Simultaneous Hand-Eye and Robot-World Calibration by Solving the AX=YB Problem without Correspondence*

Haiyuan Li[1], Qianli Ma[2], Tianmiao Wang[1] and Gregory S. Chirikjian[2]

***Abstract*— Calibration is often an important and necessary step in the use of image-guided systems. In the case of the $AX = YB$ problem, the relative hand-eye (X) and robot-world (Y) transformations must be determined to provide accurate data for use in control. As an added difficulty, the exact correspondence between the streams of sensor data (A's and B's) is typically unknown due to asynchrony in sampling rates and processing time. One common scenario is a constant shift between the two data streams. Therefore, in this paper we present a probabilistic method to simultaneously solve for X and Y without a priori knowledge of the correspondence between the streams of A's and B's. We begin by discussing probability density functions on $SE(3)$ and then use Euclidean-group invariants to obtain an exact solution for X and Y. We then present a method to simultaneously recover X and Y and the correspondence between temporally shifted data sets using a correlation method. Following this, we show how to solve the problem in the case when the data are completely scrambled, corresponding to a complete loss of temporal information. Finally, we numerically simulated the proposed method with asynchronous data and noise added to the stream of B's to verify its efficiency and robustness.**

I. Introduction

Image-guided systems have been widely employed in applications throughout robotics such as robot-assisted surgery, autonomously guided vehicles, etc. Sensors such as a camera, a laser scanner, or an ultrasound probe are usually mounted on the distal end of a robotic manipulator. For a typical "hand-eye" system as described above, the relative transformation between the sensor with respect to the end-effector should be accurately calibrated, and it is often characterized as the well-known **AX=XB** problem. A variation of this problem is the **AX=YB** problem, where both the hand-eye transformation and the pose of the robot base with respect to the world frame need to be calibrated. In a typical environment, the relationships between the sensor frame, robot frame, and world frame are variant and uncertainties exist. Therefore, simultaneous coordinate calibrations have to be determined frequently in order to enable the robot to respond to dynamic environments.

In the **AX=YB** problem, data streams of A's and B's can be respectively obtained via different sensors. The data streams may arrive in an asynchronous fashion due to the different working frequencies of the sensors. This asynchrony causes a shift between the two data streams which can obscure the correspondence between the A's and B's. Moreover, data loss can destroy information about correspondence altogether. In this paper, a novel method is presented to solve for X and Y without a priori knowledge of the correspondence between the A's and B's.

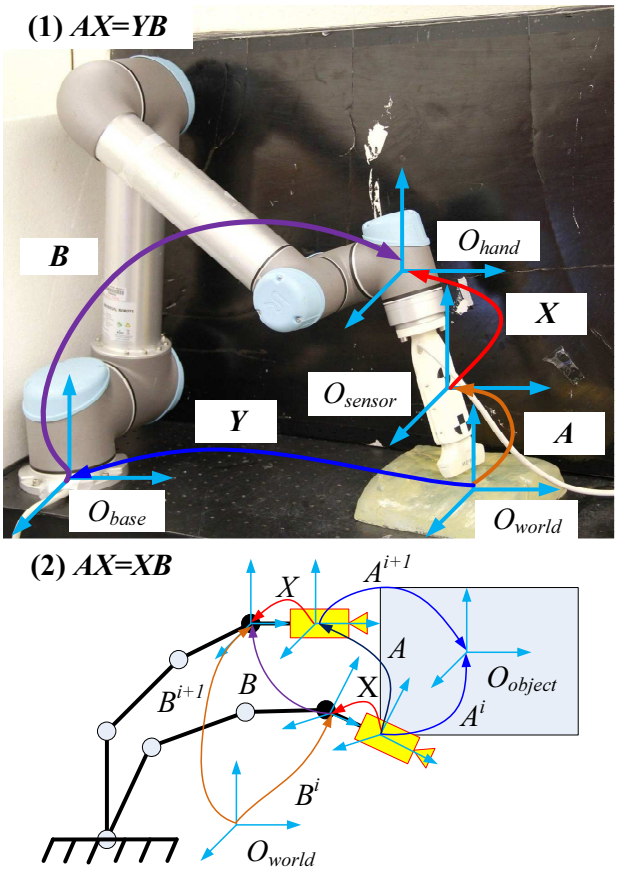


Fig. 1. (1) The hand-eye and robot-world calibration problem formulated as AX=YB. (2) The hand-eye calibration problem formulated as AX=XB. Note: matrices A and B above have different physical meanings in the $AX = XB$ and $AX = YB$ problems. (The Universal Robot pictured above is from the laboratory of Prof. Emad Boctor of Johns Hopkins University).

The hand-eye calibration problem can be modelled as **AX = XB**, where A and B are the homogeneous transformation matrices describing the relative motions of the end-effector and the sensor respectively. As shown in Fig. 1 part (2), $A = A^i(A^{i+1})^{-1}$ and $B = (B^i)^{-1}B^{i+1}$. Given multiple pairs of (A_i, B_i) with correspondence (note that (A_i, B_i)

*This work was partially supported by NSF Grant RI-Medium: 1162095. Mr. Haiyuan Li's partial work is supported by 863 Program of China: 2012AA041402.

[1]Haiyuan Li and Tianmiao Wang are with the School of Mechanical Engineering and Automation, Beihang University, Beijing 100191, China haiyuanli@hotmail.com, itm@buaa.edu.cn

[2]Qianli Ma and Gregory S. Chirikjian are with the Department of Mechanical Engineering, Johns Hopkins University, Baltimore, MD 21218, USA mqianli1@jhu.edu, gregc@jhu.edu

are the relative transformations obtained from the raw data), many deterministic methods have been proposed to solve for X. To the best of the authors' knowledge, Shiu [1] and Tsai [2] are the first to solve the $AX = XB$ sensor calibration problem. The other methods include but are not limited to the quaternion, dual quaternion, screw theory, Lie group theory, motor algebra, convex optimization and gradient descent methods [3]–[11]. All of the methods above assume a priori knowledge of the exact correspondence between A_i and B_i. For data streams $\{A_i\}$ and $\{B_j\}$ that are asynchronous, several methods have been proposed in the literature to solve for X using data without a priori knowledge of the correspondence. These methods assume that there are some corresponding pairs of data in the streams of the $\{A_i\}$ and $\{B_j\}$; however, the exact correspondence is unknown a priori [11]–[13].

Simultaneous estimation of the hand-eye and robot-world transformations has been viewed as the **AX=YB** problem. As shown in Fig. 1 part (1), Y is the transformation from the robot base to the world frame, A denotes the pose of the sensor in the world frame and B is the transformation from the end-effector to its fixed base. The A and B in **AX=YB** are different from those in **AX=XB** where the former uses absolute transformations and the latter uses relative transformations. This problem has been solved by many different methods such as the Kronecker product, quaternion, dual quaternion, and nonlinear optimization methods [14]–[21]. Simulatneous calibration of X and Y can be problematic in that all the methods above assume exact correspondence between $\{A_i\}$ and $\{B_j\}$, which is not the case in the real world, and this is why a simultaneous solution for X and Y in the **AX=YB** problem can be a challenging issue. Another similar problem involves the calibration of multiple robots in terms of hand-eye, tool-flange, and robot-robot transformations, and it is formulated as the **AXB=YCZ** problem [22] which will not be discussed in detail here. In the above methods, the correspondence between A and B is known a priori. In this paper, we focus on one case of the **AX=YB** problem where there is no a priori knowledge of the correspondence between the data streams.

The rest of the paper is organized as follows. In Section II, a novel probabilistic method is presented to solve for eight candidates of X and Y. In Section III, an algorithm involving both a temporal correlation calculation and Euclidean group invariants is proposed to recover the correspondence between $\{A_i\}$ and $\{B_j\}$, which is used to select the optimal solution among the candidates. The simulation results obtained by taking noisy data without correspondence are illustrated in Section V. In Section VI, we briefly discuss the case where one can obtain (X, Y) without recovering the correspondence between the data sets. Finally, conclusions are drawn based on the numerical results and possible future work is pointed out.

II. Solving AX=YB using a probabilistic method on motion groups

In this section, a brief introduction to the concept of probability density functions (PDFs) on the special Euclidean group $SE(3)$ is presented and the probabilistic representation of $AX = YB$ is derived.

Any rigid body transformation matrix can be viewed as a group element of $SE(3)$:

$$H(R,t) = \begin{pmatrix} R & t \\ 0^T & 1 \end{pmatrix} \in SE(3), \quad R \in SO(3) \tag{1}$$

where $SO(3)$ denotes the special orthogonal group, $t \in \mathbb{R}^3$ is a translational vector, T denotes the transpose of a vector or matrix, and H is the symbol for an element of $SE(3)$ (which is a six-dimensional Lie group) represented as a 4×4 homogeneous transformation matrix. The identity element of $SE(3)$ is then the 4×4 identity matrix, I.

Given a large set of pairs $(A_i, B_i) \in SE(3) \times SE(3)$ where $i = 1, \cdots, n$, the following equation is true if the correspondence is known as a priori:

$$A_i X = Y B_i. \tag{2}$$

For a group element $H \in SE(3)$, a Dirac delta function $\delta(H)$ is defined to be finite only at the identity and zero elsewhere,

$$\delta(H) = \begin{cases} +\infty & \text{if} \quad H = I \\ 0 & \text{if} \quad H \neq I, \end{cases} \tag{3}$$

and also satisfies the identity constraint

$$\int_{SE(3)} \delta(H) dH = 1. \tag{4}$$

A shifted Dirac delta function can be defined as $\delta_A(H) = \delta(A^{-1}H)$. Given $K, H \in SE(3)$ and two well-behaved functions f_1 and f_2, their convolution on $SE(3)$ is defined as [23], [24]

$$(f_1 * f_2)(H) = \int_{SE(3)} f_1(K) f_2(K^{-1} \circ H) dK \tag{5}$$

where $\circ$ denotes the group product for $SE(3)$, which is simply matrix multiplication. The integral over $SE(3)$ can be expressed in various coordinates, and both the bounds of the integral and the integration measure dH will take on different appearances that depend on these coordinates. For example, if H is parameterized in terms of the Cartesian coordinates $t = [x, y, z]^T$, and $R = R_{ZXZ}(\alpha, \beta, \gamma)$ is an Euler-angle description of rotations, then the integral over $SE(3)$ is the six-dimensional integral where x, y, z range over $-\infty$ to $+\infty$, α, γ range from 0 to 2π and β ranges from 0 to π. In these coordinates, the integration measure takes the form

$$dH = \sin\beta d\alpha d\beta d\gamma dx dy dz = dR dt.$$

Alternatively, exponential coordinates can be used, in which case the six-dimensional integral over $SE(3)$ and the measure take the form described in [25].

When it is clear that the argument of a function is $H \in SE(3)$, sometimes it will be convenient to abbreviate $f(H)$ as f and $(f_1 * f_2)(H)$ as $f_1 * f_2$ to avoid a proliferation of parentheses.

The convolution operation is bi-linear in the sense that

$$(a_1 f_1 + b_1 f'_1) * f_2 = a_1(f_1 * f_2) + b_1(f'_1 * f_2)$$

and

$$f_1 * (a_1 f_2 + b_1 f'_2) = a_1(f_1 * f_2) + b_1(f_1 * f'_2).$$

Moreover, convolution inherits associativity from the underlying group:

$$f_1 * (f_2 * f_3) = (f_1 * f_2) * f_3.$$

But convolution is generally not commutative, $f_1 * f_2 \neq f_2 * f_1$.

Employing the properties of the δ function, it is straightforward to see that:

$$(f * \delta)(H) = \int_{SE(3)} f(K)\delta(K^{-1} \circ H)dK = f(H). \quad (6)$$

Therefore, for each A_i and B_i, the following equations can be obtained:

$$(\delta_{A_i} * \delta_X)(H) = \delta(A_i^{-1} H X^{-1}) \quad (7a)$$
$$(\delta_Y * \delta_{B_i})(H) = \delta(Y^{-1} H B_i^{-1}). \quad (7b)$$

Using Eq. (2) and Eq. (3), the above two equations can be combined into a single equation as:

$$(\delta_{A_i} * \delta_X)(H) = (\delta_Y * \delta_{B_i})(H). \quad (8)$$

Defining the PDF of $\{A_i\}$ and $\{B_i\}$ as:

$$f_A(H) = \frac{1}{n}\sum_{i=1}^{n} \delta_{A_i}(H) \quad (9a)$$
$$f_B(H) = \frac{1}{n}\sum_{i=1}^{n} \delta_{B_i}(H), \quad (9b)$$

then by using the bi-linearity of convolution, add n instances of Eq. (8), and substitute Eq. (9) into the summation, and we will have:

$$(f_{A_i} * \delta_X)(H) = (\delta_Y * f_{B_i})(H). \quad (10)$$

The convolution of two highly-focused PDFs have some interesting properties that can be used to solve for X. In particular, the mean M and covariance Σ of a PDF $f(H)$ on $SE(3)$ are defined as:

$$\int_{SE(3)} \log(M^{-1}H) f(H) dH = \mathbb{O} \quad (11a)$$
$$\Sigma = \int_{SE(3)} \log^{\vee}(M^{-1}H)[\log^{\vee}(M^{-1}H)]^T f(H) dH \quad (11b)$$

where the explicit expression of the matrix logarithm $\log(H)$ along with its vectorized form $\log^{\vee}(H)$ are given in [26] as

$$\log(H) = \widehat{\mathbf{h}} = \begin{pmatrix} 0 & -h_3 & h_2 & h_4 \\ h_3 & 0 & -h_1 & h_5 \\ -h_2 & -h_1 & 0 & h_6 \\ 0 & 0 & 0 & 0 \end{pmatrix} \quad (12)$$

where $\mathbf{h} = \log^{\vee}(H) \in \mathbb{R}^{6\times 1}$ and $\widehat{\mathbf{h}}$ is the corresponding Lie algebra element, $\widehat{\mathbf{h}} \in se(3)$, such that $H = \exp \widehat{\mathbf{h}}$.

If $f_A(H)$ is given as in Eq. (9), then the corresponding discrete version of the mean M_A and covariance Σ_A will be:

$$\sum_{i=1}^{n} \log(M_A^{-1} A_i) = \mathbb{O} \quad (13a)$$
$$\Sigma = \sum_{i=1}^{n} \log^{\vee}(M_A^{-1} A_i)[\log^{\vee}(M_A^{-1} A_i)]^T. \quad (13b)$$

Given $\{A_i\}$ with the cloud of frames A_i clustering around M_A, an iterative formula can be used for computing M_A [25] as:

$${}^{k+1}M_A = {}^{k}M_A \circ \exp\left[\frac{1}{n}\sum_{i=1}^{n} \log({}^{k}M_A^{-1} \circ A_i)\right]. \quad (14)$$

An initial estimate of the iterative procedure can be chosen as:

$${}^{0}M_A = \exp\left(\frac{1}{n}\sum_{i=1}^{n} \log(A_i)\right).$$

Alternatively, M_A can be obtained by solving a nonlinear optimization problem with the cost function

$$C_1(M_A) = \left\| \sum_{i=1}^{n} \log(M_A^{-1} A_i) \right\|^2.$$

Note, however, that mathematically this is not the same as minimizing

$$C_2(M_A) = \sum_{i=1}^{n} \left\| \log(M_A^{-1} A_i) \right\|^2,$$

though in practice they often are minimized by very close values of M_A.

A similar procedure can be used to compute M_B. Σ_A and Σ_B are then straightforward to compute once M_A and M_B are known.

The mean and covariance for the convolution $(f_1 * f_2)(g)$ of two "highly-focused" functions f_1 and f_2 (i.e., those for which $\|\Sigma_i\| << 1$) are calculated as in [25]:

$$M_{1*2} = M_1 M_2 \quad (15a)$$
$$\Sigma_{1*2} = Ad(M_2^{-1})\Sigma_1 Ad^T(M_2^{-1}) + \Sigma_2, \quad (15b)$$

where

$$Ad(H) = \begin{pmatrix} R & O \\ \hat{t}R & R \end{pmatrix}.$$

Here the "hat" notation when applied to the three-dimensional vector t gives

$$\hat{t} = \begin{pmatrix} 0 & -t_3 & t_2 \\ t_3 & 0 & -t_1 \\ -t_2 & -t_1 & 0 \end{pmatrix} \tag{16}$$

Because X and Y are constant, their corresponding PDF will be $\delta_X(g)$ and $\delta_Y(g)$, of which the mean and covariance are $M_X = X$, $\Sigma_X = \mathbb{O}_{6\times6}$ and $M_Y = Y$, $\Sigma_Y = \mathbb{O}_{6\times6}$, respectively. Therefore, the following equations can be obtained using Eq. (15):

$$M_A X = Y M_B \tag{17a}$$
$$Ad(X^{-1})\Sigma_A Ad^T(X^{-1}) = \Sigma_B. \tag{17b}$$

This is a nonparametric result, meaning that the underlying probability density functions $f_A(H)$ and $f_B(H)$ need not be Gaussian or belong to any other family of parametric distributions. Moreover, the constraint of being highly-focused is not very restrictive because the original data sets $\{A_i\}$ and $\{B_i\}$ can always be transformed into $\{(A_i)^p\}$ and $\{(B_i)^p\}$ for some fractional power $0 < p < 1$ since

$$A_i = XB_iX^{-1} \iff (A_i)^p = X(B_i)^pX^{-1}.$$

Question: this is for AX=XB, but for AX =YB, $A_i = YB_iX^{-1}$ and the above equation doesn't apply.

The smaller the value of p is, the more concentrated the transformed data will be.

The problem of solving the above equations, Eq. (17a) is decomposed into a rotational equation and a translational equation as follows:

$$R_{M_A}R_X = R_Y R_{M_B} \tag{18a}$$
$$R_{M_A}t_X + t_{M_A} = R_Y t_{M_B} + t_Y. \tag{18b}$$

Σ_A and Σ_B can be decomposed into blocks as $\begin{pmatrix} \Sigma_A^1 & \Sigma_A^2 \\ \Sigma_A^3 & \Sigma_A^4 \end{pmatrix}$ and $\begin{pmatrix} \Sigma_B^1 & \Sigma_B^2 \\ \Sigma_B^3 & \Sigma_B^4 \end{pmatrix}$, respectively. Knowing that $X^{-1} = \begin{pmatrix} R_X^T & -R_X^T t_X \\ 0 & 1 \end{pmatrix}$, then the first two blocks of Eq. (17b) can be written as follows:

$$\Sigma_{M_B}^1 = R_X^T \Sigma_{M_A}^1 R_X \tag{19a}$$
$$\Sigma_{M_B}^2 = R_X^T \Sigma_{M_A}^1 R_X \widehat{(R_X^T t_X)} + R_X^T \Sigma_{M_A}^2 R_X. \tag{19b}$$

Because Eq. (19a) is a similarity transformation between $\Sigma_{M_B}^1$ and $\Sigma_{M_A}^1$, they share the same eigenvalues and can be eigendecomposed into $\Sigma_{M_A}^1 = Q_{M_A}\Lambda Q_{M_A}^T$ and $\Sigma_{M_B}^1 = Q_{M_B}\Lambda Q_{M_B}^T$ where Λ is a diagonal matrix whose diagonal elements are the eigenvalues of $\Sigma_{M_A}^1$ (or $\Sigma_{M_B}^1$), and Q_{M_A} (or Q_{M_B}) is a square matrix whose columns are the corresponding eigenvectors. The following equation is obtained after substituting $\Sigma_{M_B}^1$ and $\Sigma_{M_A}^1$ into Eq. (19a):

$$\Lambda = (Q_{M_A}^T R_X^T Q_{M_B})\Lambda(Q_{M_B}^T R_X Q_{M_A}) = P\Lambda P^T \tag{20}$$

where $P = Q_{M_A}^T R_X Q_{M_B}$. Since Q_{M_A} and Q_{M_B} are further constrained to be rotation matrices, the orthogonal matrix P satisfies Eq. (21).

$$\begin{cases} P^T = P^{-1} \\ \det(P) = \pm 1. \end{cases} \tag{21}$$

Combing Eq. (20) and Eq. (21), then an orthogonal matrix P can be one of $\mathcal{P}$ or $-\mathcal{P}$:

$$\mathcal{P} = \left\{ \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \begin{pmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \begin{pmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix}, \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{pmatrix} \right\}. \tag{22}$$

Therefore, there are eight candidates for R_X which can be calculated via $R_X = Q_{M_A}PQ_{M_B}^T$, and the corresponding t_X can be obtained from Eq. (19b). Given known X, Y can be solved from $Y = M_A X M_B^{-1}$. At last, eight candidate pairs of $\{X_k, Y_k\}$ can be obtained as:

$$X_k = \begin{pmatrix} R_{X_k} & t_{X_k} \\ 0^T & 1 \end{pmatrix}, \quad Y_k = \begin{pmatrix} R_{Y_k} & t_{Y_k} \\ 0^T & 1 \end{pmatrix} \tag{23}$$

where $k = 1, 2, ..., 8$.

The problem then becomes selecting the best pair of $\{X_k, Y_k\}$ from the eight candidates. Based on screw theory, it is known that a homogeneous transformation H can be expressed by the four screw parameters $(\theta, d, \mathbf{n}, \mathbf{p})$ as:

$$H = \begin{pmatrix} e^{\theta\hat{n}} & (I_3 - e^{\theta\hat{n}})p + dn \\ 0^T & 1 \end{pmatrix} \tag{24}$$

where θ is the angle of rotation, d is the translation along the rotation axis, n is the unit vector representing the axis of rotation and p is the position of a point on the line relative to the origin of a space-fixed reference frame with $p \cdot n = 0$.

Moreover, $AX_k = Y_k B$ can be written as $AX_k = X_k(X_k^{-1}Y_k B)$. Defining $B^k = X_k^{-1}Y_k B$, we have $AX_k = X_k B^k$. As discussed in [27], [28], for $AX = XB$ problem, there exist two Euclidean-group invariant relationships for each pair of $(A_i, B_i^k)(i = 1, \cdots, n; k = 1, \ldots, 8)$ as follows:

$$\theta_{A_i} = \theta_{B_i^k}, d_{A_i} = d_{B_i^k}. \tag{25}$$

Among the eight pairs (X_k, Y_k), one can find an optimal solution which minimizes the cost function defined as:

$$(X, Y) = \underset{(X_k, Y_k)}{argmin} \frac{1}{n}\sum_{i=1}^{n}(\parallel \theta_{A_i} - \theta_{B_i^k} \parallel + \parallel d_{A_i} - d_{B_i^k} \parallel). \tag{26}$$

Eight candidates of (X_k, Y_k) are calculated using the probabilistic method on $SE(3)$, which doesn't require the correspondence between A_i and B_j to be known. However, the correspondences need to be recovered to pick the optimal (X_k, Y_k). Note that the Euclidean-group invariant relationships in the context of $AX = YB$ problem are still unknown.

Therefore, $AX = YB$ is converted into $AX = XB$ problem to recover the correspondence of data using invariants.

III. Solution with unknown correspondence between A_i and B_j^k

In most cases, the two sets of homogeneous transformations $\{A_i\}$ and $\{B_j\}$ are calculated based on the data obtained from different sensors. Due to the asynchronous timing of the sensor readings, the correspondence between $\{A_i\}$ and $\{B_j\}$ is usually unknown. This section deals with the case where there is a shift between $\{A_i\}$ and $\{B_j\}$, and the Euclidean-group invariants are used to recover the correspondence between the data streams. The advantage of the above probabilistic solution lies in that X and Y can be calculated even if there is no a priori knowledge of the correspondence. However, there are still eight possible candidates of (X_k, Y_k) to choose from and by using Euclidean-group invariants, it is straightforward to determine which is the optimal pair if the correspondence between A_i and B_j^k can be known.

The Discrete Fourier Transform (DFT) decomposes a time-domain signal into its constituent frequencies. The input is a finite list of equally spaced samples of a function. Given a discrete signal consisting of a sequence of N complex numbers $x_0, x_1, \cdots, x_{N-1}$, the DFT is denoted by $X_\kappa = \mathcal{F}(x_n)$ as:

$$X_\kappa = \sum_{n=0}^{N-1} x_n \cdot exp(-i\frac{2\pi}{N} n\kappa). \tag{27}$$

where i here is the imaginary unit.

The Inverse Discrete Fourier Transform (IDFT) is denoted as:

$$x_n = \frac{1}{N} \sum_{n=0}^{N-1} X_\kappa \cdot exp(i\frac{2\pi}{N} n\kappa). \tag{28}$$

The discrete convolution of two sequences f_n and g_n is defined as:

$$(f * g)(\tau) = \sum_{j=0}^{N} f(t_j) g(t_j - \tau). \tag{29}$$

In the convolution theorem, the Fourier transform of a convolution is the product of the Fourier transforms, namely:

$$f * g = \mathcal{F}^{-1}[\mathcal{F}(f) \cdot \mathcal{F}(g)]. \tag{30}$$

The correlation theorem indicates that the correlation function, $Corr(f, g)$, will be larger for a shift vector where the two sequences f_n and g_n share more similar features. The correlation can be obtained based on the convolution theorem. The DFT of $Corr(f, g)$ is equal to the product of the DFT of f_n and the complex conjugate $\mathcal{F}^*$ of the DFT of g_n:

$$Corr(f, g) = f \star g = \mathcal{F}^{-1}[\mathcal{F}(f) \cdot \mathcal{F}^*(g)]. \tag{31}$$

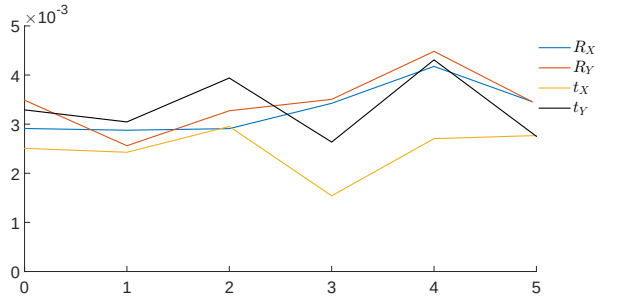


Fig. 2. The translational and rotational errors versus the shift between data streams $\{A_i\}$ and $\{B_i\}$.

Compared to the standard time-domain convolution algorithm, the complexity of the convolution by multiplication in the frequency domain is significantly reduced with the help of the convolution theorem and the Fast Fourier Transform (FFT).

Given two sequences $\{\theta_{A_i}\}$ and $\{\theta_{B_j^k}\}$ corresponding to $\{A_i\}$ and $\{B_j^k\}$, the shift that is needed to recover the data correspondence is obtained as below. Firstly, θ_{Ai} and $\theta_{B_j^k}$ are normalized as:

$$\theta_{1,k} = \frac{(\theta_{A_i} - \mu_A)}{\sigma_A}, \theta_{2,k} = \frac{(\theta_{B_j^k} - \mu_{B^k})}{\sigma_{B^k}} \tag{32}$$

where $\mu_A(\mu_{B^k})$ is the mean of $\theta_{A_i}(\theta_{B_i^k})$ and $\sigma_A(\sigma_{B^k})$ is the standard deviation.

Here, the correlation function $Corr(\theta_{1,k}, \theta_{2,k})$ is the function of the time sequence index n which describes the probability of these two sequences being separated by this particular index. The index corresponding to the maximum of $Corr(\theta_{1,k}, \theta_{2,k})$ indicates the amount of shift τ_{shift} between $\{\theta_{A_i}\}$ and $\{\theta_{B_j^k}\}$.

$$\tau_{shift} = \underset{index}{argmax}(Corr(\theta_{1,k}, \theta_{2,k})) \tag{33}$$

Therefore, the correspondence between the two sequences can be found. The data of θ_{A_i} or d_{A_i} are shifted by $-\tau_{shift}$ to obtain a sequence of new pairs $(\theta_{A_i}(i + \tau_{shift}), \theta_{B_i^k})$ and $(d_{A_i}(i + \tau_{shift}), d_{B_i^k})$, where $max(0, \tau_{shift}) \leq i \leq min(n, n + \tau_{shift})$. The data stream can be shifted back to regain correspondence to synchronize the data streams once the shift is computed, and the optimal solution of X and Y can also be recovered by minimizing the cost function in Eq. (26) using the Euclidean-group invariants as shown in Section II.

IV. Simulation Studies

For the numerical experiments in this section, the rotational and translational errors for X and Y are measured as

$$Error(R_X) = \| \ log^\vee(R_{X_{Solved}}^T R_{X_{true}}) \ \|,$$

$$Error(t_X) = \| \ t_{X_{Solved}} - t_{X_{true}} \ \|,$$

$$Error(R_Y) = \| \ log^\vee(R_{Y_{Solved}}^T R_{Y_{true}}) \ \|,$$

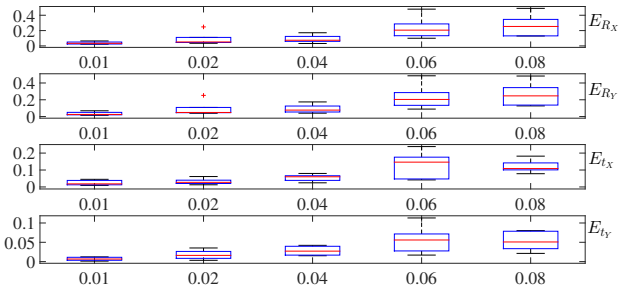
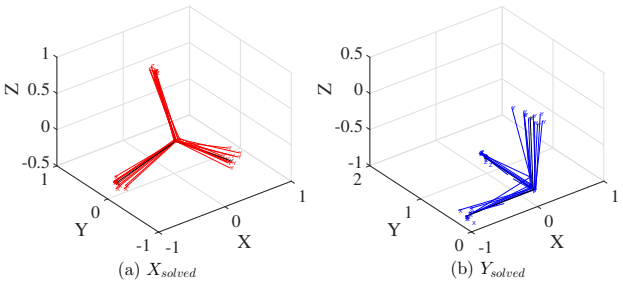


Fig. 3. Box-and-whisker plots of translational and rotational errors versus the covariance noise on data stream $\{B_i\}$.



(a) X_{solved} (b) Y_{solved}

Fig. 4. (a) The solved X (in red) and the actual X (in black) for 10 simulation trials with covariance noise of 0.05 and shift of 2. (b) The solved Y (in blue) and the actual Y (in black) for 10 simulation trials with covariance noise of 0.05 and shift of 2.

and

$$Error(t_Y) = \| t_{Y_{Solved}} - t_{Y_{true}} \|,$$

respectively.

There are multiple ways of generating the data streams $\{A_i\}$ and $\{B_i\}$. One way is to first generate $\{B_i\}$ and then map it to $\{A_i\}$ using $A = YBX^{-1}$. $\{B_i\}$ can be obtained by randomly sampling on the Lie algebra of B from a zero mean multivariate Gaussian distribution as follows:

$$\delta_i \in \mathcal{N}(\mathbf{0}; \Sigma) \subset \mathbb{R}^6 \tag{34a}$$

$$B_i = \exp(\hat{\delta}_i)\exp(\mu) \tag{34b}$$

where the mean $\mu = \mathbf{0} \in se(3)$ and the covariance matrix $\Sigma \in \mathbb{R}^{6\times 6}$ is a diagonal matrix with same diagonal elements σ. The hat operator $\hat{\delta}$ converts a 6 by 1 vector into its corresponding Lie algebra. The data stream $\{A_i\}$ can be easily obtained as described above. After employing the proposed probabilistic method, 8 sets of sequences $(\theta_{A_i}, \theta_{B_i^k})$ and $(d_{A_i}, d_{B_i^k})$ can be obtained respectively where $i = 1, \cdots, 100$ and $k = 1, \ldots, 8$.

If the data stream $\{A_i\}$ is shifted by m units relative to $\{B_i\}$, then the maximum of the cross correlation can be used to recover the shift. After that, we can shift the data stream $\{A_i\}$ back to its original position to recover the correct correspondence with $\{B_i\}$, which will be used to find a correct solution satisfying the Euclidean-group invariants as defined in Eq. (25). Therefore, a unique pair of (X_k, Y_k) $(k = 1, \cdots, 8)$ can be selected to minimize the cost function. In Fig. 2, because the shift between $\{A_i\}$ and $\{B_i\}$ is

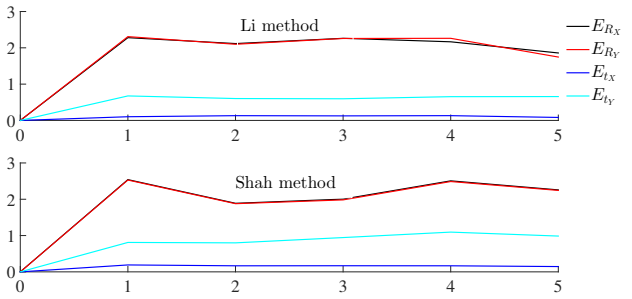


Fig. 5. Orientation and translation errors of X and Y versus shift using Li's and Shah's methods without correspondence.

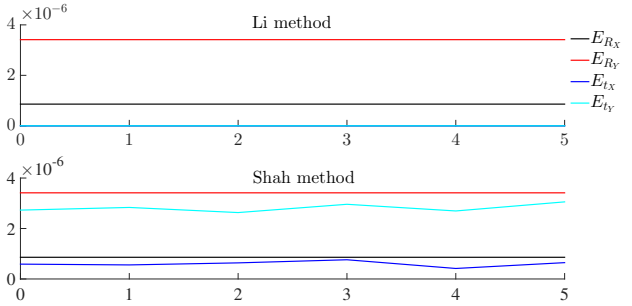


Fig. 6. Orientation and translation errors of X and Y versus shift using Li's and Shah's methods with correspondence.

calculated accurately, the translational and rotational errors fluctuate by only a small amount compared to the errors of the no-shift data streams.

To test the robustness of the proposed method, noises are exerted onto $\{B_i\}$ by employing $B_i^{noise} = B_i \exp(\widehat{\mathbf{x}}_{noise})$, where each element of the Lie Algebra $\mathbf{x}_{noise}$ belongs to the Gaussian distribution defined as $N \sim (\mu_{noise}, \sigma_{noise})$. In Fig. 3, as the covariance noise σ_{noise} increments from 0.01 to 0.08, the errors of R_X, R_Y, t_X, and t_Y increase as shown in the box-and-whisker plot. There are several outliers outside the whiskers, while the median is calculated as the final solved X and Y. Fig. 4 shows the solved (X,Y)s in red and blue with the actual (X,Y) in black with covariance noise of $\sigma = 0.05$ and shift $n = 2$.

The probabilistic method can recover the correspondence between shifted data streams, which is useful for other sensor calibration methods. In the $AX = YB$ problem, there have been many calibration methods developed for solving X and Y given data streams with correspondence. However, few of them considered the cases without correspondence. When data streams of A and B are shifted or asynchronous, most of these methods fail to give a valid solution. To further test the effectiveness of our method, we shift the data sequence of $\{A_i\}$ by $n = 0, 1, 2, 3, 4, 5$ with respect to the data sequence of $\{B_i\}$ such that A_{k+n} "matches" B_k where $k = 1, 2, \cdots, m - n$ and $i = 1, 2, \cdots, m$. We augment other $AX = YB$ solvers with our probabilistic approach by recovering the correspondence between shift data sets. In Li's method [19], X and Y are solved for at the same time, while Shah [20] solved for X and Y in a separate way. As shown in Fig. 5, when dealing with the shifted data streams

$\{A_{k+n}, B_k\}$, the errors on both rotations and translations are significant. After recovering the correspondence between data streams by using the probabilistic method, Li and Shah's methods achieve the same level of performance as shown in Fig. 6.

V. A Brief Case Study With Completely Scrambled Data

In this section, we will briefly discuss the case where $\{A_i\}$ and $\{B_j\}$ are completely scrambled. Unlike the case of shifted data, it is extremely hard to recover the correspondence between two completely scrambled data sets $\{A_i\}$ and $\{B_j\}$. The correlation theorem can't be applied because there is no shift in the scrambled data sets. Euclidean group invariants are not practical either because given $\{A_i\}$ and $\{B_j\}$ both of which have the size of m, there are $m! = m \times m-1 \times \cdots 1$ combinations between the data sets, and it is extremely computationally intensive to test all the combinations. Without recovering the correspondence between the data sets, it is impossible to choose the optimal solution from the eight candidates of $\{X_k, Y_k\}$.

In the above approach, we used Eq. (17b) to calculate X_k and Eq. (17a) to obtain the corresponding Y_k. However, we now show that one can calculate the eight candidates of Y independently and employ Eq. (17a) as a constraint to filter out the optimal $\{X, Y\}$ pair.

Given the equation $AX = YB$, apply an inverse on both sides of the equation and we will have $B^{-1}Y^{-1} = X^{-1}A^{-1}$. Following the same derivations from Eq. (6) to Eq. (17b), we have:

$$M_{B^{-1}}Y^{-1} = X^{-1}M_{A^{-1}} \tag{35a}$$

$$Ad(Y)\Sigma_{B^{-1}}Ad^T(Y) = \Sigma_{A^{-1}}. \tag{35b}$$

Similarly, Eq. (35b) can give eight candidates of Y^{-1}, or equivalently, the eight candidates of Y. Let X_{k1} where $k1 = 1, 2, \cdots, 8$ denote the Xs obtained from Eq. (17b) and Y_{k2} where $k2 = 1, 2, \cdots, 8$ denote the Ys obtained from Eq. (35b), and we can use Eq. (17a) and Eq. (35a) to form a minimization problem as:

$$\min_{k1,k2} ||M_A X_{k1} - Y_{k2} M_B||_F + ||M_B^{-1} Y_{k2}^{-1} - X_{k1}^{-1} M_A^{-1}||_F \tag{36}$$

which can give the optimal (X_{k1}, Y_{k2}) pair. We call this approach the *prob* method, and compare it with Li's method for testing its effectiveness of handling scrambled data sets. For simplicity, we use Eq. (35a) and Eq. (35b) to generate $\{B_i\}$, whereas compute $\{A_i\}$ using $A_i = X^{-1}YB_i$ without exerting noise on B_i. Then $\{A_i\}$ is scrambled at each percentage from 0% up to 100%. 50 times of simulations are performed for each percentage rate and the same error metrics are used as in Section VI. As shown in Fig. (7) and Fig. (8), as the percentage of scrambled data goes up, the errors in rotation and translation for Li's method gradually diverge, while the errors for the *prob* method are very stable and small. This shows the significant advantage of the probabilistic method in handling disordered data sets. However, Li's method is still more accurate when the exact correspondence is known between $\{A_i\}$ and $\{B_i\}$.

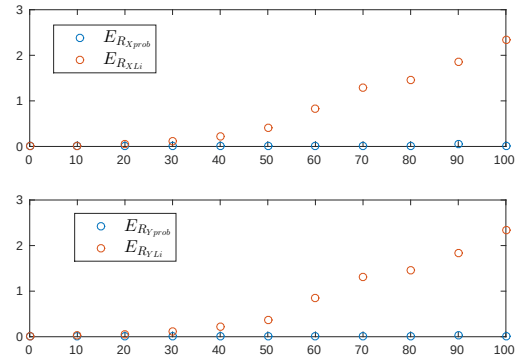


Fig. 7. Rotation error in X and Y v.s. scrambling rate for the *prob* and Li's methods

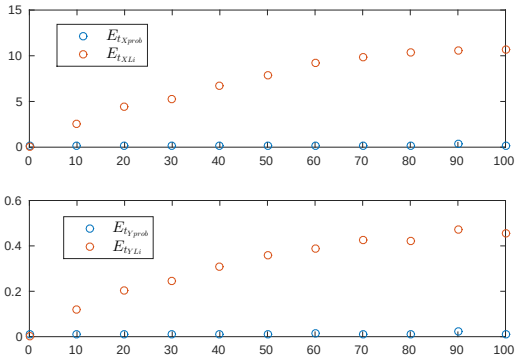


Fig. 8. Translation error in X and Y v.s. scrambling rate for the *prob* and Li's methods

VI. Conclusions

In this paper, we developed a probabilistic approach to simultaneously obtain X and Y in the $AX = YB$ sensor calibration problem. Without a priori knowledge of the correspondence between $\{A_i\}$ and $\{B_j\}$, the proposed probabilistic method on Lie groups is used to constrain the possible solutions of X and Y to eight pairs of candidates. Given shifted data streams of $\{A_{i+s}\}$ and $\{B_i\}$, using the correlation theorem with Euclidean-group invariants, the correspondence is recovered to determine the correct solution among the eight candidates. In the numerical simulation, the method performs well with different sets of data samples. Lastly, we brought up a new approach to deal with completely disordered data sets and show its effectiveness in simulation.Future work will be to improve the *prob* method and investigate on its performance dealing with noisy scrambled data sets.

ACKNOWLEDGMENT

Chirikjian's contribution to this material is based upon work supported by (while serving at) the National Science Foundation (under the IR/D Program). Any opinion, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation. We thank Mr. Joshua Davis for proofreading of this paper.

REFERENCES

[1] Y. C. Shiu and S. Ahmad, "Calibration of wrist-mounted robotic sensors by solving homogeneous transform equations of the form ax= xb," *IEEE Transactions on Robotics and Automation*, vol. 5, no. 1, pp. 16–29, 1989.

[2] R. Y. Tsai and R. K. Lenz, "A new technique for fully autonomous and efficient 3d robotics hand/eye calibration," *IEEE Transactions on Robotics and Automation*, vol. 5, no. 3, pp. 345–358, 1989.

[3] C.-C. Wang, "Extrinsic calibration of a vision sensor mounted on a robot," *IEEE Transactions on Robotics and Automation*, vol. 8, no. 2, pp. 161–175, 1992.

[4] F. C. Park and B. J. Martin, "Robot sensor calibration: solving ax= xb on the euclidean group," *IEEE Transactions on Robotics and Automation*, vol. 10, no. 5, 1994.

[5] R. Horaud and F. Dornaika, "Hand-eye calibration," *The International Journal of Robotics Research*, vol. 14, no. 3, pp. 195–210, 1995.

[6] K. Daniilidis, "Hand-eye calibration using dual quaternions," *The International Journal of Robotics Research*, vol. 18, no. 3, pp. 286–298, 1999.

[7] E. Bayro-Corrochano, K. Daniilidis, and G. Sommer, "Motor algebra for 3d kinematics: The case of the hand-eye calibration," *Journal of Mathematical Imaging and Vision*, vol. 13, no. 2, pp. 79–100, 2000.

[8] J. R. Rovelo, S. H. Garcia, and E. B. Corrochano, "Geometric hand-eye calibration for an endoscopic neurosurgery system," in *IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, 2008, pp. 1418–1423.

[9] I. Fassi and G. Legnani, "Hand to sensor calibration: A geometrical interpretation of the matrix equation ax= xb," *Journal of Robotic Systems*, vol. 22, no. 9, pp. 497–506, 2005.

[10] Z. Zhao, "Hand-eye calibration using convex optimization," in *IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, 2011, pp. 2947–2952.

[11] M. K. Ackerman, A. Cheng, E. Boctor, and G. Chirikjian, "Online ultrasound sensor calibration using gradient descent on the euclidean group," in *IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, 2014, pp. 4900–4905.

[12] M. K. Ackerman and G. S. Chirikjian, "A probabilistic solution to the ax= xb problem: Sensor calibration without correspondence," in *Geometric Science of Information*. Springer, 2013, pp. 693–701.

[13] M. K. Ackerman, A. Cheng, and G. Chirikjian, "An information-theoretic approach to the correspondence-free ax= xb sensor calibration problem," in *IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, 2014, pp. 4893–4899.

[14] H. Zhuang, Z. S. Roth, and R. Sudhakar, "Simultaneous robot/world and tool/flange calibration by solving homogeneous transformation equations of the form ax= yb," *IEEE Transactions on Robotics and Automation*, vol. 10, no. 4, pp. 549–554, 1994.

[15] F. Dornaika and R. Horaud, "Simultaneous robot-world and hand-eye calibration," *IEEE Transactions on Robotics and Automation*, vol. 14, no. 4, pp. 617–622, 1998.

[16] R. L. Hirsh, G. N. DeSouza, and A. C. Kak, "An iterative approach to the hand-eye and base-world calibration problem," in *IEEE International Conference on Robotics and Automation (ICRA)*, vol. 3. IEEE, 2001, pp. 2171–2176.

[17] F. Ernst, L. Richter, L. Matthäus, V. Martens, R. Bruder, A. Schlaefer, and A. Schweikard, "Non-orthogonal tool/flange and robot/world calibration," *The International Journal of Medical Robotics and Computer Assisted Surgery*, vol. 8, no. 4, pp. 407–420, 2012.

[18] K. H. Strobl and G. Hirzinger, "Optimal hand-eye calibration," in *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE, 2006, pp. 4647–4653.

[19] A. Li, L. Wang, and D. Wu, "Simultaneous robot-world and hand-eye calibration using dual-quaternions and kronecker product," *International Journal of Physical Sciences*, vol. 5, no. 10, pp. 1530–1536, 2010.

[20] M. Shah, "Solving the robot-world/hand-eye calibration problem using the kronecker product," *Journal of Mechanisms and Robotics*, vol. 5, no. 3, p. 031007, 2013.

[21] J. Heller, D. Henrion, and T. Pajdla, "Hand-eye and robot-world calibration by global polynomial optimization," in *IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, 2014, pp. 3157–3164.

[22] J. Wang, L. Wu, M. Q.-H. Meng, and H. Ren, "Towards simultaneous coordinate calibrations for cooperative multiple robots," in *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE, 2014, pp. 410–415.

[23] G. S. Chirikjian, *Stochastic Models, Information Theory, and Lie Groups, Volume 2: Analytic Methods and Modern Applications*. Springer Science & Business Media, 2011, vol. 2.

[24] G. Chirikjian and A. Kyatkin, *Harmonic Analysis for Engineers and Applied Scientists*. Dover, 2016.

[25] Y. Wang and G. S. Chirikjian, "Nonparametric second-order theory of error propagation on motion groups," *The International Journal of Robotics Research*, vol. 27, no. 11-12, pp. 1258–1273, 2008.

[26] R. Murray, Z. Li, and S. Sastry, *A mathematical introduction to robotic manipulation*. CRC Press, 1994.

[27] H. Chen, "A screw motion approach to uniqueness analysis of head-eye geometry," in *IEEE Conference on Computer Vision and Pattern Recognition*. IEEE, 1991, pp. 145–151.

[28] M. K. Ackerman, A. Cheng, B. Shiffman, E. Boctor, and G. Chirikjian, "Sensor calibration with unknown correspondence: Solving ax= xb using euclidean-group invariants," in *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE, 2013, pp. 1308–1313.