

Constrained hand-multiple-eyes calibration

Msuega Jnr. Iorpenda and Volker Willert

Center for Robotics (CERI). Technical University of Applied Sciences,
Würzburg-Schweinfurt,
Konrad-Geiger-Strasse 2, 97421 Schweinfurt.

Abstract This paper addresses the problem of calibrating multiple visual sensors mounted on a robotic manipulator, a task critical for accurate robot perception and interaction. We present a novel approach to hand-multiple-eyes calibration that incorporates closed-loop constraints to ensure consistency between the sensors' poses. Unlike traditional hand-eye calibration methods that handle individual sensor pairs independently, our method leverages a unified optimization framework that simultaneously optimizes the relative poses of all sensors while enforcing a loop closure constraint to each pose triplet. The core of our approach is a least squares approach to solve multiple hand-eye matrix equations of the form $\mathbf{AX} = \mathbf{XB}$, further enhanced with the method of Lagrangian multipliers to account for loop-closure constraints. We apply this idea to a minimal setup involving one hand and two eyes and demonstrate its effectiveness in improving the accuracy of pose estimation for various levels of noisy measurements.

Keywords Hand-eye calibration, multi-sensor-robot calibration, pose-graph optimization, constrained optimization

1 Introduction

The spatial relationship between a robot's end-effector (hand) and its visual sensor (eye) is critical for achieving synchronization in task execution [1]. In certain robotic applications, the use of multiple visual sensors is necessary for robust and reliable estimations, often requiring precise calibration.

A hand-eye calibration outputs the relative pose $\mathbf{X} \in SE(3)$ between a sensor (eye) and the robot end-effector (hand) comprising rotation \mathbf{R} and translation \mathbf{t} . Using N pairs of measured pose changes $\{\mathbf{A}_i, \mathbf{B}_i\}_{i=1}^N$ and minimizing the nonlinear least squares loss $\mathcal{L}(\mathbf{X}) = \frac{1}{2N} \sum_{i=1}^N \|\mathbf{A}_i \mathbf{X} - \mathbf{X} \mathbf{B}_i\|^2$ subject to \mathbf{X} via a gradient descent approach results in a very accurate estimate for the unknown pose \mathbf{X} (see also Fig. 1) outperforming non-iterative classical methods [2]. This idea has been extended to multi-sensor setups comprising K sensors by several authors [3–5] that all share the same basic idea: Simply optimizing the overall loss $\mathcal{L}_K = \sum_{j=1}^K \mathcal{L}(\mathbf{X}_j)$, whereas $\{\mathbf{X}_j\}_{j=1}^K$ are the fixed relative poses between each possible pair of sensors. This is equivalent to optimizing each pair of sensors separately because the constrained geometric relations between the relative poses are not taken into account. In [4] the constraint of equivalent rotations of the measured poses \mathbf{A}_i and \mathbf{B}_i are considered but no loop-closure constraint on the estimates \mathbf{X}_j . There are also solutions that increase accuracy for the special case of hand-cameras calibration by directly optimizing the reprojection error and considering the uncertainties of the different measures using a Gaussian-Helmert model [4,6].

2 Proposed Method

Our work follows the idea of gradient based nonlinear least squares optimization but includes additional closed loop pose-graph constraints to fulfill physical world reality for the estimates of all relative sensor poses. Here, we explore the minimal multi-sensory setup consisting of two sensors $S1$ and $S2$ rigidly attached to the end-effector EE of a serial manipulator, as illustrated in Figure 1. Here, $\{\mathbf{A}_i, \mathbf{B}_i, \mathbf{C}_i\}_{i=1}^N$ are all measured pose changes of the two sensors and the end-effector acquired by moving the robot arm accordingly. This sensory setup results in three unknown relative poses \mathbf{X} , \mathbf{Y} and \mathbf{Z} . These three relative poses form a closed pose-graph loop at any time. Hence, we can formulate an additional least squares loop closure constraint $\mathcal{L}_c = \frac{1}{2} \|\mathbf{X}\mathbf{Y} - \mathbf{Z}\|^2 \stackrel{!}{=} 0$ and add it to the hand-multiple-eyes loss \mathcal{L}_3 for $K = 3$ frames via a

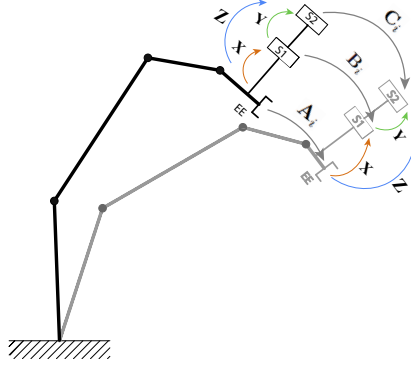


Figure 1: Schematic representation of the configuration of the robot end-effector (EE) and sensors (S1 and S2) for robot-multiple-eyes calibration. The figure illustrates two states of the robot’s motion (represented by black and gray outlines) used to measure pose changes $\{A_i, B_i, C_i\}$. Additionally, the unknown relative poses $\{X, Y, Z\}$ between the sensors and the end-effector need to form a closed pose-graph loop, where the constraint: $XY = Z$ holds in the physical real world.

Lagrangian Multiplier λ as follows: $\mathcal{L} = \mathcal{L}_3 + \lambda \mathcal{L}_c$, which reads

$$\begin{aligned} \mathcal{L} = & \frac{1}{2N} \sum_{i=1}^N \left(\|A_i X - X B_i\|^2 + \|B_i Y - Y C_i\|^2 + \|A_i Z - Z C_i\|^2 \right) \\ & + \frac{1}{2} \lambda \|XY - Z\|^2, \quad i \in [1, \dots, N]. \end{aligned} \quad (1)$$

This objective is optimized with a gradient descent approach applying constrained differential optimization [7] and the angle-axis representation for rotations.

2.1 Optimization of Rotations

We propose an optimization for estimating rotations and translations separately by decoupling the poses of the measurements and estimates from Eq. (1) into rotation matrices and translation vector components. Additionally, we incorporate the closed-loop constraint into the objective function using a Lagrange multiplier, as shown in Eqs. (3), where \mathcal{L}_{RX} , \mathcal{L}_{RY} , \mathcal{L}_{RZ} , and \mathcal{L}_{RC} represent the rotational components of the

objective function corresponding to the transformations \mathbf{X} , \mathbf{Y} , \mathbf{Z} , and the constraint, respectively. The rotation matrices are parameterized in terms of their rotation axes $\mathbf{s} = [s_0, s_1, s_2]^\top$ and denoted as $\mathbf{R}_X(\mathbf{s}_X)$, $\mathbf{R}_Y(\mathbf{s}_Y)$, and $\mathbf{R}_Z(\mathbf{s}_Z)$. The rotation objective reads

$$\mathcal{L}_R(\mathbf{s}_X, \mathbf{s}_Y, \mathbf{s}_Z, \lambda) = \mathcal{L}_{RX}(\mathbf{s}_X) + \mathcal{L}_{RY}(\mathbf{s}_Y) + \mathcal{L}_{RZ}(\mathbf{s}_Z) + \lambda \mathcal{L}_{RC}(\mathbf{s}_X, \mathbf{s}_Y, \mathbf{s}_Z), \quad (2)$$

$$\begin{aligned} \mathcal{L}_R = \frac{1}{2N} \sum_{i=1}^N & \left(\|\mathbf{R}_{Ai} \mathbf{R}_X - \mathbf{R}_X \mathbf{R}_{Bi}\|^2 + \|\mathbf{R}_{Bi} \mathbf{R}_Y - \mathbf{R}_Y \mathbf{R}_{Ci}\|^2 \right. \\ & \left. + \|\mathbf{R}_{Ai} \mathbf{R}_Z - \mathbf{R}_Z \mathbf{R}_{Ci}\|^2 \right) + \frac{1}{2} \lambda \|\mathbf{R}_X \mathbf{R}_Y - \mathbf{R}_Z\|^2. \end{aligned} \quad (3)$$

We derive the gradients of the rotational objective function with respect to the axis parameters and the Lagrange multiplier, as follows:

$$\begin{aligned} \frac{\partial \mathcal{L}_R}{\partial s_{Xk}} &= \frac{\partial \mathcal{L}_{RX}(\mathbf{R}_X(\mathbf{s}_X))}{\partial s_{Xk}} + \frac{\partial \mathcal{L}_{RC}(\mathbf{R}_X(\mathbf{s}_X), \lambda)}{\partial s_{Xk}} \\ &= \frac{1}{N} \sum_i \left\{ \text{tr} \left(\left[2\mathbf{R}_X - \mathbf{R}_{Ai}^\top \mathbf{R}_X \mathbf{R}_{Bi} - \mathbf{R}_{Ai} \mathbf{R}_X \mathbf{R}_{Bi}^\top \right]^\top \frac{\partial \mathbf{R}_X(\mathbf{s}_X)}{\partial s_{Xk}} \right) \right\} \\ &\quad + \lambda \text{tr} \left(\left[\mathbf{R}_X - \mathbf{R}_Z \mathbf{R}_Y^\top \right]^\top \frac{\partial \mathbf{R}_X(\mathbf{s}_X)}{\partial s_{Xk}} \right), \end{aligned} \quad (4)$$

$$\begin{aligned} \frac{\partial \mathcal{L}_R}{\partial s_{Yk}} &= \frac{\partial \mathcal{L}_{RY}(\mathbf{R}_Y(\mathbf{s}_Y))}{\partial s_{Yk}} + \frac{\partial \mathcal{L}_{RC}(\mathbf{R}_Y(\mathbf{s}_Y), \lambda)}{\partial s_{Yk}} \\ &= \frac{1}{N} \sum_i \left\{ \text{tr} \left(\left[2\mathbf{R}_Y - \mathbf{R}_{Bi}^\top \mathbf{R}_Y \mathbf{R}_{Ci} - \mathbf{R}_{Bi} \mathbf{R}_Y \mathbf{R}_{Ci}^\top \right]^\top \frac{\partial \mathbf{R}_Y(\mathbf{s}_Y)}{\partial s_{Yk}} \right) \right\} \\ &\quad + \lambda \text{tr} \left(\left[\mathbf{R}_Y - \mathbf{R}_X^\top \mathbf{R}_Z \right]^\top \frac{\partial \mathbf{R}_Y(\mathbf{s}_Y)}{\partial s_{Yk}} \right), \end{aligned} \quad (5)$$

$$\begin{aligned}
\frac{\partial \mathcal{L}_R}{\partial s_{Zk}} &= \frac{\partial \mathcal{L}_{RZ}(\mathbf{R}_Z(\mathbf{s}_Z))}{\partial s_{Zk}} + \frac{\partial \mathcal{L}_{RC}(\mathbf{R}_Z(\mathbf{s}_Z), \lambda)}{\partial s_{Zk}} \\
&= \frac{1}{N} \sum_i \left\{ \text{tr} \left(\left[2\mathbf{R}_Z - \mathbf{R}_{Ai}^\top \mathbf{R}_Z \mathbf{R}_{Ci} - \mathbf{R}_{Ai} \mathbf{R}_Z \mathbf{R}_{Ci}^\top \right]^\top \frac{\partial \mathbf{R}_Z(\mathbf{s}_Z)}{\partial s_{Zk}} \right) \right\} \\
&\quad + \lambda \text{tr} \left(\left[\mathbf{R}_Z - \mathbf{R}_X \mathbf{R}_Y \right]^\top \frac{\partial \mathbf{R}_Z(\mathbf{s}_Z)}{\partial s_{Zk}} \right), \tag{6}
\end{aligned}$$

$$\frac{\partial \mathcal{L}_R}{\partial \lambda} = \frac{1}{2} \|\mathbf{R}_X \mathbf{R}_Y - \mathbf{R}_Z\|^2 = \mathcal{L}_{RC}. \tag{7}$$

Compact formulas for partial derivatives of 3D rotation matrices in exponential coordinates can be found in [8]. The gradients were instrumental in the gradient descent optimization, where the update rules for the optimization parameters - rotation axes elements s_{Xk} , s_{Yk} and s_{Zk} for $k \in \{0, 1, 2\}$ and the Lagrange multiplier λ —are given like follows:

$$s_{Ek}^{i+1} = s_{Ek}^i - \alpha \frac{\partial \mathcal{L}_R}{\partial s_{Ek}^i}, \quad E \in \{X, Y, Z\}, \quad k \in \{0, 1, 2\}. \tag{8}$$

$$\lambda^{i+1} = \lambda^i + \beta \frac{\partial \mathcal{L}_R}{\partial \lambda^i} = \lambda^i + \beta \frac{1}{2} \|\mathbf{R}_X(\mathbf{s}_X^i) \mathbf{R}_Y(\mathbf{s}_Y^i) - \mathbf{R}_Z(\mathbf{s}_Z^i)\|^2. \tag{9}$$

Here, α and β are the step sizes and i representing the iteration index. It should be noted that a gradient descent is performed to find the optimum rotation parameters s_{Ek} , whereas a gradient ascent is performed to find the optimum λ [7].

2.2 Optimization of Translations

Next, we optimize Eq. (1) for the translation vectors \mathbf{t}_X , \mathbf{t}_Y and \mathbf{t}_Z assuming the rotations \mathbf{R}_X , \mathbf{R}_Y and \mathbf{R}_Z to already been optimal. This leads to the following objective:

$$\mathcal{L}_t = \mathcal{L}_{tX} + \mathcal{L}_{tY} + \mathcal{L}_{tZ} + \lambda_t \mathcal{L}_{tC}, \tag{10}$$

$$\begin{aligned}
 &= \frac{1}{2N} \sum_{i=1}^N \left(\|((\mathbf{R}_{Ai} - \mathbf{I})\mathbf{t}_X - \mathbf{R}_X\mathbf{t}_{Bi} + \mathbf{t}_{Ai})\|^2 + \|(\mathbf{R}_{Bi} - \mathbf{I})\mathbf{t}_Y - \mathbf{R}_Y\mathbf{t}_{Ci} + \mathbf{t}_{Bi}\|^2 \right. \\
 &\quad \left. + \|(\mathbf{R}_{Ai} - \mathbf{I})\mathbf{t}_Z - \mathbf{R}_Z\mathbf{t}_{Ci} + \mathbf{t}_{Ai}\|^2 \right) + \frac{1}{2} \lambda_t \|\mathbf{R}_X\mathbf{t}_Y + \mathbf{t}_X - \mathbf{t}_Z\|^2.
 \end{aligned} \tag{11}$$

The gradients for the different translation vectors read

$$\frac{\partial \mathcal{L}_{tX}}{\partial \mathbf{t}_X} = \frac{1}{N} \sum_i \left([\mathbf{R}_{Ai} - \mathbf{I}]^\top [(\mathbf{R}_{Ai} - \mathbf{I})\mathbf{t}_X - \mathbf{R}_X\mathbf{t}_{Bi} + \mathbf{t}_{Ai}] \right) + \lambda_t (\mathbf{t}_X + \mathbf{R}_X\mathbf{t}_Y - \mathbf{t}_Z), \tag{12}$$

$$\frac{\partial \mathcal{L}_{tY}}{\partial \mathbf{t}_Y} = \frac{1}{N} \sum_i \left([\mathbf{R}_{Bi} - \mathbf{I}]^\top [(\mathbf{R}_{Bi} - \mathbf{I})\mathbf{t}_Y - \mathbf{R}_Y\mathbf{t}_{Ci} + \mathbf{t}_{Bi}] \right) + \lambda_t \left(\mathbf{t}_Y + \mathbf{R}_X^\top [\mathbf{t}_X - \mathbf{t}_Z] \right), \tag{13}$$

$$\frac{\partial \mathcal{L}_{tZ}}{\partial \mathbf{t}_Z} = \frac{1}{N} \sum_i \left((\mathbf{R}_{Ai} - \mathbf{I})^\top [(\mathbf{R}_{Ai} - \mathbf{I})\mathbf{t}_Z - \mathbf{R}_Z\mathbf{t}_{Ci} + \mathbf{t}_{Ai}] \right) + \lambda_t (\mathbf{t}_Z - \mathbf{t}_X - \mathbf{R}_X\mathbf{t}_Y). \tag{14}$$

These gradients lead to the update rules for the translation parameters with step sizes γ and δ as follows

$$\mathbf{t}_E^{i+1} = \mathbf{t}_E^i - \gamma \frac{\partial \mathcal{L}_t}{\partial \mathbf{t}_E^i}, \quad E \in \{X, Y, Z\}, \tag{15}$$

$$\lambda_t^{i+1} = \lambda_t^i + \delta \frac{\partial \mathcal{L}_t}{\partial \lambda_t^i} = \lambda_t^i + \delta \frac{1}{2} \left\| \mathbf{R}_X\mathbf{t}_Y^i + \mathbf{t}_X^i - \mathbf{t}_Z^i \right\|^2. \tag{16}$$

3 Evaluation

The gradient descent approach derived in Section 2 was implemented in Matlab and tested using synthetic data. Robot and sensor poses were generated using RoboDK [9] and then perturbed with Gaussian noise to simulate real-world conditions. For all experiments the number of measurements is set to $N = 20$, the step sizes are fixed to $\alpha = 0.8$, $\beta = 0.01$, $\gamma = 0.3$ and $\delta = 10^{-7}$ and the number of iterations is $i = 1, \dots, 5000$. Each optimization run is initialized using the Tsai and Lenz method [10].

3.1 Pose-Graph Loop Closure

Our method enforces a solution for hand-multiple-eyes calibration, achieving a closed-loop pose graph by pushing the closed-loop constraints \mathcal{L}_{RC} and \mathcal{L}_{tC} close to zero, as shown in Figure 2 (a) and (b) (green lines). In contrast, the unconstrained optimization fails to meet the loop closure constraint (orange lines). The inclusion of constraints also enhances the convergence rate for both rotational and translational errors (green lines), resulting in more precise relative pose estimates (see Fig. 3) that stabilizes at a lower error level.

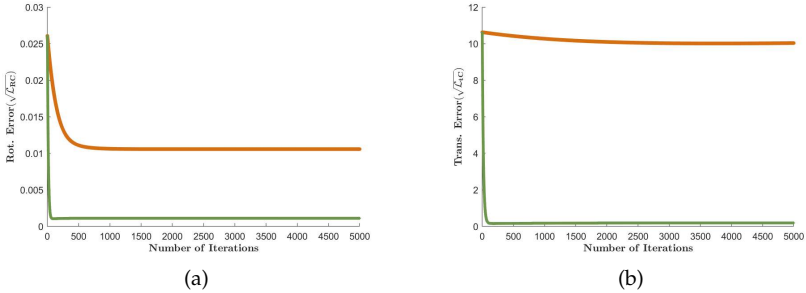


Figure 2: Rotational and translational errors for constrained (green) and unconstrained (orange) optimization. (a) Error $\sqrt{\mathcal{L}_{RC}}$ when optimizing \mathcal{L}_R including (green) or excluding \mathcal{L}_{RC} (orange). (b) Error $\sqrt{\mathcal{L}_{tC}}$ when optimizing \mathcal{L}_t including (green) or excluding \mathcal{L}_{tC} (orange).

3.2 Improved Rotation and Translation Estimates

We compared the rotation and translation estimates against the ground truth using the metrics $\mathcal{E}_R = \|\mathbf{R}_{est} - \mathbf{R}_{GT}\|$ for rotational deviation and $\mathcal{E}_t = \|\mathbf{t}_{est} - \mathbf{t}_{GT}\|$ for translational deviation. Figure 3 demonstrates the impact of applying the loop-closure constraints during the optimization process by comparing the total rotational and translational errors against ground truth with and without constraints. As can be seen the inclusion of loop-closure constraints not only enforces pose estimates that form a closed pose loop but also enhances the overall calibration accuracy, resulting in a better sensor alignment that is geometrically consistent. In contrast, the unconstrained method achieves less accurate results and does not provide a fully closed pose loop (see also Fig. 2). Additionally, we analyzed the evolution of the accuracy of the relative poses of the sensors during optimization: Sensor 1 with respect to the end effector (\mathbf{X}), sensor 2 with respect to sensor 1 (\mathbf{Y}), and sensor 2 with respect to the end effector (\mathbf{Z}), as shown in Figure 4. Improvements relative to the ground truth were observed across all system components but the improvements vary between the sensors.

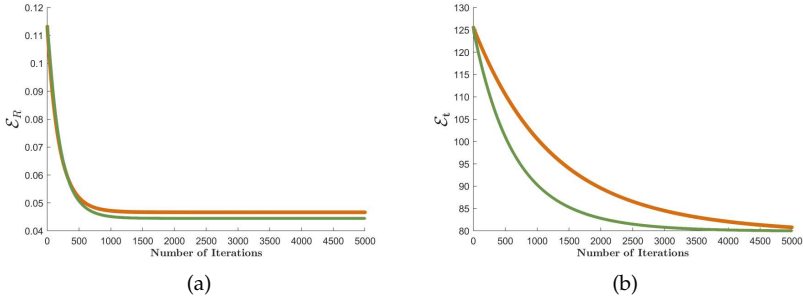


Figure 3: Rotation and translation estimates versus ground truth during unconstrained (orange) and constraint (green) optimization. (a) Overall rotation errors \mathcal{E}_R . (b) Overall translation errors \mathcal{E}_t .

3.3 Effect of Noise

We conducted 100 simulation runs per noise level, following [11], with noise sampled from a Gaussian distribution. The histograms in Fig-

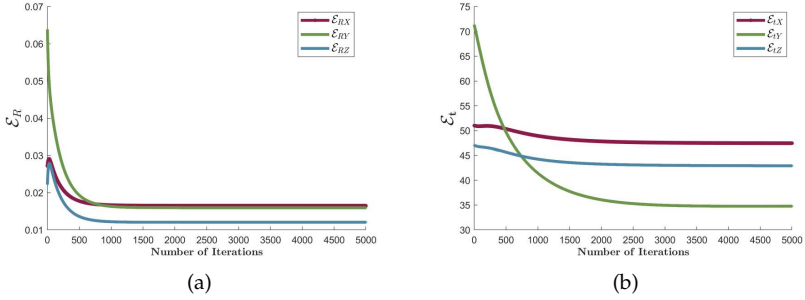


Figure 4: Individual rotation and translation estimates versus ground truth during constraint optimization. (a) Rotation errors ϵ_{RX} , ϵ_{RY} and ϵ_{RZ} . (b) Translation errors ϵ_{TX} , ϵ_{TY} and ϵ_{TZ} .

ure 5 present the results for four different noise levels (NL1 to NL4). In the rotation error histograms (top row), the noise follows a Gaussian distribution with standard deviations from 0.5 to 2.0 degrees, while in the translation error histograms (bottom row), the noise has standard deviations from 1 to 4 mm, as per [12]. As the noise levels increase (NL1 to NL4), the spread of both rotation and translation errors broadens, indicated by the larger standard deviations (green dashed lines). These results demonstrate the system’s sensitivity to increasing noise in both rotation and translation estimates.

4 Summary

We present a new extension to hand-multiple-eyes calibration by adding closed-loop constraints to ensure geometrical consistency between the poses of multiple visual sensors mounted on a robotic manipulator. Unlike traditional hand-eye calibration methods that address sensor pairs independently, our approach simultaneously optimizes the relative poses of all sensors.

First experimental results indicate that the inclusion of closed loop pose-graph constraints in the optimization process leads to estimates that form closed pose loops and each of these estimates are more accurate than if the optimization is done without adding the loop closure constraint. We have experienced that the results and the convergence

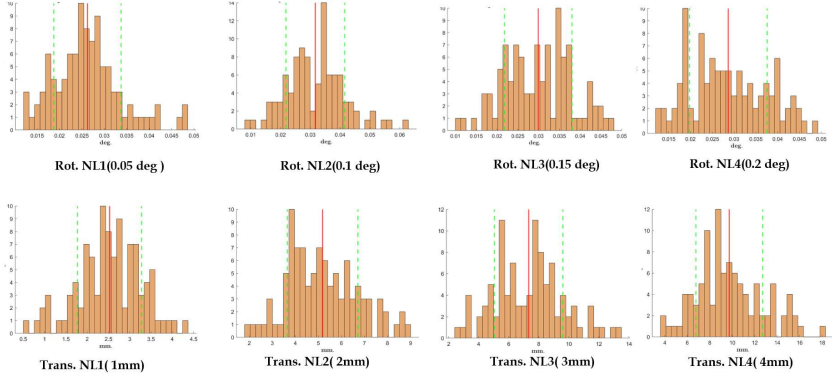


Figure 5: Histograms of rotation errors (\mathcal{E}_R , top row) and translation errors (\mathcal{E}_t , bottom row) for four different noise levels (NL1 to NL4). The orange bars show the error distribution from 100 simulation runs per noise level. The red vertical lines represent the mean errors, while the green dashed lines mark the standard deviations.

properties strongly depend on the choice of suitable step sizes. Next, an adaptive step size control should be added to take this problem into account. Further on, the dependency on the number of measurements and imbalances in the noise levels between the sensors need to be evaluated.

The Lagrangian Multiplier method allows a straight forward extension to a calibration of more than three relative poses. Also a direct optimization of the reprojection error for multiple-camera setups including the loop-closure constraints is some future work to do.

Acknowledgements

This research is co-funded by the Petroleum Technology Development Fund (PTDF) through the OSS Postgraduate Scholarship Scheme. Reference No: PTDF/ED/OSS/PHD/MI/1486/19 - 19PHD058. PTDF Towers, Plot 1058 Memorial Drive, Central Business District, Abuja, Nigeria. Email: info@ptdf.gov.ng.

References

1. B. Grossmann and V. Kruger, "Continuous hand-eye calibration using 3D points," in *15th International Conference on Industrial Informatics (INDIN)*. IEEE, 2017, pp. 311–318.
2. K. M. A. Y. Amy Tabb, "Solving the robot-world hand-eye(s) calibration problem with iterative methods," vol. 28, no. 5, pp. 569–590, 2017.
3. Q. V. Le and A. Y. Ng, "Joint calibration of multiple sensors," in *IEEE/RSJ International Conference on Intelligent Robots and Systems*, 2009, pp. 3651–3658.
4. K. Huang and C. Stachniss, "Extrinsic multi-sensor calibration for mobile robots using the gauss-helmert model," in *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 2017, pp. 1490–1496.
5. D. Zuñiga-Noël, J.-R. Ruiz-Sarmiento, R. Gomez-Ojeda, and J. Gonzalez-Jimenez, "Automatic multi-sensor extrinsic calibration for mobile robots," vol. 4, no. 3, pp. 2862–2869, 2019.
6. M. Ulrich and M. Hillemann, "Uncertainty-aware hand-eye calibration," *IEEE Transactions on Robotics*, 2023.
7. J. C. Platt and A. H. Barr, "Constrained differential optimization," in *Neural Information Processing Systems, Denver, Colorado, USA*, 1987, pp. 612–621.
8. A. Y. Guillermo Gallego, "A compact formula for the derivative of a 3-d rotation in exponential coordinates," *Journal of Mathematical Imaging and Vision*, vol. 51, pp. 378–384, 2015.
9. RoboDK, "Robodk: Robot simulation and offline programming software, version 5.4.3," <https://robodk.com>, 2023.
10. R. Y. Tsai, R. K. Lenz *et al.*, "A new technique for fully autonomous and efficient 3 d robotics hand/eye calibration," *IEEE Transactions on robotics and automation*, vol. 5, no. 3, pp. 345–358, 1989.
11. K. Koide and E. Menegatti, "General hand-eye calibration based on re-projection error minimization," *IEEE Robotics and Automation Letters*, vol. 4, no. 2, pp. 1021–1028, 2019.
12. I. Enebuse, B. K. K. Ibrahim, M. Foo, R. S. Matharu, and H. Ahmed, "Accuracy evaluation of hand-eye calibration techniques for vision-guided robots," *Plos one*, vol. 17, no. 10, p. e0273261, 2022.