Simultaneous Hand–Eye/Robot–World/Camera–IMU Calibration

Jin Wu [ID], *Member, IEEE*, Miaomiao Wang [ID], *Member, IEEE*, Yi Jiang [ID], *Member, IEEE*,
Bowen Yi [ID], *Member, IEEE*, Rui Fan [ID], *Member, IEEE*, and Ming Liu [ID], *Senior Member, IEEE*

Abstract—The problem of calibrating an extrinsic parameter between a camera and an inertial measurement unit (IMU) using an industrial robotic manipulator has been studied. This generates a result of hand–eye/robot–world/camera–IMU calibration in a simultaneous fashion. The developed method is free of inertial integration over time and, thus, is robust to uncertain IMU biases. It is derived that the problem can be solved via a simultaneous optimization of hand–eye/robot–world/camera–IMU transformations. The resulted optimization is highly nonconvex on the special Euclidean group, and we give globally optimal solutions. Experiments verify that the proposed method is capable of estimating accurate calibration parameters. Comparative studies between representatives show the global optimality of the proposed method. The new simultaneous method is capable of conducting calibration of a robot/camera/IMU combination. The designed method guarantees the global optimality; thus, the accuracy is ensured. The developed globally optimal solutions will also be computationally efficient on modern industrial computers. Finally, we show that the proposed method can give accurate calibration results for a stereo/IMU sensor combination.

Index Terms—Camera–IMU calibration, hand–eye calibration, industrial robotic manipulator, nonconvex optimization, pose estimation.

Jin Wu and Ming Liu are with the Active Vision Group, Robotics and Multiperception Laboratory, Department of Electronic and Computer Engineering, Hong Kong University of Science and Technology, Hong Kong, China (e-mail: jin_wu_uestc@hotmail.com; eelium@ust.hk).

Miaomiao Wang is with the Department of Electrical and Computer Engineering, University of Western Ontario, London, ON N6A 3K7, Canada (e-mail: mwang448@uwo.ca).

Yi Jiang was with the State Key Laboratory of Synthetical Automation for Process Industries and the International Joint Research Laboratory of Integrated Automation, Northeastern University, Shenyang 110819, China. He is now with the Department of Mechanical and Biomedical Engineering, City University of Hong Kong, Hong Kong (e-mail: jy369356904@163.com).

Bowen Yi is with the Australian Centre for Field Robotics, The University of Sydney, Sydney, NSW 2006, Australia (e-mail: b.yi@outlook.com).

Rui Fan was with the Jacobs School of Engineering and Jacobs School of Medicine and Biomedical Sciences, University of California San Diego, La Jolla, CA 92093 USA. He is now with the Department of Control Science and Engineering, College of Electronics and Information Engineering, and Shanghai Research Institute for Intelligent Autonomous Systems, Tongji University, Shanghai 200092, China (e-mail: ranger_fan@outlook.com).

Color versions of one or more figures in this article are available at https://doi.org/10.1109/TMECH.2021.3103995.

Digital Object Identifier 10.1109/TMECH.2021.3103995

I. Introduction

A. Background and Related Work

PRECISION industrial grasping tasks have high demands on the servoing accuracy. To meet such purposes, visual/inertial sensors are usually mounted to the robotic manipulator for precise inspection of the objects to be grasped. In such a system, it is required that all frames of the sensors and the manipulator should be aligned so that the processed information can be mapped from one frame to another [1]. The key step to obtain the unknown transformation from the vision sensor frame to the frame of an end-effector attached to the robotic manipulator is called the hand–eye calibration, which was first proposed by robotic scientists in the late 1980s [2], [3]. Hand–eye calibration is usually abstracted as a mathematical problem of the form $AX = XB$, in which X is the unknown hand–eye parameter to be solved. Extensive efforts have been paid to seek efficient and accurate closed-form solutions to the hand–eye calibration problem [4]–[6]. It has been later found out that, if one also takes the robot–world pose into account, the hand–eye/robot–world calibration algorithm of the type $AX = YB$ can be established, where X still represents the hand–eye relationship, while Y exactly denotes the robot–world transformation [7]. The problems $AX = XB$ and $AX = YB$ stimulate many other industrial calibration works that resemble their mathematical forms [8], [9]. Although many analytical solutions (see, e.g., [10] and [11]) have been found for these problems [12], they are still nonconvex ones indicating that a closed-form solution cannot reach the global minimum, and further iterative search is required. Therefore, in the past few years, the community concentrated on finding globally optimal solutions for these problems [13]–[15].

Recently, industrial vision sensors have been usually designed with an inertial measurement unit (IMU) in a compact manner. Such a visual/inertial sensor combination brings great convenience for very accurate and highly dynamical visual–inertial odometry [16]. The calibration for extrinsic parameter between the camera and the IMU is an important issue but may not have been solved properly. The renowned toolbox Kalibr [17] aims to offer a flexible calibration pipeline for most users. It is based on motion-based calibration concepts [18], [19] that significantly require IMU preintegration [20], which diverges quickly in the presence of large inertial biases. Historically, calibrating vision sensors using a robotic manipulator has been studied [21]. Cameras can also help the calibration of common robotic manipulators [22] or parallel mechanisms like six-degree-of-freedom Stewart platforms [23]. Unified camera–robot calibration improves the calibration accuracy for respective calibration steps [24]. Moreover, inertial sensors can also be calibrated using robotic manipulators [25]. IMUs are also benefiting the calibration of robotic manipulators [26]. For highly dynamical visual servoing tasks, the IMU also helps to gap the low-rate data sampling of the camera and the robotic manipulator [27]. Therefore, calibrating the robot, the camera, and the IMU simultaneously can consider the measurement errors in a unified fashion and, thus, enhances the calibration accuracy.

B. Contribution

The contributions of this article are listed as follows:
1) A hybrid pose determination method is derived, which considers the hand–eye/robot–world/camera–IMU transformations in a unified framework. The developed algorithm is free of IMU preintegration, so it is more robust to unknown IMU biases.
2) The pose determination is refined by a nonconvex optimization. We also give the globally optimal solutions to this problem.
3) The designed method, according to its global optimality, can obtain calibration parameters that are very close to the ground truth offered by professional calibration services.

C. Outline

The rest of this article is organized as follows. Section II contains the proposed algorithm. Section III illustrates experimental results. Finally, Section IV concludes this article.

II. SIMULTANEOUS CALIBRATION ALGORITHM

A. Problem and Solution

Related notations are described as follows. The 3-D special orthogonal group contains all 3×3 rotation matrices, subject to orthonormality, i.e., $\mathrm{SO}(3) := \{\boldsymbol{R} \in \mathbb{R}^{3\times3} | \boldsymbol{R}^\top \boldsymbol{R} = \boldsymbol{R}\boldsymbol{R}^\top = \boldsymbol{I}, \det(\boldsymbol{R}) = +1\}$, where $\boldsymbol{I}$ is the identity matrix. Combining rotation on $\mathrm{SO}(3)$ with a translation vector, the 3-D special Euclidean group is $\mathrm{SE}(3) := \{\boldsymbol{T} = \begin{pmatrix} \boldsymbol{R_T} & \boldsymbol{t_T} \\ \boldsymbol{0} & 1 \end{pmatrix} | \boldsymbol{R_T} \in$
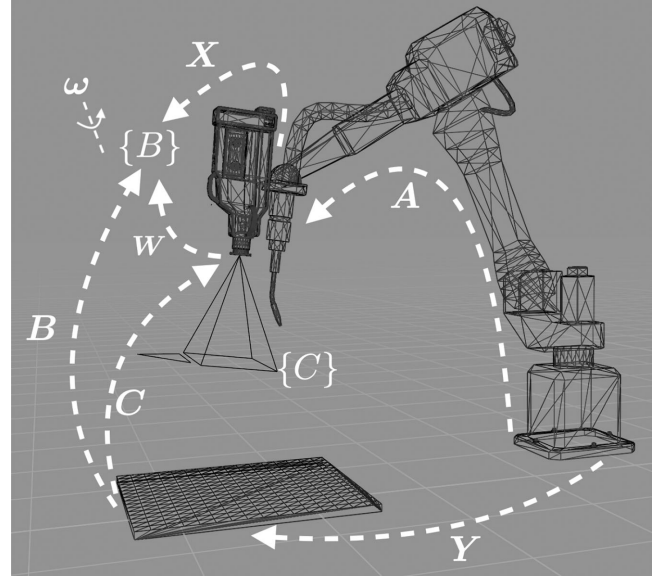
Fig. 1. Relationship between multiple poses involved in the studied problem.

$\mathrm{SO}(3), \boldsymbol{t_T} \in \mathbb{R}^3\}$, where $\boldsymbol{0}$ denotes a zero matrix with proper dimension. All the poses in this article are on $\mathrm{SE}(3)$.

The calibration problem considered in this article is illustrated in Fig. 1. The studied problem is: given kinematics of the industrial robotic manipulator $\boldsymbol{A}$, pose $\boldsymbol{C}$ from calibration board to the camera frame $\{C\}$, and angular rate measurements $\boldsymbol{\omega} \in \mathbb{R}^3$ in the IMU frame $\{B\}$, determine the unknown hand–IMU parameter $\boldsymbol{X}$, robot–world pose $\boldsymbol{Y}$, and camera–IMU extrinsic parameter $\boldsymbol{W}$, in an optimal least-squares manner. In Fig. 1, two virtual intermediate poses are illustrated, i.e., the pose $\boldsymbol{B}$ from the calibration board to the IMU and the pose $\boldsymbol{X}$ from the robot gripper to the IMU frame. The hand–eye parameter $\boldsymbol{X}$ was originally defined as the transformation between the robot gripper and the attached camera. It is assumed that the robotic manipulator has been calibrated; thus, pose $\boldsymbol{A}$ is considered to be accurate. As seen from Fig. 1, it is shown that the real *hand–eye* relationship can be restored by combining $\boldsymbol{X}$ and $\boldsymbol{W}$. To determine unknown transformations optimally shown above, we first establish the hand–eye/robot–world relation by

$$\boldsymbol{AX} = \boldsymbol{YB}. \tag{1}$$

Note that both sides of (1) can be differentiated with respect to time t, i.e.,

$$\dot{\boldsymbol{A}}\boldsymbol{X} + \boldsymbol{A}\dot{\boldsymbol{X}} = \dot{\boldsymbol{Y}}\boldsymbol{B} + \boldsymbol{Y}\dot{\boldsymbol{B}} \tag{2}$$

in which, for arbitrary pose $\boldsymbol{X} \in \mathrm{SE}(3)$, we have

$$\boldsymbol{X} = \begin{pmatrix} \boldsymbol{R_X} & \boldsymbol{t_X} \\ \boldsymbol{0} & 1 \end{pmatrix} \tag{3}$$

$$\dot{\boldsymbol{X}} = \begin{pmatrix} \dot{\boldsymbol{R}}_{\boldsymbol{X}} & \dot{\boldsymbol{t}}_{\boldsymbol{X}} \\ \boldsymbol{0} & 0 \end{pmatrix}. \tag{4}$$

For the hand–IMU parameter, it is always assumed as a constant once the installation is completed. Therefore, $\dot{\boldsymbol{X}} = \boldsymbol{0}$

characterizes such static pose. For Y, it represents the transformation between the robot base to the calibration board in the world frame, which is also static in our studied case, indicating $\dot{Y} = 0$. Therefore, (2) is simplified as

$$\dot{A}X = Y\dot{B}. \tag{5}$$

Note that since B is a virtual nonmeasurable pose, it must be related to some measurable ones. Following this purpose, the models of the angular rate ω and the camera pose C are combined. For the rigid body of the IMU, the angular rate ω contributes to the attitude dynamics [20] of

$$\dot{R}_B = R_B\omega_\times \tag{6}$$

in which $\omega_\times$ denotes the skew-symmetric matrix of ω. When imaging the calibration in the image plane, the camera captures 2-D points in its frame. Commonly, the perspective-n-point (PnP) problem seeks the camera pose by finding the transformation between 2-D points in the image plane and 3-D points in the world frame. As the calibration board is normally standard, 3-D world points are easily obtained. However, since the camera intrinsic matrix may not be accurate enough because of inaccurate estimation of radial and tangential distortions, the translational scale of the PnP pose must be adjusted. To this end, we also model the positive scale as $\varrho \in \mathbb{R}_+$ so that the camera pose C is

$$C = \begin{pmatrix} R_C & \varrho t_C \\ 0 & 1 \end{pmatrix}. \tag{7}$$

Let us write the unknown camera–IMU extrinsic parameter as

$$W = \begin{pmatrix} R_W & t_W \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} R_C^B & t_C^B \\ 0 & 1 \end{pmatrix} \tag{8}$$

which is assumed to be a fixed constant for the studied problem. Therefore, the homogeneous transformation satisfies

$$R_B = R_C R_C^B \tag{9}$$

$$t_B = R_C t_C^B + \varrho t_C. \tag{10}$$

Since R_C^B, t_C^B, and ϱ are constants, one obtains

$$\dot{R}_B = \dot{R}_C R_C^B \tag{11}$$

$$\dot{t}_B = \dot{R}_C t_C^B + \varrho \dot{t}_C. \tag{12}$$

Now, we have two categories of equations

$$\begin{cases} AX = YB(\varrho) \\ \dot{A}X = Y\dot{B}(\varrho) \end{cases} \tag{13}$$

which can be expanded as

$$\begin{cases} R_A R_X = R_Y R_C R_C^B \\ R_A t_X + t_A = R_Y \left(R_C t_C^B + \varrho t_C \right) + t_Y \\ \dot{R}_A R_X = R_Y R_C R_C^B \omega_\times \\ \dot{R}_A t_X + \dot{t}_A = R_Y \left(\dot{R}_C t_C^B + \varrho \dot{t}_C \right). \end{cases} \tag{14}$$

The first subequation in (14) is not solvable as R_C^B can be inverted to the left side and combined together with R_X. However, from the third subequation in (14), one may see that such transform does not hold any longer, i.e., the parameter R_C^B is solvable in this subequation. In experiments, engineers will gather a series of $\mathcal{N}$ measurements for error adjustment. From another aspect, the relationship in Fig. 1 indicates that $AX = YCW$, which can be transformed as

$$AU = YC \tag{15}$$

in which $U = XW^{-1}$. This shows that Y can be solved independently from (15) with given matrices A and C. Therefore, we have the following optimization:

$$\underset{U, Y \in \mathrm{SE}(3)}{\arg\min} \ \mathcal{J} = \sum_{i=1}^{\mathcal{N}} \|A_i U - YC_i\|^2 \tag{16}$$

in which $\|X\| = \sqrt{\mathrm{tr}(X^\top X)}$ for arbitrary X is the Frobenius norm and the subscript i denotes the ith pair of measurements. The problem in (16) is the typical $AX = YB$ one, which has global solution using [14], [15]. For problem (14), the optimization (17) as shown at the bottom of this page, is constructed. The problem (17) is a highly nonconvex one due to the nonlinear constraints of $\mathrm{SO}(3)$. Some parts of this problem belong to the general type of $AXB = YCZ$ [8], but are not strictly on $\mathrm{SE}(3)$. Therefore, solving (17) using analytical quaternionic methods like [8] is infeasible because there is no conversion from $\mathrm{SE}(3)$ derivatives $\dot{A}$ and $\dot{B}$ to quaternion derivatives. To decrease these nonlinear constraints, quaternions are employed. Then, the modified optimization Lagrangian is

$$\tilde{\mathcal{L}} = \mathcal{L} + \lambda^\top \begin{pmatrix} q_X{}^\top q_X - 1 \\ q_W{}^\top q_W - 1 \end{pmatrix} \tag{18}$$

where $\lambda = (\lambda_1, \lambda_2)^\top$ contains Lagrange multipliers, while q_X and q_W are attitude quaternions of R_X and R_C^B, respectively. All local minima of (17) can be obtained via

$$\nabla_y \tilde{\mathcal{L}} = 0 \tag{19}$$

in which $y = (q_X{}^\top, q_W{}^\top, t_X{}^\top, t_W{}^\top, \varrho, \lambda^\top)^\top \in \mathbb{R}^{17}$ is the optimization variable. Equation (19) indicates a polynomial system formed by components of y.

$$\underset{\substack{R_X, R_C^B \in \mathrm{SO}(3), \\ t_X, t_C^B \in \mathbb{R}^3, \varrho \in \mathbb{R}_+}}{\arg\min} \quad \mathcal{L} = \sum_i^N \begin{array}{l} \left\| R_{A,i} R_X - R_Y R_{C,i} R_C^B \right\|^2 + \left\| \dot{R}_{A,i} t_X + \dot{t}_{A,i} - R_Y \dot{R}_{C,i} t_C^B - \varrho R_Y \dot{t}_{C,i} \right\|^2 \\ \left\| \dot{R}_{A,i} R_X - R_Y R_{C,i} R_C^B \omega_{i,\times} \right\|^2 + \left\| R_{A,i} t_X + t_{A,i} - R_Y R_{C,i} t_C^B - \varrho R_Y t_{C,i} - t_Y \right\|^2. \end{array} \tag{17}$$

B. Classification of Roots

Solving (19) completely will give all local minima. There are many existing methods for solving all roots of polynomial systems accurately or approximately. Representatives include Gröbner-basis method and Wu Wenjun's method [28], [29]. However, the current system given in (19) has 17 subequations and an order of 3, indicating that the Gröbner basis of this system can hardly be obtained symbolically. As the Gröbner-basis method is essential for variable elimination, it is also hard for Wu Wenjun's method to compute all solutions to (19) within tolerable computation time. What should be noted here is that, although (19) has many solutions, only real ones are taken into account. Therefore, the kernel task is to obtain all real roots to (19). Moreover, analyzing (19) using symbolic computation engines, one can write t_X and t_W in terms of q_X, q_W, and ϱ. Therefore, inserting these formulations back into (19) gives a new system

$$f(z) = \mathbf{0} \tag{20}$$

which only consists of components in a new optimization variable $z = (q_X{}^\top, q_W{}^\top, \varrho, \lambda^\top)^\top \in \mathbb{R}^{11}$. The system (20) is not homogeneous in z, thus making it easier to be reduced. With the aid of a supercomputer from Aliyun, Inc., China, we are able to use Macaulay2 software to get a reduced system of (20). Typically, this can be done by finding out the Sylvester matrix $\mathcal{S}$ or Gröbner basis $\mathcal{G}$ of (20). As there are 11 dimensions in z, neither $\mathcal{S}$ nor $\mathcal{G}$ is hand-writable. Rather, the final result of $\mathcal{S}$ consumes a hard-disk storage of 728 GB, and for $\mathcal{G}$, the storage is 12 GB. Note that the above results are obtained by simplification after the quaternion norm constraints $q_X{}^\top q_X = 1$ and $q_W{}^\top q_W = 1$. If we extend $\mathcal{G}$ into the form of ϱ, forming a univariate polynomial as

$$g(\varrho) = 0 \tag{21}$$

the polynomial $g(\varrho)$ has at most terms of 6771, namely, for extreme cases, (21) will have 6770 complex roots. However, what we really care are those real roots that correspond to real scale factors and, furthermore, real quaternions. To identify how many real roots (21) will have, some symbolic theorems must be introduced. Let

$$g(\varrho) = a_0 \varrho^n + a_1 \varrho^{n-1} + \cdots + a_n \tag{22}$$

where $n \leq 6770$ is the order of the univariate polynomial of ϱ and a_{i-1} is the ith coefficient. The differential of (22) can be given by

$$g'(\varrho) = n a_0 \varrho^{n-1} + (n-1) a_1 \varrho^{n-2} + \cdots + a_{n-1}. \tag{23}$$

The discriminant of $g(\varrho)$ is defined as the Bezout matrix between $g(\varrho)$ and $g'(\varrho)$, say $\mathrm{discr}(g)$ [29]. The second discriminant of $g(\varrho)$ is defined as the Sylvester matrix between $g(\varrho)$ and $g'(\varrho)$, say $\mathrm{Discr}(g)$. The generation of $\mathrm{discr}(g)$ and $\mathrm{Discr}(g)$ can be fully completed by symbolic softwares like Mathematica and Maple. Here, we would like to denote $\mathcal{M}_k$ as the matrix associated with the kth subminor of $\mathrm{Discr}(g)$. In [32], it has been pointed out that $g(\varrho)$ has k distinct roots if and only if $\det(\mathcal{M}_k) \neq 0$ **and** $\det(\mathcal{M}_j)$, for $j = k+1, k+2, \ldots, n$.

Moreover, the number of real roots $\mathfrak{N}_{\varrho,\mathbb{R}}$ is equal to the variation number of signs in children components of a quadratic term $\zeta^\top \mathcal{M}_k \zeta$, where ζ is an arbitrary real vector. Therefore, it can be verified via huge amount of synthetic samples up to the quantity of 1×10^9 that $\mathfrak{N}_{\varrho,\mathbb{R}}$ has the upper bound of 256. This means that the polynomial (22) will have at most 256 real roots. Considering that any combination of $q_X, -q_X$ and $q_W, -q_W$ is valid for (20), the number of real roots z equals $256/4 = 64$.

C. Numerical Solutions

In practice, due to the characteristics of data, the scale parameter ϱ is always around 1. The values of Lagrange multiplier λ actually represent the penalty factor of the optimization. For the optimization (17), one must select the global minimum that corresponds to the least loss function value $\mathcal{L}$. Therefore, λ must be a small value because if it is too large, it can be inferred that the corresponding loss function value is also large. To effectively obtain all the real local minima of (20), we use multiple random sampled quaternions as initial conditions. To this end, via the method in [31], we are able to uniformly and randomly sample quaternions in the 4-D real unit-spherical vector space. By empirical tests, we are able to conclude that (20) normally has 4–12 local minima, but will have at most 64 local minima for some extreme cases, which coincides with previous theoretical findings. Therefore, if we perform a search of 128 uniform random samples for solving (20), it can be guaranteed that all solutions can be obtained. Here, we use the Broyden–Fletcher–Goldfarb–Shanno algorithm for unconstrained convergence of (20), which is a quasi-Newton method free of computation of analytical Jacobian. After obtaining all real solutions to (20), sorting loss function values in the ascending order and selecting the solution associated with the loss function value give the global optimum. It can be seen from the above procedures that multiple searches from different initial values will be easily implemented with parallelization. Therefore, the complete solution on a modern multicore industrial computer is computationally efficient. The final algorithmic flow is provided in Algorithm 1.

D. Sensitivity Analysis

The sensitivity analysis aims to find out the relationship between the input uncertainty and the output uncertainty. Therefore, we use the covariance as the quantitative result of sensitivity analysis. Since the covariance of measurement input is small, we use the following small-angle approximation to account for the rotation infinitesimal $\theta_G = (\theta_1, \theta_2, \theta_3)^\top$ of the rotational part of a given pose G:

$$\delta R_G = \begin{pmatrix} 1 & -\theta_3 & \theta_2 \\ \theta_3 & 1 & -\theta_1 \\ -\theta_2 & \theta_1 & 1 \end{pmatrix} \tag{24}$$

so that the rotation after perturbation is $\tilde{R}_G = (\delta R_G) R_G$. The perturbation of translation is additive, such that $\tilde{t}_G = t_G + \delta t_G$. In Algorithm 1, there are two key steps. First, we obtain Y from (16). The covariances of components of pose Y can be obtained

---

Algorithm 1: Proposed Simultaneous Hand–Eye/Robot–World/Camera–IMU Calibration Algorithm.

1. Preliminaries
1) Number of available measurements: $\mathcal{N} \geq 2$.
2) Synchronized measurements pairs:
$\{\boldsymbol{A}_i, \boldsymbol{B}_i\}, \{\dot{\boldsymbol{A}}_i, \dot{\boldsymbol{B}}_i\}$, for $i = 1, 2, \ldots, \mathcal{N}$.

2. Algorithmic Steps
1) Generate uniformly random samples of quaternions as initial start point of optimization by [31]. Use $\varrho = 1$ as initial solution.
2) Conduct parallel computation of globally optimal SE(3) poses $\boldsymbol{U}$ and $\boldsymbol{Y}$
from (16) using methods from [14], [15].
3) Conduct parallel computation of globally optimal quaternions $\boldsymbol{q_X}, \boldsymbol{q_W}$ and scale factor ϱ from (17) by solving (20) with the developed algorithm.
4) Compute the corresponding translation $\boldsymbol{t_X}$ and $\boldsymbol{t_W}$.

---

by Hessian inverse as the Fisher information

$$\boldsymbol{\Sigma_{cd}} = \left(\frac{\partial^2 \mathcal{J}}{\partial \boldsymbol{c} \, \partial \boldsymbol{d}^\top} \right)^{-1} \tag{25}$$

in which $\boldsymbol{c}$ and $\boldsymbol{d}$ represent combinations of $\boldsymbol{\theta_Y}$ and $\boldsymbol{t_Y}$. When the covariance of $\boldsymbol{Y}$ has been obtained, it can be used as input uncertainty of the next problem (17). Likewise, the covariances of components of $\boldsymbol{X}$, $\boldsymbol{W}$, and ϱ can be obtained by setting $\mathcal{J}$ as $\mathcal{L}$, and $\boldsymbol{c}$ and $\boldsymbol{d}$ as combinations of $\boldsymbol{\theta_X}$, $\boldsymbol{\theta_W}$, $\boldsymbol{t_X}$, $\boldsymbol{t_W}$, and ϱ. For detailed computation of derivatives, one may turn to [30] for analytical formulas. Moreover, in engineering applications, one may also use numerical central differentiation or Padé approximation to compute the Hessian.

III. EXPERIMENTAL RESULTS

A. Overview

The experimental setup is shown in Fig. 2. There is a Universal Robot UR5 industrial robotic manipulator with a mounted camera and a gripper. Calibration patterns, including a 12×9 chessboard with each block size of 30 mm, together with several Aruco markers of ID582 being presented on the test table. The chessboard and the markers are placed statically on the table. The installed camera is the Intel Realsense D435i with infrared/visible-light subcameras and a Bosch BMI055 IMU. The camera has the measurement uncertainty of 0.5 pixels and that of IMU is 0.005 rad/s for the gyroscope. The transformation of robot links is generated from the angular encoder of the UR5 robot. The repeatability of the UR5 robot is ± 0.1 mm, which is comparatively noiseless according to the manual. The chessboard has been set as the origin of the world frame, and the visible-light camera is selected for calibration. The BRISK descriptor is utilized for feature extraction of the corners [33]. A sample image of extracted corners after filter is shown in Fig. 3. The direct least-squares method is used for pose estimation of the camera relative to the chessboard via PnP mechanisms [34]. The `Quicksort` algorithm is employed to sort loss function
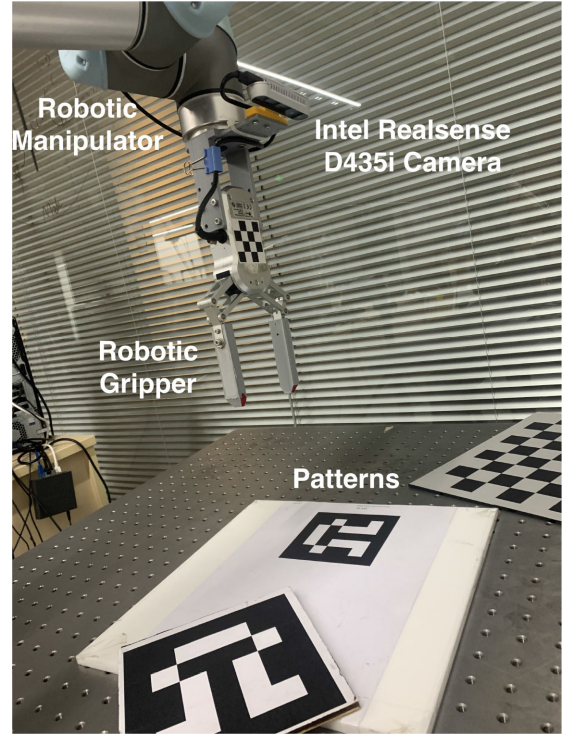

Fig. 2. Experimental setup for industrial hand–eye calibration. Related essential hardware include a robotic manipulator, an Intel Realsense camera, and several static calibration patterns. Via continuous motion of the robotic manipulator, the captured motion parameters can be used for the sensor calibration algorithm proposed in this article.
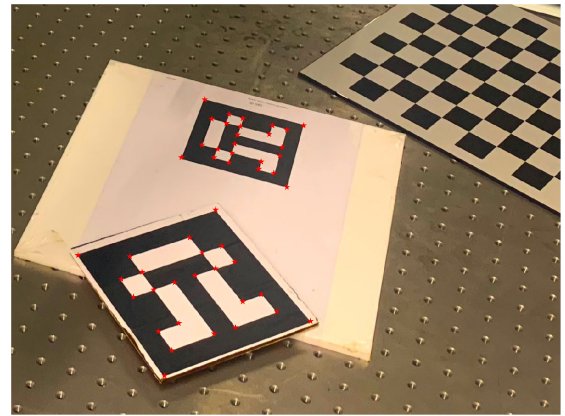

Fig. 3. Camera-captured calibration patterns and detected corners (in red).

values of different local minima [35]. We use the factory default intrinsic parameter to conduct PnP pose estimation. Several experiments have been conducted for validation of the proposed simultaneous hand–eye/robot–world/camera–IMU calibration method. The D435i is factory calibrated, and the default extrinsic parameter can be regarded as the ground truth, which is denoted as $\boldsymbol{W}_{\text{true}} = \begin{pmatrix} \boldsymbol{R}_{\text{true}} & \boldsymbol{t}_{\text{true}} \\ \boldsymbol{0} & 1 \end{pmatrix}$. The rotational and translational errors for a given pair of rotation $\boldsymbol{R}$ and translation $\boldsymbol{t}$ are

$$\varepsilon_{\boldsymbol{R}} = \arccos \left\{ \left[\text{tr} \left(\boldsymbol{R}^\top \boldsymbol{R}_{\text{true}} \right) - 1 \right] / 2 \right\} \quad (\text{rad}). \tag{26}$$

$$\varepsilon_{\boldsymbol{t}} = \|\boldsymbol{t} - \boldsymbol{t}_{\text{true}}\| / \sqrt{3} \qquad \text{(mm)}. \qquad (27)$$

Prior to experiments, we conduct a synthetic study showing the effect of measurement noise levels to the developed algorithm. In this synthetic study, the true rotation and translation of $\boldsymbol{A}$ are perturbed by noise such that

$$\boldsymbol{R_A} = \boldsymbol{R}_{\boldsymbol{A},\text{true}} \exp(\boldsymbol{\xi}_\times)$$
$$\boldsymbol{t_A} = \boldsymbol{t}_{\boldsymbol{A},\text{true}} + \boldsymbol{\gamma} \qquad (28)$$

where ture denotes true values and $\boldsymbol{\xi}$ and $\boldsymbol{\gamma}$ are noise terms that are subject to Gaussian distribution. The noise model (28) also applies for $\boldsymbol{C}$. The gyroscope readings are also perturbed via $\boldsymbol{\omega} = \boldsymbol{\omega}_{\text{true}} + \boldsymbol{\beta}$, where $\boldsymbol{\beta}$ is subject to Gaussian distribution as well. The noise level is defined as the variance of the noise terms. When evaluating noise effect from a single source, the noise from other sources will be set to zero so that the evaluation will be independent. In a single test, the noise levels of rotational and translational parts are the same. For each test, we use 1000 synthetic cases for averaged performances.

Next, via experiments, it can be concluded that the D435i camera has angular accuracy of 10 arcsec and translation accuracy of 2 mm (1280×720 resolution). The experiments are conducted on a typical MacBook Pro 2017 with processor of i7-4core 3.2 GHz, running an Ubuntu 18 system with ROS Melodic. The data rates of different sensors are downsampled to 30 Hz (30 frames/s) for synchronization. We use the `Moveit!` toolbox in the robotic operating system (ROS) to conduct motion planning of the UR5 robot. Since there is no ground truth for the hand–eye parameter $\boldsymbol{X}$ and the robot–world pose $\boldsymbol{Y}$, the only way to verify the performance is by validating the camera–IMU calibration results. We also use the `Kalibr` algorithm [17] and the method of Mirzaei and Roumeliotis [18] to conduct the camera–IMU calibration as comparative results. `Kalibr` refines the calibration parameters by considering the camera intrinsic parameter and the camera–IMU extrinsic parameter in a unified manner via the Levenberg–Marquardt algorithm (LMA). The method of Mirzaei and Roumeliotis estimates camera pose and IMU biases using a Kalman filter. The parameters of these algorithms for the presented experimental validation can be set according to the datasheets of Intel D435i and Bosch BMI055 IMU. The experimental steps are as follows:

1) Make sure that all the robotic components are self-calibrated. First, the robotic manipulator should be calibrated on the table. Second, we need an accurate intrinsic calibration of the camera. Third, the basic parameters of the IMU, including scale factor, misalignment, and offsets, should be estimated.

2) Use a planning algorithm, for instance, a rapidly exploring random tree algorithm, to generate a smooth trajectory of the robotic gripper. It should be guaranteed that throughout the trajectory, the calibration pattern is always within the field of view of the camera. The motion should be distinctive enough for observable angular rates but should not be too drastic such that the data magnitude overflows the full measurement range.
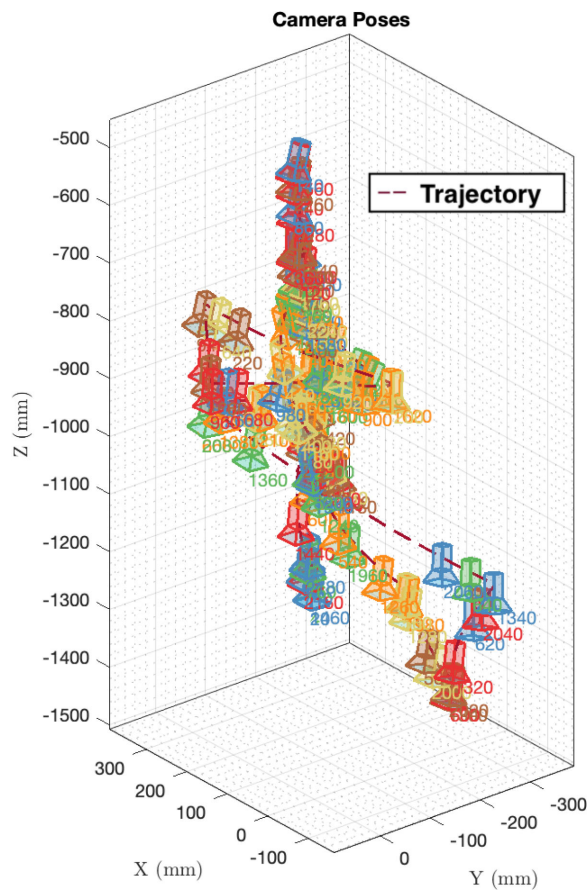

Fig. 4. Camera poses and trajectory in one experiment.

3) When gathering data, make sure that all the sensor data will be sampled at the same time instant. This sometimes requires synchronization mechanisms in hardware. For software implementation using ROS, we suggest the `Synchronizer` function to conduct synchronized measurements.[1]

B. Results

The synthetic accuracy tests for different noise levels by perturbing $\boldsymbol{A}$ are shown in Figs. 6 and 7. When perturbing $\boldsymbol{C}$, we obtain results in Figs. 8 and 9. The perturbation of the IMU model generates the results in Figs. 10 and 11.

For experimental study, the continuous motion planned can be visualized via the motion of camera (see Fig. 4). We generate distinctive motion to conduct the proposed simultaneous calibration. The inertial data of one experiment of 150 s duration are shown in Fig. 5. If there is no motion selection, i.e., small angular rates also take part in the calibration, the accuracy is not satisfactory. One may add more measurements for error adjustment, and the relationship between measurement number $\mathcal{N}$ and the accuracy is shown in Fig. 12. However, if we set a

[1]The open-source codes of this method can be accessed at https://github.com/zarathustr/SHERWCIC, where detailed procedures of experimental setup and data gathering can be found.
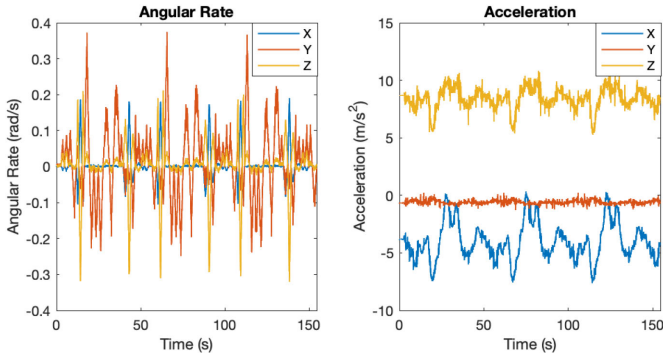
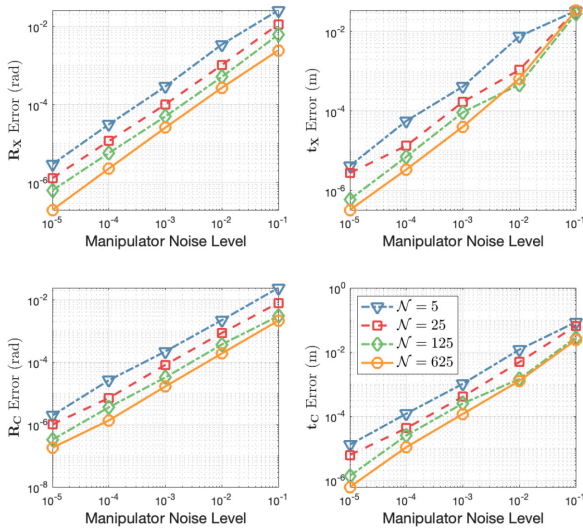Fig. 5.   Inertial data recorded for one of the experiments.

Fig. 6.   Rotational and translational errors subject to different noise levels when perturbing manipulator measurements A.
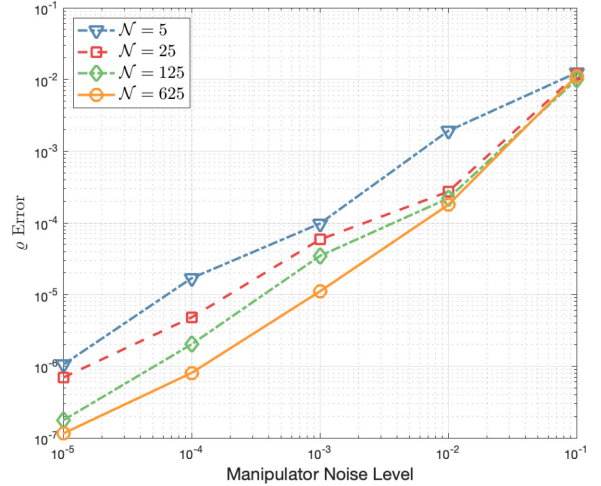

Fig. 7.   Scale error subject to different noise levels when perturbing manipulator measurements A.
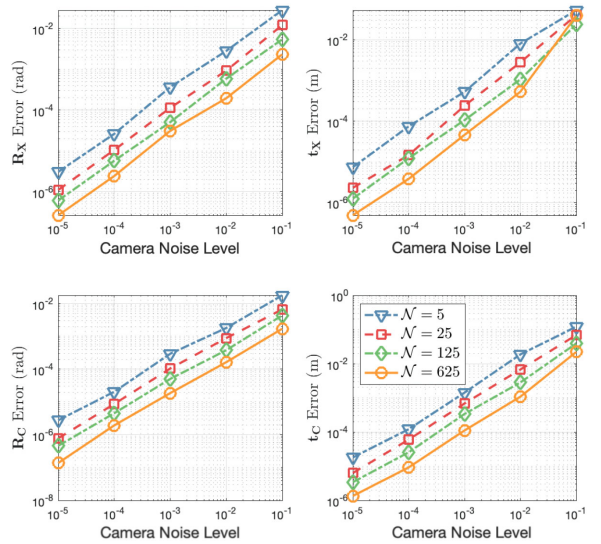

Fig. 8.   Rotational and translational errors subject to different noise levels when perturbing camera pose C.
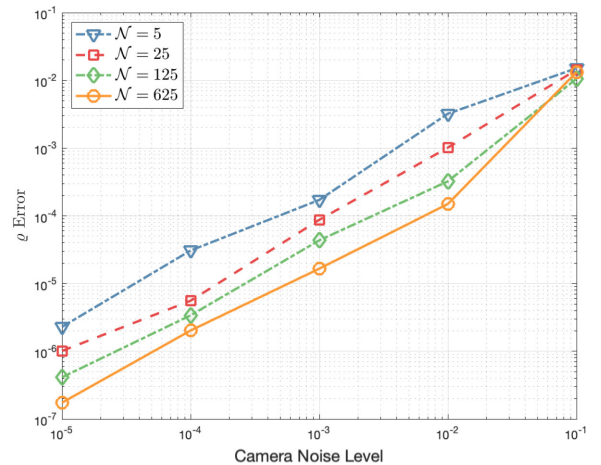

Fig. 9.   Scale error subject to different noise levels when perturbing camera pose C.

threshold η as the minimum tolerant angular rate norm, the calibration accuracy performances (100 pairs of randomly selected measurements) are shown in Fig. 13. Using $\eta = 0.05 \, \mathrm{rad/s}$, we compare the results from various algorithms, whose details are shown in Fig. 14. For such a parameter set, the scale factor is estimated as $\varrho = 0.99158 \pm 5.23 \times 10^{-5}$. The uncertainty is generated using the sensitivity analysis shown in Section II-D. We also add unknown IMU biases b_ω to the inertial data ω so that the utilized angular rate is $\tilde{\omega} = \omega + b_\omega$ and show the performance results in Table I. The run-time stats for all these algorithms are presented in Table II. The error performances of various methods are shown in Table III.

C. Discussion

From the results shown in Section III-B, one may be able to know the following:
1) From Fig. 12, larger $\mathcal{N}$ results in more accurate estimation results. This finding coincides with the basic principle of the least-squares algorithm.
2) From Fig. 13, a slightly higher angular-rate threshold η helps improve the calibration accuracy of the proposed method. This is because when $\omega \approx 0$, the value of $R_C \omega$
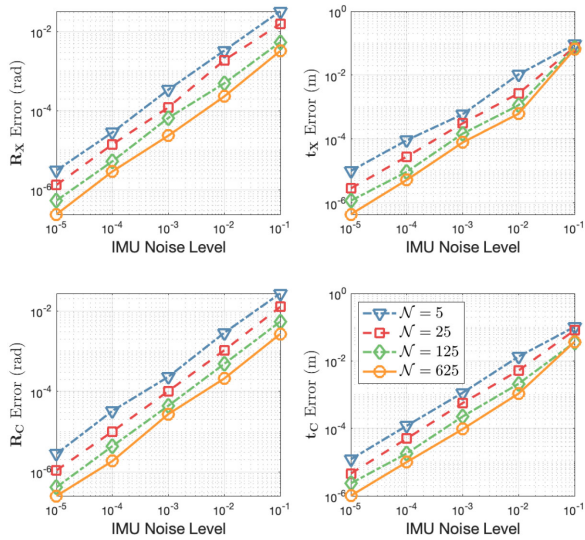
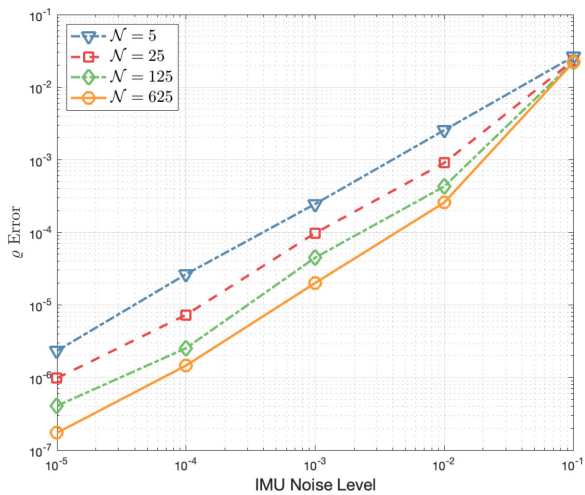Fig. 10. Rotational and translational errors subject to different noise levels when perturbing the IMU.

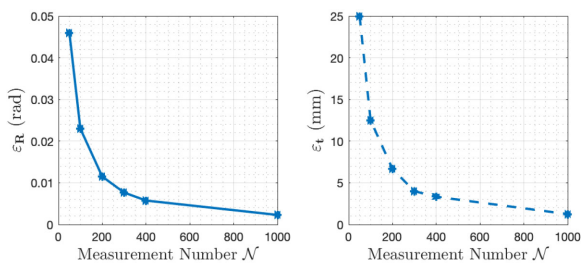Fig. 11. Scale error subject to different noise levels when perturbing the IMU.

Fig. 12. Relationship between measurement number $\mathcal{N}$ and accuracy.

approaches zero as well, which will make the equality $\dot{A}X = X\dot{B}$ singular. However, when η is too large, much useful information will be lost.
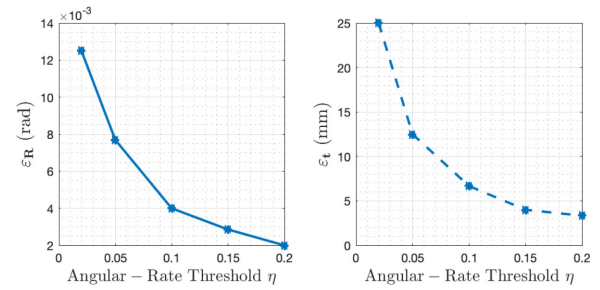

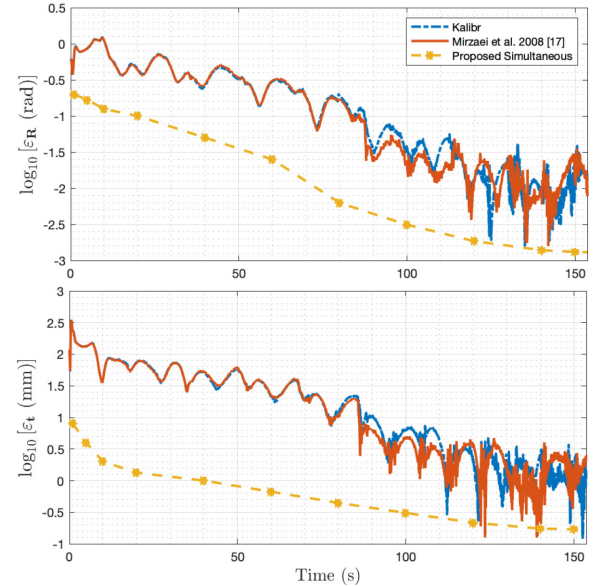Fig. 13. Relationship between angular-rate threshold η and accuracy.

Fig. 14. Comparisons between various algorithms for camera–IMU calibration.

3) From Figs. 6–11, it can be shown that the logarithm rotational and translational errors grow linearly with increasing noise levels of different error sources. It also comes to a consensus that larger measurement number $\mathcal{N}$ leads to a better calibration accuracy. However, when the noise levels become sufficiently large, for instance 10%, the algorithm becomes insensitive to the variation of $\mathcal{N}$. This shows that, to guarantee a satisfactory calibration performance, the sensor noise must be suppressed, which may require digital filters. The utilized IMU contains built-in digital low-pass filter (DLPF), which significantly decreases such an effect, and it brings convenience in further experimental validation.
4) From Fig. 14 and Table III, the proposed method achieves the best of overall accuracy. The other two algorithms converge with increasing time and own slightly better accuracies for the translational part.
5) The scale factor ϱ is not exactly 1 in practice, this shows that:
   1) the intrinsic calibration of camera may not be very accurate;

TABLE I
ACCURACY PERFORMANCES OF VARIOUS ALGORITHMS WITH UNKNOWN IMU BIASES

$\|\boldsymbol{b_\omega}\|$ rad	Kalibr $\varepsilon_{\boldsymbol{R}}$ (rad)	Mirzaei and Roumeliotis [18] $\varepsilon_{\boldsymbol{R}}$ (rad)	Proposed $\varepsilon_{\boldsymbol{R}}$ (rad)
1×10^{-3}	0.0050681	0.0056933	0.0034135
1×10^{-2}	0.0328894	0.0434408	0.0067246
1×10^{-1}	0.1352965	0.1684732	0.0894015

$\|\boldsymbol{b_\omega}\|$ rad	Kalibr $\varepsilon_{\boldsymbol{t}}$ (mm)	Mirzaei and Roumeliotis [18] $\varepsilon_{\boldsymbol{t}}$ (mm)	Proposed $\varepsilon_{\boldsymbol{t}}$ (mm)
1×10^{-3}	1.12770	1.27365	0.38901
1×10^{-2}	6.23466	8.93782	0.84683
1×10^{-1}	20.2152	32.5611	1.72936

TABLE II
RUN-TIME STATS OF VARIOUS ALGORITHMS FOR CAMERA–IMU
CALIBRATION

	Kalibr	Mirzaei and Roumeliotis [18]	Proposed
Execution Time	14368.778 s	26.52451 s	12.68926 s

TABLE III
ERROR STATISTICS OF VARIOUS ALGORITHMS

	Kalibr	Mirzaei and Roumeliotis [18]	Proposed
$\varepsilon_{\boldsymbol{R}}$	0.0049462 rad	0.0052178 rad	0.0033401 rad
$\varepsilon_{\boldsymbol{t}}$	1.02708 mm	1.03266 mm	0.37913 mm

2) the correspondence and precision of 2-D/3-D points may not be refined.

Thus, the estimation of ϱ has been validated to be necessary, and the simultaneous calibration accuracy also relies on this value.

6) From Table I, the sensitivities to the magnitudes of unknown IMU biases indicate that the proposed algorithm is robust to these uncertain biases. The reason is that the proposed method is free of IMU preintegration and, thus, is independent of attitude integral, which highly depends on the bias stability of the gyroscope measurements.

7) From Table II, the computational efficiency shows that the proposed method is fast for computation. The Kalibr processes data using Python language and refines a huge optimization via LMA, which consumes too many computational resources and can only be executed on one single core. The method of Mirzaei and Roumeliotis [18] is much faster, which uses a 21-state Kalman filter. However, as the covariance matrix is of the size 21×21, for one loop, there must have been matrix inversion for 441 elements. While the Kalman filter can only be implemented sequentially, the proposed method, which has been realized with parallelization, is the fastest one among all candidates.

D. Extended Applications: Stereo/IMU Calibration

The studied simultaneous hand–eye/robot–world/camera–IMU solver can be extended to calibration between stereo
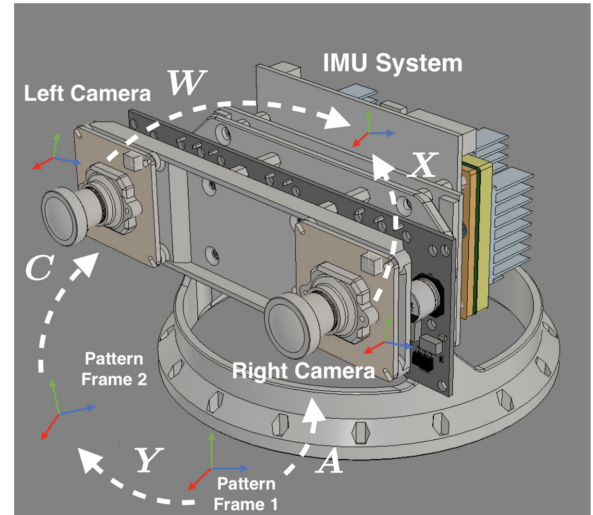

Fig. 15. Schematic of the stereo/IMU calibration.

cameras and the attached IMU. Fig. 16 depicts the schematic of the poses applied to this problem, which are very similar with that in Fig. 1. This system is important since in the 3-D reconstruction, one may require the IMU to enhance scale based on a pure stereo system. Using the proposed algorithm and two independent patterns visualized by two different cameras, we are able to obtain extrinsic parameters between left and right cameras, parameters between cameras and the IMU, and parameters between pattern frames. The stereo calibration parameters are given by $\boldsymbol{X}$ and $\boldsymbol{W}$ in Fig. 15. In this experiment, we use a pattern board with several AprilTags ([36]; see Fig. 17). Each AprilTag has a unique identifier, and the relative transformations between them can be given by the computer design system. When making the calibration board, we use very solid material so that the relative transformations are solid and can be considered as ground truth. The upper two patterns are selected for calibration. By capturing images and IMU readings at multiple positions in the 30-Hz continuous synchronized mode, we first evaluate the calibration results via the accuracy of $\boldsymbol{Y}$. The results indicate that the rotation error is within $0.01°$, and the translation error is within 1 mm. Then, we compare the proposed method with a state-of-the-art method by Huang *et al.* [37]. Huang's method aims to give online extrinsic calibration between stereo/IMU combination. By following the algorithmic steps in [37], we
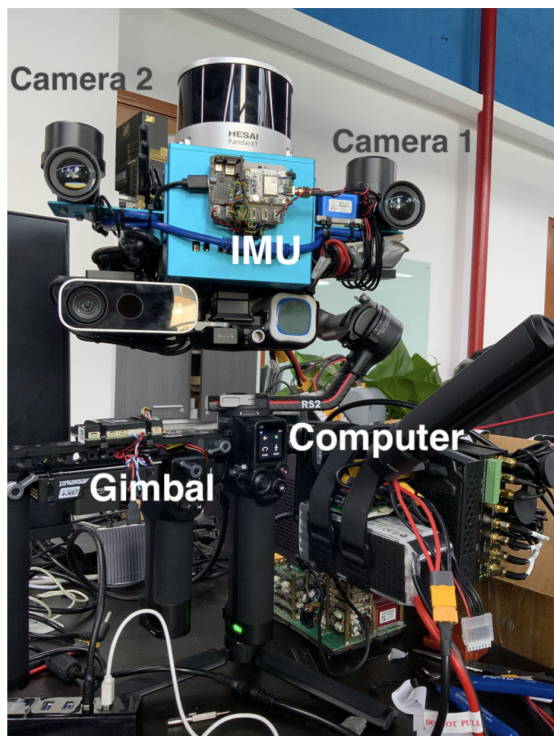
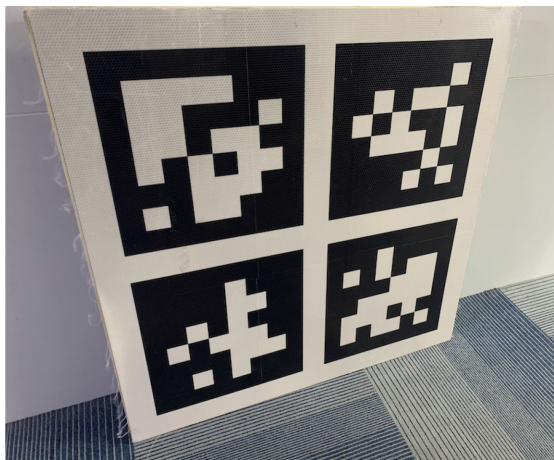Fig. 16.   Hardware of the stereo/IMU calibration.

Fig. 17.   Captured calibration patterns.

TABLE IV
REPROJECTION ERRORS OF VARIOUS METHODS

	Huang et al. [37]	Proposed
X errors	2.23 pixels	1.12 pixels
Y errors	3.01 pixels	0.92 pixels

perform the calibration and summarize the reprojection errors of calibration patterns from one camera to another in pixels. The results are shown in Table IV. Huang's method requires simultaneous calibration and visual inertial odometry. Therefore, it may be affected by long-term drift factors from the IMU, and

the accelerometer in the IMU gives the initial scale estimation, which may be affected by in-motion vibration. The proposed method does not rely on the integral of IMU measurements and, thus, is much more static than Huang's method and will give better calibration results for the designed hardware system.

IV. CONCLUSION

A new algorithm for simultaneous calibration of hand–eye/robot–world/camera–IMU relative transformations has been proposed in this article. Globally optimal estimation of these unknown transformations on SE(3) has been studied. The developed method offers a new method for calibrating the camera–IMU accurately using industrial robotic manipulator and calibration patterns. Through experiments, we verify that the method developed in this article is more robust to unknown IMU biases and is computationally efficient. Moreover, extended applications show that it is effective for stereo/IMU calibration. Future efforts will be paid to study online variants of the developed method.

REFERENCES

[1] O. Ozguner *et al.*, "Camera-robot calibration for the da vinci robotic surgery system," *IEEE Trans. Autom. Sci. Eng.*, vol. 17, no. 4, pp. 2154–2161, Oct. 2020.

[2] Y. C. Shiu and S. Ahmad, "Calibration of wrist-mounted robotic sensors by solving homogeneous transform equations of the form AX = XB," *IEEE Trans. Robot. Autom.*, vol. 5, no. 1, pp. 16–29, Feb. 1989.

[3] R. Y. Tsai and R. K. Lenz, "A new technique for fully autonomous and efficient 3D robotics hand/eye calibration," *IEEE Trans. Robot. Autom.*, vol. 5, no. 3, pp. 345–358, Jun. 1989.

[4] J. C. Chou and M. Kamel, "Finding the position and orientation of a sensor on a robot manipulator using quaternions," *Int. J. Robot. Res.*, vol. 10, no. 3, pp. 240–254, 1991.

[5] K. Daniilidis, "Hand-eye calibration using dual quaternions," *Int. J. Robot. Res.*, vol. 18, no. 3, pp. 286–298, 1999.

[6] N. Andreff, R. Horaud, and B. Espiau, "Robot hand-eye calibration using structure-from-motion," *Int. J. Robot. Res.*, vol. 20, no. 3, pp. 228–248, 2001.

[7] 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 Trans. Robot. Autom.*, vol. 10, no. 4, pp. 549–554, Aug. 1994.

[8] L. Wu, J. Wang, L. Qi, K. Wu, H. Ren, and M. Q. Meng, "Simultaneous hand-eye, tool-flange, and robot-robot calibration for comanipulation by solving the **AXB = YCZ** problem," *IEEE Trans. Robot.*, vol. 32, no. 2, pp. 413–428, Apr. 2016.

[9] L. Wu and H. Ren, "Finding the kinematic base frame of a robot by hand-eye calibration using 3D position data," *IEEE Trans. Autom. Sci. Eng.*, vol. 14, no. 1, pp. 314–324, Jan. 2017.

[10] J. Wu, Y. Sun, M. Wang, and M. Liu, "Hand-eye calibration: 4-D procrustes analysis approach," *IEEE Trans. Instrum. Meas.*, vol. 69, no. 6, pp. 2966–2981, Jun. 2020.

[11] J. Wu, M. Liu, C. Zhang, and Z. Zhou, "Correspondence matching and time delay estimation for hand-eye calibration," *IEEE Trans. Instrum. Meas.*, vol. 69, no. 10, pp. 8304–8313, Oct. 2020.

[12] F. Dornaika and R. Horaud, "Simultaneous robot-world and hand-eye calibration," *IEEE Trans. Robot. Autom.*, vol. 14, no. 4, pp. 617–622, Aug. 1998.

[13] J. Heller, M. Havlena, and T. Pajdla, "Globally optimal hand-eye calibration using branch-and-bound," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 38, no. 5, pp. 1027–1033, May 2016.

[14] J. Ha, D. Kang, and F. C. Park, "A stochastic global optimization algorithm for the two-frame sensor calibration problem," *IEEE Trans. Ind. Electron.*, vol. 63, no. 4, pp. 2434–2446, Apr. 2016.

[15] J. Wu *et al.*, "Globally optimal symbolic hand-eye calibration," *IEEE/ASME Trans. Mechatronics*, vol. 26, no. 3, pp. 1369–1379, Jun. 2021.

[16] Z. Yang and S. Shen, "Monocular visual-inertial state estimation with online initialization and camera-IMU extrinsic calibration," *IEEE Trans. Autom. Sci. Eng.*, vol. 14, no. 1, pp. 39–51, Jan. 2017.

[17] J. Rehder, J. Nikolic, T. Schneider, T. Hinzmann, and R. Siegwart, "Extending kalibr: Calibrating the extrinsics of multiple IMUs and of individual axes," in *Proc. IEEE Int. Conf. Robot. Autom.*, 2016, pp. 4304–4311.

[18] F. M. Mirzaei and S. I. Roumeliotis, "A Kalman-filter-based algorithm for IMU-camera calibration: Observability analysis and performance evaluation," *IEEE Trans. Robot.*, vol. 24, no. 5, pp. 1143–1156, Oct. 2008.

[19] J. Rehder, R. Siegwart, and P. Furgale, "A general approach to spatiotemporal calibration in multisensor systems," *IEEE Trans. Robot.*, vol. 32, no. 2, pp. 383–398, Apr. 2016.

[20] C. Forster, L. Carlone, F. Dellaert, and D. Scaramuzza, "On-manifold preintegration for real-time visual–inertial odometry," *IEEE Trans. Robot.*, vol. 33, no. 1, pp. 1–21, Feb. 2017.

[21] A. Izaguirre, P. Pu, and J. Summers, "A new development in camera calibration: Calibrating a pair of mobile cameras," *Int. J. Robot. Res.*, vol. 6, no. 3, pp. 104–116, 1987.

[22] H. Zhuang, W.-C. Wu, and Z. S. Roth, "Camera-assisted calibration of SCARA arms," *IEEE Robot. Autom. Mag.*, vol. 3, no. 6, pp. 46–53, Dec. 1996.

[23] P. Renaud, N. Andreff, J. M. Lavest, and M. Dhome, "Simplifying the kinematic calibration of parallel mechanisms using vision-based metrology," *IEEE Trans. Robot.*, vol. 22, no. 1, pp. 12–22, Feb. 2006.

[24] H. Zhuang, K. Wang, and Z. S. Roth, "Simultaneous calibration of a robot and a hand-mounted camera," *IEEE Trans. Robot. Autom.*, vol. 11, no. 5, pp. 649–660, Oct. 1995.

[25] B. Olofsson, J. Antonsson, H. G. Kortier, B. Bernhardsson, A. Robertsson, and R. Johansson, "Sensor fusion for robotic workspace state estimation," *IEEE/ASME Trans. Mechatronics*, vol. 21, no. 5, pp. 2236–2248, Oct. 2016.

[26] G. Du and P. Zhang, "Online serial manipulator calibration based on multisensory process via extended Kalman and particle filters," *IEEE Trans. Ind. Electron.*, vol. 61, no. 12, pp. 6852–6859, Dec. 2014.

[27] T. Qin, P. Li, and S. Shen, "VINS-Mono: A robust and versatile monocular visual-inertial state estimator," *IEEE Trans. Robot.*, vol. 34, no. 4, pp. 1004–1020, Aug. 2018.

[28] H. Fu, L. Yang, and J. Zhang, "A set of geometric invariants for kinematic analysis of 6R manipulators," *Int. J. Robot. Res.*, vol. 19, no. 8, pp. 784–792, 2001.

[29] W. Wu and X. Gao, "Automated reasoning and equation solving with the characteristic set method," *J. Comput. Sci. Technol.*, vol. 21, no. 5, pp. 756–764, 2006.

[30] J. Wu, M. Liu, Y. Huang, C. Jin, Y. Wu, and C. Yu, "SE(n)++: An efficient solution to multiple pose estimation problems," *IEEE Trans. Cybern.*, early access, doi: 10.1109/TCYB.2020.3015039.

[31] K. Shoemake, *Uniform Random Rotations*. New York, NY, USA: Academic, 1992.

[32] L. Yang and B. Xia, "Real solution classification for parametric semi-algebraic systems," in *Proc. Algorithmic Algebra Logic*, 2005, pp. 281–289.

[33] S. Leutenegger, M. Chli, and R. Y. Siegwart, "BRISK: Binary robust invariant scalable keypoints," in *Proc. IEEE Int. Conf. Comput. Vis.*, 2011, pp. 2548–2555.

[34] J. A. Hesch and S. I. Roumeliotis, "A direct least-squares (DLS) method for PnP," in *Proc. IEEE Int. Conf. Comput. Vis.*, 2011, pp. 383–390.

[35] C. A. R. Hoare, "Quicksort," *Comput. J.*, vol. 5, no. 1, pp. 10–16, 1962.

[36] E. Olson, "AprilTag: A robust and flexible visual fiducial system," in *Proc. IEEE Int. Conf. Robot. Autom.*, 2011, pp. 3400–3407.

[37] W. Huang, H. Liu, and W. Wan, "An online initialization and self-calibration method for stereo visual-inertial odometry," *IEEE Trans. Robot.*, vol. 36, no. 4, pp. 1153–1170, Aug. 2020.

Jin Wu (Member, IEEE) was born in Zhenjiang, Jiangsu, China, in May 1994. He received the B.S. degree from the University of Electronic Science and Technology of China, Chengdu, China. He is currently working toward the Ph.D. degree with the Robotics and Multiperception Lab, Hong Kong University of Science and Technology, Hong Kong, under the supervision of Prof. Ming Liu.

His research interests include robot navigation, multisensor fusion, automatic control, and mechatronics.

Miaomiao Wang (Member, IEEE) received the B.Sc. degree in control science and engineering from the Huazhong University of Science and Technology, Wuhan, China, in 2013, and the M.Sc. degree in control engineering from Lakehead University, Thunder Bay, ON, Canada, in 2015. He is currently working toward the Ph.D. degree with the Department of Electrical and Computer Engineering, Western University, London, ON, Canada. He is also a Research Assistant with the Department of Electrical and Computer Engineering, Western University. He is an author of technical papers in flagship journals and conference proceedings, including IEEE TRANSACTIONS ON AUTOMATIC CONTROL, *Automatica*, IEEE Conference on Decision and Control, etc. His current research interests include multiagent systems and geometric estimation and control.

Dr. Wang received the prestigious Ontario Graduate Scholarship at Western University in 2018.

Yi Jiang (Member, IEEE) was born in Ezhou, Hubei, China. He received the B.Eng. degree in automation in 2014 and the M.S. and Ph.D. degrees in control theory and control engineering in 2016 and 2020, respectively, all from the Information Science and Engineering College and State Key Laboratory of Synthetical Automation for Process Industries, Northeastern University, Shenyang, China.

From January 2017 to July 2017, he was a Visiting Scholar with the UTA Research Institute, University of Texas at Arlington, Arlington, TX, USA, and from March 2018 to March 2019, he was a Research Assistant with the University of Alberta, Edmonton, AB, Canada. He is currently a Postdoctoral Researcher with the City University of Hong Kong, Hong Kong. His research interests include networked control systems, industrial process operational control, reinforcement learning, and event-triggered control.

Dr. Jiang is an Associate Editor for *Advanced Control for Applications: Engineering and Industrial Systems*.
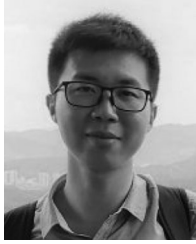
Bowen Yi (Member, IEEE) received the B.Eng. degree from Harbin Engineering University, Harbin, China, in 2014, and the Ph.D. degree from Shanghai Jiao Tong University, Shanghai, China, in 2019.

From October 2017 to April 2019, he was a Visiting Student with the Laboratoire des Signaux et Systèmes, CNRS-CentraleSupélec, Gif-sur-Yvette, France. He is currently a Postdoctoral Associate with Australian Centre for Field Robotics, The University of Sydney, Sydney, NSW, Australia. His research interests include nonlinear and adaptive control, observer design, system identification, and their applications in practical systems.

Dr. Yi received the CCTA Student Best Paper Award from the IEEE Control Systems Society in 2019.

Rui Fan (Member, IEEE) received the B.Eng. degree in automation from the Harbin Institute of Technology, Harbin, China, in 2015, and the Ph.D. degree in electrical and electronic engineering from the University of Bristol, Bristol, U.K., in 2018.

From 2018 to 2020, he was a Research Associate with the Robotics Institute and the Department of Electronic and Computer Engineering, Hong Kong University of Science and Technology, Hong Kong. Between 2020 and 2021, he was a Postdoctoral Fellow with the Department of Ophthalmology and the Department of Computer Science and Engineering, University of California San Diego, La Jolla, CA, USA. He is currently a Full Research Professor with the Department of Control Science and Engineering, College of Electronics and Information Engineering, and Shanghai Research Institute for Intelligent Autonomous Systems, Tongji University, Shanghai, China. His research interests include computer vision, machine/deep learning, image/signal processing, autonomous driving, and bioinformatics.
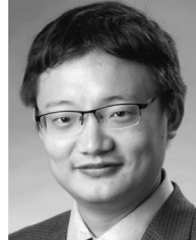
Ming Liu (Senior Member, IEEE) received the B.A. degree in automation from Tongji University, Shanghai, China, in 2005, and the Ph.D. degree from the Department of Mechanical and Process Engineering, ETH Zürich, Zürich, Switzerland, in 2013, supervised by Prof. Roland Siegwart.

During his academic study with Tongji University, he stayed one year with the University of Erlangen-Nuremberg, Erlangen, Germany, and the Fraunhofer Institute for Integrated Systems and Device Technology, Erlangen, as a Master Visiting Scholar. He is currently an Associate Professor with the Department of Electronic and Computer Engineering, Department of Computer Science and Engineering, and Robotics Institute, Hong Kong University of Science and Technology, Hong Kong. He is also a founding member of Shanghai Swing Automation, Ltd., Co., Shanghai, China. He is currently the Chairman of Shenzhen Unity Drive, Inc., China. He is coordinating and involved in National Science Foundation Projects and National 863-Hi-TechPlan Projects in China. He has authored or coauthored many popular papers in top robotics journals including IEEE TRANSACTIONS ON ROBOTICS, *International Journal of Robotics Research*, and IEEE ROBOTICS AND AUTOMATION MAGAZINE. He is a program member of the 2021 Robotics: Science and Systems Conference. His research interests include dynamic environment modeling, deep-learning for robotics, 3-D mapping, machine learning, and visual control.

Dr. Liu was a recipient of the European Micro Aerial Vehicle Competition (second place) and two awards from the International Aerial Robot Competition as a team member, the Best Student Paper Award as the first author at the 2012 IEEE International Conference on Multisensor Fusion and Information Integration, the Best Paper Award in Information at the 2013 IEEE International Conference on Information and Automation (ICIA) as the first author, the Best Paper Award Finalists as co-author, the Best RoboCup Paper Award at the 2013 (IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), the Best Conference Paper Award for the 2015 IEEE International Conference on Cyber Technology in Automation, Control, and Intelligent Systems, the Best Student Paper Finalist for the 2015 IEEE International conference on Real-time Computing and Robotics (RCAR), the Best Student Paper Finalist for the 2015 IEEE International Conference on Robotics and Biomimetics (ROBIO), the Best Student Paper for IEEE ROBIO 2019, the Best Student Paper Award for the 2017 International Conference on Advanced Robotics, the Best Paper in Automation Award for IEEE ICIA 2017, twice the innoviation contest Chunhui Cup Winning Award in 2012 and 2013, and the Wu Wenjun AI Award in 2016. He was the Program Chair of IEEE RCAR 2016 and the Program Chair of the International Robotics Conference in Foshan in 2017. He was the Conference Chair of the 2017 International Conference on Computer Vision Systems. He is an Associate Editor for IEEE ROBOTICS AND AUTOMATION LETTERS, *International Journal of Robotics and Automation*, *IET Cyber-Systems and Robotics*, and IEEE IROS Conference 2018, 2019, and 2020. He was a Guest Editor of special issues in IEEE TRANSACTIONS ON AUTOMATION SCIENCE AND ENGINEERING.