HECE-IC: An Integrated Calibration Method for Delta Robot-Based Kitchen Waste Sorting Systems

Hai Qin, Li Zhou, Songyun Deng, Xiangyu Zhang, Qiaokang Liang, *Senior Member, IEEE,* Dan Zhang, *Senior Member, IEEE,* Yaonan Wang

***Abstract*—This paper presents a multifunctional automated sorting system for kitchen waste based on a Delta robot. The sorting system is divided into three main modules: visual detection, information processing, and multifunctional robotic sorting. The visual detection module captures images of waste on a conveyor belt and transmits them in real-time to the information processing unit, where detection algorithms generate data on waste categories and grasp positions. The robotic arm, equipped with a force sensor, gripper, and suction cup, selects appropriate grasping or suction functions based on the waste type and shape to complete sorting. Additionally, this paper introduces a robust and efficient integrated calibration method for robot hand-eye-conveyor belt-encoder systems (HECE-IC), which enables simultaneous hand-eye and encoder calibration with only three simple steps. In simulations, the proposed method maintains a reconstruction error as low as 0.423 mm even under operation errors up to 0.6 mm. In practical experiments on the sorting platform, the average calibration error stabilized around 0.5 mm, achieving high calibration precision. The system achieved a sorting success rate of 90.2% and a sorting speed of 979 objects per hour. Our code is available at: https://github.com/TDA-2030/XRobot.**

***Note to Practitioners*—Due to inconsistent classification practices and human factors, kitchen waste often contains a significant amount of household waste. Currently, recyclable materials are primarily sorted from kitchen waste manually, with limited robotic systems that offer only basic sorting functions. This approach suffers from low sorting accuracy and efficiency. Most existing robotic grippers use suction or two-finger pneumatic claws, which offer limited functionality and struggle with reliable grasping, lacking adequate levels of intelligent handling. Furthermore, current hand-eye calibration methods are generally designed for single-mechanism robots and do not address the needs of integrated sorting systems for urban kitchen waste. Developing a multifunctional, reliable, and automated sorting system tailored to the specific challenges of kitchen waste handling is therefore of practical importance.**

***Index Terms*—Kitchen waste sorting; Delta robot; Integrated calibration; Reconstruction error.**

Manuscript received Month xx, 2xxx; revised Month xx, xxxx; accepted Month x, xxxx. This work was supported in part by the Project of the National Key Research and Development Program of China (2021YFC1910402), Hunan Provincial Natural Science Foundation (2026JJ90097), and Scientific Research Fund of Hunan Provincial Education Department (25A0672). *(Corresponding author: Qiaokang Liang).*

Hai Qin is with the School of Electronic Information, Hunan First Normal University, Changsha, 410205, China and National Engineering Research Center of Robot Vision Perception and Control, School of Artificial Intelligence and Robotics, Hunan University, Changsha 410082, China (email: qinhai@hnu.edu.cn).

Qiaokang Liang, Li Zhou, Xiangyu Zhang and Yaonan Wang are with the National Engineering Research Center for Robot Vision Perception and Control, School of Artificial Intelligence and Robotics, Hunan University, Changsha 410082, China (email: qiaokang@hnu.edu.cn; tda2030@hnu.edu.cn; yulle@hnu.edn.cn; yaonan@hnu.edu.cn).

Songyun Deng is with the School of Artificial Intelligence, Changsha University of Science and Technology, 410114, Changsha, China (email:songyun4747@csust.edu.cn).

Dan Zhang is with the Department of Mechanical Engineering The Hong Kong Polytechnic University Kowloon, Hong Kong (email:dan.zhang@polyu.edu.hk).

I. Introduction

URBAN kitchen waste is a significant source of solid waste and one of the main factors contributing to environmental degradation such as climate change, soil pollution, and groundwater contamination. Studies have found that under conventional management pathways, urban solid waste will emit 32 to 35 billion tons of carbon dioxide equivalent greenhouse gases from 2020 to 2050 [1]. It is estimated that by 2050, global waste generation will nearly double compared to 2016, and by 2100, it will double again [2]. The zero waste approach advocated by the United Nations means maximizing the reuse and recycling of resources. China's "Zero Waste City" Construction Pilot Work Plan issued in 2018 proposes to promote solid waste reduction and resource utilization in a green manner [3]. The publication of the "National Catalogue of Advanced Pollution Prevention and Control Technologies (Solid Waste and Soil Pollution Prevention and Control Field)" by China in 2023 indicates a shift from traditional landfilling and incineration to green and low-carbon waste classification, recycling, and reuse for solid waste management. Kitchen waste refers to easily perishable household waste containing organic matter, including household kitchen waste, catering waste, and other kitchen waste [4]. Among them, household kitchen waste contains a large amount of recyclable waste such as plastic bottles, aluminum cans, glass bottles, tetra pak, and hard paper. Therefore, scientifically and reasonably sorting garbage can separate impurities from kitchen waste, leaving biomass waste for resourceful and harmless treatment such as high-temperature aerobic composting and efficient anaerobic fermentation [5] [6]. Furthermore, by recycling valuable waste [7] [8], the purpose of reducing energy consumption and emissions can be achieved, promoting the development of a green circular economy.

Currently, garbage sorting mainly relies on manual labor. However, in the actual operational environment of urban kitchen waste sorting and processing, factors such as dirty environments, irritating gases, and repetitive labor significantly affect the subjective judgment of sorting workers, resulting in high misclassification rates and low sorting efficiency. Moreover, there is a shortage of workers willing to engage in this job. In addition, kitchen waste sorting presents several inherent

challenges, including irregular shapes, wet and deformable surfaces, severe occlusions and stacking among objects, and the coexistence of multiple material types requiring different grasping strategies. Furthermore, the continuous motion of conveyor belts imposes strict requirements on real-time detection accuracy and robot–vision coordination. These factors make reliable automated sorting of kitchen waste considerably more difficult than that of conventional solid waste. Robots have found many applications in solid waste identification and sorting tasks, including household waste, construction waste, and industrial waste. For instance, in research on household waste identification and sorting systems, Guo et al. [9] developed a garbage sorting robot system using a material identification method guided by vision and touch. Kong et al. [10] developed an intelligent surface cleaning robot system for collecting floating plastic garbage. Chen et al. [11] proposed a self-moving robot system based on artificial intelligence and the Internet of Things for garbage collection. In the research on construction waste sorting applications, Chen et al. [12] introduced a robot for real-time navigation in construction and demolition waste recycling using simultaneous localization and mapping (SLAM). In industrial waste processing, Koskinopoulou et al. [13] discussed an autonomous robot system for sorting and physically separating recyclables based on material type, implemented in an industrial waste processing plant. These methods have shown promising results in their respective sorting scenarios.

However, in the field of kitchen waste, which is an important source of solid waste, efficient robot sorting systems are still lacking according to our research. Nevertheless, such systems continue to encounter challenges in grasping efficiency and stability when handling kitchen waste sorting tasks. To address these challenges, the sorting system must be designed with regard to three key aspects: the visual servo system, grasp detection, and the end-effector [14]. Among these, grasp detection can be broadly classified into generative methods and discriminative methods [15]. The end-effector is typically implemented using either two-finger grippers or suction cups. Although some articles [16] [17] have made improvements to grippers to enhance grabbing efficiency, there has been a lack of structural innovation, leading to a limited variety of sorting methods for garbage. Existing grasp detection methods often struggle to integrate the detection of object-specific characteristics. The integration of intelligent classification and recognition into automated intelligent systems is the mainstream trend for achieving refined sorting of kitchen waste. At present, research in this area primarily focuses on intelligent classification of household waste [18] [19] [20]. Therefore, developing a multifunctional robotic hand system for kitchen waste sorting is highly necessary.

A typical garbage sorting system consists of hardware modules such as robots, cameras, conveyors, controllers, and corresponding control software modules. As a necessary step for robots to achieve precise operations under visual guidance, hand-eye calibration [21] [22] [23] [24] is a prerequisite for robots to grasp target objects in visual scenes. In conventional sorting systems, the robotic sorting area is confined to the machine's visual range. Notably, in kitchen waste sorting scenarios, driven by the challenging dirty environment and the requirement for high-speed operation, robotic visual servo systems are typically integrated with conveyor belts. The visual detection zone is positioned upstream of the sorting area. Items are captured by the visual servo system before reaching the sorting zone and are subsequently conveyed to the robotic sorting area, where sorting operations are executed.When a conveyor belt is introduced into the system, additional calibration of the conveyor belt is required [25]. In existing research, hand-eye calibration and conveyor belt calibration are typically treated as two independent processes, resulting in complex workflows, low deployment efficiency, and unstable accuracy. These limitations are particularly pronounced in intelligent sorting tasks for unstructured kitchen waste, where the system must handle a wide variety of object types, severe stacking occlusions, and large variations in object shapes. Such challenges place high demands on the stability and responsiveness of the entire system. In kitchen waste sorting, which demands high load capacity, rapid motion, and stable operation, parallel manipulators offer distinct advantages over their serial counterparts. Among common industrial parallel robots, the Delta parallel manipulator stands out for its relatively high load capacity, compact size, and fast movement, making it particularly well-suited for such sorting applications. However, there is a lack of efficient and integrated calibration solutions tailored to Delta robot-based architectures involving camera–conveyor–manipulator coordination.

To address these issues, this paper presents a Delta robot-based kitchen waste sorting system and proposes a unified high-precision integrated calibration method. Different from conventional hand–eye calibration approaches, the proposed integrated calibration method is developed from a system-level perspective, where the robot, camera, conveyor belt, and encoder are jointly modeled and calibrated within a unified framework. This formulation enables simultaneous estimation of the spatial relationships among all modules, thereby avoiding error accumulation caused by multi-stage calibration procedures. Moreover, the proposed method requires only a small number of calibration points and completes parameter estimation through three simple operations, which significantly improves calibration efficiency and practical applicability. The main contributions are as follows:

1) We propose a robust and streamlined integrated calibration method that simultaneously calibrates the robot, camera, conveyor belt, and encoder within a unified framework. The proposed method performs hand–eye calibration and encoder synchronization using only three key operations, avoiding error accumulation caused by multi-stage calibration. The overall process is fast and user-friendly, with the final spatial mapping error controlled within 0.5 mm, meeting the accuracy requirements of engineering applications.
2) A dedicated calibration board is designed for vision–mechanics integrated calibration tasks. The board consists of an 8×11 square checkerboard with 20 mm grid lengths and can be printed directly on A4 paper. Cross-shaped circular auxiliary markers are embedded

in the four corner grids to facilitate high-precision alignment of the robot end-effector during calibration, thereby enhancing calibration consistency and accuracy.

3) We present a high-efficiency Delta robot-based kitchen waste sorting system that integrates visual detection, information processing, and a multifunctional end-effector. Combined with the human–machine interaction control platform "XRobot," a closed-loop automation pipeline from image recognition to grasping control is realized. Experimental results on a physical testbed demonstrate a sorting success rate of 90.2%, showcasing the system's excellent performance in real-world deployment and its practical potential. The video footage of our experimental tests can be found at https://github.com/TDA-2030/XRobot?tab=readme-ov-filet#demonstration-video.

II. Related works

The significance of hand-eye calibration in robot sorting systems lies in ensuring the accuracy and reliability of the system [26] [27] [28] [29]. Hand-eye calibration is the process of aligning the coordinate systems of the robot's end-effector (hand) with the visual system (eye). The methods for hand-eye calibration generally fall into two categories $AX = XB$ [22] [30] or $AX = YB$ [31] [32]. The former involves solving for the homogeneous transformation matrix X between the hand coordinate system and the eye coordinate system, given the relative motions A of the hand in the robot coordinate system and B of the calibration object in the camera eye coordinate system. Nguyen et al. [33] provided a rigorous derivation of the covariance of the solution X when A and B are random perturbation matrices. The latter simultaneously calculates the hand-eye transformation relationship X and the pose transformation Y between the calibration object and the robot base coordinate system. Pan et al. [34] proposed a novel calibration method using LMI-SDP (Linear Matrix Inequality and Semidefinite Programming) optimization to solve the $AX = YB$ problem. Wu et al. [35] proposed solving the hand-eye calibration problem using only 3D position measurements, while Li et al. [36] introduced a probabilistic method to solve for X and Y without prior knowledge of the correspondences between the flows of A and B. Currently, there are calibration methods based on deep learning, where the principle is typically to identify the robot's joint key points in images [37] [38]. Bahadir et al. [39] achieved hand-eye calibration through a regression architecture of deep learning by finding the transformation between the robot end-effector and the external camera. Lu et al. [23] proposed an end-to-end pose estimation framework for online camera-to-robot calibration. Li et al. [24] presented an industrial robot hand-eye calibration method combining data augmentation strategies and Actor-Critic networks, representing the parameters in the hand-eye transformation matrix with fixed X-Y-Z angles and formulating parameter optimization strategies to reduce workload and improve accuracy. Odeyemi et al.[40] explored an automated method for fine grasp control of prosthetic fingers using computer vision and machine learning algorithms. However, most existing calibration methods focus primarily on conventional hand–eye calibration for single robotic systems. When a conveyor belt separates the visual workspace of the camera and the robot, these methods become impractical. Kim et al. [41] addressed this issue by using deep learning to recover blurred images caused by defocusing, thereby reducing errors in hand–eye calibration. Cheng et al. [25] proposed a robot hand–eye calibration method for conveyor-based systems; however, the problem was simplified to a planar case, which requires strict parallel relationships among system components and is difficult to satisfy in practical applications. Moreover, most existing studies calibrate only the robot–camera relationship, while the conveyor motion and encoder information are handled separately. When additional components such as conveyors and encoders are introduced, the calibration of the entire sorting system remains insufficiently studied [42]. Therefore, from a system-level perspective, there is still a lack of simple, fast, and efficient integrated calibration methods that simultaneously consider the robot, camera, conveyor belt, and encoder. Developing such an integrated calibration approach can reduce error accumulation caused by multi-stage calibration, improve calibration efficiency, and enhance the overall accuracy and reliability of robotic sorting systems.

III. Kitchen Waste Robot Sorting System

A. Overall Structure

The kitchen waste robot sorting platform designed in this paper mainly consists of a Delta parallel robot, an end-effector multifunctional mechanical hand with a force-controlled electric gripper and integrated suction cup, a camera with a resolution of 1600×1200 and a frame rate of 60 fps, and a visual detection unit equipped with a light source. It also includes a conveyor belt with a 1:10 speed ratio, adjustable speed ranging from 10 to 30 meters per minute, controllers, and a host. The Delta robot has three independent motors controlling the movement of the robot's end-effector along the X, Y, and Z axes. Although the parallel structure limits the flexibility of motion, its parallel-driven characteristics give it a high load capacity at the end-effector and incredible motion speed, making it suitable for high-speed sorting of garbage on kitchen waste processing lines. Current garbage sorting robots are typically equipped with only grippers or suction cups at the end-effector, with limited functionality. In this paper, we designed an end-effector multifunctional mechanical hand with a force-controlled electric gripper and integrated suction cup, capable of selecting gripping or suction sorting methods according to different types of garbage and controlling the gripping force. The overall structure of the system is illustrated in Fig. 1. The designed human-machine interaction robot control software XRobot is installed on the host. By deploying target detection algorithms such as YOLO, it can generate detection boxes and predict categories based on the detected garbage images, and send sorting instructions to the robot.

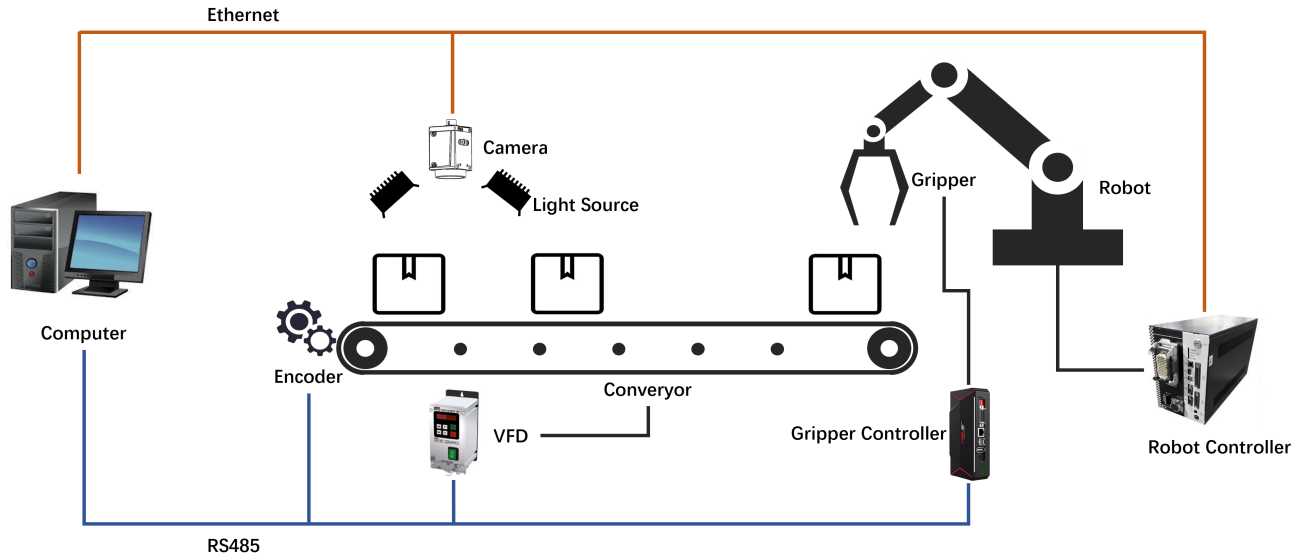

Fig. 1: Schematic diagram of kitchen waste robot sorting system structure.

B. System Sorting Operation Principle

The kitchen waste robot sorting system consists of a visual unit, processing unit, and execution unit, which are responsible for image acquisition, target identification and control, and target grabbing, respectively. The main components of the system are connected to the host in the processing unit via two buses: Ethernet and RS485.

The sorting control flowchart is shown in Fig. 2. The specific sorting control process is as follows: Firstly, the camera captures images of kitchen waste on the conveyor belt at a fixed frequency and sends them to the target detection model for object detection. Subsequently, all targets in each output image are input into a target filter for screening. Targets are filtered based on their position and size on the conveyor belt, and duplicates or excessively small targets are removed before being output to the target queue. Then, the detected targets undergo position delay compensation based on the conveyor belt speed and system processing delay time before being sent to the sorting controller. Through this controller, the targets are sorted, and appropriate tools (graspers or suction cups) are selected based on the target category, followed by the sequential generation of grabbing paths. Finally, the robot is driven to grab the targets. Simultaneously, targets in the target queue are sent to the visualization module for real-time display of objects on the conveyor belt and to the statistics module for analysis.

C. Human-Machine Interaction Module

We have designed the human-machine interaction module software XRobot for the kitchen waste robot sorting system. XRobot is a robot sorting control software developed based on PyQt5, aimed at identifying and sorting targets such as plastic bottles, glass bottles, cans, Tetra Paks, and plastic bags from kitchen waste. The control software consists mainly of four parts: robot equipment control, intelligent image recognition, data statistics, and system calibration.

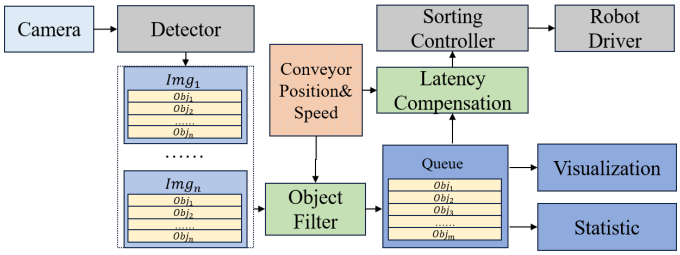

Fig. 2: Flowchart of sorting control process in robot system.

The robot-related equipment includes a high-speed four-axis parallel robot, a conveyor belt with an encoder, and an industrial camera, all connected to the control software via Ethernet. The intelligent image recognition utilizes advanced deep learning algorithms to preprocess RGB images from the robot's front camera and predict the location and type of garbage in the image. The data statistics section counts and stores the number of objects sorted by the system over a period of time. The system calibration part measures and calculates the parameters of the components in the system through a series of calibration steps and stores them for future use. The software's functional modules and main interface are shown in Fig. 3.

IV. Integrated Calibration Method

A. Calibration Task Abstraction and Description

components of the system: the robot, conveyor belt, and camera. Fig. 4 describes the five coordinate systems used for calibration tasks in this paper. Among them, F_C represents the camera coordinate system established within the camera calibration chessboard, with its XOY plane coinciding with the conveyor belt plane; F_B represents the conveyor belt coordinate system, which is derived from the camera coordinate system by rotating about the Z-axis by angle θ to align with the X-axis in the direction of conveyor belt motion, and it does

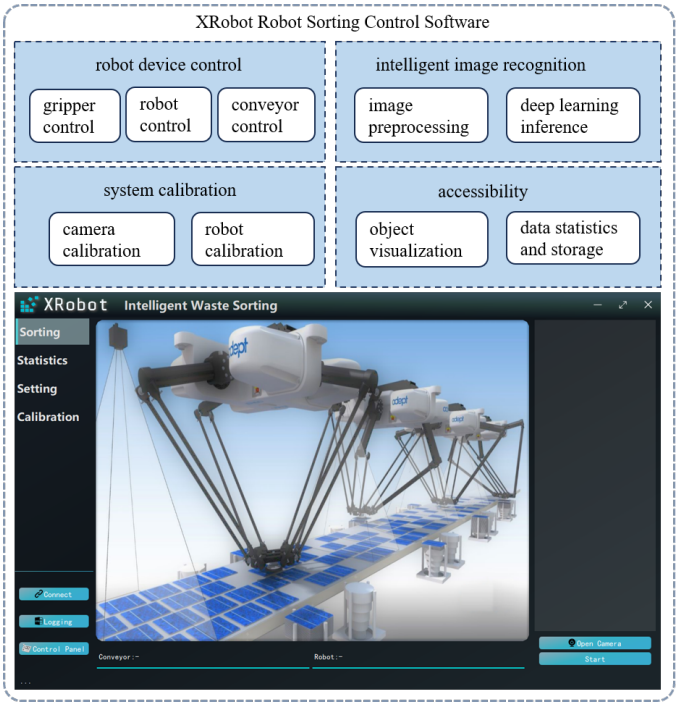

Fig. 3: Functional modules and main interface of the human-machine interaction control software XRobot.

not move with the conveyor belt, i.e., it is the world coordinate system; F_U represents the camera coordinate system translated along the conveyor belt to the upstream line coordinate system, where only objects past the upstream line are picked up by the robot; F_D represents the downstream line coordinate system, ensuring operational safety by preventing objects beyond the downstream line from being picked up by the robot; F_R represents the robot base coordinate system, which is a right-handed coordinate system, while the other coordinate systems, for the sake of correspondence with image coordinates, are all left-handed coordinate systems.

During operation of the sorting system, objects are conveyed from the right end to the left end of the conveyor belt at a constant speed. In this process, the objects on the conveyor belt pass through the camera capture area A_1 and the robot working area A_2 sequentially. These two areas are connected by the conveyor belt and do not overlap. Additionally, an encoder is installed on the conveyor belt to measure the distance traveled by the conveyor belt. The camera captures images of the objects on the conveyor belt at a fixed frame rate and processes them to obtain the pixel coordinates representing the object positions P^*. These pixel coordinates are then transformed into the robot coordinate system, and finally, the robot grasps or adheres to the targets accordingly.

$$\mathbf{t} = H(P^*) \tag{1}$$

$$\mathbf{R} = Rot(\begin{bmatrix}0 & 0 & 1\end{bmatrix}^T, \alpha) \tag{2}$$

$$\mathbf{P_R} = T_B^R T_C^B(\Delta E) \begin{bmatrix}\mathbf{R} & \mathbf{t} \\ \mathbf{0} & 1\end{bmatrix} \tag{3}$$

where $H(P^*)$ represents the transformation function from the image coordinate system to the camera coordinate system. α is the deviation angle of the object in the camera coordinate system. $Rot(\mathbf{k}, \alpha)$ represents a rotation transformation of angle α around axis $\mathbf{k}$. $\mathbf{P_R}$ represents the coordinates of objects on the conveyor belt in the robot coordinate system. T_B^R represents the homogeneous transformation matrix from the conveyor belt coordinate system to the robot coordinate system. $T_C^B(\Delta E)$ represents the homogeneous transformation matrix from the camera coordinate system to the conveyor belt coordinate system, where parameter ΔE is the displacement measured by the encoder. Therefore, the goal of calibration is to obtain T_B^R, $T_C^B(\Delta E)$ and $H(P^*)$.

B. Calibration Process

The integrated calibration proposed in this paper involves only three simple steps conducted on the sorting platform to calibrate the system's camera, robot, conveyor belt, and encoder. The calibration procedure is illustrated in Fig. 5, and the specific process is as follows:

* **Calibrating the Camera:** Place the designed calibration chessboard flat on the conveyor belt within the camera's field of view (FOV). Subsequently, the chessboard should be repositioned to another location on the conveyor belt and another photo should be taken. Repeat this process until at least 10 different photos are captured, as shown in Fig. 6(b). To ensure proper calibration of camera distortions, the chessboard should cover the FOV as much as possible. After the final capture, keep the chessboard stationary relative to the conveyor belt until the end of calibration. Simultaneously, record the current encoder count as E_0.
* **Calibrating the Robot:** Initiate the conveyor belt to transport the chessboard from the last capture of the previous step to the robot operation area. Stop the conveyor belt when the identified point "1" on the chessboard grid aligns with the manually marked upstream line. Then, using the teach pendant, maneuver the robot's end effector to touch the four identified points on the chessboard grid in turn, as shown in Fig. 6(a). Obtain the coordinates of these four points in the robot coordinate system, as shown in Fig. 6(c). Also, record the current encoder count as E_1.
* **Calibrating the Conveyor Belt:** Restart the conveyor belt to allow the chessboard to continue downstream until point "1" on the chessboard grid aligns with the manually marked downstream line. Stop the conveyor belt once the alignment is achieved. Then, using the teach pendant, maneuver the robot's end effector to touch point "1" on the identified chessboard grid, as depicted in Fig. 6(d). Obtain its coordinates in the robot coordinate system, denoted as P_d, and simultaneously record the current encoder count as E_2.

C. Encoder Calibration

The encoder installed on the conveyor belt is used to count the distance traveled by the conveyor belt. The count

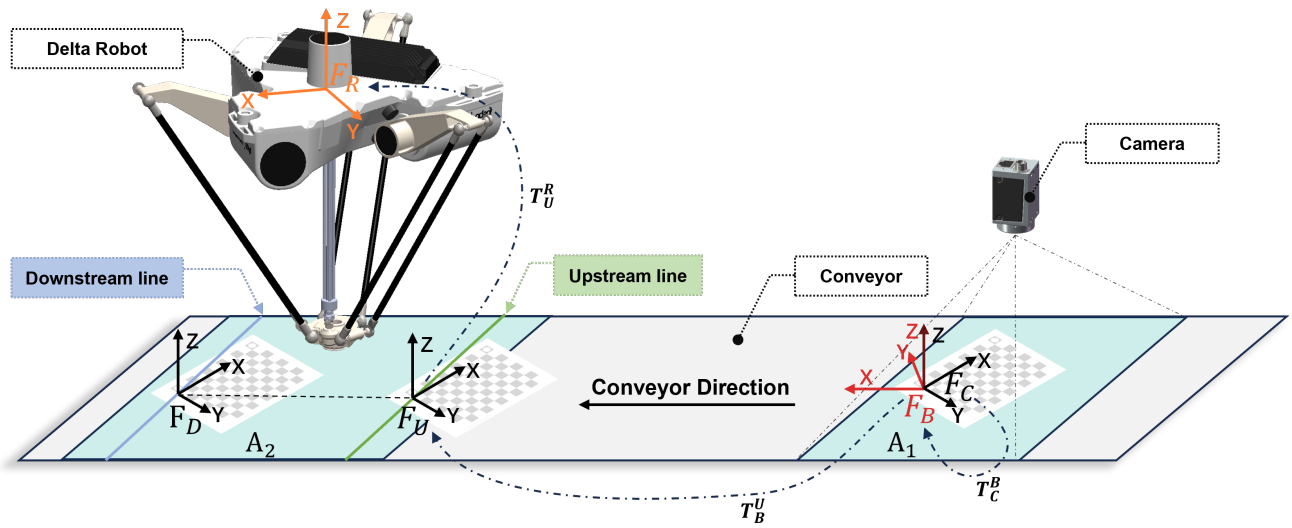

Fig. 4: The five coordinate systems used for calibration tasks in this paper.

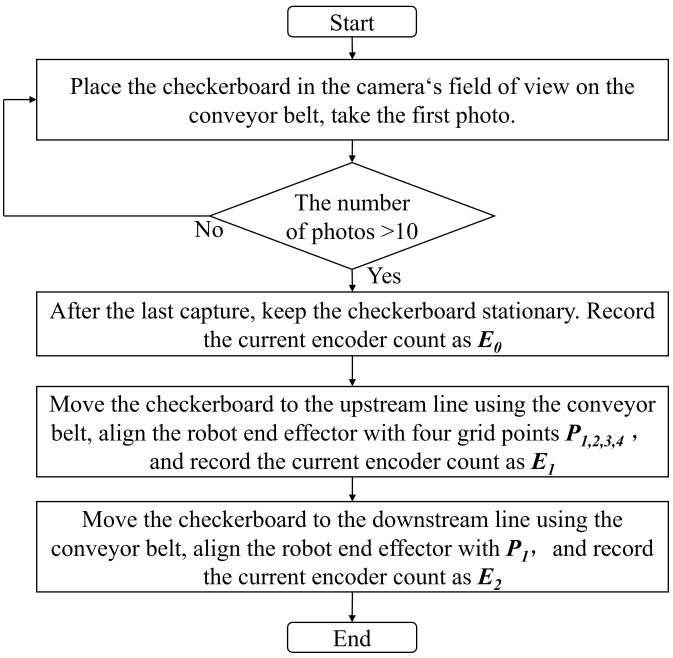

Fig. 5: Flowchart of integrated calibration.

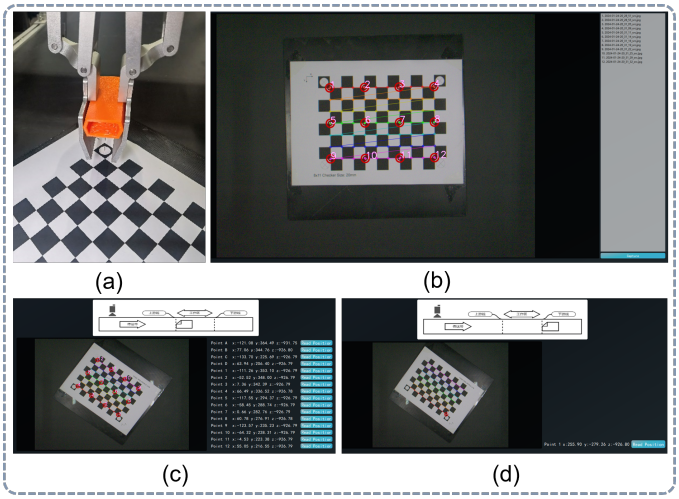

Fig. 6: Integrated Calibration Procedure. (a) Alignment of the robot end effector with a chessboard grid corner point; (b) Identification and numbering of corner points on the chessboard grid; (c) Robot end effector sequentially touching the four identified grid points on the chessboard; (d) Robot end effector touching point "1" identified on the chessboard grid.

value E and the distance L moved have the following linear relationship:

$$L = K_e \times E \tag{4}$$

where K_e is the encoder coefficient, and in the last two calibration steps, we respectively obtained the corresponding encoder count values E_1 and E_2 of the robot coordinates at the upstream and downstream positions on the conveyor belt. Therefore, we have:

$$K_e = \frac{\|P_{U_1} - P_d\|_2}{E_2 - E_1} \tag{5}$$

Furthermore, we can easily obtain the vertical distance between the upstream line and the origin of the conveyor belt coordinate system:

$$L_1 = K_e(E_1 - E_0) \tag{6}$$

As well as the distance between the upstream line and the downstream line:

$$L_2 = K_e(E_2 - E_1) \tag{7}$$

D. Calibration of Robot to Conveyor Belt

In the kitchen waste sorting system, the robot is fixedly installed above the conveyor belt. Therefore, there is a fixed

relationship between the robot coordinate system F_R and the conveyor belt coordinate system F_B, which can be expressed as follows:

$$P_R = T_B^R P_B \tag{8}$$

here, T_B^R is the homogeneous transformation matrix between the two coordinate systems. P_R is a point in the robot coordinate system. P_B is a point in the conveyor belt coordinate system. Since the conveyor belt coordinate system does not have an actual reference object, it is difficult to directly measure T_B^R. It is known that F_U is derived from F_B by rotating around the Z-axis and then translating along the X-axis by L_1, while F_U is a coordinate system defined in the chessboard grid. We can easily determine the coordinates of each grid point in the chessboard grid. Therefore, we can first measure T_U^R and then obtain T_B^R through transformation.

In the second step of the calibration operation, we obtained four pairs of points $P_{R_i}, P_{U_i}; i = 1, 2, 3, 4$. The calculation of T_U^R can be expressed as follows:

$$T_U^R = \begin{bmatrix} \mathbf{R_u^r} & \mathbf{t_u^r} \\ \mathbf{0} & 1 \end{bmatrix} \tag{9}$$

in this equation, T_U^R is the homogeneous transformation matrix from the upstream coordinate system F_U to the robot coordinate system F_R, $\mathbf{R_u^r}$ is a 3×3 rotation matrix, $\mathbf{t_u^r}$ is a 3-dimensional translation vector, and the objective is to minimize the squared error $\hat{E}$.

$$\hat{E} = \frac{1}{4}\sum_{i=1}^{4} \|\mathbf{R_u^r} P_{U_i} - P_{R_i} + \mathbf{t_u^r}\|_2^2 \tag{10}$$

According to the classical Kabsch algorithm [43], we can calculate the optimal rotation matrix R_u^r between these two sets of coordinate points.

$$\overline{d} = \frac{1}{4}\sum_{i=1}^{4} P_{R_i} \tag{11}$$

$$\overline{m} = \frac{1}{4}\sum_{i=1}^{4} P_{U_i} \tag{12}$$

Minimizing $\hat{E}$ is equivalent to maximizing $trace(R_u^r M)$, where the correlation matrix M is defined by:

$$M = \sum_{i=1}^{4} m_{c_i}\, d_{c_i}^T \tag{13}$$

where $d_{c_i} = P_{R_i} - \overline{d}$, $m_{c_i} = P_{U_i} - \overline{m}$. Given that the Singular Value Decomposition (SVD) of matrix M is represented by $M = U\Sigma V^T$, where U and V are orthogonal matrices, and Σ is a diagonal matrix. The optimal rotation matrix R_u^r that maximizes the desired trace is denoted by:

$$R_u^r = VU^T \tag{14}$$

after obtaining the rotation matrix, the translation vector can be easily derived as:

$$t_u^r = \overline{d} - R_u^r \overline{m} \tag{15}$$

In the second step of the calibration process, we obtained the coordinate points P_d, and through T_U^R transformation, we obtain their coordinates in the upstream coordinate system F_U, denoted as $\hat{P}_d$:

$$\hat{P}_d = \begin{bmatrix} \hat{x} \\ \hat{y} \\ \hat{z} \\ 1 \end{bmatrix} = T_U^{R^{-1}} \begin{bmatrix} \mathbf{P_d} \\ 1 \end{bmatrix} \tag{16}$$

In the XOY plane of the upstream coordinate system F_U, connecting the origin of F_D with the origin of F_U, the rotation angle θ between coordinate systems F_U and F_B satisfies the following relationship:

$$\tan\theta = \frac{\hat{y}}{\hat{x}} \tag{17}$$

Thus, the relationship between the upstream coordinate system and the conveyor belt coordinate system is determined.

$$T_B^U = \begin{bmatrix} \mathbf{R_b^u} & \mathbf{t_b^u} \\ \mathbf{0} & 1 \end{bmatrix} = \begin{bmatrix} \cos\theta & -\sin\theta & 0 & -L_1 \\ \sin\theta & \cos\theta & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \tag{18}$$

Finally, the homogeneous transformation matrix between the two coordinate systems is obtained.

$$T_B^R = T_U^R T_B^U \tag{19}$$

E. Camera-to-Conveyor Calibration

In the calibration process from the camera to the conveyor belt, the image captured by the camera is represented by coordinates $P^* = [u, v]^T$, indicating a unique pixel. A point in the camera coordinate system is denoted as $P_C = [x, y, z]^T$ (where z is a constant), expressed in augmented vector form as $\hat{P}^* = [u, v, 1]^T$ and $\hat{P}_C = [x, y, z, 1]^T$. Camera calibration refers to the commonly used Zhang's calibration method [44]. To reduce errors and facilitate integrated calibration more conveniently, we designed the calibration board shown in Fig. 7. This design consists of an 8x11 grid of 20mm square chessboard patterns laid out on an A4 sheet. We intentionally placed a cross-circle marker in each of the four corners of the grid to provide precise alignment points required for the robot end effector during the calibration process. This simple and practical design aims to meet the requirements of our experiments, ensuring the reliability and accuracy of the calibration.

The calibration of the camera to the conveyor belt is achieved by capturing images of the chessboard pattern placed on the conveyor belt, as depicted in Equation 20:

$$\widetilde{P_C} = H(P^*) = R^{-1}A^{-1}s\widetilde{P^*} - R^{-1}t \tag{20}$$

where $[R \quad t]$ represents the extrinsic parameters of the camera, comprising the rotation R and translation t of the

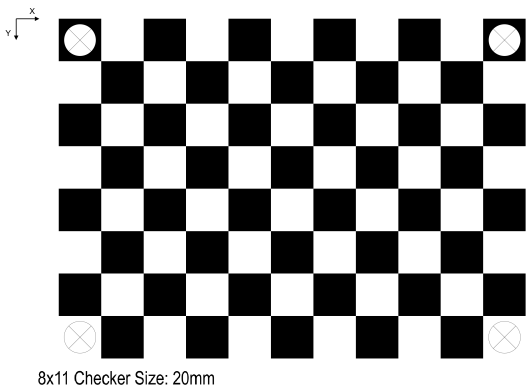

Fig. 7: High-precision calibration board for integrated calibration.

camera relative to the conveyor belt coordinate system; A is a camera intrinsic matrix 3×3 used to rectify the internal offsets and distortions of the camera; s is an arbitrary scaling factor. Since the camera is fixed at a specific position upstream of the conveyor belt, there is only a translational relationship between F_C and F_U, thus:

$$\mathbf{R_c^b} = \mathbf{R_b^u}^T \tag{21}$$

Taking into account that kitchen waste on the conveyor belt is continuously moving, meaning the objects are constantly in motion, the position of the objects captured by the camera is related to the movement of the conveyor belt. The direction of movement of the conveyor belt is in the positive direction of the X axis. Therefore, the relationship between the camera and the conveyor belt can be expressed as:

$$T_C^B(\Delta E) = \begin{bmatrix} \mathbf{R_c^b} & \mathbf{t_c^b} \\ \mathbf{0} & 1 \end{bmatrix} = \begin{bmatrix} \cos\theta & \sin\theta & 0 & \Delta E \\ -\sin\theta & \cos\theta & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \tag{22}$$

here, $\Delta E = K_e(E_c - E_s)$ represents the displacement of the conveyor belt, where E_c and E_s are the current and recorded encoder count values respectively, and K_e represents the encoder coefficient.

V. Experiment and Result Analysis

A. Simulation Experiments

The ground truth of the calibration results cannot be obtained in real-world scenarios. To validate and evaluate the accuracy and efficiency of our method, we conducted two numerical simulation experiments:

- **Experiment 1:** Assessing the stability of the proposed method under Gaussian noise conditions applied to the pose data retrieved from the robot.
- **Experiment 2:** Investigating how the quantity and distribution of grid points used in the second step of calibration affect the accuracy of the calibration results.

The parameters of the sorting system were set to approximate our actual system, as shown in Table I. All described

TABLE I: Settings of the Parameters for Sorting System in Simulation Experiments

T_B^R	$\begin{bmatrix} Rot([0 & 0 & 1]^T, \frac{\pi}{6}) & [-500, 100, -900]^T \\ & \mathbf{0} & & 1 \end{bmatrix}$
θ	$\frac{\pi}{3}$
$L1$	600
$L2$	500
K_e	1.0

TABLE II: NOISE LEVEL Parameter Description

Noise level	very low	low	middle	high	very high
σ	0.1	0.15	0.2	0.25	0.3
l	0.2	0.3	0.4	0.5	0.6

experiments were conducted using Python 3.8 on a computer with an Intel i5-9400f CPU, 16GB DDR4 RAM, and GTX1060 6GB GPU.

1) Data Generation: Following the method described in reference [44], the calibration results of the camera are considered mature and stable. Therefore, in the simulation experiments, we only simulated the calibration process from the conveyor belt to the robot. To simulate the actual calibration process, it is necessary to introduce error noise into the ideal sample data. Inspired by reference [45], the normal distribution is suitable for describing positioning errors caused by manual operations. Thus, we assume that the distribution of positioning errors due to manual operation is a truncated normal distribution. Corresponding to the 12 simulated data coordinate points of the robot end-effector in Fig. 6(b), each coordinate true value can be obtained by adding noise using Equation 23.

$$\mathbf{P_{si}} = \mathbf{P_{gti}} + \begin{bmatrix} c\left(\mathcal{N}(\mu, \sigma), -l, l\right) \\ c\left(\mathcal{N}(\mu, \sigma), -l, l\right) \\ c\left(\mathcal{N}(\mu, \sigma), -l, l\right) \end{bmatrix} \tag{23}$$

where $N(\mu, \sigma)$ represents the mean μ and standard deviation σ of the normal distribution, $c(n, low, high)$ represents truncating n in the interval [low,high]. In the experiments, μ is consistently set to 0, and the noise levels are categorized into five levels: "very low," "low," "middle," "high," and "very high," based on the parameters σ and l. The parameter values for each level are described in Table II. Data at each noise level are generated in 500 sets for experimentation, and the average is taken as the error result.

In Experiment 2, to investigate how the number and spatial distribution of calibration points used during the calibration process affect the accuracy of the results, the 12 corner points labeled in Fig. 6(b) were divided into 9 different subsets. The specific configurations of these subsets are illustrated in Fig. 8. For example, in subset S1 (top-left corner), each of the 12 grid cells corresponds to one corner point. The black-labeled points indicate those used in the calibration calculation, while

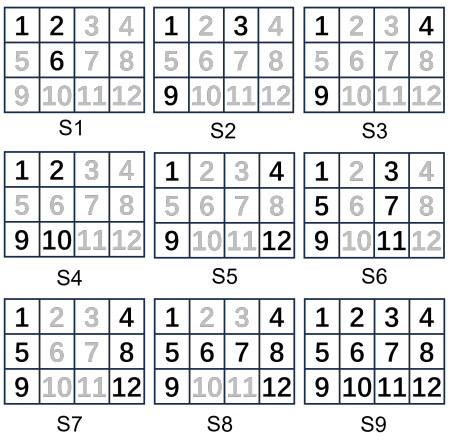

Fig. 8: Nine different sets of points representing various numbers and distributions of calibration points. Black numbers indicate points used for calibration, while gray indicates points not used for calibration.

the gray-labeled points indicate those excluded. Thus, subset S1 uses only points 1, 2, and 6 for calibration, and the setup for other subsets follows a similar pattern. To reduce the influence of external variables, the noise level for data generation in Experiment 2 was fixed at the "medium" level throughout. Additionally, two new types of noise distributions were introduced: triangular distribution and uniform distribution. Their standard deviation and mean were set to match those of the normal distribution to ensure comparability. Each data group was repeated 500 times in the experiment.

2) Error Metrics: In the case of simulation experiments, we can obtain the accurate pose of the robot relative to the conveyor belt. We can utilize this information to obtain the absolute rotation error and absolute translation error. The absolute translation error is calculated by equation 24, and the absolute rotation error is calculated by equation 25. Here, $\widehat{t_B^R}$ and $\widehat{R_B^R}$ represent the true values of the translation vector and rotation matrix; t_B^R and R_B^R represent the calculated values of the translation vector and rotation matrix; $\|\cdot\|_2$ is the L2 norm, and $\theta(\cdot)$ is a function that converts the rotation matrix to Euler angles.

$$E_t = \left\| \widehat{t_B^R} - t_B^R \right\|_2 \tag{24}$$

$$E_r = \left\| \theta(\widehat{R_B^R}^{-1} R_B^R) \right\|_2 \tag{25}$$

Furthermore, to make the results of the simulation experiments comparable to the actual experimental results, we also use the reconstruction error [46] as a metric to evaluate the calibration quality in the absence of ground truth information. The reconstruction error reflects the absolute error and provides an intuitive understanding in practical tasks. Equation 26 provides an estimation of the robot end-effector position $\widetilde{P}_i$ from the transformed coordinates ${P^C}_i$ obtained through the calibration process. The reconstruction error is defined as the translation error with respect to the robot end-effector position

TABLE III: Average translation error, rotation error, reconstruction error, and encoder coefficient error at different noise levels

Noise level	Mean E_t(mm)	Mean E_r(degree)	Mean $\widetilde{E_t}$(mm)	Mean E_{K_e}
Very low	0.374	0.036	0.141	2.42E-04
Low	0.596	0.059	0.212	3.60E-04
Middle	0.759	0.076	0.282	4.77E-04
High	0.969	0.092	0.353	6.22E-04
Very high	1.147	0.114	0.423	7.14E-04

${P^R}_i$ used in the calibration process, calculated according to equation 27. In the simulation scenario, ${P^R}_i = P_{si}$ is calculated, while in actual experiments, ${P^R}_i$ is obtained from real data collection.

$$\widetilde{P}_i = T_B^R T_C^B(L_1) {P^C}_i \tag{26}$$

$$\widetilde{E_{ti}} = \left\| \widetilde{P}_i - {P^R}_i \right\|_2 \tag{27}$$

3) Experimental Results: The results of Experiment 1 are shown in Table III. The results indicate that the errors are generally linearly increasing with the errors introduced by manual operation. It is worth noting that the rate of increase in reconstruction error is significantly lower than that of translation error. According to equations 15 to 19, this is because the reconstruction error includes both $L1$ error and accumulated transmitted rotation error, making it more sensitive to noise. In other words, it is influenced to a greater extent by the parameters of the system, achieving smaller values when $L2$ is longer and $L1$ is shorter, and vice versa. With errors in operation reaching a maximum of 0.6 mm under parameters close to our actual system, the reconstruction error remains at a low level of 0.423 mm. In our implemented system, the errors introduced by the robot and gripper are ±0.1mm and ±0.2mm, respectively, resulting in a total error range of ±0.3mm for the robot sorting platform.

The results of Experiment 2 are as expected, as shown in Fig. 9, where errors decrease as the number of target points used for calibration increases. Comparing S4 with S5, it is evident that using points with wider spacing can significantly reduce errors compared to using adjacent points. When comparing S7, S8 and S9 with S5, an increase in workload only results in a slight decrease in errors. Therefore, using the S5 set of calibration points appears to be a suitable choice, achieving a good balance between manual operation workload and calibration accuracy. The design of the calibration board in Fig. 7 is also based on this consideration.

B. Real-world Experiments

To effectively evaluate the proposed method in practical tasks, we designed a robot sorting platform as shown in Fig. 1 and an integrated system control program incorporating the described method. In the system, the robot is an OMRON HORNET 565 (4 degrees of freedom, repeatability accuracy of 0.1mm), and the camera is a Basler acA1600 with a resolution of 1600x1200, installed at a height of 0.8m above the conveyor

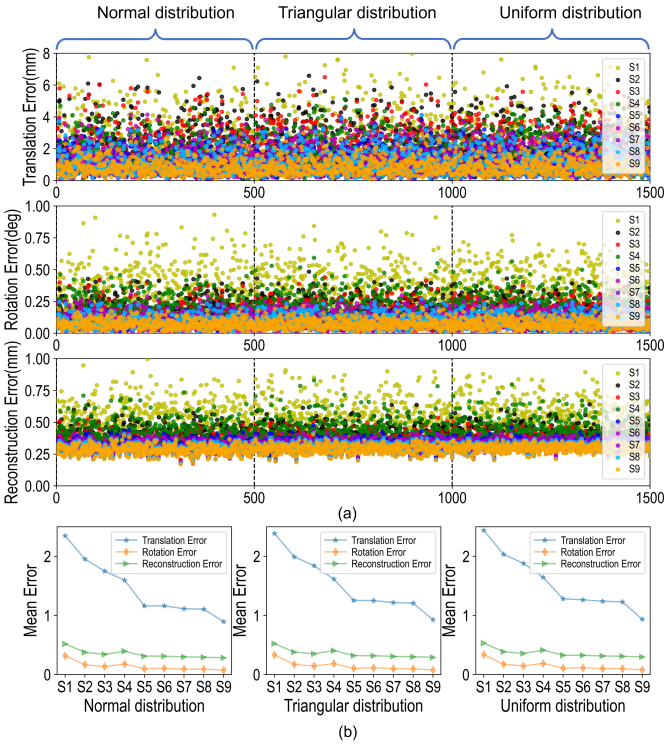

Fig. 9: Impact of different calibration point distributions and counts on calibration results. (a) The dispersion of errors under different point sets. There are three segments, respectively, demonstrate the impact of noise distributed as a normal distribution, a triangular distribution, and a uniform distribution on the error. (b) The variations in the mean error across distinct sets of points when subjected to three different noise distributions.

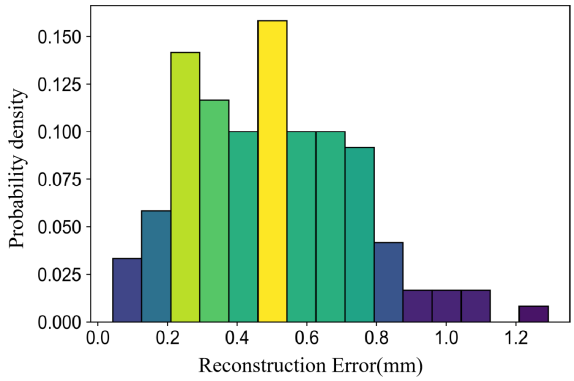

Fig. 10: Histogram of the reconstruction errors in the actual scene calibration experiment.

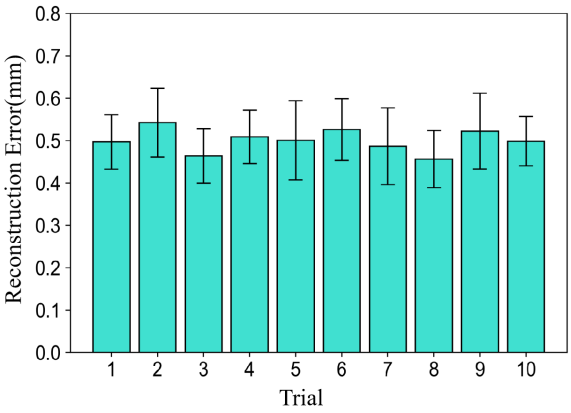

Fig. 11: Average reconstruction error of the real scene experiment, values are presented as Mean ± SEM. The average reconstruction error of 10 calibrations is 0.501mm.

belt. The conveyor belt encoder is a 10-bit resolution absolute encoder. During the experiment, the conveyor belt runs at a constant speed. The homogeneous transformation matrix of the end effector tool of the robot was measured from CAD drawings, and the transformation matrices for the gripper and suction cup can be represented respectively as follows:

$$T_E^G = \begin{bmatrix} Rot([0 \quad 0 \quad 1]^T, \frac{\pi}{6}) & [0,0,0] \\ \mathbf{0} & 1 \end{bmatrix}$$
$$T_E^S = \begin{bmatrix} Rot([0 \quad 0 \quad 1]^T, \frac{\pi}{6}) & [157,0,-20] \\ \mathbf{0} & 1 \end{bmatrix} \tag{28}$$

In the real-world experiments, we conducted a calibration experiment and a grasping experiment. The calibration experiment reflected the performance of the proposed calibration method on the physical system, while the grasping experiment further verified the effectiveness of the integrated calibration method in performing target grasping with the entire sorting system.

1) Calibration Experiment: To facilitate precise calibration in the experiment, we designed and fabricated a calibration tool, as shown in Fig. 6(a), which could be securely held by the gripper. We defined the Tool Center Point (TCP) of the robot at the tip of this calibration tool. During the calibration process, we aligned the tip of the calibration tool with the 12 points shown in Fig. 6(b) on the calibration board as well as four additional circular markers. The first 12 points were used to calculate the reconstruction error, while the latter four points were used as input for the calibration algorithm. To verify the reliability of our proposed integrated calibration method, we repeated the calibration experiment 10 times. In the actual experiment, considering the calibration error of the camera, the reconstruction error for the i-th experiment can be represented by Equation 29.

$$\widetilde{E_{pi}} = \left\| T_B^R T_C^B(L_1) H(P^*{}_i) - P^R{}_i \right\|_2 \tag{29}$$

The reconstruction error distribution of the actual scene calibration experiment is shown in Fig. 10. From the error distribution, it can be observed that the majority of errors are between 0.3 to 0.8 mm. Through Anderson-Darling test, it is indicated that the reconstruction errors follow a normal distribution. For the extremely few errors exceeding 1 mm, in comparison with the simulation experiment results, reconstruction errors greater than the Low Noise level are attributed not only to inherent system errors but also to the instability in manual operations and the errors caused by insufficient rigidity of the robot TCP resulting in oscillations. The average reconstruction

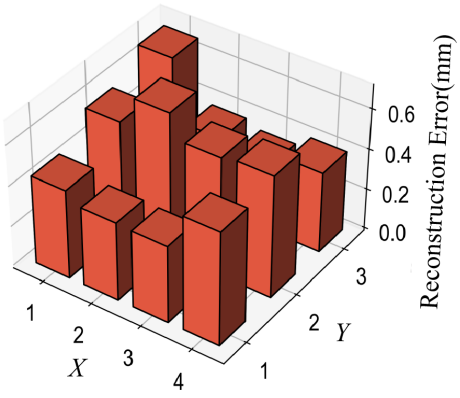

Fig. 12: Average error at 12 different positions on the calibration board.

errors from multiple repeated measurements are shown in Fig. 11, with the average error stabilizing around 0.5 mm. ANOVA repeated measurements on the data yielded a P-value of 0.998, indicating that at a 99% confidence level, there were no significant differences between the mean values of each measurement, demonstrating the stability of the calibration results and the reliability of the integrated calibration method proposed in this paper. After completing the calibration using the S5 grid points, error verification was conducted for all 12 points, and the results are shown in Fig. 12. The average reconstruction errors for the 12 points on the chessboard grid are concentrated between 0.3 to 0.5 mm, further illustrating the effectiveness of the calibration method proposed in this paper. It is noteworthy that by utilizing only the four corner points of the chessboard grid (S5) for calibration, lower reconstruction errors were achieved, with no significant differences in errors across different spatial positions, thereby balancing complexity and calibration efficiency.

To further validate the effectiveness of the proposed integrated calibration method, we conducted comparative experiments on the sorting platform against the traditional stepwise calibration approach, which uses an industrial camera, conveyor belt, and manual point collection. Under the same hardware platform and system configuration, two types of experiments were carried out using both calibration methods: (1) Calibration Error Evaluation, in which positioning errors were measured at 12 spatial locations and reported as Mean ± SEM; and (2) Sorting Success Rate Evaluation, where 10 types of objects were randomly grasped in 10 rounds (10 objects per round), resulting in a total of 100 grasping attempts for each method. The experimental results are presented in Table IV.

As shown in Table IV, the proposed method significantly reduces both the maximum positioning error and the variation in error, thereby enhancing the overall consistency of system localization. In the grasping experiments, it also achieved a higher sorting success rate. An analysis of the failed cases revealed that, after applying our method, most grasping failures were attributed to object detection inaccuracies, target rolling,

TABLE IV: Comparison of calibration methods: positioning error and sorting success rate.

Method	Max Positioning Error (mm)	Mean Error ± SEM (mm)	Sorting Success Rate
Traditional Stepwise Calibration	2.2	1.91 ± 0.52	87.6%
Proposed Integrated Calibration	0.8	0.42 ± 0.18	90.2%

or mechanical execution errors such as gripper misalignment, rather than system localization errors. This indicates that the proposed calibration approach provides practically effective localization accuracy for the sorting system. In the final system deployment, the overall grasping success rate reached 90.2%, demonstrating that the calibration accuracy meets the operational requirements of the system.

2) Online Grasping Experiment: The robot sorting platform constructed in this paper is shown in Fig. 13 (a). The sorting objects, training set and scene configurations are derived from the KWG2024 dataset [15] and related literature [8]. We integrated detection algorithms such as YOLOv5s [47], AL-DETR [8], and GG-CNN [48] into the system to conduct grabbing experiments on different categories of garbage, as illustrated in Fig. 13 (b). Utilizing the integrated calibration results following the steps described in Fig. 2, the detected object bounding boxes or grasp rectangles, as shown in Fig. 13 (c), were transformed into spatial grasp poses in the robot coordinate system and sent to the sorting controller. We employed a perception-planning-execution loop [49] to continuously process objects on the conveyor belt. To achieve stable grasping during conveyor belt motion, we employed an open-loop controller to move the gripper directly above the object's position, then tracked the object while descending to the object's position for grasping, and finally placed it in the corresponding location based on its category. Since the object detection algorithm cannot provide the object's rotation angle, we set the gripper to always grasp the short edge of the target bounding box. The maximum width of the robot gripper is 130 mm. In practice, the gripper opens only to a width exceeding the target object's grasping width by no more than 10%. This approach helps maintain high grasping efficiency and minimizes collisions with other items. In the experiment, the end-effector of the robot was equipped with a multifunctional gripper integrated with a force sensor. The gripper is used for sorting solid waste with certain weight, while suction cups are employed for handling lightweight and flat objects. For fragile objects such as glass bottles, the robot adjusted the gripping force through the control software to prevent breakage during the sorting process. Specifically, force control parameters were configured for both gripping and suction operations. A threshold of 20 N was set as the standard for adjusting the gripping force when the detected object was classified as a glass bottle. To simplify the process, fixed gripping forces of 10 N and 13 N were assigned for plastic bottles and other waste categories, respectively. For waste items such as bags and cartons, vacuum suction was applied with a pressure range between -60 kPa and -80 kPa

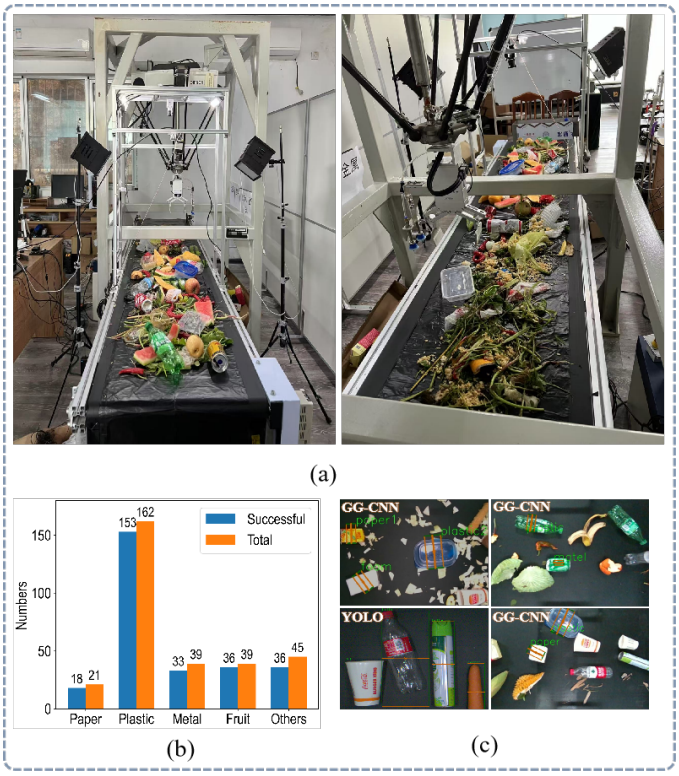

Fig. 13: (a) Physical robot sorting system, (b) Objects statistic result for grasping experiment, and (c) Object detection results from sorting software.

for sorting.

In the actual grasping experiments in this section, we evaluate the efficiency of the sorting system based on the grasping success rates ω_i for each garbage category, the average grasping success rate of the sorting system $\bar{\omega}$, and the sorting rate η. The relevant definitions are given by equations (30)-(32).

$$\omega_i = \frac{\sum G_i}{\sum N_i} \times 100\% \tag{30}$$

$$\bar{\omega} = \frac{\sum G}{\sum N} \times 100\% \tag{31}$$

$$\eta = \frac{\sum G}{T} \tag{32}$$

In the above equation, N represents the detected object on the conveyor belt, G represents the successfully grasped objects, i denotes the i-th garbage category object, and T is the unit of time. To closely resemble real-world scenarios of urban kitchen waste, we designed the experiments by taking into account the conditions of our laboratory space and available equipment. Specifically, we created complex experimental environments similar to those in waste sorting centers by simulating factors such as diverse waste categories, conveyor belt speed, interference from discarded fruit and vegetable leaves, and object stacking. Due to laboratory constraints, highly oily, watery, or decayed waste was not included in these experiments. This is because the success rate of robotic grasping is more strongly influenced by the detection accuracy of the perception model than by the execution precision of the robotic manipulator. In the experiment, ten objects were randomly placed on a conveyor belt moving at a speed of 104 mm/s and tested over 10 repeated trials. The sorting results under two unstructured and complex scenarios are presented in Table V.

TABLE V: Sorting performance of the robot for different waste categories under two types of unstructured complex scenarios.

Categories		Paper	Plastic	metal	Fruit	Other
Condition 1	Total	21	162	39	39	45
	Successful	18	153	33	36	36
	ω_t	85.7%	94.4%	84.6%	92.3%	80.0%
	$\bar{\omega}$			90.2%		
Condition 2	Total	21	162	39	39	45
	Successful	13	147	32	30	31
	ω_t	61.9%	90.7%	82.1%	76.92%	68.9%
	$\bar{\omega}$			82.7%		

Note: Condition 1 refers to an unstructured scenario with a relatively clean background; Condition 2 represents an unstructured scenario with a complex background.

According to the experimental results in Table V, the sorting success rate in Scenario 1 exceeds 90%, while Scenario 2 achieves a success rate of 82.7%, indicating that the complex background introduces a certain degree of interference in visual detection. The sorting success rates for plastic bottles in Scenario 1 and Scenario 2 are 94.4% and 90.7%, respectively. This demonstrates that bottles, due to their distinct visual features and high frequency in the KWG2024 dataset, are well-trained and can be reliably sorted by the platform. For paperboard waste, the sorting success rates vary significantly between the two scenarios, suggesting that small paper scraps and tissues in the cluttered background negatively impact the detection of this category. For metal cans, the performance is consistent across both scenarios, implying that the model has effectively learned the visual characteristics of this category. In the case of fruits and vegetables, the sorting success rate drops considerably in Scenario 2, mainly because discarded leaves and vegetable debris introduce strong interference, leading to misclassification by the detection model and resulting in failed grasps. Futhormore, sorting accuracy remains unaffected by slight deformations in object appearance. Mis-sorting occurs only when the object undergoes significant deformation.

In the sorting tests, we recorded the specific causes of 30 grasping failures under Scenario 1. Among them, 18 failures were due to detection errors, 5 were caused by grasp slippage, 3 resulted from system calibration inaccuracies, and 4 were attributed to other factors such as conveyor belt inertia. It was observed that the majority of failures occurred due to grasp attempts based on inaccurate target localization or misclassification by the detection module. The limitations of the dataset also contribute to sorting failures.

The proposed sorting system and integrated calibration method are capable of accurately sorting objects of varying shapes and materials on a conveyor belt in real time, achieving a sorting success rate of 90.2% in a simple scenario and 82.7%

in a complex scenario, with a throughput of 979 items per hour. In practical applications, our system can be adapted to various sorting-related domains by switching to appropriate detection models. However, future work will focus on improving the structure of the multifunctional end-effector and developing more accurate object detection models.

VI. Conclusion

In this study, we designed a multifunctional sorting system for kitchen waste based on a Delta robot, consisting of three main modules: visual detection, information processing, and multifunctional robotic arm sorting. The visual detection module captures images of waste on the conveyor belt and transmits them to the information processing unit in real time. The information processing is performed on the designed robot sorting control system XRobot, which integrates a robotic arm with grippers and suction cups equipped with force sensors. The robotic arm selects between grippers and suction cups based on the category and morphology of the waste, enabling multifunctional sorting operations. Additionally, we proposed a robust and easy-to-use integrated calibration method for the system, based on our innovative calibration board. With just three simple operations, achieving high-precision calibration of the system. Both simulation and real-world experiments demonstrated that our proposed calibration method achieves low error and high accuracy. Finally, the reliability of the system was validated through practical experiments in waste sorting, providing a feasible solution for actual kitchen waste sorting applications.

References

[1] Z. X. Hoy, K. S. Woon, W. C. Chin, Y. Van Fan, and S. J. Yoo, "Curbing global solid waste emissions toward net-zero warming futures," *Science*, vol. 382, no. 6672, pp. 797–800, 2023.

[2] W. Ferdous, A. Manalo, R. Siddique, P. Mendis, Y. Zhuge, H. S. Wong, W. Lokuge, T. Aravinthan, and P. Schubel, "Recycling of landfill wastes (tyres, plastics and glass) in construction–a review on global waste generation, performance, application and future opportunities," *Resources, Conservation and Recycling*, vol. 173, p. 105745, 2021.

[3] L. Chen and M. Gao, "Novel information interaction rule for municipal household waste classification behavior based on an evolving scale-free network," *Resources, Conservation and Recycling*, vol. 168, p. 105445, 2021.

[4] Q. Yu and H. Li, "Moderate separation of household kitchen waste towards global optimization of municipal solid waste management," *Journal of cleaner production*, vol. 277, p. 123330, 2020.

[5] J. Yu, L. Tang, Y. Pang, G. Zeng, H. Feng, J. Zou, J. Wang, C. Feng, X. Zhu, X. Ouyang *et al.*, "Hierarchical porous biochar from shrimp shell for persulfate activation: A two-electron transfer path and key impact factors," *Applied Catalysis B: Environmental*, vol. 260, p. 118160, 2020.

[6] H. Liao, X. Lu, C. Rensing, V. P. Friman, S. Geisen, Z. Chen, Z. Yu, Z. Wei, S. Zhou, and Y. Zhu, "Hyperthermophilic composting accelerates the removal of antibiotic resistance genes and mobile genetic elements in sewage sludge," *Environmental science & technology*, vol. 52, no. 1, pp. 266–276, 2018.

[7] L. Fang, Q. Tang, L. Ouyang, J. Yu, J. Lin, S. Ding, and L. Tang, "Long-tailed object detection of kitchen waste with class-instance balanced detector," *Science China Technological Sciences*, vol. 66, no. 8, pp. 2361–2372, 2023.

[8] H. Qin, L. Shu, L. Zhou, S. Deng, H. Xiao, W. Sun, Q. Liang, D. Zhang, and Y. Wang, "Active learning-detr: Cost-effective object detection for kitchen waste," *IEEE Transactions on Instrumentation and Measurement*, vol. 73, pp. 1–15, 2024.

[9] D. Guo, H. Liu, B. Fang, F. Sun, and W. Yang, "Visual affordance guided tactile material recognition for waste recycling," *IEEE Transactions on Automation Science and Engineering*, vol. 19, no. 4, pp. 2656–2664, 2021.

[10] S. Kong, M. Tian, C. Qiu, Z. Wu, and J. Yu, "Iwscr: An intelligent water surface cleaner robot for collecting floating garbage," *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, vol. 51, no. 10, pp. 6358–6368, 2020.

[11] L.-B. Chen, X.-R. Huang, W.-H. Chen, W.-Y. Pai, G.-Z. Huang, and W.-C. Wang, "Design and implementation of an artificial intelligence of things-based autonomous mobile robot system for cleaning garbage," *IEEE Sensors Journal*, 2023.

[12] X. Chen, H. Huang, Y. Liu, J. Li, and M. Liu, "Robot for automatic waste sorting on construction sites," *Automation in Construction*, vol. 141, p. 104387, 2022.

[13] M. Koskinopoulou, F. Raptopoulos, G. Papadopoulos, N. Mavrakis, and M. Maniadakis, "Robotic waste sorting technology: Toward a vision-based categorization system for the industrial robotic separation of recyclable waste," *IEEE Robotics & Automation Magazine*, vol. 28, no. 2, pp. 50–60, 2021.

[14] T. Kiyokawa, J. Takamatsu, and S. Koyanaka, "Challenges for future robotic sorters of mixed industrial waste: A survey," *IEEE Transactions on Automation Science and Engineering*, vol. 21, no. 1, pp. 1023–1040, 2024.

[15] S. Deng, R. Pei, L. Zhou, H. Qin, W. Sun, and Q. Liang, "An efficient generative intelligent multiobjective grasping model for kitchen waste sorting," *IEEE Transactions on Instrumentation and Measurement*, vol. 74, pp. 1–10, 2025.

[16] R. Sadeghian, S. Shahin, and S. Sareh, "Vision-based self-adaptive gripping in a trimodal robotic sorting end-effector," *IEEE Robotics and Automation Letters*, vol. 7, no. 2, pp. 2124–2131, 2022.

[17] M. Leveziel, G. J. Laurent, W. Haouas, M. Gauthier, and R. Dahmouche, "A 4-dof parallel robot with a built-in gripper for waste sorting," *IEEE Robotics and Automation Letters*, vol. 7, no. 4, pp. 9834–9841, 2022.

[18] S. Jin, Z. Yang, G. Królczykg, X. Liu, P. Gardoni, and Z. Li, "Garbage detection and classification using a new deep learning-based machine vision system as a tool for sustainable waste recycling," *Waste Management*, vol. 162, pp. 123–130, 2023.

[19] J. Li, J. Chen, B. Sheng, P. Li, P. Yang, D. D. Feng, and J. Qi, "Automatic detection and classification system of domestic waste via multimodel cascaded convolutional neural network," *IEEE transactions on industrial informatics*, vol. 18, no. 1, pp. 163–173, 2021.

[20] Z. Wu, H. Li, X. Wang, Z. Wu, L. Zou, L. Xu, and M. Tan, "New benchmark for household garbage image recognition," *Tsinghua Science and Technology*, vol. 27, no. 5, pp. 793–803, 2022.

[21] S. Qiu, M. Wang, and M. R. Kermani, "A new formulation for hand–eye calibrations as point-set matching," *IEEE Transactions on Instrumentation and Measurement*, vol. 69, no. 9, pp. 6490–6498, 2020.

[22] Q. Ma, H. Li, and G. S. Chirikjian, "New probabilistic approaches to the ax= xb hand-eye calibration without correspondence," in *2016 IEEE international conference on robotics and automation (ICRA)*. IEEE, 2016, pp. 4365–4371.

[23] J. Lu, F. Richter, and M. C. Yip, "Markerless camera-to-robot pose estimation via self-supervised sim-to-real transfer," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2023, pp. 21 296–21 306.

[24] L. Li, L. Zhou, T. Zhang, and J. Zheng, "Industrial robot hand-eye calibration combining data augmentation and actor-critic network," *IEEE Transactions on Instrumentation and Measurement*, 2023.

[25] H. Cheng, Z. Zhang, and W. Li, "Efficient hand eye calibration method for a delta robot pick-and-place system," in *2015 IEEE International Conference on Cyber Technology in Automation, Control, and Intelligent Systems (CYBER)*. IEEE, 2015, pp. 175–180.

[26] M. Cai, H. Liu, and M. Dong, "Easy pose-error calibration for articulated serial robot based on three-closed-loop transformations," *IEEE Transactions on Instrumentation and Measurement*, vol. 70, pp. 1–11, 2021.

[27] W.-L. Li, H. Xie, G. Zhang, S.-J. Yan, and Z.-P. Yin, "Hand–eye calibration in visually-guided robot grinding," *IEEE transactions on cybernetics*, vol. 46, no. 11, pp. 2634–2642, 2015.

[28] L. Zhang, J.-Z. Zhang, X. Jiang, and B. Liang, "Error correctable hand–eye calibration for stripe-laser vision-guided robotics," *IEEE Transactions on Instrumentation and Measurement*, vol. 69, no. 10, pp. 8314–8327, 2020.

[29] M. Ulrich and M. Hillemann, "Uncertainty-aware hand-eye calibration," *IEEE Transactions on Robotics*, vol. 40, pp. 573–591, 2024.

[30] 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.

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

[32] J. Ha, "Probabilistic framework for hand–eye and robot–world calibration $ax = yb$," *IEEE Transactions on Robotics*, vol. 39, no. 2, pp. 1196–1211, 2022.

[33] H. Nguyen and Q.-C. Pham, "On the covariance of $\boldsymbol{X}$ in $\boldsymbol{AX} = \boldsymbol{XB}$," *IEEE Transactions on Robotics*, vol. 34, no. 6, pp. 1651–1658, 2018.

[34] J. Pan, Z. Fu, H. Yue, X. Lei, M. Li, and X. Chen, "Toward simultaneous coordinate calibrations of ax=yb problem by the lmi-sdp optimization," *IEEE Transactions on Automation Science and Engineering*, vol. 20, no. 4, pp. 2445–2453, 2023.

[35] L. Wu and H. Ren, "Finding the kinematic base frame of a robot by hand-eye calibration using 3d position data," *IEEE Transactions on Automation Science and Engineering*, vol. 14, no. 1, pp. 314–324, 2016.

[36] H. Li, Q. Ma, T. Wang, and G. S. Chirikjian, "Simultaneous hand-eye and robot-world calibration by solving the $ax = yb$ problem without correspondence," *IEEE Robotics and Automation Letters*, vol. 1, no. 1, pp. 145–152, 2015.

[37] K. Pachtrachai, F. Vasconcelos, P. Edwards, and D. Stoyanov, "Learning to calibrate-estimating the hand-eye transformation without calibration objects," *IEEE Robotics and Automation Letters*, vol. 6, no. 4, pp. 7309–7316, 2021.

[38] L. Li, L. Zhou, T. Zhang, and J. Zheng, "Industrial robot hand-eye calibration combining data augmentation and actor-critic network," *IEEE Transactions on Instrumentation and Measurement*, vol. 72, pp. 1–12, 2023.

[39] O. Özgüner, T. Shkurti, S. Huang, R. Hao, R. C. Jackson, W. S. Newman, and M. C. Çavuşoğlu, "Camera-robot calibration for the da vinci robotic surgery system," *IEEE Transactions on Automation Science and Engineering*, vol. 17, no. 4, pp. 2154–2161, 2020.

[40] J. Odeyemi, A. Ogbeyemi, K. Wong, and W. Zhang, "On automated object grasping for intelligent prosthetic hands using machine learning," *Bioengineering*, vol. 11, no. 2, p. 108, 2024.

[41] H.-S. Kim, T.-Y. Kuc, and K.-H. Lee, "Hand-eye calibration using images restored by deep learning," in *2020 IEEE International Conference on Consumer Electronics-Asia (ICCE-Asia)*. IEEE, 2020, pp. 1–4.

[42] D. Shang, Y. Wang, Z. Yang, J. Wang, and Y. Liu, "Study on comprehensive calibration and image sieving for coal-gangue separation parallel robot," *Applied Sciences*, vol. 10, no. 20, p. 7059, 2020.

[43] W. Kabsch, "A solution for the best rotation to relate two sets of vectors," *Acta Crystallographica Section A: Crystal Physics, Diffraction, Theoretical and General Crystallography*, vol. 32, no. 5, pp. 922–923, 1976.

[44] Z. Zhang, "A flexible new technique for camera calibration," *IEEE Transactions on pattern analysis and machine intelligence*, vol. 22, no. 11, pp. 1330–1334, 2000.

[45] G. Wang, W.-l. Li, C. Jiang, D.-h. Zhu, H. Xie, X.-j. Liu, and H. Ding, "Simultaneous calibration of multicoordinates for a dual-robot system by solving the axb= ycz problem," *IEEE Transactions on Robotics*, vol. 37, no. 4, pp. 1172–1185, 2021.

[46] J. Jiang, X. Luo, Q. Luo, L. Qiao, and M. Li, "An overview of hand-eye calibration," *The International Journal of Advanced Manufacturing Technology*, vol. 119, no. 1-2, pp. 77–97, 2022.

[47] G. Jocher, A. Chaurasia, A. Stoken, J. Borovec, Y. Kwon, K. Michael, J. Fang, Z. Yifu, C. Wong, D. Montes *et al.*, "ultralytics/yolov5: v7. 0-yolov5 sota realtime instance segmentation," *Zenodo*, 2022.

[48] D. Morrison, P. Corke, and J. Leitner, "Learning robust, real-time, reactive robotic grasping," *The International journal of robotics research*, vol. 39, no. 2-3, pp. 183–201, 2020.

[49] F. Islam, O. Salzman, A. Agarwal, and M. Likhachev, "Provably constant-time planning and replanning for real-time grasping objects off a conveyor belt," *The International Journal of Robotics Research*, vol. 40, no. 12-14, pp. 1370–1384, 2021.