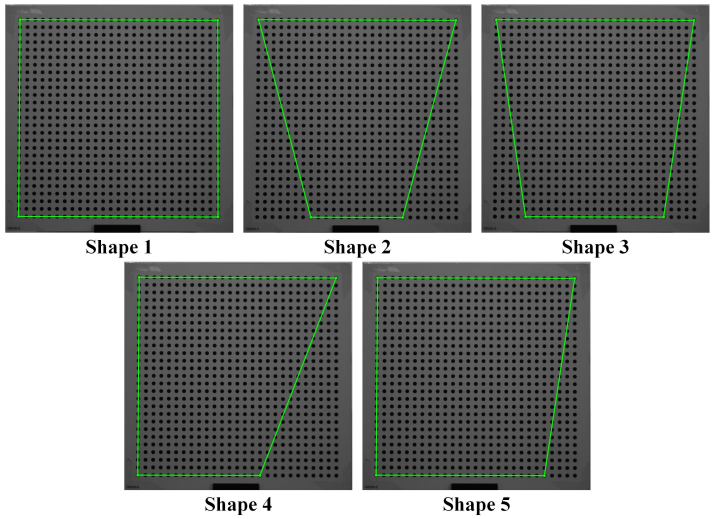


Fig. 8. Different shapes of the array.

The mean re-projection errors corresponding to these 24 positions were calculated. They are shown in Fig. 9. As seen from the figure, compared to the other pose estimation methods, the measurement results of our method are very stable and hardly affected by the shape of the target array. Compared to the DLT method, the non-iterative part of our method considers the line constraints formed by every two feature points. The solution stability is strengthened. Adding only one iteration, the solution stability of our method is further strengthened. The re-projection error of our method is also the lowest.

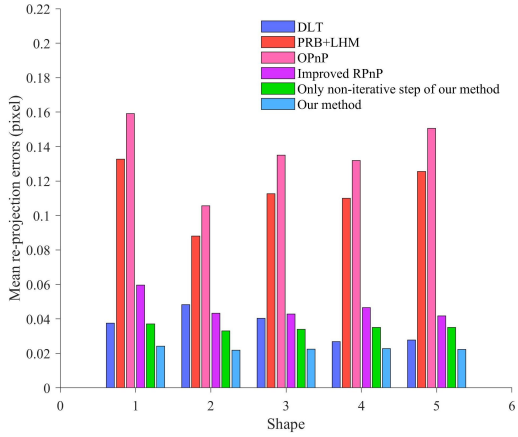


Fig. 9. Re-projection errors using different shapes of the array.

Pose estimation experiments – rotation angle measurement

The target was fixed and rotated at 5° intervals within the range of −45° to 45°. The captured images are shown in Fig. 10.

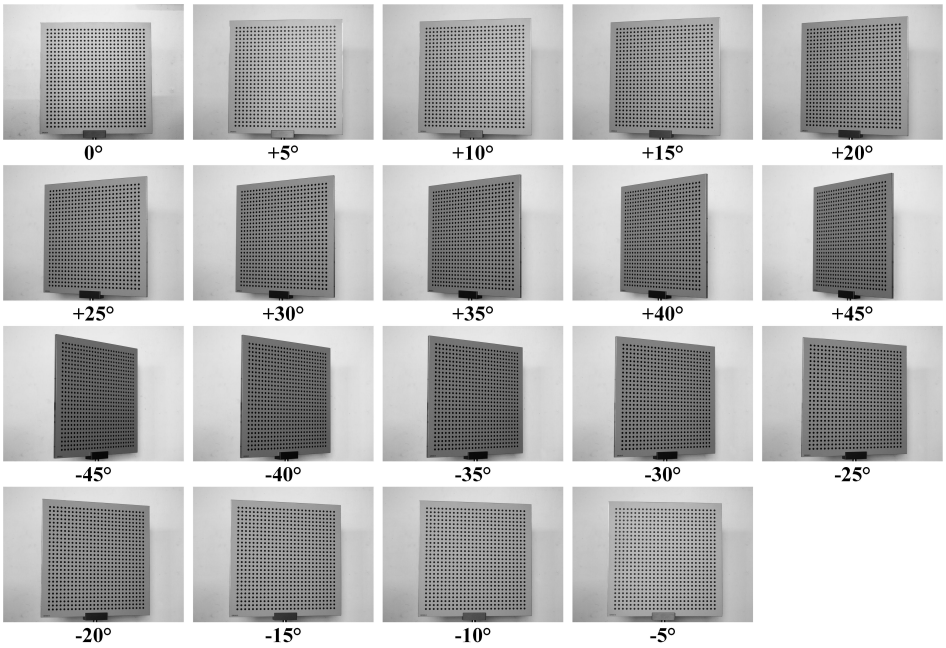


Fig. 10. Images captured for the rotation angle measurement.

The target's pose [**R T**] was estimated using our and the other 5 methods. The rotation angle θ of each position relative to position 0° was obtained. The measurement errors, which were $|\theta - \theta_{\text{true}}|$, are shown in Fig. 11. As can be seen from the figure, the PRB+LHM method is iterative, even though it is based on the incident-ray-camera model. When the initial solution is unsuitable, it is easy to fall into a local optimum. The corresponding measurement error is significant. The DLT method is based only on the constraints of each single feature point. There is various noise in real measurement environments caused by image processing, non-concentric errors of circular feature points, uneven illumination, etc. The measurement results are not stable. The simulation experiment results of the OPnP method are good. However, the real experimental results of the OPnP method are worse. This is because the noise in the real experiments is more complex, as illustrated above than in the simulation experiments. The OPnP method uses the Gröbner basis method to solve the equations, which takes lots of time and does not guarantee the solution stability. The improved RPnP method is suboptimal. It solves the pose based on the three-point (P3P) subsets formed by the two points of the rotation axis and any other feature point to avoid the unit orthogonality problem of the matrix. Therefore, it can obtain relatively better results. The non-iterative step of our method fully considers the constraints formed by the feature points, including the correspondences of the 3D feature points with their 2D projections and line constraints formed by every two feature points. A pretty good pose solution can be obtained, and the measurement accuracy and stability are further improved after one iteration. The measurement accuracy and stability of our method are the best. In the −45° to +45° measuring range, the maximum measurement error is no more than 0.039°, and the average measurement error is no more than 0.016°. Compared with the DLT, PRB+LHM, OPnP, and Improved RPnP methods, our method lowered their measurement errors by 83%, 91%, 92%, and 69%, respectively.

Metrol. Meas. Syst., Vol. 30 (2023), No. 2, pp. 235–258
DOI: 10.24425/mms.2023.144871

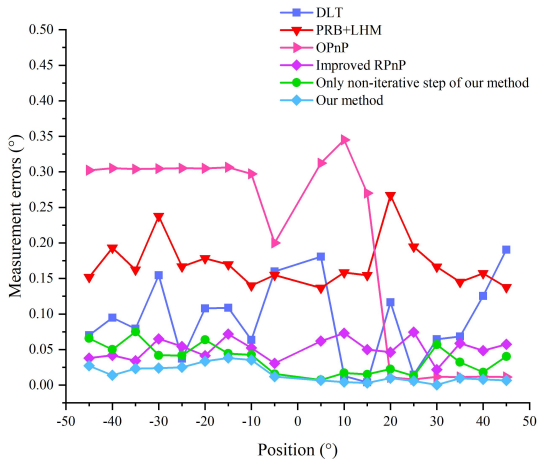


Fig. 11. Results of measuring the rotation angle.

The camera was moved backward, and the previous steps were repeated for some additional experiments. The captured images of the target are shown in Fig. 12.

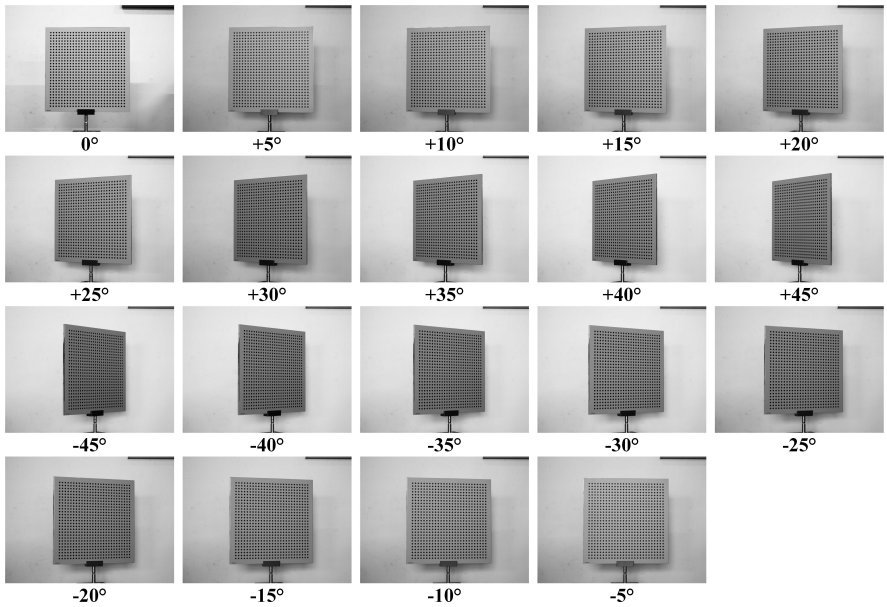


Fig. 12. Images captured for the rotation angle measurement at a long measurement distance.

The measurement errors are shown in Fig. 13. The experimental results are similar to the previous ones. They also demonstrate our method's advantages in measuring rotation angle. The line constraints formed by every two feature points and the redundant feature points help ensure our method's solution accuracy. Compared with the DLT, PRB+LHM, OPnP, and Improved RPnP methods, our method lowered their measurement errors by 91%, 96%, 97%, and 81%, respectively.

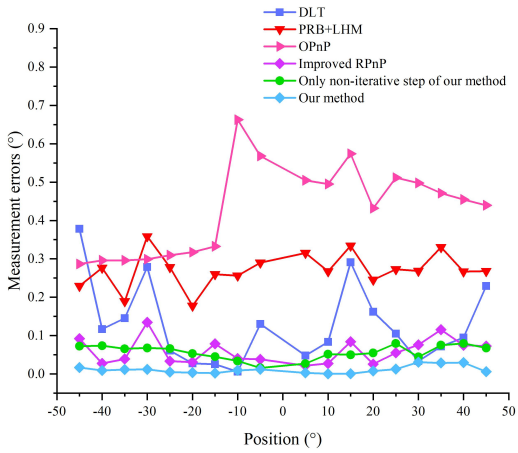


Fig. 13. Results of measuring rotation angle at a long distance.

The mean measurement errors at different distances of the camera from the array are shown in Fig. 14.

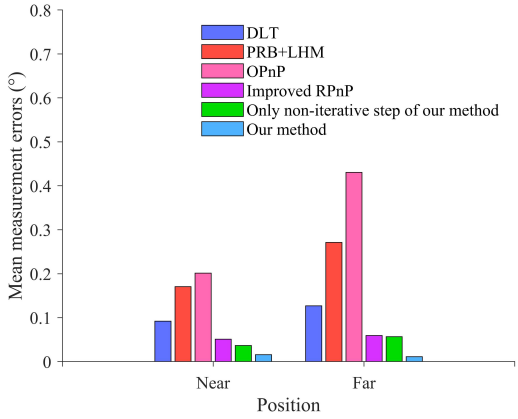


Fig. 14. Mean measurement errors at different distances of the camera from the array.

The DLT, PRB+LHM, and OPnP methods are easily affected by the distance between the camera and the target array. When the measurement distance becomes larger, the measurement error increases. The robustness of the Improved RPnP method is better. However, our method lowered its measurement errors by 69% and 81%, corresponding to the near and far distances, respectively. Our method is not easily affected by the measurement distance. This is because our method considers more constraint information and has a one-iteration step.

Pose estimation experiments – displacement measurement

The target was fixed and moved at 2 mm intervals from 0 mm to 30 mm. The captured images are shown in Fig. 15.

The target's pose [**R T**] was estimated using our and the other 5 methods. The linear displacement t between each position and the position 0 mm was obtained. The measurement errors, which

Metrol. Meas. Syst., Vol. 30 (2023), No. 2, pp. 235–258
DOI: 10.24425/mms.2023.144871

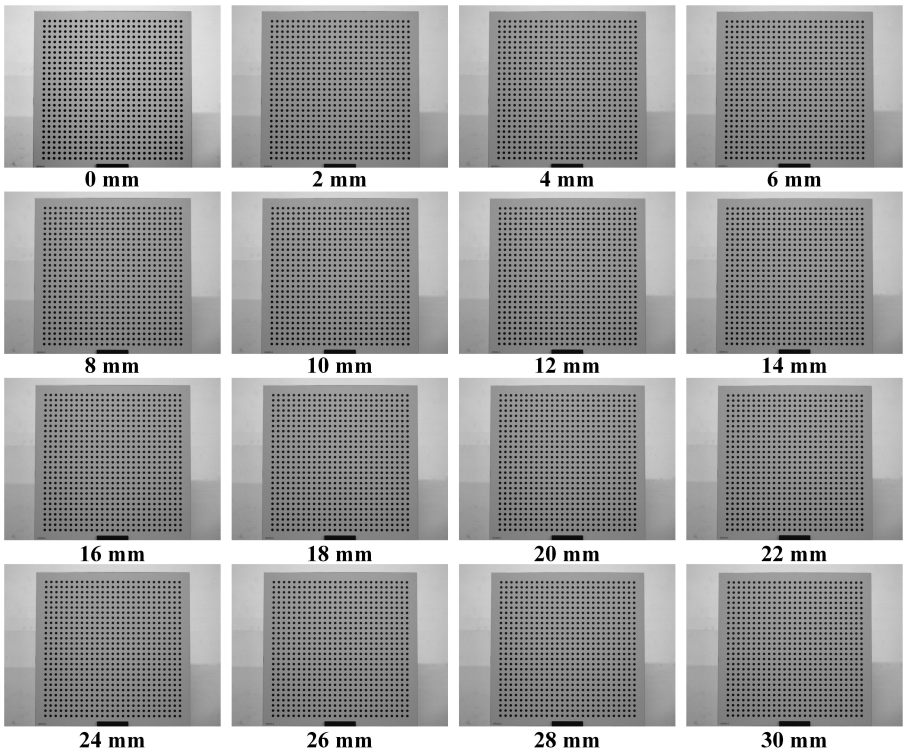


Fig. 15. Images captured for the displacement measurement.

were $|t - t_{\text{true}}|$, are shown in Fig. 16. The Improved RPnP method completes the pose estimation by constructing a new coordinate frame, and it is unnecessary to choose the two feature points with the longest 2D projection to be the rotation axis. The solving process is stable. Therefore, the measurement accuracy of the Improved RPnP method is better than that of the DLT method. Also, there was various noise in the real measurement environment. For binocular vision calibration, the length or angle constraints formed between the calibration points are added to the iterative function to ensure calibration accuracy and robustness [27, 28]. The non-iterative methods are better than iterative ones for pose estimation, as illustrated in our paper. However, for non-iterative methods, adding the constraints provided between feature points to increase the accuracy and stability of the solution still constitutes a problem. The non-iterative step of our method solves this problem and fully considers the constraints. Its measurement accuracy is almost equivalent to that of the Improved RPnP method. One iteration is needed to further improve the measurement accuracy and stability. However, the measurement accuracy of the OPnP method is low compared with that of the DLT method. This is caused by its complex solution process and complex noise in real measurement experiments. The PRB+LHM method is improved from the LHM method. It is a non-iterative method, and the inappropriate initial solution makes its measurement results unstable. As to our method's measurement accuracy and stability, in the 0 mm to 30 mm measuring range, the maximum measurement error is no more than 0.049 mm, and the average measurement error is no more than 0.012 mm. Compared with the DLT, PRB+LHM, OPnP, and Improved RPnP methods, our method lowered their measurement errors by 81%, 82%, 93%, and 67%, respectively.

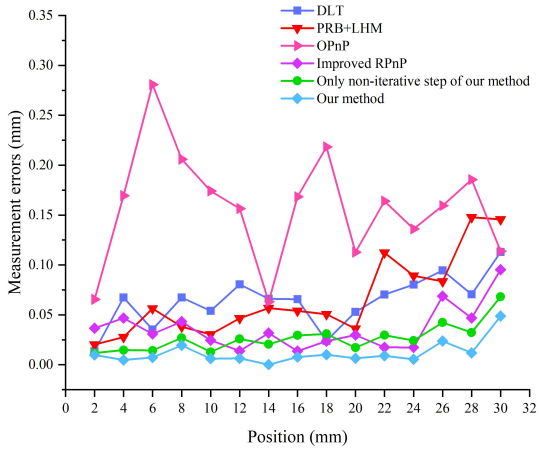


Fig. 16. Measurement errors of the displacement.

The position of the camera was moved to the left and right accordingly. The previous steps were repeated for some additional experiments. The measurement errors are shown in Fig. 17. The measurement error of the PRB+LHM method is still unstable. The measurement accuracy of positioning the camera to the left is better than setting the camera to the right. The experimental results of other non-iterative methods like the DLT, OPnP, and Improved RPnP methods are similar to the previous. This proves that the measurement accuracy of the non-iterative methods is more stable. The experimental results also demonstrate the high accuracy of our method in terms of displacement measurements. Compared with the DLT, PRB+LHM, OPnP, and Improved RPnP methods, our method lowered their measurement errors by 86%, 81%, 92%, and 72%, respectively.

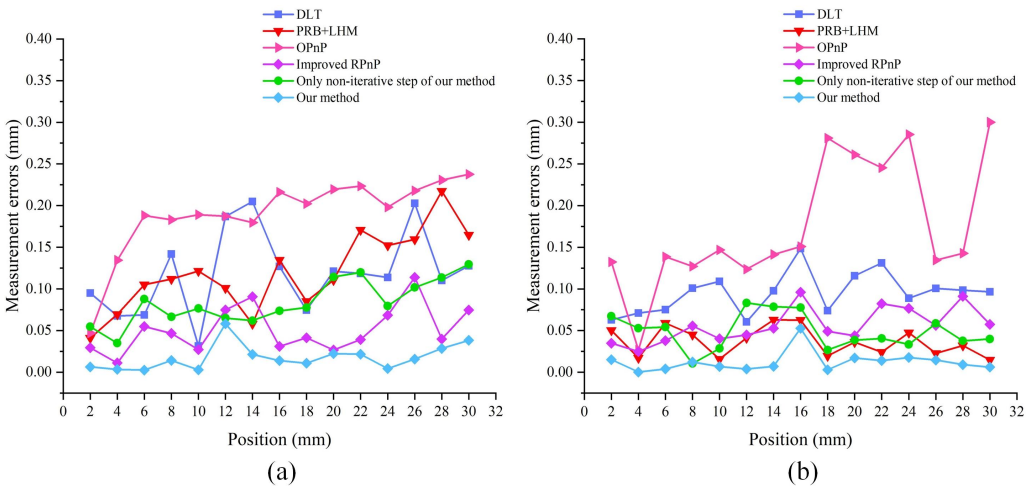


Fig. 17. Measurement errors of the displacement with a long measurement distance: (a) The measurement errors set the camera to the right. (b)The measurement errors set the camera to the left.

Metrol. Meas. Syst., Vol. 30 (2023), No. 2, pp. 235–258
DOI: 10.24425/mms.2023.144871

Pose estimation experiments – arbitrarily selected place the target

Placed the target at 12 positions arbitrarily. The captured images are shown in Fig. 18. The target's pose [**R T**] was estimated using our and the other 5 methods.

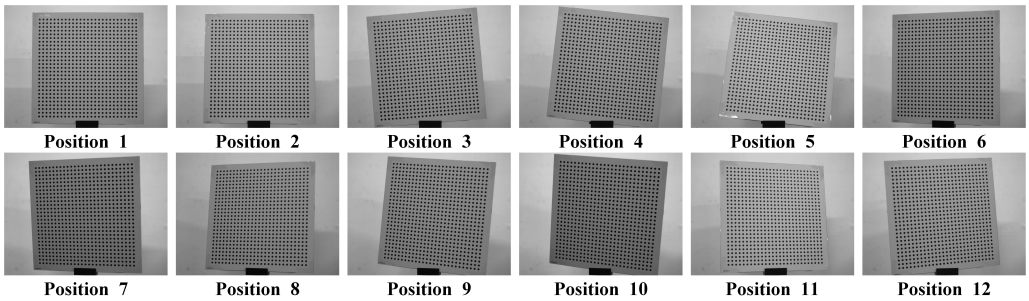


Fig. 18. Images of the arbitrarily placed target for pose estimation.

The re-projection errors were then calculated. They are shown in Fig. 19. As seen from Fig. 19, the re-projection error of our method is the lowest. The accuracy of the PRB+LHM method depends on the choice of the initial pose. The OPnP method has a complex solution process. The re-projection errors of these two methods fluctuate more than the other methods. The DLT method established the pose estimation model only according to the constraints formed by each single feature point, and there is no iterative step. It is easily affected by the various noise in the real environment. The re-projection error of the Improved RPnP method is lower than that of the DLT method. This is because its solving process avoids the unit orthogonality of the rotation matrix. However, it is a suboptimal method, and the accuracy can still be improved. The non-iterative step of our method considers the line constraints formed by every two feature points in addition to the correspondences of the feature points with their image points. The re-projection error is reduced compared to the DLT method. The re-projection error of the non-iterative step of our method is higher than that of the Improved RPnP method at some measurement positions.

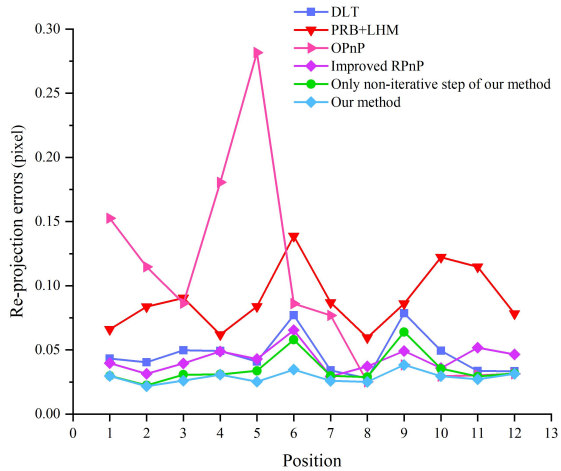


Fig. 19. Re-projection errors of the experiments.

After one iteration, the re-projection error is reduced. The pose estimation accuracy and stability of our method are the best. Compared with the DLT, PRB+LHM, OPnP, and Improved RPnP methods, our method lowered their errors by 38%, 68%, 70%, and 33%, respectively.

The position of the camera was moved backward. The previous steps were repeated for some additional experiments. The captured images of the target are shown in Fig. 20.

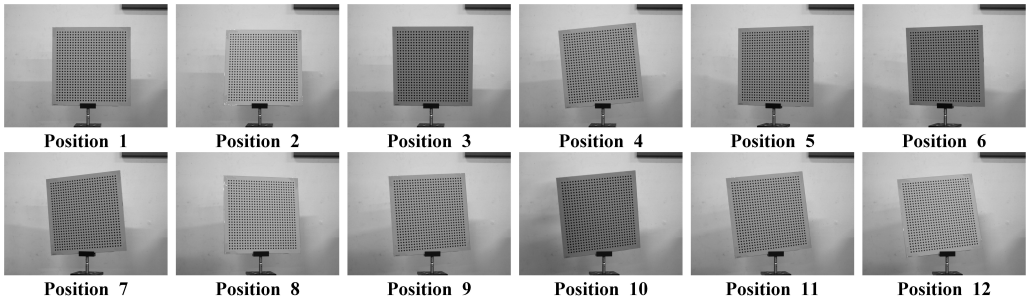


Fig. 20. Images of the arbitrarily placed target at a long measurement distance for pose estimation.

Figure 21 depicts the corresponding re-projection errors. The re-projection errors of the PRB+LHM and OPnP methods are unstable. The experimental results of other non-iterative methods, like the DLT and Improved RPnP, are similar to the previous ones. The re-projection error of our method is the lowest. Compared with the DLT, PRB+LHM, OPnP, and Improved RPnP methods, our method lowered their measurement errors by 57%, 78%, 83%, and 39%, respectively.

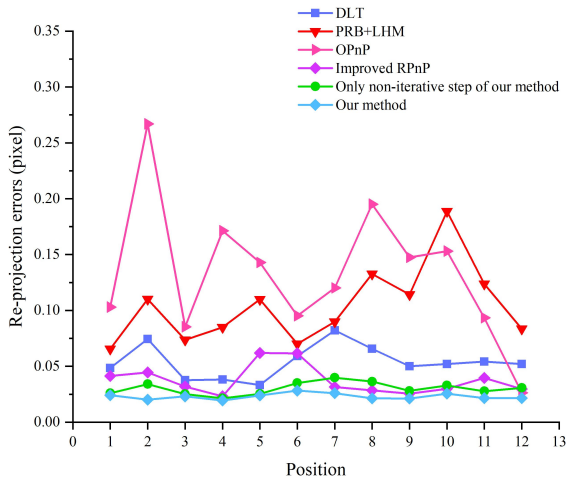


Fig. 21. Re-projection errors of experiments at a long measurement distance.

The mean re-projection errors with different distances of the camera from the array are shown in Fig. 22. It can be seen that we obtain similar experimental results and our method is not easily affected by the measurement distance.

Metrol. Meas. Syst., Vol. 30 (2023), No. 2, pp. 235–258
DOI: 10.24425/mms.2023.144871

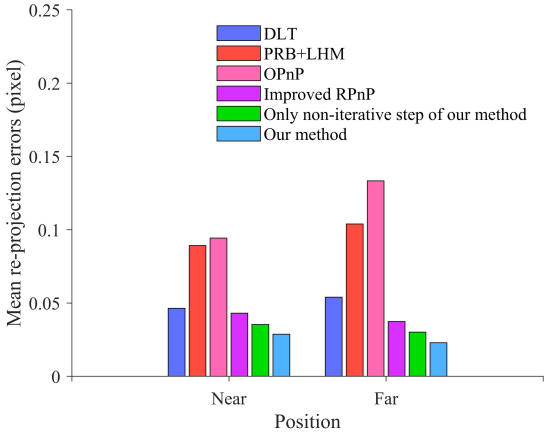


Fig. 22. Re-projection errors at different distances of the camera from the array.

Pose estimation experiments – pose estimation of the real objects

We estimated the pose of a book. The images for pose estimation are shown in Fig. 23. The re-projections of pose estimation with different methods are shown in Fig. 24.

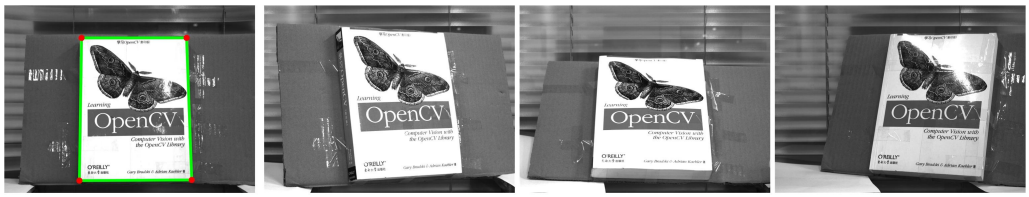


Fig. 23. Images of a book: (a) Reference image, (b) Images for pose estimation.

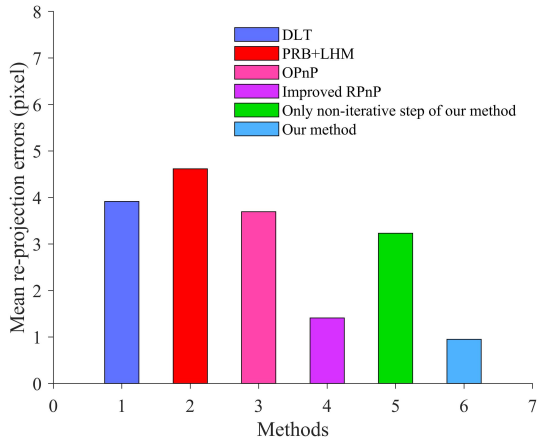
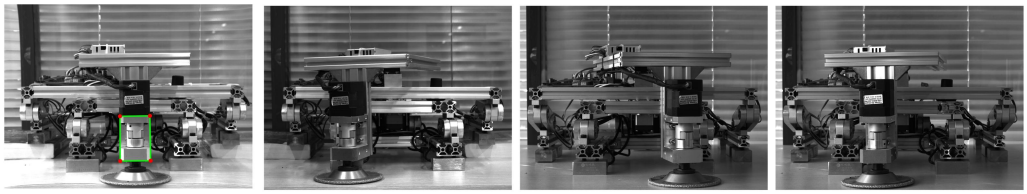


Fig. 24. Re-projection errors with different methods.

We estimated the pose of the suction cup of a climbing robot. The images for pose estimation are shown in Fig. 25. The re-projections of pose estimation with different methods are shown in Fig. 26.



(a) (b)

Fig. 25. Images of the suction cup of a climbing robot: (a) Reference image, (b) Images for pose estimation.

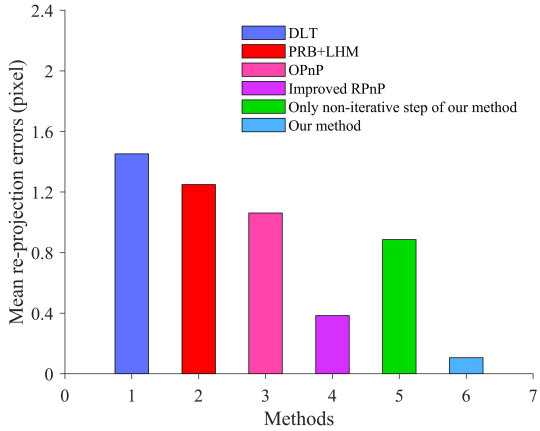


Fig. 26. Re-projection errors with different methods.

As shown in Fig. 24 and Fig. 26, the number of feature points is 4. At the same time, the extraction of the corner points of the object is more easily affected by the noise in the measurement environment than the circular feature points of the test board, which are specially machined. Therefore, the re-projection errors of estimating the real objects increase for all the methods compared with those of estimating the test board pose. However, due to the incorporation of line constraints provided between the feature points in our method and the subsequent one-step iteration, our method yields more accurate results than the other methods. Also, our method outperforms the other methods in real-life applications. This further proves our previous claims.

Pose estimation experiments – computational time

The computational time of each method for estimating the target's pose with the images shown in Fig. 18 was calculated and is shown in Fig. 27. The PRB+LHM method is iterative and therefore the efficiency is the lowest. The OPnP method is based on matrix synthesis technology, thus the solution process takes more time and the solution efficiency is low. On the other hand, the DLT method only considers the correspondences of the 3D feature points with their 2D projections and its computational time is the least. Compared with the DLT method, the improved RPnP method is suboptimal, and the computational time increases. The non-iterative step of our

Metrol. Meas. Syst., Vol. 30 (2023), No. 2, pp. 235–258
DOI: 10.24425/mms.2023.144871

method considers more constraints than the DLT method and the computational time slightly increases. The one-iteration step refines the pose obtained in the non-iterative step and takes a little extra time. As for the current computer platform, the average computational time of our method was 0.00189s. However, either the accuracy of the non-iterative step of our method or that of our entire process is significantly better than that of the DLT method. The results obtained in the non-iterative step are satisfactory enough for some applications that emphasize real-time. On the premise of ensuring measurement accuracy, our method is the most efficient compared with the other existing pose estimation methods.

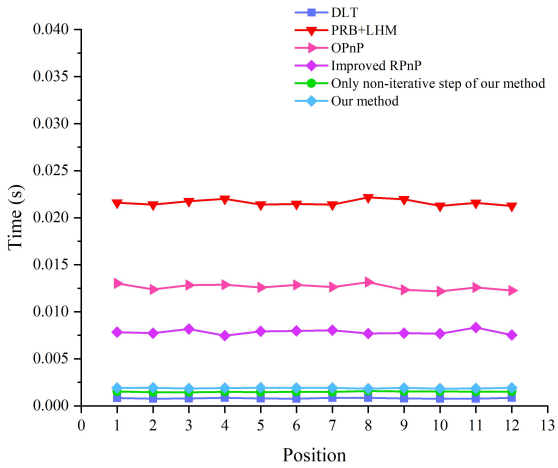


Fig. 27. Computational time of each method for estimating the target's pose.

In summary, simulation and real experiment results demonstrated the superiority of our method in accuracy, stability, and efficiency. On the premise of ensuring measurement accuracy, our method is the most efficient compared with the other existing pose estimation methods.

5. Conclusions

This paper proposed an accurate and stable pose estimation method for planar cases. A novel non-iterative pose estimation model is established considering the imaging constraint provided by each feature point and the line constraints formed by every two feature points. Only one iteration is needed to refine the solved pose because we fully consider all the constraints in the non-iterative step. The measurement accuracy and stability are improved compared to the other non-iterative methods. Based on the non-iterative step's efficiency, adding one iteration still guarantees the solution's efficiency.

Acknowledgements

This research was supported by the Tianjin University Science and Technology Development Fund Project (2022ZD029).

References

[1] Fischler, M. A., & Bolles, R. C. (1981). Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography. *Communications of the ACM*, *24*(6), 381–395. https://doi.org/10.1145/358669.358692

[2] Dong, Y., Zhang, G., Chang, S., Zhang, Z., & Li, Y. (2021). A pose measurement algorithm of space target based on monocular vision and accuracy analysis. *Acta Photonica Sinica*, *50*(8), 1112003. https://doi.org/10.3788/gzxb20215011.1112003

[3] Zhu, Z., Wang, S., Zhang, H., & Zhang, F. (2020). Camera–projector system calibration method based on optimal polarization angle. *Optical Engineering*, *59*(3), 035104–035104. https://doi.org/10.1117/1.OE.59.3.035104

[4] Zhang, Z. (2000). A flexible new technique for camera calibration. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *22*(8), 1330–1334. https://doi.org/10.1109/34.888718

[5] Cui, H., Sun, R., Fang, Z., Lou, H., Tian, W., & Liao, W. (2020). A novel flexible two-step method for eye-to-hand calibration for robot assembly system. *Measurement and Control*, *53*(9-10), 2020–2029. https://doi.org/10.1177/0020294020964842

[6] Yuan, K., Guo, Z., & Wang, Z. J. (2020). RGGNet: Tolerance aware LiDAR-camera online calibration with geometric deep learning and generative model. *IEEE Robotics and Automation Letters*, *5*(4), 6956–6963. https://doi.org/10.1109/LRA.2020.3026958

[7] Enayati, N., De Momi, E., & Ferrigno, G. (2015). A quaternion-based unscented Kalman filter for robust optical/inertial motion tracking in computer-assisted surgery. *IEEE Transactions on Instrumentation and Measurement*, *64*(8), 2291–2301. https://doi.org/10.1109/TIM.2015.2390832

[8] Fu, B., Han, F., Wang, Y., Jiao, Y., Ding, X., Tan, Q., ... & Xiong, R. (2021). High-precision multicamera-assisted camera-IMU calibration: Theory and method. *IEEE Transactions on Instrumentation and Measurement*, 70, 1–17. https://doi.org/10.1109/TIM.2021.3051726

[9] Guo, X., Tang, J., Li, J., Shen, C., & Liu, J. (2019). Attitude measurement based on imaging ray tracking model and orthographic projection with iteration algorithm. *ISA Transactions*, 95, 379–391. https://doi.org/10.1016/j.isatra.2019.05.009

[10] Zhang, Z., Liu, B., & Jiang, Y. (2015). A two-step pose estimation method based on four non-coplanar points. *Optik*, *126*(17), 1520–1526. https://doi.org/10.1016/j.ijleo.2015.04.039

[11] DeMenthon, D. F., & Davis, L. S. (1995). Model-based object pose in 25 lines of code. *International Journal of Computer Vision*, 15, 123–141. https://doi.org/10.1007/BF01450852

[12] David, P., Dementhon, D., Duraiswami, R., & Samet, H. (2004). SoftPOSIT: Simultaneous pose and correspondence determination. *International Journal of Computer Vision*, 59, 259-284. https://doi.org/10.1023/B:VISI.0000025800.10423.1f

[13] Oberkampf, D., DeMenthon, D. F., & Davis, L. S. (1996). Iterative pose estimation using coplanar feature points. *Computer Vision and Image Understanding*, *63*(2), 495–511. https://doi.org/10.1006/cviu.1996.0037

[14] Lu, C. P., Hager, G. D., & Mjolsness, E. (2000). Fast and globally convergent pose estimation from video images. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *22*(6), 610–622. https://doi.org/10.1109/34.862199

[15] Schweighofer, G., & Pinz, A. (2006). Robust pose estimation from a planar target. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *28*(9), 2024–2030. https://doi.org/10.1109/TPAMI.2006.252

Metrol. Meas. Syst., Vol. 30 (2023), No. 2, pp. 235–258
DOI: 10.24425/mms.2023.144871

[16] Dong, H., Sun, C., Zhang, B., & Wang, P. (2019). Simultaneous pose and correspondence determination combining softassign and orthogonal iteration. *IEEE Access*, 7, 137720–137730. https://doi.org/10.1109/ACCESS.2019.2939020

[17] Sun, C., Dong, H., Zhang, B., & Wang, P. (2018). An orthogonal iteration pose estimation algorithm based on an incident ray tracking model. *Measurement Science and Technology*, *29*(6), 095402. https://doi.org/10.1088/1361-6501/aad014

[18] Wu, P. C., Tseng, H. Y., Yang, M. H., & Chien, S. Y. (2018). Direct pose estimation for planar objects. *Computer Vision and Image Understanding*, *172*, 50–66. https://doi.org/10.1016/j.cviu.2018.03.006

[19] Lepetit, V., Moreno-Noguer, F., & Fua, P. (2009). EP n P: An accurate O (n) solution to the PnP problem. *International Journal of Computer Vision*, *81*, 155-166. https://doi.org/10.1007/s11263-008-0152-6

[20] Hesch, J. A., & Roumeliotis, S. I. (2011, November). A direct least-squares (DLS) method for PnP. In *2011 International Conference on Computer Vision* (pp. 383–390). IEEE. https://doi.org/10.1109/ICCV.2011.6126266

[21] Li, S., Xu, C., & Xie, M. (2012). A robust O(n) solution to the perspective-n-point problem. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *34*(7), 1444–1450. https://doi.org/10.1109/TPAMI.2012.41

[22] Wang, P., Xu, G., Cheng, Y., & Yu, Q. (2018). A simple, robust and fast method for the perspective-n-point problem. *Pattern Recognition Letters*, *108*, 31–37. https://doi.org/10.1016/j.patrec.2018.02.028V

[23] Zheng, Y., Sugimoto, S., & Okutomi, M. (2013). ASPnP: An accurate and scalable solution to the perspective-n-point problem. *IEICE Transactions on Information and Systems*, *96*(7), 1525–1535. https://doi.org/10.1587/transinf.E96.D.1525

[24] Zheng, Y., Kuang, Y., Sugimoto, S., Astrom, K., & Okutomi, M. (2013). Revisiting the PnP problem: A fast, general and optimal solution. *In Proceedings of the IEEE International Conference on Computer Vision* (pp. 2344–2351). https://doi.org/10.1109/ICCV.2013.291

[25] Kneip, L., Li, H., & Seo, Y. (2014). UPnP: An optimal O(n) solution to the absolute pose problem with universal applicability. *In Computer Vision – ECCV 2014: Computer Vision – ECCV 2014. Lecture Notes in Computer Science* (pp. 127–142). Springer International Publishing. https://doi.org/10.1007/978-3-319-10590-1_9

[26] Liu, Y., Lei, B., Fan, B., Bian, J. (2020). Target positioning technology and its structural parameter optimization based on vision measurement. *Infrared and Laser Engineering*, *49*(S02).

[27] Liu, X., Liu, Z., Duan, G., Cheng, J., Jiang, X., & Tan, J. (2018). Precise and robust binocular camera calibration based on multiple constraints. *Applied Optics*, *57*(18), 5130–5140. https://doi.org/10.1364/AO.57.005130

[28] Yang, P., Yin, Y., Lu, R., & Zhu, H. (2022). Binocular camera calibration based on directional target and multi-constraint optimization. *Acta Optica Sinica*, *42*(8), 0815002. https://doi.org/10.3788/AOS202242.0815002 (in Chinese)

[29] Zhang Z., Xu K., Wu Y., Zhang S., et al., A simple and precise calibration method for binocular vision. *Measurement Science and Technology*, 2022; 33(6): 065016. https://doi.org/10.1088/1361-6501/ac4ce5

[30] Zimiao, Z., Kai, X., Yanan, W., Shihai, Z., & Yang, Q. (2022). A simple and precise calibration method for binocular vision. *Measurement Science and Technology*, *33*(6), 065016. https://doi.org/10.1080/09500340.2017.1397218

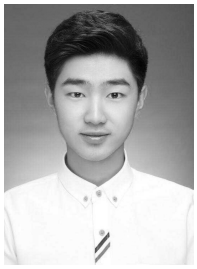
[31] Huo, J., Cui, J. S., & Wang, W. X. (2014). Error analysis of monocular visual position measurement based on coplanar feature points. *Acta Photonica Sinica*, *43*(5), 144–150. https://doi.org/10.3788/gzxb20144305.0512003

[32] Qu, Y., & Hou, W. (2019). Attitude accuracy analysis of PnP based on error propagation theory. *Optics and Precision Engineering*, *27*(2), 479–487. https://doi.org/10.3788/OPE.20192702.0479 (in Chinese)

[33] Qu, Y., Liu, J., & Hou, W., (2020). Graphics Design of Cooperative Targets on Monocular Vision High Precision Measurement. *Acta Optica Sinica*, *40*(10). https://doi.org/10.3788/AOS202040.1315001 (in Chinese)

Zhang Zimiao received his Ph.D. from the Tianjin University, Tianjin, China, in 2012. He is currently an associate professor at the School of Mechanical Engineering, Tianjin University of Technology and Education. His current research interests include machine vision and measurement technology.

Zhang Fumin received his Ph.D. from the Tianjin University, Tianjin, China, in 2009. He is currently a professor at the School of Precision Instrument and Optoelectronic Engineering, Tianjin University. His current research interests include high-precision laser absolute ranging methods and visual image measurement technology.



Zhang Hao received his bachelor's degree from Qingdao University of Technology, Qingdao, China, in 2019. He is currently a graduate student at the School of Mechanical Engineering, Tianjin University of Technology and Education. His current research interests include vision measurement and image processing.

Zhang Shihai received his Ph.D. from the Beijing University of Technology, Beijing, China, in 2012. He is currently a professor at the School of Mechanical Engineering, Tianjin University of Technology and Education. His current research interests include measurement technology and dynamic balance technology.