



Hand-eye calibration with a new linear decomposition algorithm*

Rong-hua LIANG[†], Jian-fei MAO

(College of Information Engineering, Zhejiang University of Technology, Hangzhou 310032, China)

[†]E-mail: rhliang@zjut.edu.cn

Received Apr. 25, 2008; revision accepted Aug. 6, 2008

Abstract: To solve the homogeneous transformation equation of the form $AX=XB$ in hand-eye calibration, where X represents an unknown transformation from the camera to the robot hand, and A and B denote the known movement transformations associated with the robot hand and the camera, respectively, this paper introduces a new linear decomposition algorithm which consists of singular value decomposition followed by the estimation of the optimal rotation matrix and the least squares equation to solve the rotation matrix of X . Without the requirements of traditional methods that A and B be rigid transformations with the same rotation angle, it enables the extension to non-rigid transformations for A and B . The details of our method are given, together with a short discussion of experimental results, showing that more precision and robustness can be achieved.

Key words: Homogeneous transformation equation, Singular value decomposition (SVD), Optimal rotation matrix, Rigid transformations

doi:10.1631/jzus.A0820318

Document code: A

CLC number: TP242.62

INTRODUCTION

Hand-eye calibration (HEC) is a process used in the field of robotics for determining the relative orientation of a camera (or sensor) in relation to a robot hand. HEC is essential for a robot's vision system, e.g., in guiding its motion and locating positional information in the surrounding world. The classical HEC approach is to employ an *a priori* defined object, called the 'calibration object' (Zhang, 2000). Fig.1a shows a general hand-eye system in which a sensor (or camera) is mounted on one joint of a robotic arm. To achieve HEC translates to the problem of solving the equation $AX=XB$, where X is a matrix representing an unknown transformation from the camera to the robot hand, and A and B are the transformation matrices representing the movement transformation of the robot hand and the camera, respectively (Shiu and Ahmad, 1989). A brief description of the equation

is shown in Fig.1b.

Fig.1b shows the coordinate transformations involved in the calibration process: the planar calibration object is moved from location i to location j in front of the camera by the robot. The key realization is that the two coordinate transformations $c \rightarrow b \rightarrow g^{(i)} \rightarrow t^{(i)} \rightarrow c$ and $c \rightarrow b \rightarrow g^{(j)} \rightarrow t^{(j)} \rightarrow c$ form a closed loop separately, and thus we have the following equation:

$$\left(H_{gb}^{(i)}\right)^{-1} H_{gb}^{(j)} H_{tg} = H_{tg} \left(H_{tc}^{(i)}\right)^{-1} H_{tc}^{(j)}. \quad (1)$$

From Eq.(1) we have

$$C_{ij} H_{tg} = H_{tg} D_{ij}, \quad (2)$$

or

$$E_{ij} H_{cb} = H_{cb} F_{ij}, \quad (3)$$

where H_{tg} is a transformation matrix from the calibration object to the robot gripper while H_{cb} is the transformation matrix from the camera to the robot base. Therefore, to solve the transformation matrix for H_{cb} or H_{tg} is equivalent to solving the unknown X in $AX=XB$.

* Project (No. 60703002) supported by the National Natural Science Foundation of China

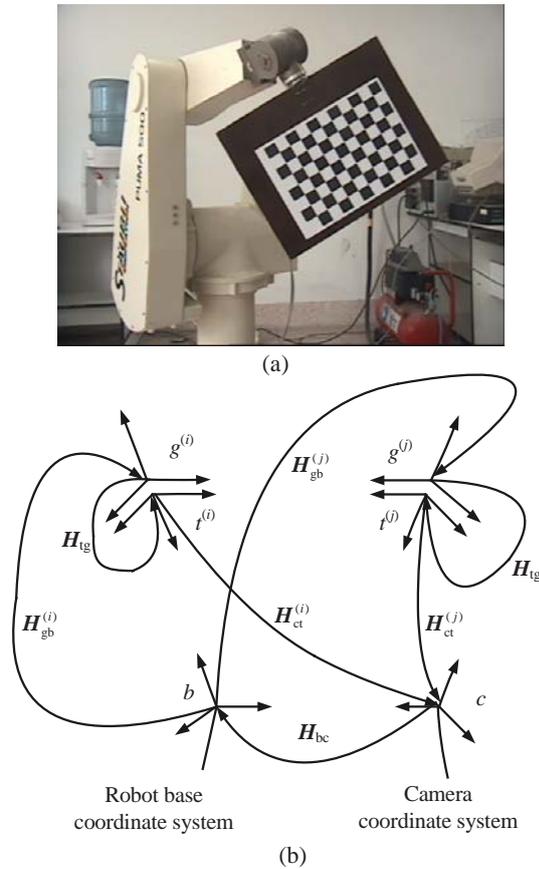


Fig.1 (a) Hand-eye calibration process; (b) Coordinate transformations involved in the process (The gripper, planar pattern and robot base coordinate systems are denoted by g , t and b , respectively. H_{xy} transforms from the x to the y coordinate system)

Typically, to solve the equation, either quaternions (Shiu and Ahmad, 1989; Tsai and Lenz, 1989; Zhuang and Roth, 1991; Daniilidis, 1999; Fassi and Legnani, 2005) or matrix screw theory (Zhao and Liu, 2006) are employed. Such classic algorithms consist of two stages: initially the rotational axis of X is solved linearly according to the rotational axes (or quaternion vectors) of A and B , and the result is then used to solve the translation vector by a linear approach. The limitation of classic approaches is that A and B must represent rigid transformations with the same rotation angle, while in a real experimental environment in the presence of noise, the rotation angles of A and B are a little different, and A and B can even be non-rigid transformations. A further limitation is that as a two-stage linear method is employed, the estimation rotation error in the first stage can propagate to the second stage. An optimal method,

called the 'Jacobian equation', was described in (Zhuang and Roth, 1995; Zhuang, 1998) for limiting this propagation error. Nevertheless, this approach still assumes that A and B represent rigid transformations, which may lead to a great error for the real experimental environment.

Based on singular value decomposition (SVD) and optimal orthogonal estimation, this paper presents a novel linear method that greatly minimizes the error caused by the noise while does not require A and B be rigid transformations with an identical rotation angle. We first describe the details of our method for an optimal orthogonal estimation of a generic matrix, and then calculate X by integrating SVD, the estimation of an optimal rotation matrix and the least squares method. A short discussion follows the experimental results using both simulation and real-world datasets. Our approach features the following two characteristics: (1) we use linear decomposition operations on A and B , and do not need to calculate the screw vector by decomposing A and B , which enables our approach to be extended to non-rigid A and B ; (2) we propose a method to estimate the optimal orthogonal matrix for λB with $\det(\lambda B) \neq 0$, which minimizes the error caused by the noise.

The remainder of this paper is organized as follows. The detailed steps and some proofs are described in Section 2. Our experimental results with discussions are presented in Section 3. Finally Section 4 contains a summary and pointers to future developments.

ALGORITHM DETAILS

Let us concentrate on the equation $AX=XB$, where A , B and X express the homogeneous transformation matrices:

$$A = \begin{pmatrix} R_a & T_a \\ \mathbf{0} & 1 \end{pmatrix}, B = \begin{pmatrix} R_b & T_b \\ \mathbf{0} & 1 \end{pmatrix}, X = \begin{pmatrix} R_x & T_x \\ \mathbf{0} & 1 \end{pmatrix}, \quad (4)$$

where R_a, R_b, R_x are 3×3 rotation matrices, and T_a, T_b, T_x are 3×1 translation vectors. Therefore, $AX=XB$ can be equivalently expressed by the following two equations:

$$R_a R_x = R_x R_b, \quad (5)$$

$$(R_a - I_3) T_x = R_x T_b - T_a, \quad (6)$$

where \mathbf{I}_3 is a 3×3 identity matrix. The key problem is to solve the general condition for rotation matrices \mathbf{R}_a and \mathbf{R}_b in Eq.(5). The research in (Zhuang and Roth, 1995) demonstrated that to use more than two equations for Eq.(6) will produce the translation vector \mathbf{T}_x , and that if more equations are used, the results of \mathbf{T}_x can be more robust. \mathbf{T}_x can be acquired using QR factorization or SVD in Eq.(6), see (Zhuang and Roth, 1995) for details. The following formulations and lemmas are proposed to solve the rotation matrix. Because \mathbf{A} and \mathbf{B} are real matrices, subsequent operations are performed in a real closed field.

Lemma 1 Eq.(5) can be represented as

$$(\mathbf{R}_a \otimes \mathbf{I}_3 - \mathbf{I}_3 \otimes \mathbf{R}_b^T) \text{vec}(\mathbf{R}_x) = \mathbf{0}, \quad (7)$$

where vec is a linear operator, and \otimes is a Kronecher product operator.

Proof Using Kronecher product to expand Eq.(5), we can have $\mathbf{R}_a \mathbf{R}_x \mathbf{I}_3 - \mathbf{I}_3 \mathbf{R}_x \mathbf{R}_b = \mathbf{0}$, which is another form of Eq.(6).

Eq.(7) is equivalent to $\mathbf{A}\mathbf{X}=\mathbf{0}$, that is, the least squares solution of $\|\mathbf{A}\mathbf{X}\|$. We apply SVD on \mathbf{A} , and have $\mathbf{A}=\mathbf{U}\mathbf{S}\mathbf{V}^T$, where \mathbf{U} and \mathbf{V} are orthogonal matrices, and \mathbf{S} is a diagonal matrix. The eigenvectors of \mathbf{V} are the solution to Eq.(7). The approximate solution of rotation matrix \mathbf{R}_x can be obtained by the following Lemmas 2 and 3:

Lemma 2 For an $n \times n$ matrix \mathbf{B} , if $\det \mathbf{B} \neq 0$, the SVD of \mathbf{B} is $\mathbf{B}=\mathbf{U}\mathbf{S}\mathbf{V}^T$, the most approximate orthogonal matrix of $\lambda \mathbf{B}$ is $\mathbf{R}=\pm \mathbf{U}\mathbf{V}^T$, where $\lambda = \pm \text{tr} \mathbf{S} / \text{tr}(\mathbf{B}\mathbf{B}^T)$.

Proof Optimal estimation of \mathbf{R} meets the requirement for minimization of $\|\lambda \mathbf{B} - \mathbf{R}\|_F$ (denoted by $\mathbf{R} = \{\mathbf{R} \mid \min \|\lambda \mathbf{B} - \mathbf{R}\|_F\}$), where

$$\begin{aligned} \|\lambda \mathbf{B} - \mathbf{R}\|_F &= \text{tr}[(\lambda \mathbf{B} - \mathbf{R})^T (\lambda \mathbf{B} - \mathbf{R})] \\ &= \text{tr}(\lambda^2 \mathbf{B}^T \mathbf{B} - \lambda \mathbf{B}^T \mathbf{R} - \lambda \mathbf{R}^T \mathbf{B} + \mathbf{I}_3) \\ &= 3 + \lambda^2 \text{tr}(\mathbf{B}^T \mathbf{B}) - 2\lambda \text{tr}(\mathbf{B}^T \mathbf{R}), \end{aligned}$$

and finally it can be converted into

$$3 + \text{tr}(\mathbf{B}\mathbf{B}^T)[\lambda - \text{tr}(\mathbf{B}^T \mathbf{R}) / \text{tr}(\mathbf{B}\mathbf{B}^T)] - \text{tr}^2(\mathbf{B}^T \mathbf{R}) / \text{tr}(\mathbf{B}\mathbf{B}^T). \quad (8)$$

For $\text{tr}(\mathbf{B}^T \mathbf{B}) > 0$, to minimize Eq.(8) is equivalent to $\mathbf{R} = \{\mathbf{R} \mid \max(\text{tr}^2(\mathbf{B}^T \mathbf{R}) / \text{tr}(\mathbf{B}\mathbf{B}^T))\}$, that is, $\mathbf{R} = \{\mathbf{R} \mid \max(\text{tr}(\mathbf{B}^T \mathbf{R}))\}$, which can be denoted by

$$\mathbf{R} = \{\mathbf{R} \mid \max(\text{tr}(\mathbf{Z}\mathbf{S}))\}. \quad (9)$$

For $\text{tr}(\mathbf{B}^T \mathbf{R}) = \text{tr}(\mathbf{U}\mathbf{S}\mathbf{V}^T \mathbf{R}) = \text{tr}(\mathbf{U}^T \mathbf{R} \mathbf{V} \mathbf{S})$, and $\mathbf{Z} = \mathbf{U}^T \mathbf{R} \mathbf{V}$, where $\mathbf{Z}\mathbf{Z}^T = \mathbf{I}$, $\|Z_{ii}\| = 1$, and $|Z_{ii}| \leq 1$, we have

$$|\text{tr}(\mathbf{Z}\mathbf{S})| = \left| \sum_i Z_{ii} S_{ii} \right| \leq \sum_i |Z_{ii}| S_{ii} \leq \sum_i S_{ii}. \quad (10)$$

We can easily acquire $\mathbf{Z} = \pm \mathbf{I} \Rightarrow \pm \mathbf{I} = \mathbf{Z} = \mathbf{U}^T \mathbf{R} \mathbf{V} \Rightarrow \mathbf{R} = \pm \mathbf{U}\mathbf{V}^T$. Finally we have

$$\lambda = \pm \text{tr}(\mathbf{V}\mathbf{S}\mathbf{U}^T \mathbf{U}\mathbf{V}^T) / \text{tr}(\mathbf{B}\mathbf{B}^T) = \pm \text{tr} \mathbf{S} / \text{tr}(\mathbf{B}\mathbf{B}^T). \quad (11)$$

Lemma 2 offers the estimation orthogonal matrix of $\lambda \mathbf{B}$, which is, however, not exactly the rotation matrix. To estimate the rotation matrix, we will propose Lemma 3.

Definition 1 For any odd real square matrix, define

$$|\mathbf{A}| = \begin{cases} \mathbf{A}, & \det \mathbf{A} \geq 0, \\ -\mathbf{A}, & \det \mathbf{A} < 0. \end{cases}$$

Lemma 3 If \mathbf{B} is a 3×3 matrix, $\mathbf{B} \neq \mathbf{0}$, and $\mathbf{B} = \mathbf{U}\mathbf{S}\mathbf{V}^T$ by using SVD, the most approximate rotation matrix $\lambda \mathbf{B}$ is $\mathbf{R} = |\mathbf{U}\mathbf{V}^T|$.

Proof Assume that the columns of the rotation matrix \mathbf{R} are $\{\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3\}$, we have $\mathbf{r}_3 = \mathbf{r}_1 \times \mathbf{r}_2$, and $\det \mathbf{R} = \mathbf{r}_3 \cdot (\mathbf{r}_1 \times \mathbf{r}_2) = 1$, where ‘ \times ’ is the cross product and ‘ \cdot ’ is the dot product.

From $\mathbf{I} = (\mathbf{U}\mathbf{V}^T)(\mathbf{U}\mathbf{V}^T)^T$ we have $1 = \det^2(\mathbf{U}\mathbf{V}^T)$, then $\det(\mathbf{U}\mathbf{V}^T) = \pm 1$. As \mathbf{B} is an odd real square matrix, we have $\mathbf{R} = |\mathbf{U}\mathbf{V}^T|$.

Finally, we conclude our approach to obtain the most approximate rotation matrix as follows:

(1) From Eq.(7), we acquire some equations using SVD on \mathbf{A} , and obtain the eigenvector with respect to the minimal eigenvalue and then obtain \mathbf{B} .

(2) According to Lemmas 2 and 3, the optimal solution of the rotation matrix of \mathbf{B} is $\mathbf{R}_x = |\mathbf{U}\mathbf{V}^T|$.

To solve the translation vector \mathbf{T}_x is similar to Zhuang’s approach (Zhuang and Roth, 1995) by using QR factorization on Eq.(6).

EXPERIMENTS AND DISCUSSION

We first produced some simulation data to validate the relationship between the noise level and the number of equations with our approach, and then

compared our algorithm with classic algorithm at the same noise level. Our experiment was finally implemented with a real environment and then compared with the previous algorithm.

Both A and B were assumed to be rigid matrices. We generated random data for X and A , and computed $B=X^{-1}AX$, prior to adding white noises to A and B to simulate the computation of X .

The following two rules were given for the simulation:

(1) Simulation data should range in the space of all the real data;

(2) Simulation data should be uniformly distributed in the whole data space.

According to the rules, we devised 100 simulation data for X with mean distribution, thus producing 25 equations for each datum, and 21 noise levels range from 0 to 1 equally (that is, the noise levels are in the order 0, 0.05, 0.10, ..., 1.00).

From Eqs.(6) and (7), we have

$$F(i) = (R_a(i) \otimes I_3 - I_3 \otimes R_b^T(i))vec(R_x), \quad (12)$$

$$G(i) = (R_a(i) - I_3)T_x - R_x T_b(i) + T_a(i), \quad (13)$$

where $F(i)$ is a 9×1 vector to represent the mean error of the rotation matrix, and $G(i)$ is a 3×1 vector to represent the mean error of the translation vector. We can obtain the three statistics—mean of rotation error norm, mean of translation error norm, and mean residual norm, as follows:

$$avg(F) = \frac{1}{mn} \sum_{i=1}^{mn} \|F(i)\|,$$

$$avg(G) = \frac{1}{mn} \sum_{i=1}^{mn} \|G(i)\|,$$

$$\|H\|_m = \left[\frac{1}{mn} \sum_{i=1}^{mn} (\|F(i)\|^2 + \|G(i)\|^2) \right]^{1/2},$$

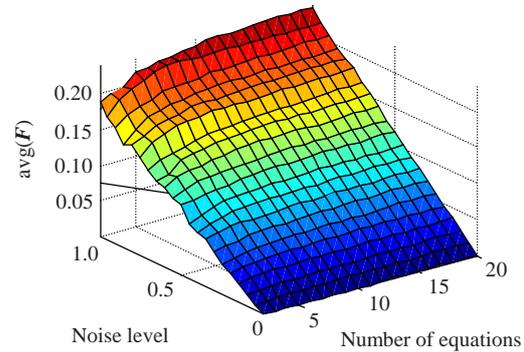
where m is the number of X 's, and n is the number of equations for each X . So there are mn pairs of $(\|F(i)\|, \|G(i)\|)$ for Eqs.(12) and (13).

The mean errors of the rotation matrix and the translation matrix are shown in Figs.2a and 2b, respectively. We then compared our algorithm with the classic approach at the same noise level. Figs.3a and

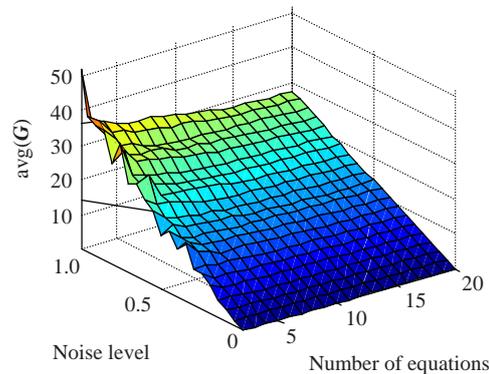
3b show the comparisons at the noise levels of 0.45 and 0.95, respectively. From the experiments we can conclude that: (1) when the number of equations is larger than 4, the error is approximately proportional to the noise level; (2) the solution is robust if the number of equations is larger than 4. The comparison reveals that our approach has a higher precision.

We implemented the real experiment with a calibration object (see Fig.1) and simulated the robot to grasp the object. Then our approach was again compared with the classic algorithm, and we found that if the times for the captured pictures of the calibration object (in another word, the number of equations) are greater, the results will be more stable. The results indicate that our algorithm has a much higher precision than the classic approach (Fig.4).

In addition, the experiments show that our algorithm can directly use linear decomposition without the requirement that A and B have the same rotation angle; that is, our algorithm can also be applied in non-rigid transformations for A and B .



(a)



(b)

Fig.2 Error relationship between the noise level and the number of equations. (a) Analysis of the rotation error; (b) Analysis of the translation error

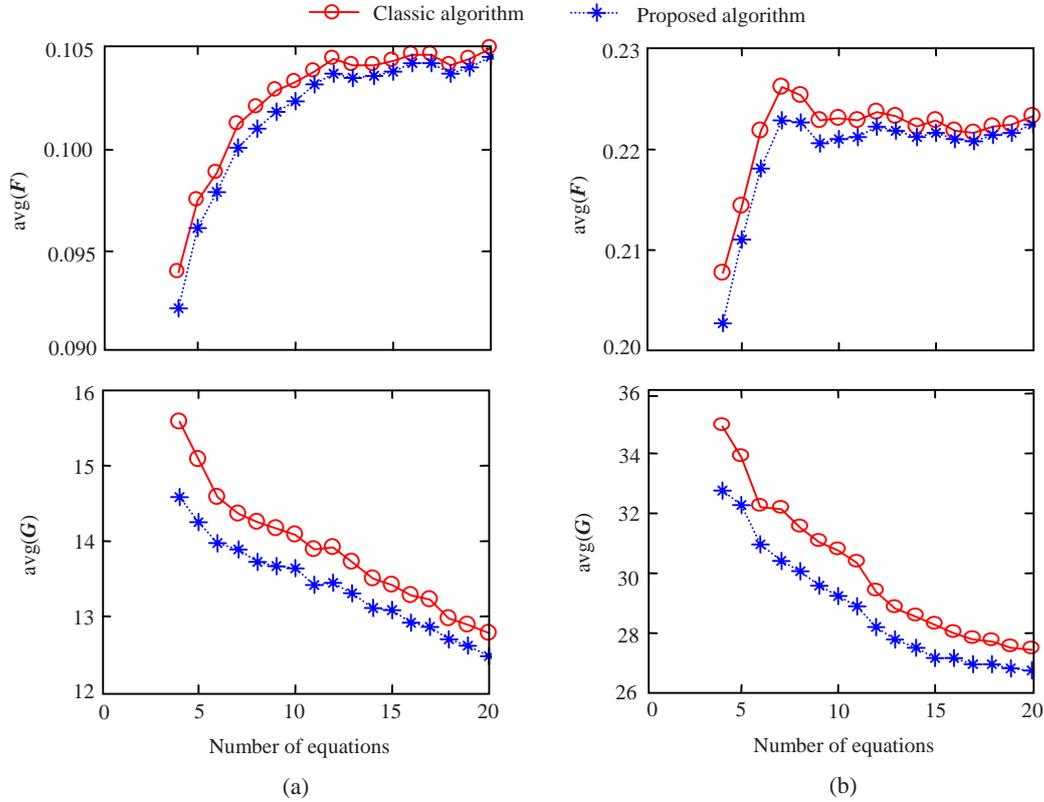


Fig.3 Comparison of the mean of rotation error and translation error between our algorithm and the classic approach at the same noise level. (a) At noise level 0.45; (b) At noise level 0.95

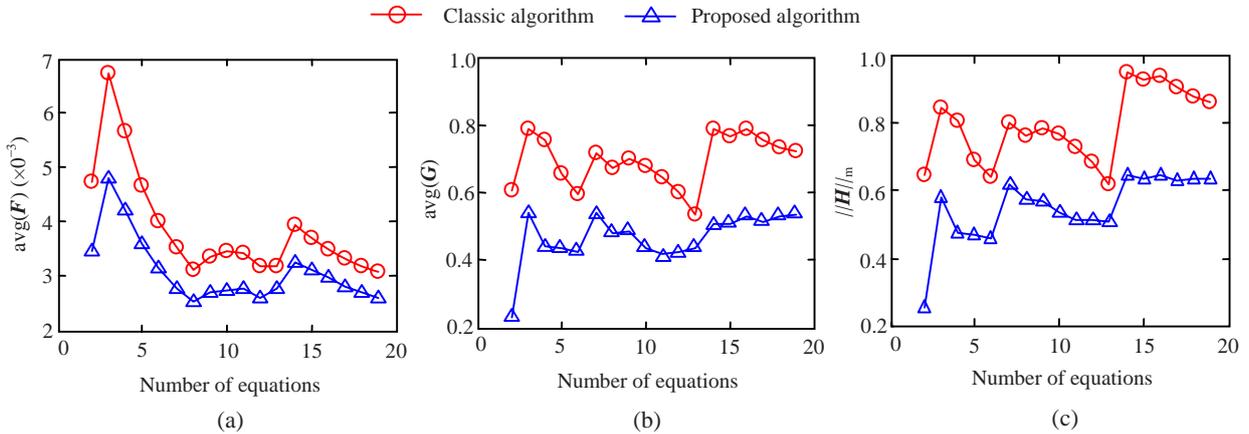


Fig.4 Comparison of the classic approach and ours. (a) Mean of the rotation error norm; (b) Mean of the translation error norm; (c) Mean residual norm

We also analyzed another computationally efficient approach—Park’s method (Park and Martin, 1994), which is a two-step linear solution, using the Lie theory. The first step is to use linear least squares to achieve the rotation part of matrix \mathbf{X} , and the second step is to apply the result of the first step to solve

the translation vector by a linear approach. The advantages of Park’s method are three-fold. Firstly, this method is also applicable when \mathbf{A} and \mathbf{B} are measured in the presence of noise, that is, \mathbf{A} and \mathbf{B} can be rigid transformations with different rotation angles. Secondly, the method is computationally efficient in

obtaining the linear closed-form solution in Lie group. Thirdly, the approach is immune to measure scales such as inch, centimeters. However, Park's approach requires that A and B should be rigid transformations; otherwise, $\log A$ and $\log B$ cannot be calculated from A and B . Like other classic methods, the rotation error in the first stage can propagate to the second stage. Therefore, Park's method cannot be considered the most optimal approximation.

CONCLUSION

In this paper, we use Lemmas 1~3 to give the detailed procedure of the solution to $AX=XB$. By employing optimal orthogonal estimation for the rotation matrix B , the whole procedure does not require the rotation angles of A and B to be the same. Notably this meets the needs in a real environment where there is some difference between A and B . The reason lies in that we use linear decomposition operations on A and B and do not need to calculate the screw vector by decomposing A and B , and thus our approach can be easily extended to non-rigid matrices A and B . We propose an approach to estimating the optimal orthogonal matrix for λB with $\det(\lambda B) \neq 0$, which leads to great reduction in the error caused by noise. Experimental results show that our algorithm is more precise and robust than the previous approaches, especially when A and B are non-rigid transformations and measured in the presence of noise.

Our algorithm has been successfully applied to a robot grasping project supported by the National Natural Science Foundation of China. Our future research will focus on solving the rotation and

translation parts of the matrix X at the same time to further reduce the rotation and translation errors; for example, we can initialize appropriate optimization objectives for the rotation and translation parts of X .

References

- Daniilidis, K., 1999. Hand-eye calibration using dual quaternions. *Int. J. Rob. Res.*, **18**(3):286-298. [doi:10.1177/02783649922066213]
- Fassi, I., Legnani, G., 2005. Hand to sensor calibration: a geometrical interpretation of the matrix equation $AX=XB$. *J. Rob. Syst.*, **22**(9):497-506. [doi:10.1002/rob.20082]
- Park, F.C., Martin, B.J., 1994. Robot sensor calibration: solving $AX=XB$ on the Euclidean group. *IEEE Trans. on Rob. Autom.*, **10**(5):717-721. [doi:10.1109/70.326576]
- Shiu, Y.C., Ahmad, S., 1989. Calibration of wrist-mounted robotics sensors by solving homogeneous transform equations of the form $AX=XB$. *IEEE Trans. on Rob. Autom.*, **5**(1):16-27. [doi:10.1109/70.88014]
- Tsai, R., Lenz, R., 1989. A new technique for fully autonomous and efficient 3D robotics hand/eye calibration. *IEEE Trans. on Rob. Autom.*, **5**(3):345-358. [doi:10.1109/70.34770]
- Zhang, Z.Y., 2000. A flexible new technique for camera calibration. *IEEE Trans. on Pattern Anal. Machine Intell.*, **22**(11):1330-1334. [doi:10.1109/34.888718]
- Zhao, Z., Liu, Y., 2006. Hand-eye Calibration Based on Screw Motions. 18th Int. Conf. on Pattern Recognition, p.1022-1026. [doi:10.1109/CPR.2006.620]
- Zhuang, H., 1998. Hand/eye calibration for electronic assembly robots. *IEEE Trans. on Rob. Autom.*, **14**(4):612-616. [doi:10.1109/70.704231]
- Zhuang, H., Roth, Z.S., 1991. Comments on "Calibration of wrist-mounted robotic sensors by solving homogeneous transform equations of the form $AX=XB$ ". *IEEE Trans. on Rob. Autom.*, **7**(6):877-878. [doi:10.1109/70.105398]
- Zhuang, H., Roth, Z.S., 1995. Simultaneous calibration of a robot and a hand-mounted camera. *IEEE Trans. on Rob. Autom.*, **11**(5):649-660. [doi:10.1109/70.466601]