Calibration-Free Robot–Sensor Calibration approach based on Second-Order Cone Programming

Li Wei^{1,*}, LU Naiguang^{1,2}, Dong Mingli², and Lou Xiaoping²

¹Beijing University of Posts & Telecommunications, Beijing 100876, China

²Beijing Information Science & Technology University, Beijing 100192, China

Abstract. In order to overcome the restrictions of traditional robot-sensor calibration method which solve the tool-camera transformation and robot-world transformation rely on calibration target, a calibration-free approach that solve the robot-sensor calibration problem of the form $AX = YB$ based on Second-Order Cone Programming is proposed. First, a Structure-From-Motion approach was used to recover the camera motion matrix up to scaling. Then rotation and translation matrix in calibration equation were parameterized by dual quaternion theory. Finally, the Second-Order Cone Programming method was used to simultaneously solve the optimal solution of camera motion matrix scale factor, the robot-world and hand-eye relation. The experimental results indicate that the calibration precision of rotation relative error is 3.998% and the translation relative error is 0.117% in defect of calibration target as 3D benchmark. Compared with similar methods, the proposed method can effectively improve the calibration accuracy of the robot-world and hand-eye relation, and extend the application field of robot-sensor calibration method.

1 Introduction

In order to relate the measurements made by a camera mounted on a robotic gripper to the tool's coordinate frame, a homogeneous transformation from the tool to the camera needs to be determined. This problem usually called robot-sensor calibration and it is divided into two forms: $AX = XB$ and $AX = YB$ [1]. Figure 1 illustrates schematically the geometry of $AX = YB$ robot-sensor calibration, which defines a problem in which the relationship between the world frame and robot base frame is unknown in addition to the unknown relationship between the robot tool and the rigidly mounted sensor, and it has been well studied in [2–6]. The common aspect of all the existing methods is that they do not work with the camera measurements directly, but rather with the camera poses derived from them by observing a known calibration target. Since the calibration target has known dimensions, camera poses with correct scale can be obtained [7]. However, there are many situations when using an accurately manufactured calibration target is not convenient or is not possible at all. Indeed, using a calibration target in applications such as mobile robotics or endoscopy may be unacceptable due to the restrictions in the limited on-board weight or in the strict sterile condition [8].

To solve this tough problem, later on, a method for calibration-free robot-sensor calibration based on the Structure-From-Motion (SFM) was proposed by Andreff [9]. The modified method which approached a wider problem named extended robot-sensor calibration, which made use of SFM to recover the unknown camera poses.

Since SFM can recover camera poses up to scale, the formulation introduced an explicit scaling factor to the robot-sensor calibration equation. Some similar methods were presented in [10, 11], where a scaling factor was included into branch and bound traversal and angular reprojection error formulation [12].

Recently, a number of globally optimal solutions to common structure and motion geometrical problems in computer vision were introduced [13]. Second-Order Cone Programming (SOCP) is actually a kind of global optimization, which is not necessary to find an initial value for the optimization process. In this paper we presented the form $AX = YB$ robot-sensor calibration approach without calibration target based on the SFM. First, we estimated the rotational part of the robot-sensor calibration separately using dual quaternion method. Next, we used SOCP to estimate the translational part based on SFM. Finally, relevant experiment system was established to validate the calibration accuracy. The experiment proved that our method has the advantages of simple operation, low cost and wide applicability.

2 Problem formulation

Suppose a camera has been rigidly attached to a robot's gripper, as shown in Figure 1, A represents the transformation from camera to target world, B represents the transformation from tool to robot base. The unknown X represents the fixed homogeneous transformation between tool and camera, and the unknown Y represents the fixed homogeneous transformation from robot base

* Corresponding author: liweikilary@163.com

to target world. The robot-sensor calibration model can be written as

$$AX = YB \quad (1)$$

Fig. 1. Experimental setup for robot-sensor calibration.

Let R_A , R_B , R_X and R_Y denote the respective 3×3 rotation matrices of A , B , X and Y , and let t_A , t_B , t_X and t_Y denote the respective three-dimensional translational vectors. Equation (1) can thus be rewritten as.

$$\begin{bmatrix} R_A & t_A \\ 0^T & 1 \end{bmatrix} \begin{bmatrix} R_X & t_X \\ 0^T & 1 \end{bmatrix} = \begin{bmatrix} R_Y & t_Y \\ 0^T & 1 \end{bmatrix} \begin{bmatrix} R_B & t_B \\ 0^T & 1 \end{bmatrix} \quad (2)$$

One may easily decompose equation (2) into a rotation equation and a position equation

$$\begin{aligned} R_A R_X &= R_Y R_B \\ R_A t_X + t_A &= R_Y t_B + t_Y \end{aligned} \quad (3)$$

In the method proposed in [2], the first equation is solved by least square minimization of a linear system, which obtained by using the quaternion algebra representation of the rotation matrices. Once R_Y is known, the second equation about t_X and t_Y is easily solved with linear least squares techniques.

3 Method modification

3.1. Formulation by orthonormal matrix

Consider camera poses with up to scaling factor λ using SFM, equation (1) homogeneous transformations $A(\lambda)$, B , X and Y are rewritten by the following relation

$$\begin{aligned} R_A R_X &= R_Y R_B \\ R_A t_X + \lambda t_A &= R_Y t_B + t_Y \end{aligned} \quad (4)$$

Note that Equation (4) can be solved for unknown rotational part R_X and R_Y regardless of the value of λ , rotational part can be obtained by standard methods. Once R_Y is known, we can restructure above equation for all robot motions ($i=1, 2, \dots, N$) as the linear system

$$\underbrace{\begin{bmatrix} R_A & -I_{3 \times 3} & t_A \end{bmatrix}}_C \begin{bmatrix} t_X \\ t_Y \\ \lambda \end{bmatrix} = \underbrace{\begin{bmatrix} R_Y t_B \\ d \end{bmatrix}}_d \quad (5)$$

x

Then we can build the error function about unknown translational part and scaling factor

$$f(x) = \|Cx - d\|_2 \quad (6)$$

It is easily solvable using non-linear optimization such as (Levenberg Marquardt, LM). Considering the presence of the measurement noise may be increase the risk of convergence to the local minima, we have to introduce dual quaternions theory parameterize the above error function.

3.2. Formulation by dual quaternions

This section shows how the estimation of rotation, translation and scale factor can be formulated using dual quaternions. As quaternions are only a representation for 3-D rotations, but dual quaternions treat rotations and translations in a unified way. By using dual quaternions parameterization, the equation (4) can be transformed to

$$\begin{aligned} a \bullet q_x &= q_y \bullet b \\ \lambda a' \bullet q_x + a \bullet q'_x &= q_y \bullet b' + q'_y \bullet b \end{aligned} \quad (7)$$

Where all symbols are denoted by quaternions, and there are relationships in dual quaternions that $2q' = q \bullet tq$. Using the multiplication of quaternions, and performing some algebraic operations, the first rotation equation can be concluded that

$$\underbrace{\begin{bmatrix} M(a) & -M(b)^+ \end{bmatrix}}_{V(a,b)} \underbrace{\begin{bmatrix} q_x \\ q_y \end{bmatrix}}_{q(x,y)} = 0 \quad (8)$$

The above 4×4 quaternion matrix $M(a)$ and $M(b)^+$ can be respectively presented as

$$M(a) = \begin{bmatrix} a_0 & -a_1 & -a_2 & -a_3 \\ a_1 & a_0 & -a_3 & a_2 \\ a_2 & a_3 & a_0 & -a_1 \\ a_3 & -a_2 & a_1 & a_0 \end{bmatrix} \quad M(b)^+ = \begin{bmatrix} b_0 & -b_1 & -b_2 & -b_3 \\ b_1 & b_0 & b_3 & -b_2 \\ b_2 & -b_3 & b_0 & b_1 \\ b_3 & b_2 & -b_1 & b_0 \end{bmatrix}$$

Where $V(a, b)$ is a 4×8 matrix defined for each pair of quaternions (a, b) . The unit quaternion q_x and q_y for the rotation is determined in the least square sense by the SVD decomposition. The translation equation for quaternion q'_x, q'_y and scaling factor λ formulation is also obtained as follows

$$\underbrace{\begin{bmatrix} M(a) & -M(b)^+ & M(a') \end{bmatrix}}_{C(a',b')} \begin{bmatrix} q'_x \\ q'_y \\ \lambda \end{bmatrix} = \underbrace{M(b')^+}_{d_i} q_y \quad (9)$$

x

Where $C(a', b')$ is a 4×9 matrix defined for each pair of quaternions (a', b') , d_i is a three-dimensional vector.

4 Solution

This section will show how to convert equation (9) to a problem of Second-Order Cone Program. It was observed in [14] that general SOCP can be expressed in the form

$$\begin{aligned} & \text{minimize } c^T x \\ & \text{subject to } F_i x = g_i, x \in \mathbf{K}, i = 1, 2, \dots, N \end{aligned} \quad (10)$$

Where $\mathbf{K}$ is assumed to be quadratic cones such as equation (11) and equation (12), and x_i denote a variable quantity of unknown parameter x . Moreover, we assume that coefficient matrix $F_i \in \mathbf{R}_{m \times n}$, $g_i \in \mathbf{R}_m$ and $c \in \mathbf{R}_n$ have conforming dimensions. A primal solution x is said to be feasible if it satisfies all the constraints of equation (10), otherwise it is infeasible.

Quadratic cone

$$\mathbf{K} = \left\{ x_i \in \mathbf{R}_n \mid x_1^2 \geq \sum_{i=2}^n x_i^2, x_1 \geq 0 \right\} \quad (11)$$

Rotated quadratic cone

$$\mathbf{K} = \left\{ x_i \in \mathbf{R}_n \mid 2x_1x_2 \geq \sum_{i=3}^n x_i^2, x_1, x_2 \geq 0 \right\} \quad (12)$$

Many optimization problems can be expressed in the above form. Some examples are linear, convex quadratic, and convex quadratic constrained optimization. Other examples are the problem of minimizing the sum of norms and robust linear optimization. In this paper we used it to solve the robot-sensor calibration convex optimization problem. Suppose L_2 norm optimization problems (9) can be written in the following form

$$\begin{aligned} & \min_x \max_i \|C_i x - d_i\|_2 \\ & \text{subject to } q_x^T q_x' = 0, q_y^T q_y' = 0 \end{aligned} \quad (13)$$

It is obviously seen that this is a convex optimization problem and it can be transformed by introducing an additional variable δ into a Second-Order Cone Program problem of the below form

$$\begin{cases} \text{minimize } \delta \\ \text{subject to } \|C_i x - d_i\|_2 \leq \delta, i = 1, \dots, N \\ q_x^T q_x' = 0, q_y^T q_y' = 0 \end{cases} \quad (14)$$

It is easily solvable using commonly available Matlab toolboxes, for example SeDuMi or CVX [15].

5 Experiments

Figure 2(a) shows an overview of the setup used for the real data experiment. DENS0 VS-G serial manipulator

with DAHENG MER-500-14GM/GC Industrial CMOS camera and COMPUTAR 12mm lens is used to acquire the data for the experiment. The robot was instructed to move the gripper along the surface of a sphere, the sphere's radius was approximately 600mm, and was centred in the middle of the scene objects. The position of the gripper is adjusted with ten different locations. The camera resolution is 2592×1944 pixels. Two image sets for two different scenes were acquired—a scene with a calibration target used for obtaining internal camera calibration and a scene with general objects to show the precision differences of the calibration-free method over the classical hand-eye calibration approach that rely on a known calibration target.

(a) DENS0 manipulator (b) 3D model output
Fig. 2. Real data experiment.

(a) Calibration (b) Toys (c) Books
Fig. 3. Sample images of test scenes taken by the camera mounted on the gripper of the DENS0 robot.

5.1. Calibration scene

Because of no ground truth value is available for verification, to evaluate the correctness of the solution obtained by our proposed method, we have to compare it with those obtained by calibration target with the same data. Calibration scene, as Figure 3(a) shows, was firstly used for internal camera calibration [16] and method [3] (M1) was used to finish robot-sensor calibration that relied on a known calibration target. However since the calibration procedure outputs the matched feature points

as its by-product, then the camera motion homogeneous transformation $A(\lambda)$ was obtained from the correspondences and the robot homogeneous transformation B was obtained from the known tool-to-base transformation. Finally, we used the same correspondence points for robot-sensor calibration base on structure-from-motion (SFM) approach with linear kronecker product [4] (M2), nonlinear dual quaternion [5] (M3) and our SOCP method (M4), then compared their results to initial robot-sensor transformation (M1). To measure the errors in rotation and translation, we choose the common relative error as

$$e_r = \left(\sum \|R_A R_X - R_Y R_B\|^2 \right)^{1/2}$$

$$e_t = \left(\frac{\sum \|R_A t_X + \lambda t_A - R_Y t_B - t_Y\|^2}{\sum \|R_A t_X + \lambda t_A\|^2} \right)^{1/2} \quad (15)$$

The average results obtained for ten repeated experiments are given in Figure 4. It can be seen that M1 gives the smallest error in rotation and translation due to the exact spatial constraints. As for M2 and M3, they give larger errors as expected since the 3D model is not used. However, the second-order cone programming method (M4) yields a higher accuracy that very closest to the result of the method M1, the calibration precision of rotation relative error is 3.998% and the translation relative error is 0.117%, which show validity of the proposed method.

(a) Relative rotation error (b) Relative translation error
Fig. 4. Relative error in rotation and translation with images of checkerboard target for four methods.

5.2. General scene

In the second experiment, as Figure 3(b-c) show, calibration target was removed, and general scene was used to show the performance of the method in real-world conditions. SIFT image features and Bundler algorithms [17] were used to obtain the camera poses. Figure 2(b) show the resulting 3D model output from bundler contained 104712 triangulated points and the poses of all the ten cameras. The camera motion homogeneous transformations $A(\lambda)$ were output from Bundler. Due to a high number of correspondences, only every hundredth member of the set of corresponding image points was used. The robot homogeneous transformations B were again obtained from the known tool to base transformations. The 5.1 section calculation results (X_{bar} and Y_{bar}) with calibration target was believed

as true value, and we used linear kronecker product (M2), nonlinear dual quaternion (M3) and our SOCP method (M4) to calculated homogeneous transformation X_{scene} and Y_{scene} , then compared their results to the true value X_{bar} and Y_{bar} . The errors of robot-world and hand-eye relation are defined as follows

$$E_x = \|X_{\text{bar}} - X_{\text{scene}}\|_2 \quad E_y = \|Y_{\text{bar}} - Y_{\text{scene}}\|_2 \quad (16)$$

The results of general scene experiment are given in Table 1. It can be seen that M4 is nearest to the results of the method M1 for the two general scenes among the three methods. Even though there are no space constraints in the general scene, compared with other similar methods our proposed method can perform well.

Table 1. General scene data experiment robot-sensor calibration without calibration target (Unit: mm).

Method	hand-eye relation E_x		robot-world relation E_y	
	Books	Toys	Book	Statue
M2	8.758	5.408	3.601	2.362
M3	7.456	5.018	3.024	2.141
M4	3.979	2.881	0.826	1.161

6 Conclusion

Using a precise calibration target is not possible for many applications of robot-sensor calibration. In such situations, we proposed a robot-sensor self-calibration method, which allow computing the tool-camera transformation and robot-world transformation without the necessity for using a calibration target in order to obtain the camera poses. Instead, SFM approach was used to recover the camera poses up to scaling. However, due to its inherent scale ambiguity, SFM technique will bring an additional over-parameterization to the solution of nonlinear equation. This paper addressed this drawback by formulating the estimation of the robot-sensor displacement as a second-order cone programming optimization problem. Compared with similar robot-sensor calibration methods without a calibration target, our proposed method has maximally guaranteed to obtain the optimal solution, although the cost of releasing the space constraint is a small degradation of the numerical accuracy, the advantages of the proposed method may outweigh this drawback, which break the restrictions of mobile robot and endoscopic surgery robot in the limited on-board weight or in the strict sterile condition.

References

- Shah, Mili, Roger D. Eastman, and Tsai Hong. An overview of robot-sensor calibration methods for evaluation of perception systems. *Proceedings of the Workshop on Performance Metrics for Intelligent Systems. ACM*, 15-20 (2012).

2. Zhuang H Q, Roth Z S, Sudhakar R. Simultaneous robot/world and tool/flange calibration by solving homogeneous transformation equations of the form $AX=YB$. J. *IEEE Transactions on Robotics and Automation*, **10(4)**:549-554 (1994).
3. Shah M. Solving the robot-world/hand-eye calibration problem using the Kronecker product. J. *Journal of Mechanisms and Robotics*, **5(3)**, 031007 (2013).
4. Li A, Wang L, Wu D. Simultaneous robot-world and hand-eye calibration using dual-quaternions and Kronecker product. J. *International Journal of Physical Sciences*, **5(10)**, 1530-1536 (2010).
5. Dornaika F, Horaud R. Simultaneous robot-world and hand-eye calibration. J. *IEEE transactions on Robotics and Automation*, **14(4)**, 617-622 (1998).
6. Heller J, Henrion D, Pajdla T. Hand-eye and robot-world calibration by global polynomial optimization. *IEEE International Conference on Robotics and Automation (ICRA)*, 3157-3164 (2014).
7. Wang Z, Liu Z, Ma Q, et al. Vision-Based Calibration of Dual RCM-Based Robot Arms in Human-Robot Collaborative Minimally Invasive Surgery. J. *IEEE Robotics and Automation Letters*, **3(2)**, 672-679 (2017).
8. Pachtrachai K, Allan M, Pawar V, et al. Hand-eye calibration for robotic assisted minimally invasive surgery without a calibration object. *Intelligent Robots and Systems (IROS)*, 2485-2491 (2016).
9. Andreff N, Horaud R, Espiau B. Robot hand-eye calibration using structure-from-motion. J. *The International Journal of Robotics Research*, **20(3)**, 228-248 (2001).
10. Heller J, Havlena M, Sugimoto A, et al. Structure-from-motion based hand-eye calibration using L_∞ minimization. *Computer Vision and Pattern Recognition (CVPR)*, 3497-3503 (2011)
11. Park Y, Choi Y, Seo Y. Globally optimal camera-and-rotation-sensor calibration with a branch-and-bound algorithm. J. *Applied Optics*, **56(12)**, 3462-3469 (2017).
12. Hartley R I, Kahl F. Global optimization through rotation space search. J. *International Journal of Computer Vision*, **82(1)**, 64-79 (2009).
13. Heller J, Pajdla T. World-base calibration by global polynomial optimization. *3D Vision (3DV)*, **1**, 593-600 (2014).
14. Andersen E D, Roos C, Terlaky T. On implementing a primal-dual interior-point method for conic quadratic optimization. J. *Mathematical Programming*, **95(2)**, 249-277 (2003).
15. Lasserre J B, Löfberg J. GloptiPoly 3: moments, optimization and semidefinite programming. J. *Optimization Methods & Software*, **24(4-5)**, 761-779 (2009).
16. Zhang Z. A flexible new technique for camera calibration. J. *IEEE Transactions on pattern analysis and machine intelligence*, **22(11)**, 1330-1334 (2000).
17. Wu C. Towards linear-time incremental structure from motion. *3DTV-Conference*, 127-134 (2013).