

# On Flange-based 3D Hand-Eye Calibration for Soft Robotic Tactile Welding

Xudong Han<sup>1</sup>, Ning Guo<sup>1</sup>, Yu Jie<sup>1</sup>, He Wang<sup>2</sup>, Fang Wan<sup>2,\*</sup>, Chaoyang Song<sup>1,\*</sup>

<sup>1</sup>Department of Mechanical and Energy Engineering, Southern University of Science and Technology (SUSTech), 1088 Xueyuan Road, Shenzhen, Guangdong 518055, China

<sup>2</sup>School of Design, SUSTech, 1088 Xueyuan Road, Shenzhen, Guangdong 518055, China

\*Correspondence to: Fang Wan ([wanf@sustech.edu.cn](mailto:wanf@sustech.edu.cn)) and Chaoyang Song ([songcy@ieee.org](mailto:songcy@ieee.org))

**Abstract:** This paper investigates the direct use of standardized designs on the robot to perform robot hand-eye calibration using industrial and commercial grade 3D scanners with cobots. We exploit the well-established geometric features of the industrial robot flange by directly capturing its point cloud data. In particular, we propose an iterative method to facilitate point cloud processing toward a refined calibration outcome. We conducted extensive experiments using industrial-grade Photoneo Phoxi S & M models as well as a commercial-grade Microsoft Azure Kinect DK over a range of collaborative manipulators, including Universal Robots UR5 & UR10 e-series, Franka Emika, and Aubo i5. Simulation and experiment results show that our proposed method can effectively and efficiently achieve a hand-eye calibration accuracy as high as the camera's, probing the hardware limit. We also propose a soft robotic tactile welding system via multi-modal fusion, integrating flange-based hand-eye calibration and soft tactile sensing in weld seam tracking. The experiment results show that our presented tactile welding system enables the robot to adjust its motion in real time, ensuring consistent weld quality and paving the way for more efficient and adaptable manufacturing processes.

**Keywords:** 3D Vision, Hand-Eye Calibration, Measurement Standards, Robotic Welding

**CRediT authorship contribution statement:** **Xudong Han:** Data curation; Formal analysis; Investigation; Methodology; Software; Validation; Visualization; Roles/Writing – original draft. **Ning Guo:** Data curation; Formal analysis; Investigation; Methodology; Software; Validation; Visualization; Roles/Writing – original draft. **Yu Jie:** Data curation; Formal analysis; Investigation; Software; Validation; Visualization. **He Wang:** Methodology. **Fang Wan:** Conceptualization; Data curation; Formal analysis; Funding acquisition; Investigation; Methodology; Project administration; Resources; Software; Supervision; Validation; Visualization; Roles/Writing – original draft; Writing – review & editing. **Chaoyang Song:** Conceptualization; Formal analysis; Funding acquisition; Investigation; Methodology; Project administration; Resources; Supervision; Visualization; Roles/Writing – original draft; Writing – review & editing

**Acknowledgment:** This work was partly supported by the Science, Technology, and Innovation Commission of Shenzhen Municipality [JSGG20220831110002004], the National Natural Science Foundation of China [62206119], Shenzhen Long-Term Support for Higher Education at SUSTech [20231115141649002], and SUSTech Virtual Teaching Lab for Machine Intelligence Design and Learning [Y01331838].

# On Flange-based 3D Hand-Eye Calibration for Soft Robotic Tactile Welding

XXXX

---

## Abstract

This paper investigates the direct use of standardized designs on the robot to perform robot hand-eye calibration using industrial and commercial grade 3D scanners with cobots. We exploit the well-established geometric features of the industrial robot flange by directly capturing its point cloud data. In particular, we propose an iterative method to facilitate point cloud processing toward a refined calibration outcome. We conducted extensive experiments using industrial-grade Photoneo Phoxi S & M models as well as a commercial-grade Microsoft Azure Kinect DK over a range of collaborative manipulators, including Universal Robots UR5 & UR10 e-series, Franka Emika, and Aubo i5. Simulation and experiment results show that our proposed method can effectively and efficiently achieve a hand-eye calibration accuracy as high as the camera's, probing the hardware limit. We also propose a soft robotic tactile welding system via multi-modal fusion, integrating flange-based hand-eye calibration and soft tactile sensing in weld seam tracking. The experiment results show that our presented tactile welding system enables the robot to adjust its motion in real time, ensuring consistent weld quality and paving the way for more efficient and adaptable manufacturing processes.

*Keywords:*

3D vision, Hand-eye Calibration, Measurement Standards, Robotic Welding

---

<sup>1</sup> **1. Introduction**

<sup>2</sup> Depth sensor provides a versatile perception of the physical world with re-  
<sup>3</sup> fined details through three-dimensional (3D) measurements. Since Microsoft's  
<sup>4</sup> Kinect (Sarbolandi et al., 2015), a wide range of consumer-grade 3D scanners  
<sup>5</sup> has lowered the entry barriers when integrating robotic vision in research and  
<sup>6</sup> applications (Jarabo et al., 2017). Through optical perception, depth-sensing  
<sup>7</sup> technologies translate the geometric details in the physical world into three-  
<sup>8</sup> dimensional point cloud data concerning the camera frame (Kolb et al., 2010).  
<sup>9</sup> In areas including robotic welding (Fan et al., 2019; Lü et al., 2023), material  
<sup>10</sup> handling (Zhang et al., 2020; Peng et al., 2021), and human-robot collabora-  
<sup>11</sup> tions (Shi et al., 2012), robotics researchers have shown a growing acceptance of  
<sup>12</sup> adopting depth-sensing technologies (Hyatt et al., 2018), yet the robot-camera,  
<sup>13</sup> or so-called hand-eye, calibration remains the first problem in practice (Rout  
<sup>14</sup> et al., 2019; Wu et al., 2016, 2020).

<sup>15</sup> Industrial manipulators are usually built with excellent repeatability but rel-  
<sup>16</sup> atively low accuracy, often requiring calibration using machine vision (Huang  
<sup>17</sup> et al., 2023; Jiang et al., 2017). The repeatability problem is commonly solved  
<sup>18</sup> by directly reading the sensor data saved at each joint (Kluz and Trzepieciński,  
<sup>19</sup> 2014). In contrast, the accuracy problem involves the inverse kinematic com-  
<sup>20</sup> putation of a specific or target pose in the Cartesian space at the end-effector  
<sup>21</sup> (Zhang et al., 2021). The hand-eye calibration enhances the manipulator's accu-  
<sup>22</sup> racy by compensating the errors between the robot controller's computed pose  
<sup>23</sup> and the camera's measured pose by the camera (Kahn et al., 2014).

<sup>24</sup> With the growing market of robotic engineering, the standardization of ma-  
<sup>25</sup> nipulator design and manufacturing provides a rich set of geometric features  
<sup>26</sup> that are directly measurable by the depth sensors. The International Standard  
<sup>27</sup> Organization (ISO) 9409-1:2004 defines the main dimensions, designations, and  
<sup>28</sup> markings for a circular plate and a cylindrical shaft on the tool flange as the me-  
<sup>29</sup> chanical interface to ensure the exchangeability and orientation of end-effectors  
<sup>30</sup> (ISO, 2004). However, there remains to be more research on using depth-sensing

31 technologies to directly measure such standardized mechanical interfaces in 3D,  
32 which is the focus of this paper.

33 Adopting high-precision 3D vision scanners in robotic welding enables robust  
34 and efficient weld seam tracking and fully autonomous welding (Geng et al.,  
35 2024). Accurate hand-eye calibration is the first step in the general process of  
36 seam tracking via a vision system (Rout et al., 2019; Geng et al., 2023). Tactile  
37 sensors are also used when vision systems are not proficient, e.g., the weld seam  
38 is not entirely in the view of a camera or polluted environments (Lei et al., 2020;  
39 Tannous et al., 2020).

40 Hand-eye calibration is a 3D problem that can be solved using classical 2D  
41 cameras or by incorporating emerging 3D depth sensors. The classical problem  
42 of kinematic calibration has been well-studied over the years (Shah et al., 2012),  
43 which usually involves a manipulator as the *robot*, a camera as the *eye*, an end-  
44 effector as the *hand*, and a high-precision calibration marker or object as the  
45 *world* (Wu et al., 2016).

46 • The **hand-eye** calibration commonly refers to the relationship between  
47 the Tool Center Point (TCP) and the “eye” camera. This is especially  
48 true in scenarios with a fixed end-effector as the hand to simplify the  
49 expression. However, in many academic and engineering applications,  
50 robotic researchers also choose to use the default TCP on the tool flange  
51 as the hand to enhance the re-usability of the calibration results.

52 • The **tool-flange** calibration specifies the relationship between the default  
53 TCP at the tool flange and the actual TCP at the end-effector. The de-  
54 fault TCP is directly accessible in most robot controllers, which the man-  
55 ufacturer already calibrates. When an end-effector is attached, one can  
56 directly refer to the technical data sheet for the tool-flange relationship.  
57 However, many end-effectors are customized according to the specific use,  
58 which may require further calibration to determine the tool-flange rela-  
59 tionship.

60 • The **robot-robot** calibration refers to the case when multiple manipula-

61       tors are used for collaborative tasks, such as dual-arm robots and robot-  
62       assisted surgeries. In practice, many multi-robot systems are designed  
63       with one manipulator attached with an “eye” camera and the other with  
64       an end-effector. The relative positioning between the manipulator base  
65       frames must be calibrated before use.

66       Recent work by Wu et al. (2020) proposed a 4D Procrustes Analysis Ap-  
67       proach for the hand-eye calibration problem, where standardized objects are  
68       still required for implementation. Hu (2013) proposed a hand-to-eye calibra-  
69       tion method using a Bursa coordinate transform model through depth sensing.  
70       Kahn et al. (2014) designed a 3D calibration object with a curved surface such  
71       that its pose can be uniquely estimated using the iterative closest point (ICP)  
72       algorithm, demonstrating that 3D calibration provides more accurate results on  
73       average. Zhang et al. (2017) addressed hand-eye calibration using a surgical  
74       robot with a stereo laparoscope by proposing a computationally efficient itera-  
75       tive method. Yang et al. (2018) adopted a sphere model as the calibration object  
76       and reformulated the hand-eye calibration problem to use only the calibration  
77       object’s translation (3-DoF) data.

78       Mathematically, the calibration object can be any object in the camera’s  
79       view as long as the object’s pose can be estimated (Li et al., 2018). In in-  
80       dustrial applications such as assembly and manipulation, where high accuracy  
81       is required, standard calibration objects with high manufacturing precision are  
82       required for hand-eye calibration, which is usually expensive (An et al., 2016).  
83       In the eye-to-hand scenario, installing and removing the calibration object from  
84       the robot arm adds an extra burden to the already time-consuming deployment  
85       of industrial robots.

86       In this paper, we propose a novel method using high-fidelity 3D scanners  
87       to directly measure the standardized geometric features on the tool flange of a  
88       manipulator for hand-eye calibration, as shown in Fig. 1.

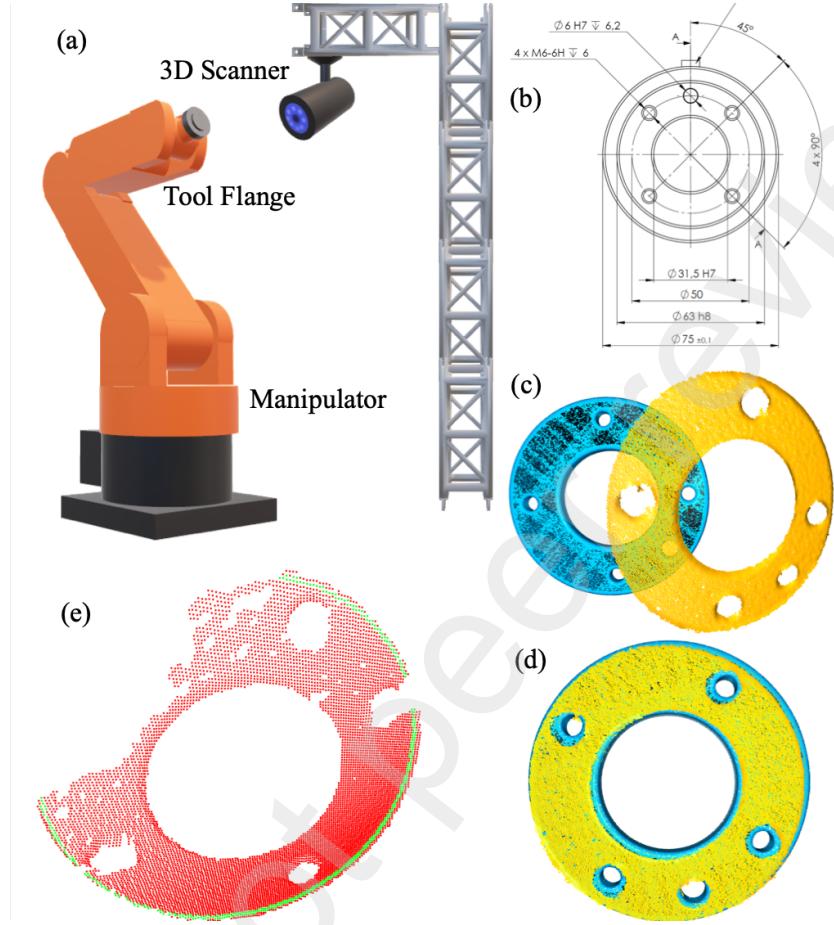


Figure 1: **A direct hand-eye calibration based on three-dimensional measurement of the standardized geometrical features on the manipulator's tool flange:** (a) The setup of direct flange-based hand-eye calibration; (b) The flange design of the UR5 robot following ISO standards; (c) The point cloud of the flange captured by the CAD model in blue and the 3D scanner in yellow; (d) Sample result of flange-based hand-eye calibration using 3D scanners; and (e) The proposed iterative method for circular feature calibration with partial point clouds.

Unlike previous approaches that mainly measure calibration objects external to the robot system, the proposed method focuses on a calibration process using a 3D measurement of the geometric features within the robot system under international standardization. The proposed method effectively reduces

93 system errors by removing unnecessary estimations and transformations during  
94 the calculation process of hand-eye calibration. With the growing adoption of  
95 3D scanners in robotic welding, the proposed method can potentially reduce  
96 the complexities in setting up the vision-based robot system for integration, as  
97 demonstrated in an original robotic welding system integrating 3D vision and  
98 soft tactile sensing. Contributions of this paper are listed as the following:

- 99     • Proposed a novel hand-eye calibration method by measuring the intrinsic  
100       design features of the robot system, i.e., tool flange and base mount, using  
101       high-fidelity 3D scanners;
- 102     • Implemented an iterative algorithm that effectively and efficiently optimizes  
103       the calibration accuracy as high as the camera's, probing the hardware limit;
- 104     • Conducted a quantitative and systematic evaluation of the calibration accuracy  
105       using several collaborative robots, industrial-grade, and consumer-grade  
106       3D scanners;
- 107     • Proposed a safe and adaptable weld seam tracking system that combines  
108       vision-based detection and soft tactile-based compensation.

109     The rest of this paper is structured as follows. Section 2 proposes the hand-eye  
110     calibration method via standardized robotic flanges and soft robot tactile  
111     welding method via multi-modal fusion. Section 3 presents the simulation and  
112     experiment results using the flange-based calibration method and its application  
113     in robotic welding. Section 4 discusses and evaluates the proposed method's  
114     effectiveness. Conclusions, limitations, and future works are summarized in  
115     section 5.

116

## 2. Methods

117

### 2.1. Hand-Eye Calibration via Standardized Robotic Flanges

118     The hand-eye calibration is a kinematic calibration problem, which usually  
119     involves four coordinate systems, including the base of the manipulator, the

121 tool-mounting flange of the manipulator, the camera frame, and the calibra-  
122 tion object frame, which are denoted by {Base}, {Flan}, {Cam} and {Mark}  
123 respectively. In this paper, we denote  ${}^A_B H$  as the homogeneous transformation  
124 matrix of frame  $B$  relative to  $A$ . The rest of the paper will focus mainly on the  
125 hand-eye calibration problem to demonstrate the proposed method, which can  
126 be further extended to other calibration configurations. The following notes are  
127 usually considered before analysis.

- 128 • The transformation between a manipulator's {Base} and {Flan}, i.e.,  
129  ${}^{\text{Base}}_{\text{Flan}} H$ , is usually known depending on the manipulator's specifications.  
130 • The transformation between the camera {Cam} and object {Mark}, i.e.,  
131  ${}^{\text{Cam}}_{\text{Mark}} H$ , is a calculated matrix based on the camera's optical measurement  
132 of the object in the form of a 2D image or 3D point cloud.  
133 • Depending on the object and camera placement relative to the manipula-  
134 tor, there are two common configurations of hand-eye calibration, namely  
135 Eye-in-Hand and Eye-on-Base configurations in Fig. 2.  
136 • Another more advanced case of co-manipulation is when two manipulators  
137 are involved, equivalent to a combination of the Eye-in-Hand and Eye-  
138 on-Base. Furthermore, we will introduce another configuration of Eye-on-  
139 Arm calibration, where the camera can be fixed at any convenient location  
140 on the arm.

141 *2.1.1. Configuration of Hand-Eye Calibration*

142 For Eye-in-Hand Configuration, the camera is mounted on the manipulator's  
143 wrist near the tool flange, which can be expressed as  ${}^{\text{Flan}}_{\text{Cam}} \hat{H}$ . On the other  
144 hand, the object is placed at a fixed location concerning the manipulator base,  
145 expressed as  ${}^{\text{Mark}}_{\text{Base}} \hat{H}$ . The *hat* on top denotes that this transformation matrix will  
146 be calculated for calibration. Therefore, a closed-loop coordinate transformation  
147 can be formed as follows:

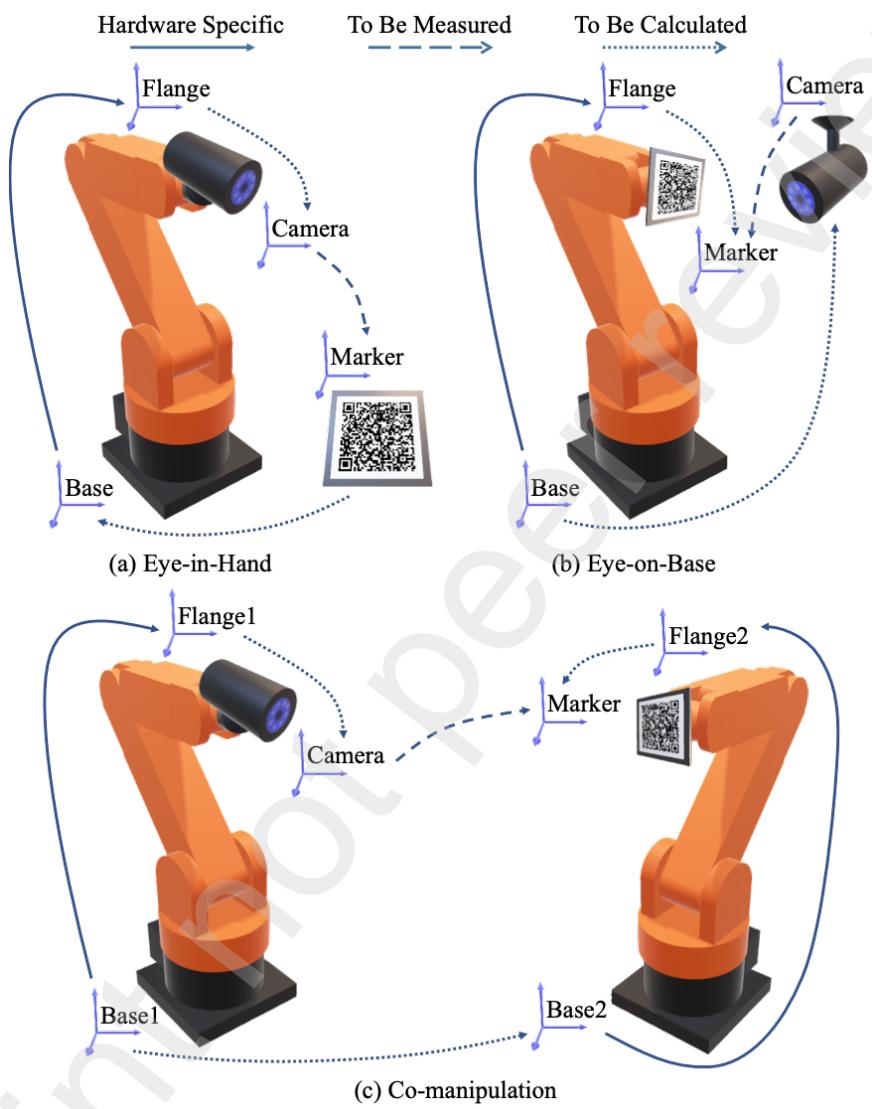


Figure 2: **Common configurations of robot hand-eye calibration:** (a) Eye-in-hand configuration, where the camera is fixed at the manipulator's wrist near the flange; (b) Eye-on-Base configuration, where the camera is fixed outside the manipulator, concerning the base; and (c) Co-manipulation configuration, where the camera is fixed to a robot1's wrist near flange1 and a marker is fixed to a robot2's wrist near flange2.

$$\frac{\text{Base}}{\text{Flan}} H \cdot \frac{\text{Flan}}{\text{Cam}} \hat{H} \cdot \frac{\text{Cam}}{\text{Mark}} H \cdot \frac{\text{Mark}}{\text{Base}} \hat{H} = I. \quad (1)$$

148 Note that in Eq. (1),  $\frac{\text{Base}}{\text{Flan}} H$  is a known matrix based on the manipulator's  
 149 joint configuration, and  $\frac{\text{Cam}}{\text{Mark}} H$  is also a known one based on the camera's mea-  
 150 surement. The iterative method is a standard solution to Eq. (1) by sampling  
 151 multiple points within the camera's view range and the robot's dexterity space.  
 152 For example, by moving the manipulator from point  $p_1$  to  $p_2$  in the configu-  
 153 ration space, the following two equations can be obtained concerning  $\{\text{Flan}\}_{p_1}$   
 154 and  $\{\text{Flan}\}_{p_2}$ .

$$\frac{\text{Base}}{\text{Flan}_{p_2}} H \cdot \frac{\text{Flan}_{p_1}}{\text{Cam}_{p_1}} \hat{H} \cdot \frac{\text{Cam}_{p_1}}{\text{Mark}} H = \frac{\text{Base}}{\text{Flan}_{p_2}} H \cdot \frac{\text{Flan}_{p_2}}{\text{Cam}_{p_2}} \hat{H} \cdot \frac{\text{Cam}_{p_2}}{\text{Mark}} H, \quad (2)$$

$$\frac{\text{Flan}_{p_1}}{\text{Cam}_{p_1}} \hat{H} = \frac{\text{Flan}_{p_2}}{\text{Cam}_{p_2}} \hat{H} = \frac{\text{Flan}}{\text{Cam}} \hat{H}. \quad (3)$$

155 One can further rewrite Eqs. (2) and (3) as  $AX = XB$  to solve for  $\frac{\text{Flan}}{\text{Cam}} \hat{H}$ ,  
 156 where  $X = \frac{\text{Flan}}{\text{Cam}} \hat{H}$ ,  $A = \frac{\text{Base}}{\text{Flan}_{p_2}} H^{-1} \cdot \frac{\text{Base}}{\text{Flan}_{p_1}} H$  and  $B = \frac{\text{Cam}_{p_2}}{\text{Mark}} H \cdot \frac{\text{Cam}_{p_1}}{\text{Mark}} H^{-1}$ . In  
 157 practice, the points to be sampled can be as many as 10~30 points to improve  
 158 the calibration accuracy.

159 In this paper, we use the term "Eye-on-Base" instead of "Eye-to-Hand" to  
 160 differentiate it from the Eye-in-Hand configuration further, as the camera is  
 161 usually mounted at a fixed location in the world frame concerning the manip-  
 162 ulator's base frame, expressed as  $\frac{\text{Base}}{\text{Cam}} \hat{H}$ . A high-precision calibration object  
 163 is usually fixed on the manipulator's wrist near the tool flange, expressed as  
 164  $\frac{\text{Flan}}{\text{Mark}} \hat{H}$ . Similar to the method used in the Eye-in-Hand configuration, we can  
 165 express the transformation of the Eye-on-Base configuration like the following,  
 166 which can be symbolically expressed as  $AX = YB$  and solved using a similar  
 167 iterative method as the Eye-in-Hand configuration.

$$\frac{\text{Base}}{\text{Flan}} H \cdot \frac{\text{Flan}}{\text{Mark}} \hat{H} = \frac{\text{Base}}{\text{Cam}} \hat{H} \cdot \frac{\text{Cam}}{\text{Mark}} H. \quad (4)$$

168 Collaborative robots recently emerged as a viable solution for human-robot  
 169 collaborative tasks such as co-manipulation, where multiple manipulators work  
 170 collaboratively for an integrated task. Following the above notation, we can

172 express the transformation of the co-manipulation as follows: Solving Eq. (5)  
 173 becomes more challenging as three unknown matrices need to be determined  
 174 simultaneously.

$$\frac{\text{Base}^1}{\text{Flan}1} H \cdot \frac{\text{Flan}^1}{\text{Cam}} \hat{H} \cdot \frac{\text{Cam}}{\text{Mark}} H = \frac{\text{Base}^1}{\text{Base}2} \hat{H} \cdot \frac{\text{Base}^2}{\text{Flan}2} H \cdot \frac{\text{Flan}^2}{\text{Mark}} \hat{H} . \quad (5)$$

175 However, one can decompose the co-manipulation problem into a simultaneous  
 176 calculation of an Eye-in-Hand problem and an Eye-on-Base one. The left  
 177 side of Eq. (5) can be rewritten as an Eye-in-Hand problem for robot1 if the  
 178 robot2 with a calibration object is at a fixed pose using Eq. (6)a). Similarly,  
 179 the right side of Eq. (5) can be rewritten as an Eye-on-Base problem for robot2  
 180 when robot 1 with a camera is at a fixed pose in space using Eq. (6)b). As a  
 181 result, Eq. (5) becomes the following.

$$\begin{cases} \frac{\text{Base}^1}{\text{Flan}1} H \cdot \frac{\text{Flan}^1}{\text{Cam}} \hat{H} \cdot \frac{\text{Cam}}{\text{Mark}} H \cdot \frac{\text{Mark}}{\text{Base}1} \hat{H} = I & (a) \\ \frac{\text{Base}^2}{\text{Flan}2} H \cdot \frac{\text{Flan}^2}{\text{Mark}} \hat{H} = \frac{\text{Base}^2}{\text{Cam}} \hat{H} \cdot \frac{\text{Cam}}{\text{Mark}} H & (b) \\ \frac{\text{Base}^1}{\text{Base}2} \hat{H} = \frac{\text{Base}^1}{\text{Flan}1} H \cdot \frac{\text{Flan}^1}{\text{Cam}} \hat{H} \cdot \frac{\text{Base}^2}{\text{Cam}} \hat{H}^{-1} & (c) \end{cases} \quad (6)$$

182 Recent work by (Wu et al., 2016) provides a comprehensive solution to the  
 183 co-manipulation problem similar to Eq. (5), in which the  $\frac{\text{Flan}^1}{\text{Cam}} \hat{H}$  represents the  
 184 Hand-Eye calibration problem for robot1; the  $\frac{\text{Flan}^2}{\text{Mark}} \hat{H}$  represents the Tool-Flange  
 185 calibration problem for robot2; and the  $\frac{\text{Base}^1}{\text{Base}2} \hat{H}$  represents the Robot-Robot cal-  
 186 ibration problem between robot1 and robot2. These three problems can be  
 187 integrated into a matrix equation of  $AXB = YCZ$  for a simultaneous solution.  
 188 Due to the complexity of the problem, the developed algorithm remains chal-  
 189 lenging to implement due to the uncertainties of the sensor noise (Ma et al.,  
 190 2018).

191 The relative placements of the camera and the marker differentiate the con-  
 192 figurations of hand-eye calibration above. This naturally leads to the possibility  
 193 of fixing the camera somewhere in the middle of the robot arm {Arm}, namely  
 194 the Eye-on-Arm configuration in Fig. 3. Depending on the placement of the  
 195 calibration object, the first case of the Eye-on-Arm configuration is when the  
 196 object is fixed on the manipulator's wrist near the robot flange, which is similar

<sup>197</sup> to the Eye-in-Hand configuration in Eq. (1) as  $\text{Base } H \cdot \text{Arm } \hat{H} \cdot \text{Cam } H \cdot \text{Mark } \hat{H} = I$ .

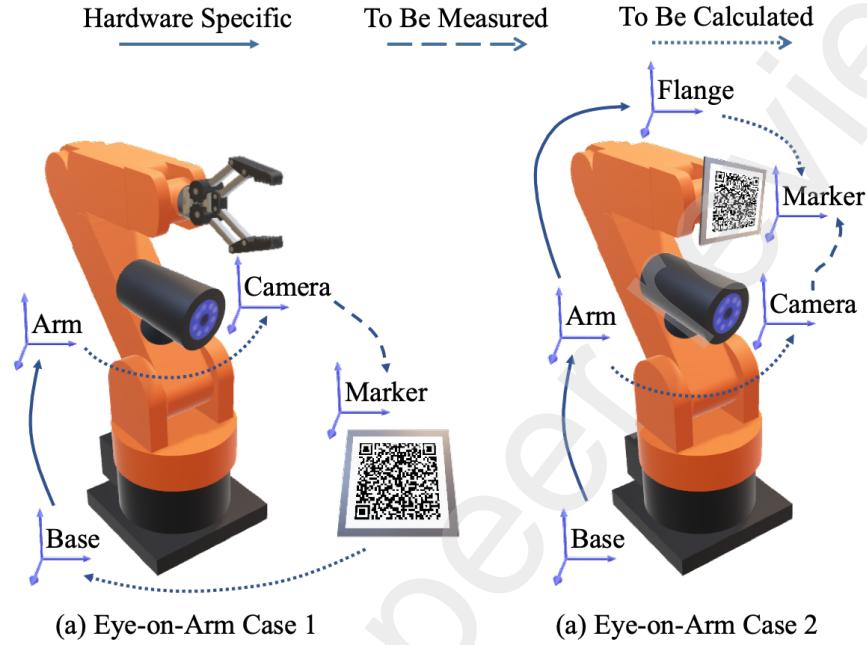


Figure 3: **Eye-on-Arm configuration for robot hand-eye calibration:** (a) Case 1 is equivalent to the Eye-in-Hand configuration; and (b) Case 2 is equivalent to the Eye-on-Base configuration.

<sup>198</sup>

<sup>199</sup> The second case is similar to the Eye-on-Base configuration, where the  
<sup>200</sup> marker object is fixed to a point in space relative to the manipulator's base  
<sup>201</sup> mounting flange. The coordinate transformation of this case can be written as  
<sup>202</sup>  $\text{Base } H \cdot \text{Flan } \hat{H} = \text{Base } H \cdot \text{Arm } \hat{H} \cdot \text{Cam } H \cdot \text{Mark } \hat{H}$ , which is similar to Eq. (4).

#### <sup>203</sup> 2.1.2. Standardized Design of Robot Flanges

<sup>204</sup> The preliminary statistics of the World Robotics Report show a total of  
<sup>205</sup> 3,477,000 units of operational stock of industrial robots worldwide in 2021,  
<sup>206</sup> growing at an average of 14% since 2016. Standardizing robot interfaces at  
<sup>207</sup> various levels is critical to the reusability and exchangeability of robot systems,  
<sup>208</sup> including mechanical, electrical, and communication. Among the International

209 Standard Organization's catalog 25.040.30 industrial robots and manipulators,  
 210 ISO 9404-1 specifies the design standardization of the mechanical interfaces or  
 211 the fixture design on the tool flange. Fig. 4 is adapted from the latest version  
 212 released in 2004, which specifies the critical mechanical interfaces, including the  
 213 threaded holes referencing circle diameter in  $d_1$ , the flange's outer circle diameter  
 214 in  $d_2$ , the number of threaded holes  $N$  to be used for fixture, the size of the  
 threaded holes  $d_4$ , etc.

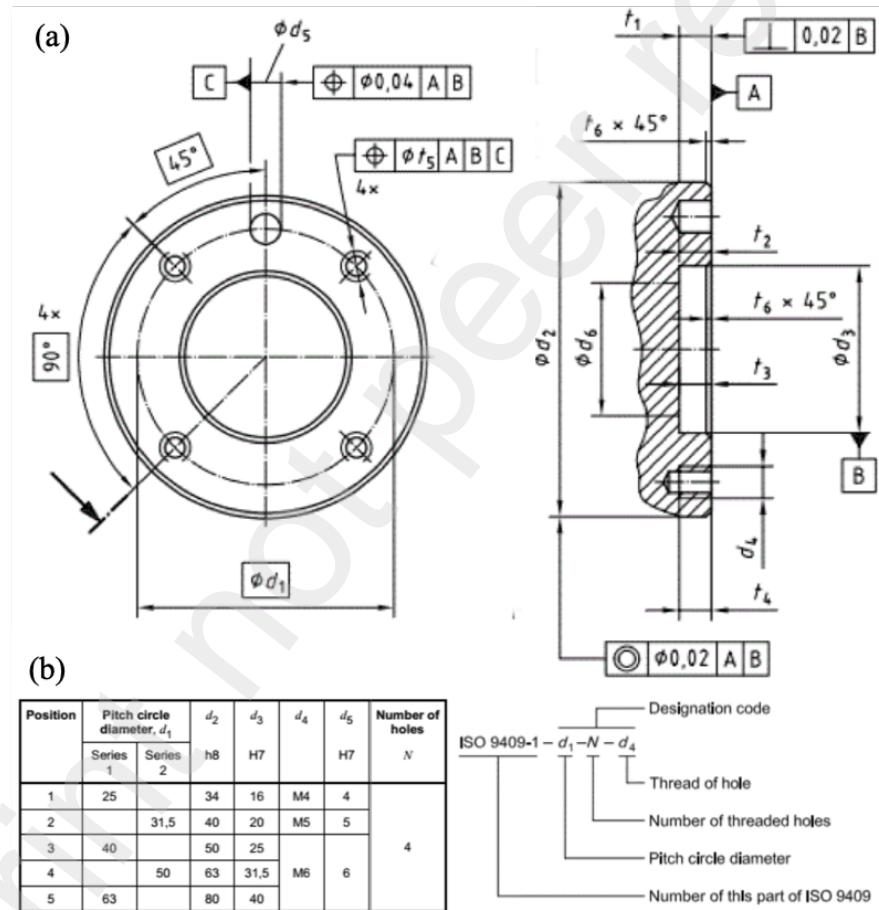
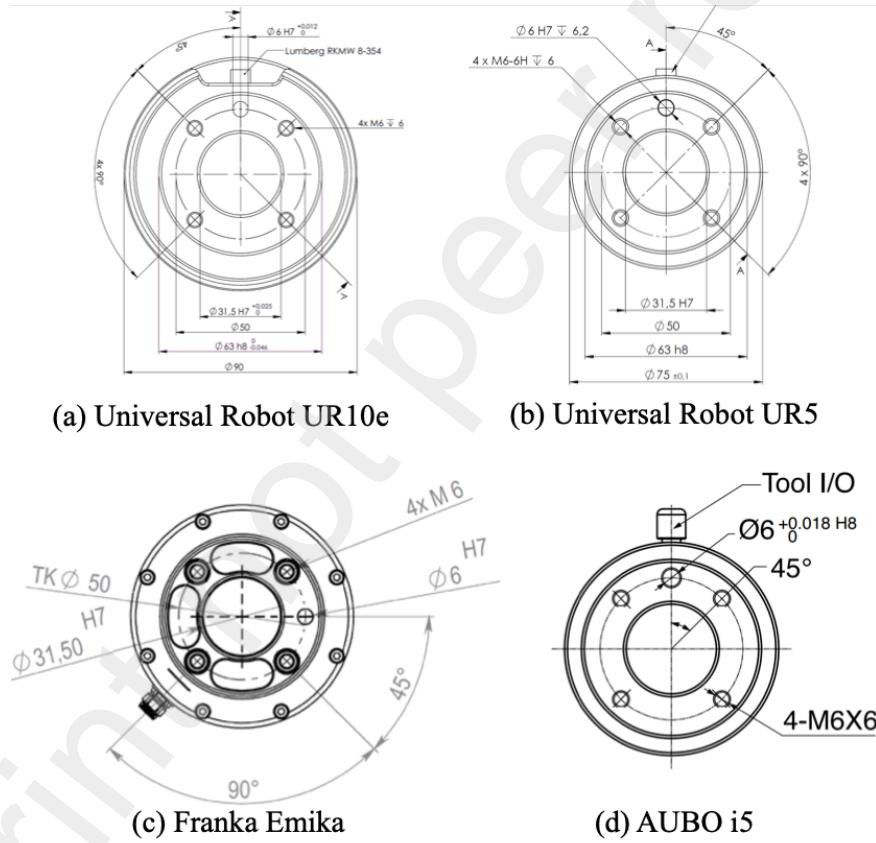


Figure 4: The standardized geometric features on the flange of an industrial manipulator following ISO 9409-1: 2004: (a) mechanical drawings; and (b) recommended dimensions and designation code format.

215

216 A few flange design examples following the ISO 9409-1-50-4-M6 are repro-  
 217 duced in Fig. 5 among common collaborative robot brands. Manipulator and  
 218 end-effector manufacturers following the same designation code can be easily  
 219 attached to accommodate different configurations of the robot systems for vari-  
 220 ous applications. Such standardization also requires the manufacturers to meet  
 221 specific manufacturing qualities to facilitate exchangeability. This study pro-  
 222 poses to utilize such standardized design features for direct hand-eye calibration  
 using high-fidelity 3D scanners, which will be explained next.



223 **Figure 5: The flanges of a few collaborative manipulators following ISO 9409-1-50-4-M6 standard:** (a) Universal Robots' UR5, (b) UR10 e-series, (c) Franka's Emika, and (d) AUBO's i5.

224     2.1.3. Flange-based Hand-Eye Calibration

225     In this section, the Eye-on-Base configuration is used to demonstrate the  
 226     proposed method of flange-based hand-eye calibration. To the authors' best  
 227     knowledge, it is the first time that the tool-mounting flange of the manipulator  
 228     is directly used as a calibration reference, especially when 3D depth sensing is  
 229     used as the "eye" camera. As regulated by ISO 9409-1, the tool flanges are  
 230     usually designed in a circular shape. Therefore, the center of the flange, which  
 231     is also the manipulator's Tool Center Point (TCP), is selected as the referencing  
 232     point for hand-eye calibration. As a result, the hand-eye calibration problem is  
 233     reformulated as

$$\begin{bmatrix} {}_{\text{Base}}p_i \\ 1 \end{bmatrix} = {}_{\text{Cam}}^{\text{Base}}H \cdot \begin{bmatrix} {}_{\text{Cam}}p_i \\ 1 \end{bmatrix}, \quad (7)$$

234     where  $\{({}_{\text{Base}}p_i, {}_{\text{Cam}}p_i) \mid i = 1, 2, \dots, n\}$  is a set of coordinate pair about  
 235     the TCP relative to  $\{\text{Base}\}$  and  $\{\text{Cam}\}$ , respectively. As the marker point is  
 236     also the TCP of the robot arm,  ${}_{\text{Base}}p_i$  can be directly obtained in the robot  
 237     controller, whereas  ${}_{\text{Cam}}p_i$  can be calculated by point cloud algorithms. Then,  
 238     the hand-eye calibration problem becomes finding the least-squares estimation  
 239     of transformation parameters between two sets of 3-DoF data as the following,  
 240      $\min \frac{1}{n} \sum_{i=1}^n \| {}_{\text{Cam}}R \cdot {}_{\text{Cam}}p_i + {}_{\text{Cam}}t - {}_{\text{Base}}p_i \|$ , which has an SVD-based optimal  
 241     solution of  ${}_{\text{Base}}^{\text{Cam}}\hat{H}$  as

$$\begin{cases} {}_{\text{Base}}^{\text{Cam}}\hat{R} = USV^T & (a) \\ {}_{\text{Base}}^{\text{Cam}}\hat{t} = {}_{\text{Base}}\mu - c \cdot {}_{\text{Base}}^{\text{Cam}}\hat{R} \cdot {}_{\text{Cam}}\mu & (b) \end{cases}, \quad (8)$$

242     where

$$\begin{cases} {}_{\text{Base}}\mu = \frac{1}{n} \sum_{i=1}^n {}_{\text{Base}}p_i \\ {}_{\text{Cam}}\mu = \frac{1}{n} \sum_{i=1}^n {}_{\text{Cam}}p_i \\ \Sigma = \frac{1}{n} \sum_{i=1}^n ({}_{\text{Base}}p_i - {}_{\text{Base}}\mu)({}_{\text{Cam}}p_i - {}_{\text{Cam}}\mu)^T \\ c = \frac{n \times \text{tr}(DS)}{\sum_{i=1}^n \| {}_{\text{Cam}}p_i - {}_{\text{Cam}}\mu \|^2} \end{cases}, \quad (9)$$

<sup>243</sup> and let the singular value decomposition of  $\Sigma$  be  $UDV^T$  and

$$S = \begin{cases} I & \text{if } \det(\Sigma_{\text{cov}}) \geq 0 \\ \text{diag}(1, 1, \dots, 1, -1) & \text{if } \det(\Sigma_{\text{cov}}) < 0 \end{cases}. \quad (10)$$

<sup>244</sup> At least four non-coplanar points are required to estimate a unique transfor-  
<sup>245</sup> mation matrix (Wong et al., 2018). The optimal solution essentially represents  
<sup>246</sup> the transformation matrix  ${}_{\text{Base}}^{\text{Cam}'} \hat{H}$  from  $\{\text{Base}\}$  to a calculated camera frame  
<sup>247</sup>  $\{\text{Cam}'\}$ .

<sup>248</sup> *2.1.4. Solvability Analysis*

<sup>249</sup> During calibration, the robot arm is moved to poses such that the tool flange  
<sup>250</sup> plane faces the 3D scanner. To obtain a high-quality point cloud of the tool  
<sup>251</sup> flange, the angle between the normal vector of the tool flange and the optical  
<sup>252</sup> axis should be less than a desirable threshold  $\theta_{\max}$ . The position of the TCP  
<sup>253</sup> relative to the 3D scanner is estimated in two steps.

- <sup>254</sup> • First, pass-through and statistical filters are applied to the original point  
<sup>255</sup> cloud of the scene to remove backgrounds such as table and floor and to  
<sup>256</sup> remove noises. The cloud point of the flange plane can be, therefore, iso-  
<sup>257</sup> lated using basic geometry segmentation algorithms in point cloud library  
<sup>258</sup> (Rusu and Cousins, 2011) together with geometric constraints, such as the  
<sup>259</sup> segmented cluster can not have a range more extensive than the diameter  
<sup>260</sup> of the tool flange.
- <sup>261</sup> • Second, the center of the flange plane is estimated using the RANSAC  
<sup>262</sup> algorithm (Raguram et al., 2012). Moreover, we applied a model check so  
<sup>263</sup> that only circles within the desired radius range would go to the verifica-  
<sup>264</sup> tion stage, improving the search algorithm's efficiency.

<sup>265</sup> While the re-projection error is usually adopted as the error metric for 2D  
<sup>266</sup> hand-eye calibration, it does not apply to the proposed method in this paper as  
<sup>267</sup> our method inherently includes the fitting error of the flange circle. It is easy to

<sup>268</sup> acquire the 3D model of the tool flange to generate a ground true point cloud  
<sup>269</sup> in the tool flange coordinate  ${}^{\text{Flan}}P_{\text{true}}$  in Fig. 6.

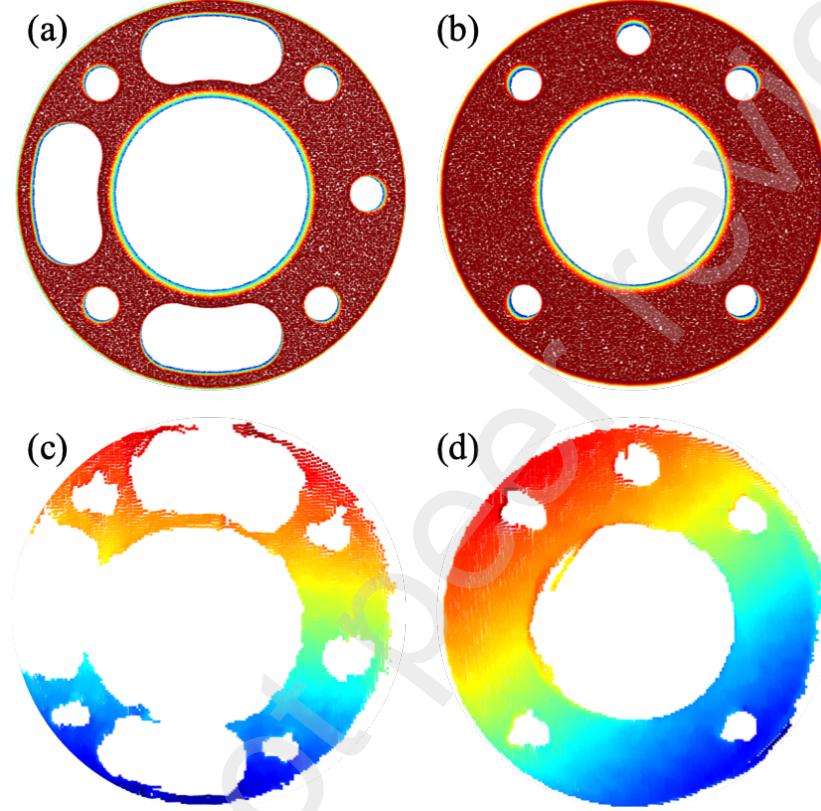


Figure 6: **Point clouds of tool flanges from CAD model and the respective point clouds measured by a 3D scanner:** (a) and (c) are for Franka Emika, and (b) and (d) are for UR5. The color is plotted according to the  $z$ -axis values of the point cloud.

<sup>270</sup> Using a known pose of the robot arm  ${}^{\text{Base}}H_v$  for verification, we can obtain  
<sup>271</sup> the ground true point cloud in the robot coordinate. Given the hand-eye trans-  
<sup>272</sup> formation matrix, the corresponding measured point cloud  ${}^{\text{Cam}}P_v$  can also be  
<sup>273</sup> transformed into the robot coordinate. Then, one can align the ground true  
<sup>274</sup> point cloud and the measured point cloud using the ICP algorithm (Chen and

275 Medioni, 1992). Hence, we define the calibration error as

$$e_{\text{icp}} = \begin{cases} \left\| \begin{bmatrix} {}^{\text{Cam}}P_v \\ 1 \end{bmatrix} - {}^{\text{Base}}\hat{H} \cdot {}^{\text{Base}}H_v \cdot \begin{bmatrix} {}^{\text{Flan}}P_{\text{true}} \\ 1 \end{bmatrix} \right\|_{\text{icp}}, & \text{if ICP fails} \\ +\infty, & \text{if ICP fails} \end{cases}, \quad (11)$$

276 where  ${}^{\text{Cam}'}\hat{H}$  is the transformation matrix from calculated  $\{\text{Cam}'\}$  to  $\{\text{Base}\}$ ,  
277  $\|{}^{\text{Cam}}P_v - {}^{\text{Cam}'}P_v\|_{\text{icp}}$  calculate the rotation error  $\delta R \in SO(3)$  and translation  
278 error  $\delta t \in \mathbb{R}^3$  on the Euclidean group of rigid-body motions  $SE(3)$  such that the  
279 two point clouds are registered. The ICP error metric  $e_{\text{icp}}$  essentially registered  
280 the ground true camera frame  $\{\text{Cam}\}$  and the calculated camera frame  $\{\text{Cam}'\}$ ,  
281 namely  ${}^{\text{Cam}'}H$ , which is more appropriate and informative than traditional 2D  
282 error metric. In 2D hand-eye calibration, it is common to find that the hand-  
283 eye calibration error in position is small at the center of the camera's field  
284 of view and increases in marginal areas. The rotation error of the hand-eye  
285 calibration causes this phenomenon. With the help of a 3D scanner, the reason  
286 can be immediately verified and visualized. Before the calibration starts, the  
287 verification point cloud  $P_v$  is suggested to be collected at a pose where the robot  
288 arm will mostly workaround.

289 A common problem during hand-eye calibration in 2D or 3D is the sampling  
290 quality, which usually requires further optimization. Typical issues include par-  
291 tial sampling, occlusion, and inaccurate circle fitting using the standardized  
292 RANSAC algorithm. Therefore, the analytical solution must collect more than  
293 four points to ensure a high-quality calibration.

294 Therefore, we propose an online iterative calibration method described in  
295 Algorithm 1 to collect as few points as possible and increase the efficiency of  
296 the calibration process.

297 The main difference is that it includes a self-verification mechanism such  
298 that the online calibration process becomes a closed loop. The goal is to  
299 maintain an optimized pool of four pairs of point cloud and robot pose to  
300 minimize the ICP error metric in  $SE(3)$ , which requires a real-valued cost

---

**Algorithm 1** The iterative flange calibration algorithm.

---

**Input:** Ground true point cloud  $P_{\text{true}}$ , verification point cloud  $P_v$ , flange radius

$R$ , calibration threshold  $e_{\text{required}}$ , maximum iteration  $k < k_{\max}$ ;

**Output:** Hand-eye transformation  $\hat{H}_{\text{optimal}}$  and ICP error metric  $e_{\text{icp}}$

```
1:  $k \leftarrow 0$ ;
2: Collect a set of four initial pairs of point cloud and robot pose  $S_k =$ 
    $\{{}^{\text{Cam}}P_i, {}^{\text{Base}}\text{pose}_i \mid i = 1, 2, 3, 4\}$ ;
3: Calculate  $\hat{H}$  and  $e_{\text{icp}}$  and  $\hat{H}_{\text{optimal}} \leftarrow \hat{H}$ ,  $e_{\text{optimal}} \leftarrow e_{\text{icp}}$ ;
4: while  $\|e_{\text{optimal}}\|_{\text{cost}} > e_{\text{required}}$  and  $k \leq k_{\max}$  do
5:   Collect a new pair of point cloud and robot pose  $S'_k$ ;
6:   for  $i$  in  $\{1, 2, 3, 4\}$  do
7:     Replace the  $i$ th pair in  $S_k$  by the new pair of data;
8:     Calculate  $\hat{H}$  and ICP error  $e_{\text{icp}}$ ;
9:     if  $\|e_{\text{icp}}\|_{\text{cost}} \geq \|e_{\text{optimal}}\|_{\text{cost}}$  then
10:       Undo the replacement
11:     else
12:        $\hat{H}_{\text{optimal}} \leftarrow \hat{H}$ 
13:        $e_{\text{optimal}} \leftarrow e_{\text{icp}}$ 
14:     end if
15:   end for
16: end while
```

---

301 metric  $\|\cdot\|_{\text{cost}} : SE(3) \mapsto \mathbb{R}$ . The calibration process keeps adding new  
302 data pairs to the pool and retaining the optimal four pairs with the least  
303 cost. The online calibration process stops once the cost metric has achieved  
304 a target error  $e_{\text{required}}$ . In practice, the design of the cost metric  $\|\cdot\|_{\text{cost}}$   
305 can be flexible according to the application scenario. Codes are available at  
306 [https://github.com/ancorasir/flange\\_handeye\\_calibration](https://github.com/ancorasir/flange_handeye_calibration).

307 *2.2. Soft Robotic Tactile Welding via Multi-Modal Fusion*

308 *2.2.1. Touch-based Welding Tool Design*

309 We proposed a touch-based welding tool with soft robotic metamaterial  
310 (SRM) based on our previous work (Wu et al., 2024). As shown in Fig. 7(a),  
311 the welding tool consists of a flange plate, a camera, a mounting base, an ArUco  
312 tag, a soft robotic metamaterial (SRM), and a welding torch. The whole tool  
313 can be mounted on the robot with the left flange. The key features are the  
314 right flange-like geometric feature facing towards the 3D scanner on the top of  
315 the plate for hand-eye calibration and a soft robotic metamaterial (SRM) with  
316 tactile sensing capability. As shown in 7(b), The SRM has omni-directional de-  
317 formation capability when contacting with surrounding objects, such as welding  
318 seam, which is suitable for local seam path detection. During welding, the SRM  
319 deforms when there is a deviation between the robot and the intended paths.  
320 This deformation could be predicted by tracking the displacement and pose of  
321 the ArUco tag.

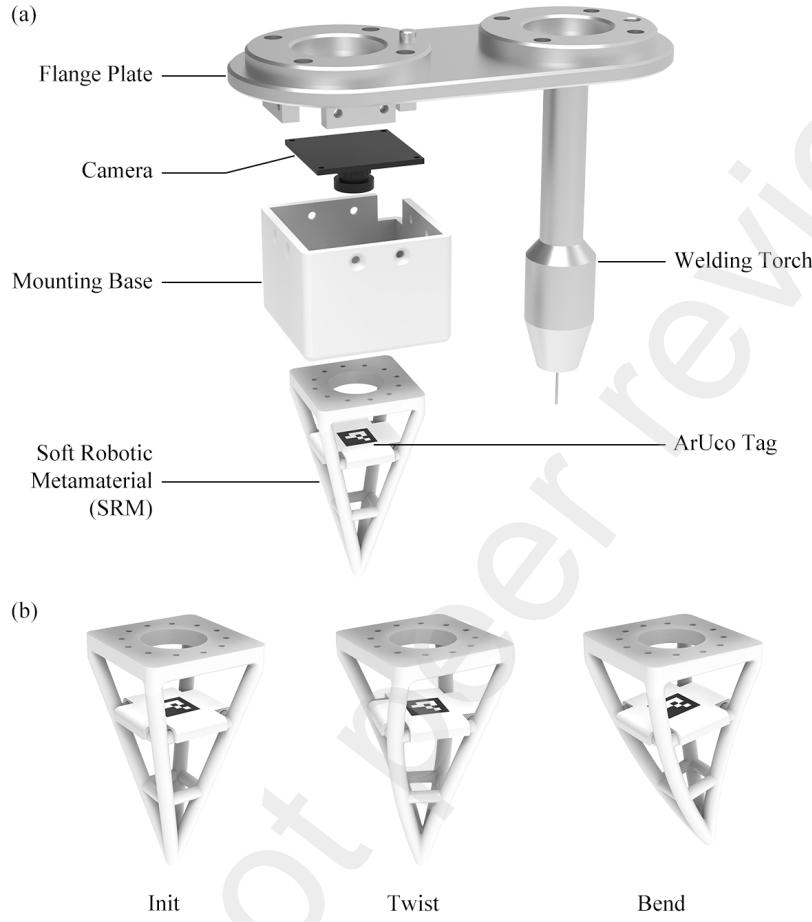


Figure 7: **Welding tool with soft robotic metamaterial (SRM).** (a) The welding tool consists of a flange plate, a camera, a mounting base, an ArUco tag, an SRM, and a welding torch. (b) The SRM has omni-directional deformation capability when twisting or bending, and the position of the tip can be predicted by tracking the ArUco tag.

*322 2.2.2. Welding Seam Tracking by Soft Touch*

*323 Before welding starts, a welding seam path can be captured and extracted  
 324 using 3D vision sensors. However, due to unavoidable error sources, such as  
 325 welding seam feature extraction or hand-eye calibration, the welding path for  
 326 the robot to execute can be unpredictable. It may lead to significant deviations  
 327 from the intended path. Therefore, the fusion of tactile welding seam tracking*

328 with 3D vision could offer a comprehensive solution to address these challenges.

329 The SRM introduces a novel approach for tracking the welding seam by  
330 leveraging a simple mechanism of deformation servo. Commanding the SRM  
331 tip to maintain contact with one side of the welding seam effectively follows  
332 the planned welding path from 3D vision, thus compensating for potential noise  
333 and errors inherent in the vision data. Specifically, this is achieved by direct-  
334 ing the soft torch to follow the planned path only in its tangential direction  
335 while maintaining a predefined deformation along its normal direction. This  
336 strategy ensures that the SRM consistently identifies and tracks the side of the  
337 welding seam without losing its position. Additionally, the tip of the SRM  
338 is continuously guided to remain within the V-shaped welding seam, allowing  
339 for the recording of its position and generating accurate real torch path points  
340 based on this recorded information. This innovative approach demonstrates a  
341 promising method for reliable welding seam tracking with the SRM, mitigating  
potential errors and ensuring precise and consistent weld quality.

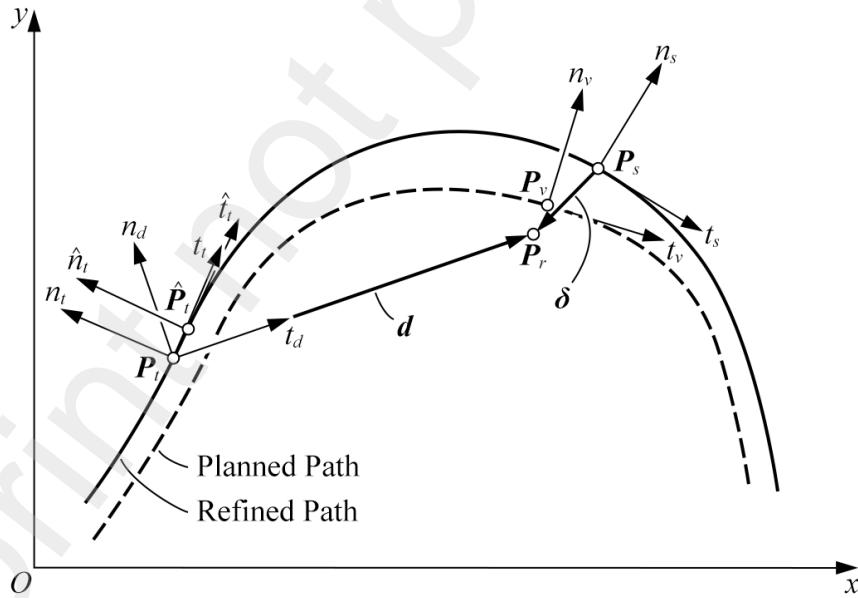


Figure 8: **Online robotic welding trajectory generation enhanced by tactful weld seam tracking**

342 Here, we describe the online robotic welding trajectory generation enhanced  
 343 by tactual weld seam tracking. As shown in Fig 8, we consider the online  
 344 welding trajectory generation problem in a 2D plane. Denote  $\mathbf{P}_r = [x_r, y_r]$  as  
 345 the position vector of the robot end effector in a fixed reference frame  $\{O\}$ ,  
 346  $\mathbf{P}_s = [x_s, y_s]$  as the position vector of SRM tip in reference frame  $\{O\}$ , and  
 347  $\mathbf{P}_t = [x_t, y_t]$  as the position vector of welding torch in the same reference frame.

348 The direction of a touch-based welding tool can be described by orthonormal  
 349 vectors obtained from the rotation angle  $\alpha$  of the robot end effector. These  
 350 vectors are defined as  $\vec{t}_d = [\cos \alpha, \sin \alpha]$  and  $\vec{n}_d = [-\sin \alpha, \cos \alpha]$ , respectively.  
 351 Because of the rigid connection between the position vector of the robot end effector  
 352 and the welding torch, we can express this relationship using the following  
 353 equation:

$$\begin{cases} \mathbf{d} = \|\mathbf{d}\| \vec{t}_d \\ \mathbf{P}_r = \mathbf{P}_t + \mathbf{d} \end{cases}, \quad (12)$$

354 where,  $\mathbf{d}$  represents the position difference vector between the robot end effector  
 355 and the welding torch, and  $\mathbf{P}_r$  can be obtained from the robot controller.

356 At each point along the welding path planned using 3D vision data, when  
 357 the SRM makes contact with one side of the welding seam, the deformation  
 358 can be sensed and calculated by utilizing the displacement of the Aruco marker  
 359  $\delta \in \mathbb{R}^2$  detected through the in-finger camera of the SRM. This computation is  
 360 represented as:

$$\begin{cases} \mathbf{P}_r^{t_s} = \mathbf{P}_s^{t_s} + \delta^{t_s} \\ \mathbf{P}_r^{n_s} = \mathbf{P}_s^{n_s} + \delta^{n_s} \end{cases}, \quad (13)$$

361 where,  $\mathbf{P}_r^{t_s}$ ,  $\mathbf{P}_s^{t_s}$  and  $\delta^{t_s}$  represent the corresponding position vectors projected  
 362 along the tangential direction at each path point, while  $\mathbf{P}_r^{n_s}$ ,  $\mathbf{P}_s^{n_s}$  and  $\delta^{n_s}$  denote  
 363 vectors along the normal direction.

364 To uphold the SRM's deformation along the normal direction, we initially  
 365 establish a desired deformation along the normal direction, denoted as  $\delta_d^{n_s}$ .  
 366 Subsequently, to maintain this directional deformation, the robot end effector

<sup>367</sup> velocity can be commanded using the following formula:

$$\begin{cases} \mathbf{V}_r^{t_s} = \mathbf{V}_{const} \\ \mathbf{V}_r^{n_s} = -k_p(\delta^{n_s} - \delta_d^{n_s}) \end{cases} \quad (14)$$

<sup>368</sup> where, the robot end effector velocity along the tangential direction  $\mathbf{V}_r^{t_s}$  main-  
<sup>369</sup> tains  $\mathbf{V}_{const}$ , which is a constant norm vector but changes direction. While in  
<sup>370</sup> the normal direction, the velocity of the robot end effector  $\mathbf{V}_r^{n_s}$  opposes the dif-  
<sup>371</sup> ference between the current detected Aruco marker displacement in the normal  
<sup>372</sup> direction and the pre-defined deformation. This action serves as a resilience  
<sup>373</sup> measure to maintain the deformation at the desired level, with the action being  
<sup>374</sup> governed by the parameter  $k_p$ .

<sup>375</sup> During the deformation servo, the position of the SRM tip, denoted as  $\mathbf{P}_s$   
<sup>376</sup> is recorded as a refined welding seam path to be subsequently executed by the  
<sup>377</sup> real welding torch. And this path will be executed by the real welding torch.  
<sup>378</sup> As the welding torch is rigidly coupled to the robot end effector, the implicated  
<sup>379</sup> velocity of the torch can be calculated as:

$$\begin{cases} \mathbf{V}_t^{t_d} = \mathbf{V}_r^{t_d} \\ \mathbf{V}_t^{n_d} = \mathbf{V}_r^{n_d} + \omega \|\mathbf{d}\| \vec{n}_d \end{cases}, \quad (15)$$

<sup>380</sup> where, the implicated velocity of the real welding torch  $\mathbf{V}_t$  due to the velocity of  
<sup>381</sup> the robot end effector  $\mathbf{V}_r$  is projected along the normal direction and tangential  
<sup>382</sup> direction at the point of the welding torch.  $\|\mathbf{d}\|$  represents the distance between  
<sup>383</sup> the robot end effector and welding torch in the defined 2D plane, while  $\omega$  denotes  
<sup>384</sup> the angular velocity of the robot end effector.

<sup>385</sup> The requirement for the welding torch velocity to align with the refined  
<sup>386</sup> welding path dictated by the SRM tip implies that the normal component of the  
<sup>387</sup> torch's implicated velocity should be zero. This relationship can be expressed  
<sup>388</sup> as:

$$\mathbf{V}_t^{n_d} \cdot \vec{n}_t + \mathbf{V}_t^{t_d} \cdot \vec{n}_t = 0, \quad (16)$$

<sup>389</sup> where,  $\vec{n}_t$  denotes the normal direction at the refined path point of the welding  
<sup>390</sup> torch.

391 The algorithm of online robotic welding trajectory generation enhanced by  
 392 tactual weld seam tracking is shown in Algorithm 2.

393 **3. Results**

394 *3.1. Simulation Results for Flange-based Hand-Eye Calibration*

395 Our simulation of the Eye-on-Base configuration in the Gazebo involves a  
 396 6-DOF manipulator of UR5 and a depth camera. The ground-true value of the  
 397 hand-eye transformation matrix is set as the following, similar to a real scenario,  
 398 including a roll-pitch-yaw of [3.1415, 0, -1.57] in radians between the robot base  
 399 and camera and a translational vector of [0.6, -0.0125, 1] in meters.

$$\text{Base}_{\text{Cam}}^{\text{Cam}} H_{\text{true}} = \begin{bmatrix} 0 & -1 & 0 & 0.6 \\ -1 & 0 & 0 & -0.0125 \\ 0 & 0 & -1 & 1 \\ 0 & 0 & 0 & 1 \end{bmatrix}. \quad (17)$$

400 A total sample of seventy-five tool flange poses was recorded to test the ro-  
 401 bustness of the proposed method. The manipulator's tool flange moved in a  
 402 grid pattern within a workspace of 0.3 m × 0.3 m × 0.2 m with random orienta-  
 403 tions. The actual positions of the TCP of the robot concerning the camera were  
 404 obtained using TCP concerning the robot base and the ground true hand-eye  
 405 transformation matrix. To investigate the proposed iterative calibration method  
 406 under various levels of disturbance to the 3D scanner, different Gaussian noise  
 407  $N(0, \sigma_{\text{noise}})$  were added directly to the actual values. The disturbed point is  
 408 defined by

$$\text{Cam}_{\text{Base}}^{\text{Cam}} p_i = \text{Base}_{\text{Base}}^{\text{Cam}} R_{\text{true}} \cdot \text{Base}_{\text{Base}}^{\text{Cam}} p_i + \text{Base}_{\text{Base}}^{\text{Cam}} t_{\text{true}} + N(0, \sigma_{\text{noise}}), \quad (18)$$

409 where  $\text{Base}_{\text{Base}}^{\text{Cam}} R_{\text{true}}$  and  $\text{Base}_{\text{Base}}^{\text{Cam}} t_{\text{true}}$  are the rotational and translational parts of the  
 410 ground true hand-eye transformation matrix. The standard deviation of the  
 411 Gaussian noise  $\sigma_{\text{noise}}$  ranged from 0.2 mm to 10 mm at a step of 0.2 mm,  
 412 representing the range of precision of industrial-grade and consumer-grade 3D

---

**Algorithm 2** Online welding trajectory generation via multi-modal fusion

---

**Require:** Welding seam path points  $\{\mathbf{P}_v^1, \mathbf{P}_v^2, \dots, \mathbf{P}_v^n\}$  planned from 3D vision sensor, servo parameter  $k_p$ ;

- 1: Initialize desired deformation along normal direction  $\delta_d$ ;
- 2: **while** True **do**
- 3:      $\alpha \leftarrow$  current rotation angle of the robot end effector;
- 4:      $\vec{n}_d, \vec{t}_d \leftarrow$  orthonormal vectors of touch-based welding tool;
- 5:      $\mathbf{P}_r \leftarrow$  current position of robot end effector;
- 6:      $\delta^{n_s}, \delta^{t_s} \leftarrow$  current Aruco marker displacement along each direction;
- 7:      $\{\hat{\mathbf{P}}_t^1, \hat{\mathbf{P}}_t^2, \dots, \hat{\mathbf{P}}_t^m\} \leftarrow \text{REFINE}(\mathbf{P}_r, \delta^{n_s}, \delta^{t_s})$ ;  $\triangleright$  refined path points
- 8:      $\mathbf{V}_r \leftarrow \text{SERVO}(\mathbf{P}_r, \delta^{n_s})$ ;  $\triangleright$  linear velocity of robot end effector
- 9:      $\mathbf{P}_t \leftarrow \mathbf{P}_r - \|\mathbf{d}\| \vec{t}_d$ ;  $\triangleright$  position of torch
- 10:     $\hat{\mathbf{P}}_t \leftarrow \arg \min_{j \in \{1, 2, \dots, m\}} \|\mathbf{P}_t - \hat{\mathbf{P}}_t^j\|$   $\triangleright$  nearest point to  $\mathbf{P}_t$  in refined path
- 11:     $\vec{n}_t, \vec{t}_t \leftarrow$  orthonormal vectors at point  $\hat{\mathbf{P}}_t$
- 12:     $\mathbf{V}_r^{n_d}, \mathbf{V}_r^{t_d} \leftarrow$  project  $\mathbf{V}_r$  along  $\vec{n}_d, \vec{t}_d$
- 13:     $\omega \leftarrow -\frac{\mathbf{V}_r^{t_d} \cdot \vec{n}_t + \mathbf{V}_r^{n_d} \cdot \vec{t}_t}{\|\mathbf{d}\| \vec{n}_d \cdot \vec{n}_t}$   $\triangleright$  angular velocity of robot end effector
- 14:    apply velocity  $(\mathbf{V}_r, \omega)$  to robot end effector;
- 15: **end while**
- 16: \_\_\_\_\_
- 17: **function** REFINE( $\mathbf{P}_r, \delta^{n_s}, \delta^{t_s}$ )
- 18:      $\mathbf{P}_s \leftarrow \mathbf{P}_r^{n_s} - \delta^{n_s} + \mathbf{P}_r^{t_s} - \delta^{t_s}$   $\triangleright$  refine welding position vector
- 19:      $\hat{\mathbf{P}}_t^j \leftarrow \mathbf{P}_s$   $\triangleright$  append point to refined path
- 20:     **return**  $\{\hat{\mathbf{P}}_t^1, \hat{\mathbf{P}}_t^2, \dots, \hat{\mathbf{P}}_t^j\}$
- 21: **end function**
- 22: **function** SERVO( $\mathbf{P}_r, \delta^{n_s}$ )
- 23:      $\mathbf{P}_v \leftarrow \arg \min_{i \in \{1, 2, \dots, n\}} \|\mathbf{P}_r - \mathbf{P}_v^i\|$   $\triangleright$  nearest point to  $\mathbf{P}_r$  in planned path
- 24:      $\vec{n}_v, \vec{t}_v \leftarrow$  orthonormal vectors at point  $\mathbf{P}_v$
- 25:      $\mathbf{V}_r \leftarrow V_{const} \vec{t}_v - k_p (\delta^{n_s} - \delta_d)$   $\triangleright$  resultant linear velocity
- 26:     **return**  $\mathbf{V}_r$
- 27: **end function**

---

413 scanners. For each level of Gaussian noise, the calibration results were evaluated  
414 over 100 random realizations of the noise. The rotation error  $\delta R$  is expressed  
415 in terms of roll, pitch, and yaw vectors ( $\delta_{\text{roll}}$ ,  $\delta_{\text{pitch}}$ ,  $\delta_{\text{yaw}}$ ). The ICP error metric  
416 can be directly calculated by  $e_{\text{icp}} = \frac{\text{Cam}'}{\text{Base}} \hat{H} \cdot \frac{\text{Base}}{\text{Cam}} H_{\text{true}}$  and the cost metric  
417 is defined as the Euclidean norm of the translation error vector  $\|e_{\text{icp}}\|_{\text{cost}} =$   
418  $\|\delta t\|_2 = \|(\delta_x, \delta_y, \delta_z)\|_2$ .

419 The results in Fig. 9 show the calibration error results using all data points  
420 versus the proposed iterative method, in which the translational and rotational  
421 errors are plotted separately concerning the various noise levels. All mean values  
422 of the calibration errors remain close to zero as there is no systematic error.

423 The standard deviations of the calibration errors for both methods grow  
424 linearly concerning the amplitude of the Gaussian noise. However, as shown in  
425 the left column, the iterative method is much less sensitive to noise, and the  
426 growth rate of the standard deviation of the translational error is reduced to  
427 1/7 compared to using all data. It was also found that the translational errors  
428 are mainly contributed to by the  $x$  and  $y$  components. Hence, when minimizing  
429 the Euclidean norm of the translational error, the  $z$ -axis translational errors of  
430 both methods remain at the same level, as shown in Fig. 9. In the right column  
431 of Fig. 9, rotational error is mainly contributed by the yaw component.

432 Fig. 10 shows the calibration errors concerning the iterative steps when  
433  $\sigma_{\text{noise}}$  is set at 1 mm, and all the means of the calibration errors remain close  
434 to zero.

435 In Fig. 10(a), all standard deviations of the translation errors converge  
436 to less than 1 mm within ten sampled points and finally converge to 0.2 mm  
437 within 20 sampled points. In Fig. 10(b), all the standard deviations of the  
438 rotational errors converge to less than half of the initial values even though only  
439 the translational errors are optimized. This supports the fact the translational  
440 and rotational errors are coupled, meaning that the decrease of one will lead to  
441 a decrease of the other. The primary rotational error is a yaw angle of about  
442 0.2 degrees. This corresponds to a 1 mm displacement in the  $x$ - $y$  plane at a  
443 distance  $r$  of 300 mm away from the optical axis in our simulation setup, which

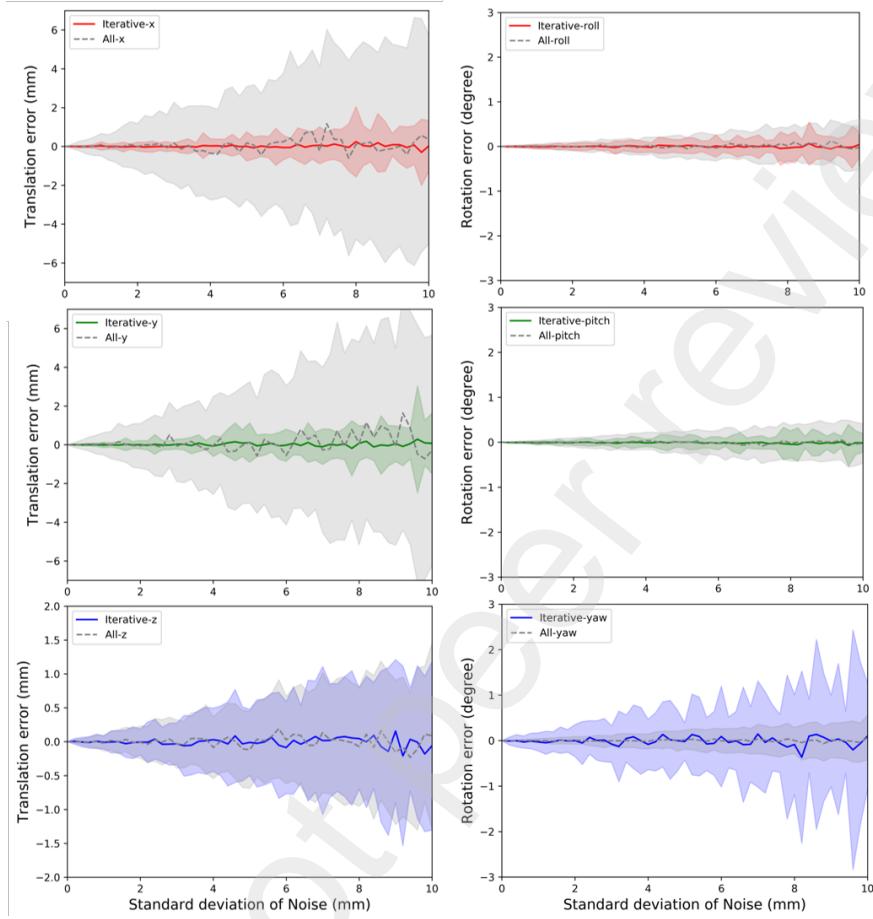


Figure 9: **Comparison of the calibration error of using all data points (gray) and that of the proposed iterative method (colored).** The gray dashed line and gray shaded area denote the mean and standard deviation of the calibration error metric using all 75 data points evaluated over 100 random realizations of Gaussian noises. The solid colored line and shaded area denote the mean and standard deviation of the calibration error metric using the proposed iterative method.

is estimated as  $\delta_{\text{yaw}} \cdot r$ .

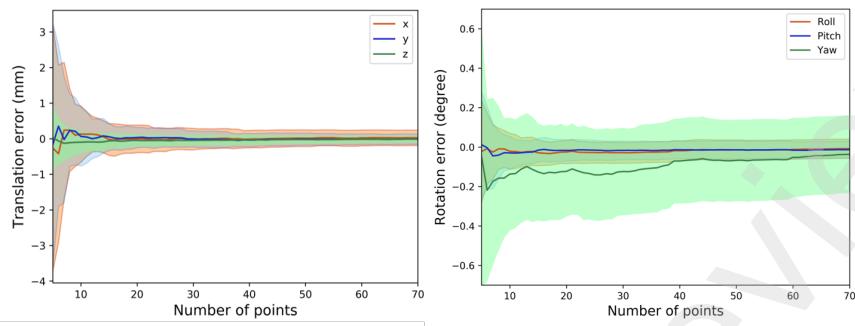


Figure 10: **The calibration error metric concerning the iterative steps.** The results are evaluated over 100 random realizations. Means are plotted in lines, and standard deviations are plotted as shaded areas.

445     3.2. *Hardware Results for Flange-based Hand-Eye Calibration*

446     3.2.1. *Setup*

447     The proposed iterative calibration method was applied to the hand-eye cal-  
 448     ibration between two industrial-grade 3D scanners (PhoXi S model and PhoXi  
 449     M model) and four robot arms consisting of UR5, UR10e, Aubo-i5, and Franka.  
 450     The Phoxi series scanners are based on structured light and produce up to 3.2  
 451     million 3D points. The working distances of PhoXi S and M are 384 to 520 mm  
 452     and 458 to 1,118 mm, respectively. Their temporal noise and calibration accu-  
 453     racy are 0.05 mm and 0.1 mm, respectively. The 3D scanner was mounted about  
 454     1 meter above the table facing downward, and the robot arms were mounted on  
 455     the table. The 3D scanner and robot arms were controlled by a laptop, which  
 456     was also responsible for collecting and processing all the data. The computer  
 457     has an Intel 2.80 GHz i7 CPU and 16 GB RAM.

458     During the experiment, the robot arms moved in a grid pattern in random  
 459     orientations within the workspace while the roll, pitch, and yaw concerning the  
 460     robot base were all less than 0.3 radians. After segmentation and RANSAC  
 461     fitting, the TCP positions in the 3D scanner frame were computed from the  
 462     point cloud. We collected 60 to 80 point clouds for each robot arm, and the  
 463     valid pairs of point cloud and robot pose are slightly less than these numbers

464 as the flange segmentation failed occasionally. For analysis, we collected all the  
465 data and calculated the hand-eye calibration. However, in practice, the user  
466 should use the iterative calibration method online and do the calibration while  
467 collecting data. We shuffled the order of point clouds 50 times for each robot  
468 arm to mimic different data collection processes in which ill-conditioned points  
469 came at different sequences. The calibration results were then evaluated over  
470 these 50 calculations.

471 *3.2.2. Raw Data Processing*

472 During the segmentation process for UR5, three false segmentations of the  
473 tool flange were purposely retained in the calibration point cloud set to demon-  
474 strate the influence of significant outliers on the performance of the proposed  
475 iterative calibration methods. The false segmented point clouds correspond to  
476 the other parts of the robot arm instead of the tool flange.

477 The quality of the point cloud of tool flanges varied among robot arms. The  
478 point clouds of UR’s tool flange were of good quality. For Aubo-i5, part of  
479 the tool flange was missing in some point clouds, partly due to its black color  
480 finish. Franka’s point clouds were incomplete due to the unique interface design.  
481 However, the RANSAC circle fitting algorithm was robust even for incomplete  
482 point clouds of tool flanges. The parameters of the RANSAC are as follows: The  
483 sample size was three, which is the minimum number to define a circle. The  
484 radius of flange  $R = 31$  mm. Distance threshold  $e_d = 0.3$  mm, which should be  
485 appropriately chosen according to the precision of the 3D scanner. Radius error  
486 tolerance  $e_r = 1$  mm, max iteration  $k_{\max} = 10,000$ .

487 *3.2.3. Results*

488 Table 1 lists the calibration results with all the sample points using the SVD  
489 method. In the UR5 case, where the calibration error is so significant that the  
490 ICP algorithm fails, the translation between the TCP position measured from  
491 the point cloud and the TCP read from the robot arm controller was calculated  
492 first. This translation was used as the initial guess for the transformation of the

493 ICP algorithm. Although only 3 out of 54 sample point clouds are false segmen-  
 494 tations, the translation error is over 10 mm, leading to a practically unusable  
 495 hand-eye matrix. Franka’s calibration errors are also significant, although all  
 496 the flanges are correctly segmented. This is due to the geometry feature of  
 497 Franka’s flange, where the circle fitting is more challenging. The calibration  
 498 results of UR10e are relatively better, with translational errors of less than 2  
 499 mm. Aubo i5 and Phoxi S achieved the best calibration results with less than  
 500 0.4 mm translation errors. The tool flange of Franka and that of Aubo i5 are of  
 501 the same geometry and correctly segmented in the experiment. The calibration  
 502 results from Aubo i5 and Phoxi S suggest that 3D scanners with higher precision  
 lead to higher calibration accuracy.

Table 1: Calibration errors using all sample points without iterative optimization.

Robot Setups	<i>x</i> (mm)	<i>y</i> (mm)	<i>z</i> (mm)	roll (°)	pitch (°)	yaw (°)
UR5 PhoXi M	11.60	2.81	5.55	0.49	-0.89	-0.47
UR10e PhoXi M	0.60	1.75	-0.41	0.10	-0.04	-0.23
Franka PhoXi M	4.34	-4.06	-0.13	-0.30	-0.34	-0.07
Aubo i5 PhoXi S	-0.26	-0.27	-0.35	0.14	0.04	0.01

503  
 504 Fig. 11 shows the calibration errors concerning the iterative steps for robot  
 505 UR5, UR10e, Aubo-i5, and Franka. In all four scenarios, the translation errors  
 506 are mainly contributed by *x* and *y* components, and the rotation errors are  
 507 mainly by the yaw component, which agrees well with the simulation results.  
 508 UR5, UR10e, and Aubo i5 converge to less than 0.28 mm, while their rotational  
 509 errors converge to less than 0.25 degrees. Despite the challenge in circle fitting,  
 510 the translation errors of Franka converge to less than 0.4 mm at the expense of

511 growth in yaw-component rotation error converging to less than 0.6 degrees.

512 *3.2.4. Accuracy and Efficiency*

513 The hand-eye calibration accuracy based on 3D geometry is highly depen-  
514 dent on the precision of 3D scanners, the geometry of the tool flanges, and the  
515 robustness of the feature extraction process. The proposed iterative calibration  
516 method achieves much better accuracy than only using point data, especially  
517 when there are significant outliers in the data. The convergence speed and  
518 amplitude of calibration errors depend on the point cloud quality of the tool  
519 flange and the robustness of the circle fitting algorithms. The iterative method  
520 is statistically meaningful as it resists various noise levels. In practice, a more  
521 efficient online iterative calibration process is to perform error compensation  
522 on the calculated hand-eye matrix once a valid ICP error metric is obtained in  
523 steps 3 or steps 12 and 13 in Algorithm 1. The error compensation is formulated  
524 as

$$\hat{H}_{\text{compensated}} = \hat{H}_{\text{optimal}} \cdot e_{\text{optimal}}. \quad (19)$$

525 In the experiment results, the initial  $e_{\text{optimal}}$  in step 3 in Algorithm 1 is  
526 sometimes infinite due to noisy data. However, after the first iteration, the  
527 chance of  $e_{\text{optimal}}$  remaining infinity is found to be small. In Fig. 11, all 50  
528 random calibration calculations conducted for each robot arm have finite  $e_{\text{optimal}}$   
529 within five pairs of data, which can be referred from the limited means and  
530 standard deviations of components of  $e_{\text{optimal}}$ .

531 We also included the hand-eye calibration result between the 3D scanner  
532 PhoXi M and UR10e using commercial software developed by Photoneo. The  
533 software uses a calibration sphere mounted on the robot tool flange and cal-  
534 culates the hand-eye matrix after collecting a minimum of four pairs of robot  
535 pose and point cloud. A few calibration trials with 4 and 16 pairs of data are  
536 conducted respectively for comparison, and their calibration errors are calcu-  
537 lated using the ICP metric defined in Eq. (11) with the same verification point  
538 cloud  ${}^{\text{Cam}}P_v$ . The results in Table 2 show that more sample points lead to fewer  
539 errors. However, it is also found that the calibration result is sensitive to the

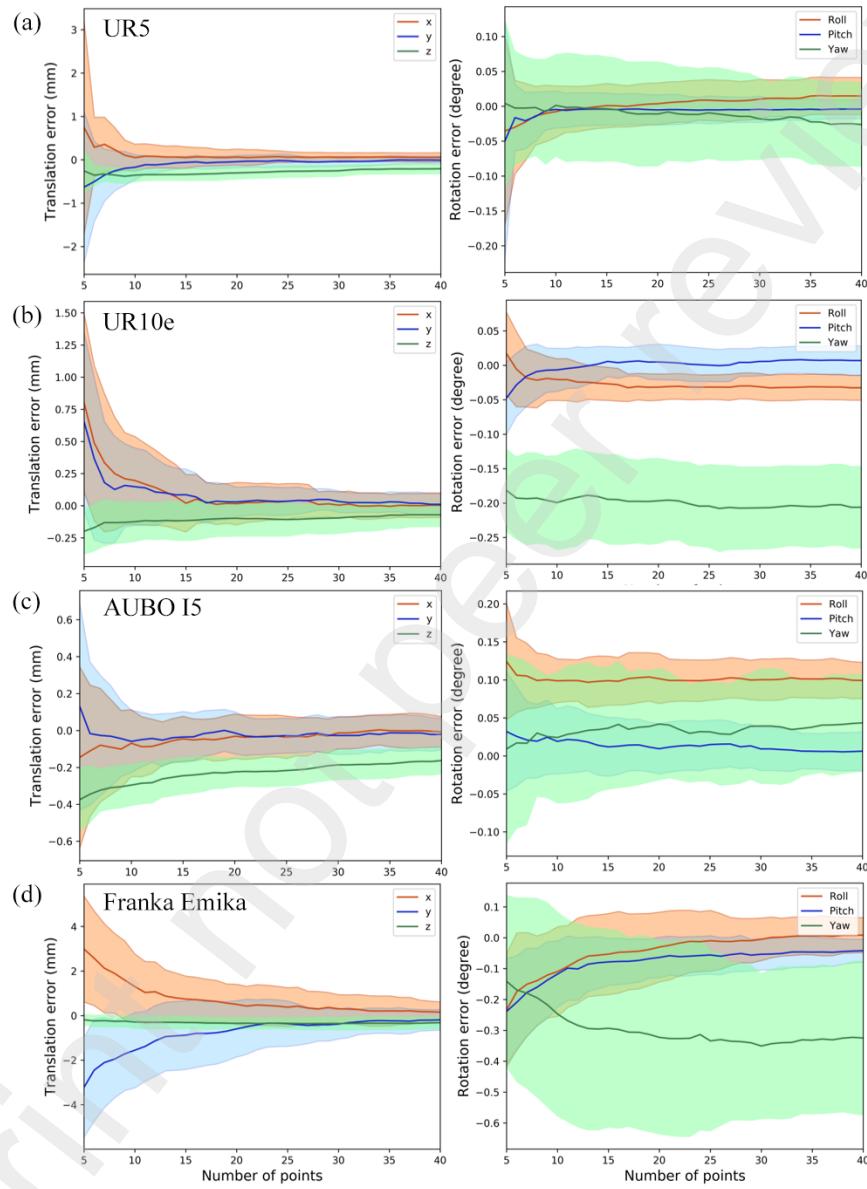


Figure 11: **The calibration error metric concerning the iterative steps.** The results are evaluated over 50 calculations, and means are plotted in lines and standard deviations as shaded areas.

Table 2: The calibration errors using Photoneo software with a spherical calibration object.

Number of points	<i>x</i> (mm)	<i>y</i> (mm)	<i>z</i> (mm)	roll (°)	pitch (°)	yaw (°)
4 pairs	-1.94	-0.90	-0.86	-0.07	0.04	-0.10
16 pairs	-0.98	-1.01	-0.83	-0.07	0.09	-0.11

540 diversity of the sampled points. In one of the few trials where the sample points  
 541 did not vary enough, the translational error is about 6 mm, which is unusable.

542 *3.3. Demonstrated Applications in Soft Robotic Tactile Welding*

543 This section presents a robotic welding system within a laboratory environ-  
 544 ment, as shown in Fig. 12.

545 The hardware platform comprises a PhoXi M 3D scanner (resolution at  
 546 0.3 mm), an industrial collaborative robot (UR10e, repeatability accuracy at  
 547 0.05 mm), and the proposed touch-based welding tool mounted to the robot’s  
 548 end-effector. All hardware is connected to a computer to process point clouds,  
 549 real-time tactile signals, and robot control.

550 In Fig. 12, the system leverages the proposed flange-based method to obtain  
 551 the hand-eye transformation. After capturing the workpiece’s point cloud from  
 552 the 3D scanner, we registered it with the workpiece’s CAD model to obtain the  
 553 position of the welding seam in the scanner’s coordinate. The position is then  
 554 converted to the robot base’s coordinate using the hand-eye transformation.  
 555 The SRM tip is designed to glide ahead of the welding torch, tracing the weld  
 556 seam extracted and planned by the 3D scanner. Utilizing the deformation servo  
 557 mechanism featured in Section 2.2.2, the system corrects the path followed by  
 558 the robot end effector, applying an offset  $\delta$  that compensates for the deformation  
 559 detected by the SRM at the relevant location. Consequently, the welding torch  
 560 is instructed to carry out the adjusted trajectory.

561 Fig. 13 shows the SRM tip, welding torch, and robot paths. During the  
 562 welding experiment we describe, corrective adjustments are limited to those  
 563 within a two-dimensional plane orthogonal to the z-axis. The corrective offset

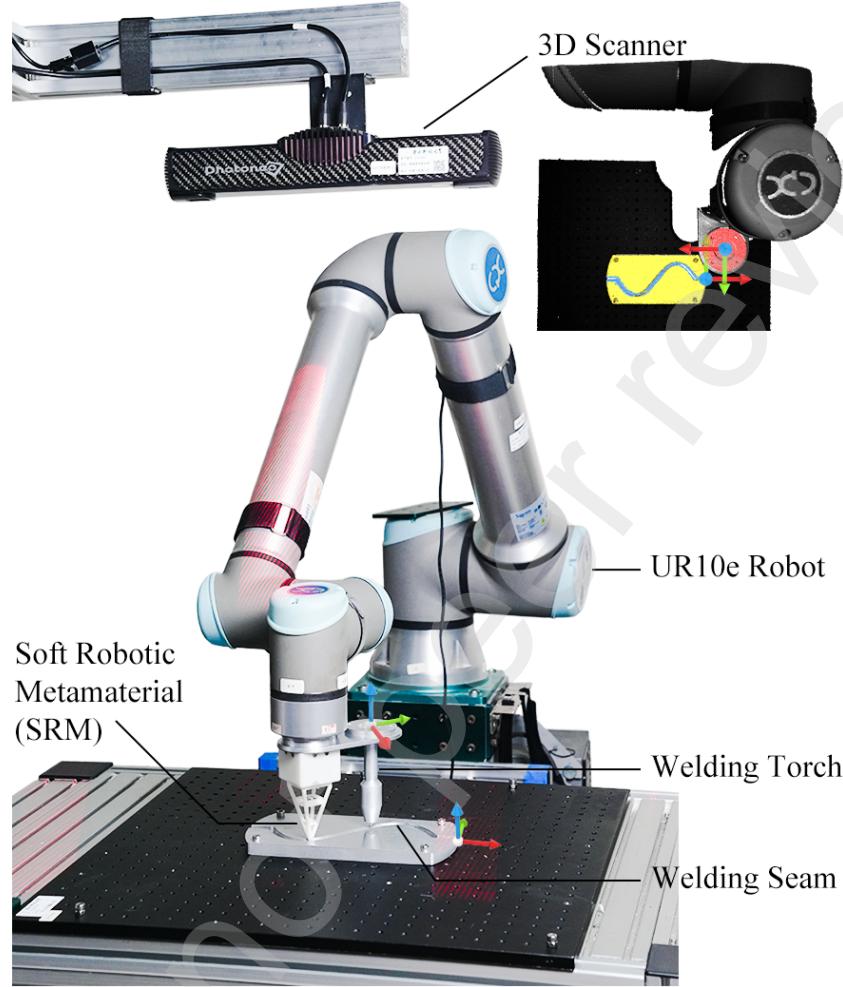


Figure 12: Experiment setup of flange-based calibration for tactile-enhanced robotic welding

564 in the direction perpendicular to the SRM tip  $\delta^{n_s}$  is maintained at a nearly  
 565 constant magnitude throughout the trajectory, in alignment with the initially  
 566 specified value  $\delta_d$ . An exception occurs at the end of the welding path, where  
 567 the SRM unexpectedly loses contact with the seam, prompting us to reset the  
 568 deformation servo reference value  $\delta_d$  to zero. This adjustment results in a  
 569 noticeable jitter at the end portion of the path as seen in the plot.

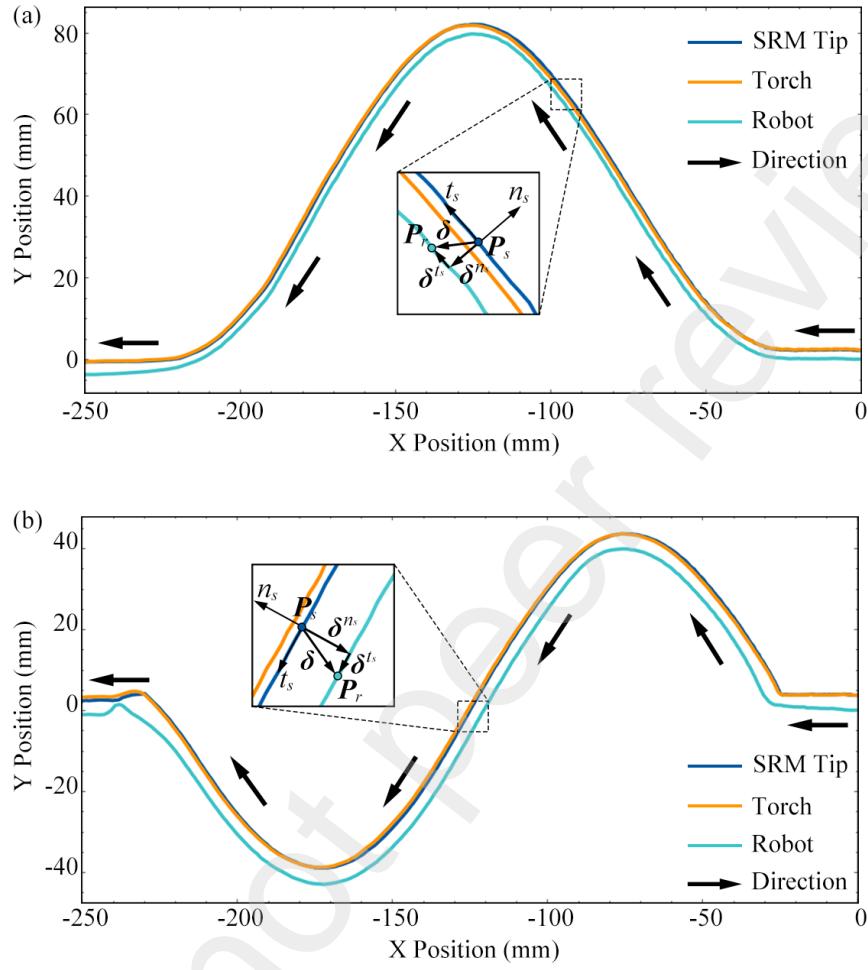


Figure 13: Executed paths of SRM tip, welding torch, and robot for (a) and (b) two different welding paths.

570 Assisted by the straightforward flange hand-eye calibration method and tac-  
 571 tile information provided by the low-cost SRM, the robotic welding system ac-  
 572 complishes autonomous weld seam tracking for two distinct workpieces. See  
 573 Movie S1 in the Supplementary Materials for a video demonstration.

574    **4. Discussion**

575    *4.1. Towards a Depth-based Hand-Eye Calibration*

576    Figs. 14(a) to (e) are graphical representations summarizing the above con-  
577    figurations of hand-eye calibration. The proposed method in this paper is to  
578    exploit high-quality, three-dimensional perception against standardized design  
579    and manufacturing of the mechanical components on the manipulator, which  
580    fall within a comparable level of accuracy and tolerance. Therefore, one can  
581    reconfigure these established methods for hand-eye calibration into those shown  
582    in Figs. 14(f) to (j) by removing the {Mark} frame in each case. Note that in  
583    the reconfigured Eye-in-Hand case in Fig. 14(f), the camera can alternatively  
584    inspect the geometric features, i.e., circles, on the base mounting flange com-  
585    monly found in most robots to proceed with the proposed calibration method.  
586    We suggest that one can freely choose any geometric features on the manu-  
587    pulator as long as a direct reading or deduction of the referencing feature's pose  
588    information can be obtained from the manipulator's controller or teach pendant.

589

590    *4.2. Industrial Robot Flanges for Hand-Eye Calibration*

591    In both simulation and experiment results, the cost metric minimized was  
592    chosen as the Euclidean norm of the translational error vector. The reason is  
593    that the impact of the translational error on applying the hand-eye transforma-  
594    tion matrix is uniform in a 3D scanner. However, the rotational error is found  
595    to be small enough, and its impact on the application is minimized around the  
596    optical axis of the 3D scanner, which linearly increases when moving away from  
597    the optical axis. In most scenarios, the workspace is usually located around  
598    the optical axis of the 3D scanner. Users can choose the appropriate elements  
599    to minimize according to their needs. For example, in the pick and place sce-  
600    nario, one might only minimize the translation errors in the  $x$  and  $y$ -axis if the  
601    end-effector has flexibility in the  $z$ -axis. If both translation and rotation errors  
602    are critical, the  $\|\cdot\|_{\text{cost}}$  can be defined as  $\|(\delta_x, \delta_y, \delta_z, \delta_{\text{roll}} \cdot r, \delta_{\text{pitch}} \cdot r, \delta_{\text{yaw}} \cdot r)\|_2$ ,  
603    where  $r$  is the work radius from the optical axis of the 3D scanner.

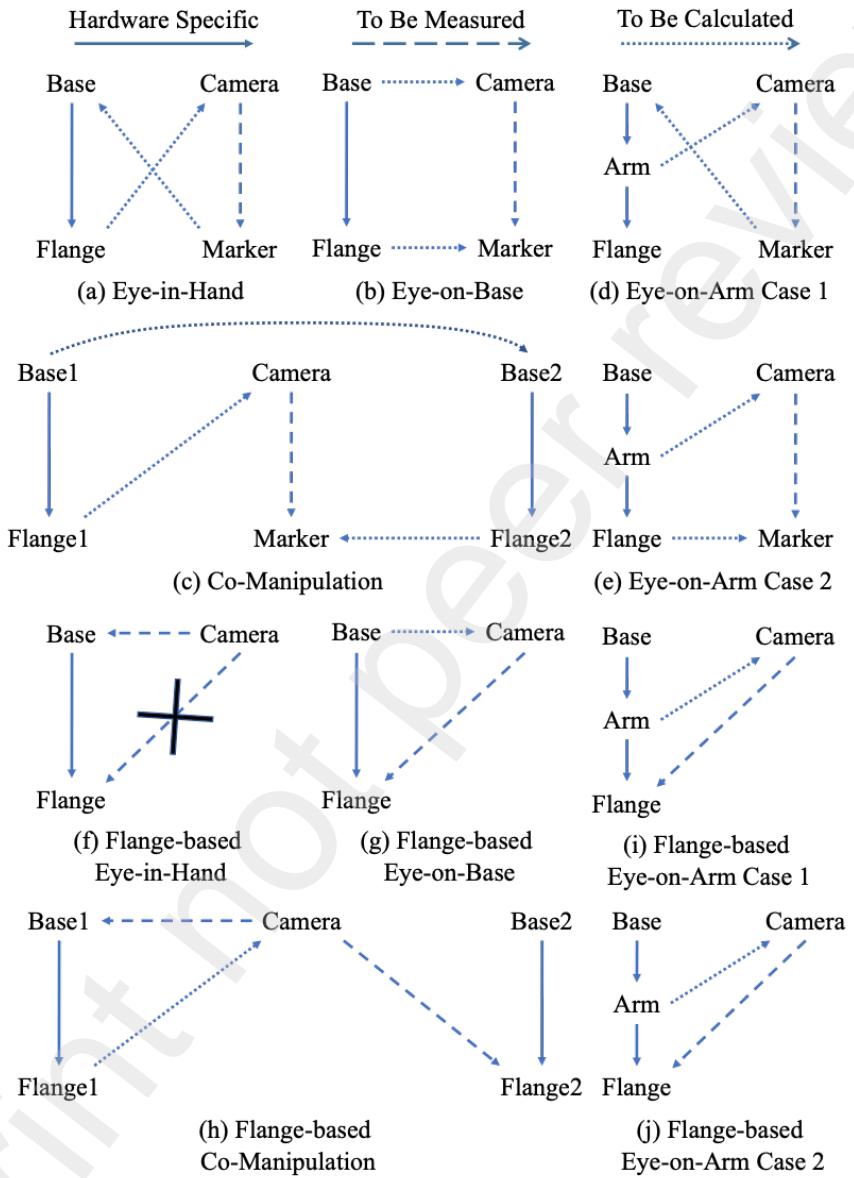


Figure 14: **A graphical summary of calibration methods:** (a) to (e) Review of the common configurations of hand-eye calibration. (f) The updated configuration when reconfigured using the proposed flange-based calibration method.

We also experimented with our proposed method using a commercial-grade 3D scanner to test its performance. The latest release of Microsoft’s Kinect sensor, the Azure Kinect DK, contains a 1-MP time-of-flight (ToF) depth sensor. Following a similar process, we experimented using a UR10e and an Azure Kinect DK. Due to the technical limitation of the ToF mechanism on metal surfaces, the quality of the point cloud of the tool flange is rather noisy. Instead of looking for the outer circle of the tool flange, we switch to the outer circle on UR10e’s wrist joint, which is 90 mm, as in Fig. 5(b). Fifty pairs of point clouds and the corresponding robot poses are recorded. Verification point clouds and robot pose are used to calculate the ICP error metric. The translational and rotational errors using all the data points are  $[-11.11, 3.90, 7.15]$  in mm and  $[-0.15, 0.92, 1.56]$  in degrees. The compensated hand-eye matrix  $\hat{H}_{\text{compensated}}$  can be obtained from Eq. (19). Nevertheless, with high-quality depth sensors entering consumer-grade applications, our method holds the potential for a broader range of applications, where a raw measurement of the depth data can be used directly for calibration and interaction.

#### 4.3. Multi-Modal Fusion for Soft Robotic Tactile Welding

In robot welding applications, hand-eye calibration is crucial to ensure precision in positioning the welding apparatus, which directly influences the welds’ quality and integrity. On the other hand, by employing tactile sensors, the welding process gains the ability to receive real-time feedback, allowing the system to adjust, adapt, and refine its operations based on the feedback received. Integrating global visual planning and local tactile feedback provides comprehensive supervision and control over the welding process, enhancing the system’s precision and adaptability. Our presented robotic welding system offers opportunities for improvement, such as refining spatial welding path (Lei et al., 2020) based on the designed tactile sensor, improving the accuracy of corrective actions through the use of machine learning algorithms (Mahadevan et al., 2021), and more.

633    **5. Conclusion, Limitations, and Future Work**

634    In this paper, we proposed an iterative hand-eye calibration method based  
635    on the 3D measurement of the robot tool flange. Using the Tool Center Point  
636    of the robot arm as the referencing point, the hand-eye calibration is simplified  
637    to fitting two 3D point sets with the least-squares analytical solution. Our  
638    method adopts the 3D error metric based on Iterative Closest Point registration  
639    to monitor and optimize the online calibration process. Once the desired hand-  
640    eye calibration accuracy is achieved, the calibration process is stopped, requiring  
641    only a minimum set of point clouds to be processed. The proposed method was  
642    applied to calculate the hand-eye transformation between an industrial-grade  
643    3D scanner and four robot arms, including UR5, UR10e, Aubo i5, and Franka.  
644    The results demonstrated that the iterative method converged quickly and is  
645    robust with tested robot arms.

646    The proposed online welding trajectory generation Algorithm 2 heavily relies  
647    on the geometric relationship between the weld seam and the welding tool.  
648    In practice, using this novel touch-based welding tool can lead to kinematic  
649    singularities when  $\vec{n}_d \perp \vec{n}_t$ , making it impossible to adjust the angular velocity  
650     $\omega$  of robot end effector to meet the velocity requirements of the seam in Eq. (16).  
651    Furthermore, there is a concern about the sudden change in the deformation  
652    servo reference value  $\delta_d$ , which leads to a significant deviation from the welding  
653    seam at the end of the path. These issues seem to hinder the effectiveness of  
654    the algorithm. It might be worth considering ways to address these problems  
655    to improve the quality of soft robotic tactile welding processes.

656    Our future work will investigate further details on applying the proposed  
657    approach to scenarios where the 3D scanner is mounted on the robot arm.  
658    This flexible setup allows robots to perform hand-eye calibration while moving,  
659    making it well-suited for challenging mobile welding tasks in unstructured envi-  
660    ronments. Additionally, we will further explore research into utilizing machine  
661    learning algorithms for real-time calculation of corrected welding trajectories  
662    based on the output data from SRM in more intricate scenarios.

663      **Supplementary Materials**

- 664      • **Movie S1. Evaluation of tactile-enhanced autonomous weld seam  
665      tracking.** In this movie, we showcase the proposed robotic welding system  
666      achieving autonomous weld seam tracking for two different workpieces in  
667      Section 3.3.

668      **References**

- 669      An, Y., Bell, T., Li, B., Xu, J., Zhang, S., 2016. Method for Large-Range  
670      Structured Light System Calibration. *Applied Optics* 55, 9563–9572.
- 671      Chen, Y., Medioni, G., 1992. Object Modelling by Registration of Multiple  
672      Range Images. *Image and Vision Computing* 10, 145–155.
- 673      Fan, J., Jing, F., Yang, L., Teng, L., Tan, M., 2019. A Precise Initial Weld  
674      Point Guiding Method of Micro-Gap Weld Based on Structured Light Vision  
675      Sensor. *IEEE Sensors Journal* 19, 322–331.
- 676      Geng, Y., Lai, M., Tian, X., Xu, X., Jiang, Y., Zhang, Y., 2023. A  
677      Novel Seam Extraction and Path Planning Method for Robotic Welding of  
678      Medium-Thickness Plate Structural Parts based on 3D Vision. *Robotics and  
679      Computer-Integrated Manufacturing* 79, 102433.
- 680      Geng, Y., Zhang, Y., Tian, X., Zhou, L., 2024. A Novel 3D Vision-based Robotic  
681      Welding Path Extraction Method for Complex Intersection Curves. *Robotics  
682      and Computer-Integrated Manufacturing* 87, 102702.
- 683      Hu, F., 2013. A Rapid Eye-to-Hand Coordination Method of Industrial Robots.  
684      *Journal of Information & Computational Science* 10, 1489–1496.
- 685      Huang, Y., Ke, J., Zhang, X., Ota, J., 2023. Dynamic Parameter Identification  
686      of Serial Robots Using a Hybrid Approach. *IEEE Transactions on Robotics*  
687      39, 1607–1621.

- 688 Hyatt, P., Kraus, D., Sherrod, V., Rupert, L., Day, N., Killpack, M.D., 2018.  
689 Configuration Estimation for Accurate Position Control of Large-Scale Soft  
690 Robots. *IEEE/ASME Transactions on Mechatronics* 24, 88–99.
- 691 ISO, 2004. ISO 9409-1:2004 - Manipulating industrial robots – Mechanical  
692 interfaces – Part 1: Plates. Technical Report. International Organization for  
693 Standardization.
- 694 Jarabo, A., Masia, B., Marco, J., Gutierrez, D., 2017. Recent Advances in  
695 Transient Imaging: A Computer Graphics and Vision Perspective. *Visual  
696 Informatics* 1, 65–79.
- 697 Jiang, Y., Li, T., Wang, L., Chen, F., 2017. Kinematic Accuracy Improvement  
698 of a Novel Smart Structure-Based Parallel Kinematic Machine. *IEEE/ASME  
699 Transactions on Mechatronics* 23, 469–481.
- 700 Kahn, S., Haumann, D., Willert, V., 2014. Hand-Eye Calibration with a Depth  
701 Camera: 2D or 3D?, in: International Conference on Computer Vision Theory  
702 and Applications (VISAPP), IEEE. pp. 481–489.
- 703 Kluz, R., Trzepieciński, T., 2014. The Repeatability Positioning Analysis of the  
704 Industrial Robot Arm. *Assembly Automation* 34, 285–295.
- 705 Kolb, A., Barth, E., Koch, R., Larsen, R., 2010. Time-of-Flight Cameras in  
706 Computer Graphics, in: *Computer Graphics Forum*, Wiley Online Library.  
707 pp. 141–159.
- 708 Lei, T., Huang, Y., Shao, W., Liu, W., Rong, Y., 2020. A Tactual Weld  
709 Seam Tracking Method in Super Narrow Gap of Thick Plates. *Robotics and  
710 Computer-Integrated Manufacturing* 62, 101864.
- 711 Li, W., Dong, M., Lu, N., Lou, X., Sun, P., 2018. Simultaneous Robot–World  
712 and Hand–Eye Calibration without a Calibration Object. *Sensors* 18, 3949.
- 713 Lü, X., Xie, C., He, X., Li, S., Xu, Y., He, S., Fang, J., Zhang, M., Yang, X.,  
714 2023. Automatic Recognition of Multiple Weld Types Based on Structured

- 715 Light Vision Sensor Using Deep Transfer Learning. IEEE Sensors Journal 23,  
716 7142–7152.
- 717 Ma, Q., Goh, Z., Ruan, S., Chirikjian, G.S., 2018. Probabilistic Approaches to  
718 the AXB = YCZ Calibration Problem in Multi-Robot Systems. Autonomous  
719 Robots 42, 1497–1520.
- 720 Mahadevan, R., Jagan, A., Pavithran, L., Shrivastava, A., Selvaraj, S.K., 2021.  
721 Intelligent Welding by using Machine Learning Techniques. Materials Today:  
722 Proceedings 46, 7402–7410.
- 723 Peng, J., Xu, W., Wang, F., Han, Y., Liang, B., 2021. A Hybrid Hand-Eye Cal-  
724ibration Method for Multilink Cable-Driven Hyper-Redundant Manipulators.  
725 IEEE Transactions on Instrumentation and Measurement 70, 1–13.
- 726 Raguram, R., Chum, O., Pollefeys, M., Matas, J., Frahm, J.M., 2012. USAC:  
727 A Universal Framework for Random Sample Consensus. IEEE Transactions  
728 on Pattern Analysis and Machine Intelligence 35, 2022–2038.
- 729 Rout, A., Deepak, B., Biswal, B., 2019. Advances in Weld Seam Tracking Tech-  
730 niques for Robotic Welding: A Review. Robotics and Computer-Integrated  
731 Manufacturing 56, 12–37.
- 732 Rusu, R.B., Cousins, S., 2011. 3D is Here: Point Cloud Library (PCL), in:  
733 IEEE International Conference on Robotics and Automation (ICRA), IEEE.  
734 pp. 1–4.
- 735 Sarbolandi, H., Lefloch, D., Kolb, A., 2015. Kinect Range Sensing: Structured-  
736 light versus Time-of-Flight Kinect. Computer Vision and Image Understand-  
737 ing 139, 1–20.
- 738 Shah, M., Eastman, R.D., Hong, T., 2012. An Overview of Robot-Sensor Cali-  
739 bration Methods for Evaluation of Perception Systems, in: Proceedings of the  
740 Workshop on Performance Metrics for Intelligent Systems, ACM. pp. 15–20.

- 741 Shi, J., Jimmerson, G., Pearson, T., Menassa, R., 2012. Levels of Human and  
742 Robot Collaboration for Automotive Manufacturing, in: Proceedings of the  
743 Workshop on Performance Metrics for Intelligent Systems, pp. 95–100.
- 744 Tannous, M., Miraglia, M., Inglese, F., Giorgini, L., Ricciardi, F., Pelliccia, R.,  
745 Milazzo, M., Stefanini, C., 2020. Haptic-based Touch Detection for Collab-  
746 orative Robots in Welding Applications. *Robotics and Computer-Integrated*  
747 *Manufacturing* 64, 101952.
- 748 Wong, X.I., Singla, P., Lee, T., Majji, M., 2018. Optimal Linear Attitude  
749 Estimator for Alignment of Point Clouds, in: IEEE/CVF Conference on  
750 Computer Vision and Pattern Recognition Workshops (CVPRW), IEEE. pp.  
751 1577–15778.
- 752 Wu, J., Sun, Y., Wang, M., Liu, M., 2020. Hand-Eye Calibration: 4-D Pro-  
753 crustes Analysis Approach. *IEEE Transactions on Instrumentation and Mea-*  
754 *surement* 69, 2966–2981.
- 755 Wu, L., Wang, J., Qi, L., Wu, K., Ren, H., Meng, M.Q.H., 2016. Simultaneous  
756 Hand-Eye, Tool-Flange, and Robot-Robot Calibration for Comanipulation  
757 by Solving the  $AXB = YCZ$  Problem. *IEEE Transactions on Robotics* 32,  
758 413–428.
- 759 Wu, T., Dong, Y., Liu, X., Han, X., Xiao, Y., Wei, J., Wan, F., Song, C., 2024.  
760 Vision-based Tactile Intelligence with Soft Robotic Metamaterial. *Materials*  
761 & Design 238, 112629.
- 762 Yang, L., Cao, Q., Lin, M., Zhang, H., Ma, Z., 2018. Robotic Hand-Eye Cal-  
763 ibration with Depth Camera: A Sphere Model Approach, in: International  
764 Conference on Control, Automation and Robotics (ICCAR), IEEE. pp. 104–  
765 110.
- 766 Zhang, L., Zhang, J.Z., Jiang, X., Liang, B., 2020. Error Correctable Hand-Eye  
767 Calibration for Stripe-Laser Vision-Guided Robotics. *IEEE Transactions on*  
768 *Instrumentation and Measurement* 69, 8314–8327.

- <sup>769</sup> Zhang, X., Song, Y., Yang, Y., Pan, H., 2017. Stereo Vision based Autonomous  
<sup>770</sup> Robot Calibration. *Robotics and Autonomous Systems* 93, 43–51.
- <sup>771</sup> Zhang, Y., Qiu, Z., Zhang, X., 2021. A Simultaneous Optimization Method of  
<sup>772</sup> Calibration and Measurement for a Typical Hand-Eye Positioning System.  
<sup>773</sup> *IEEE Transactions on Instrumentation and Measurement* 70, 1–11.