

Bimanual Regrasp Planning and Control for Active Reduction of Object Pose Uncertainty

Ryuta Nagahama, Weiwei Wan , Zhengtao Hu , *Member, IEEE*, and Kensuke Harada , *Fellow, IEEE*

Abstract—Precisely grasping an object is a challenging task due to pose uncertainties. Conventional methods have used cameras and fixtures to reduce object uncertainty. They are effective but require intensive preparation, such as designing jigs based on the object geometry and calibrating cameras with high-precision tools fabricated using lasers. In this study, we propose a method to reduce the uncertainty of the position and orientation of a grasped object without using a fixture or a camera. Our method is based on the concept that the flat finger pads of a parallel gripper can reduce uncertainty along its opening/closing direction through flat surface contact. Three approximately orthogonal grasps by parallel grippers with flat finger pads collectively constrain an object's position and orientation to a unique state. Guided by the concepts, we develop a regrasp planning and admittance control approach that sequentially finds and leverages three approximately orthogonal grasps of two robotic arms to actively reduce uncertainties in the object pose. We evaluated the proposed method on different initial object uncertainties and verified that it had good repeatability. The deviation levels of the experimental trials were on the same order of magnitude as those of an optical tracking system, demonstrating strong relative inference performance.

Index Terms—Grasping uncertainty, regrasping, bimanual manipulation.

I. INTRODUCTION

A SIGNIFICANT challenge in robotic manipulation lies in addressing the uncertainties associated with object grasping. The uncertainties often arise from errors in environmental registration, inaccuracies in object pose recognition, and unbalanced contact during grasping that leads to pose deviations. The uncertainties can result in discrepancies between the actual and expected poses of objects or tools, potentially causing task failures. Existing methods used to address uncertainty problems typically fall into two categories: Sensor-based correction methods [1], [2] and fixture-based constraint methods [3], [4]. Sensor-based methods leverage additional hardware, such as force

Received 28 February 2025; accepted 1 June 2025. Date of publication 16 June 2025; date of current version 30 June 2025. This article was recommended for publication by Associate Editor F. Shi and Editor O. Stasse upon evaluation of the reviewers' comments. The work was supported by JSPS KAKENHI Grant Number JP22K12205 and partially a joint project with Panasonic Holdings Corporation. Zhengtao Hu's contribution was supported by National Natural Science Foundation of China (62403297). (*Corresponding author: Weiwei Wan.*)

Ryuta Nagahama, Weiwei Wan, and Kensuke Harada are with the Graduate School of Engineering Science, Osaka University, Osaka 565-0871, Japan (e-mail: wan.weiwei.es@osaka-u.ac.jp).

Zhengtao Hu is with the School of Mechatronic Engineering and Automation, Shanghai University, Shanghai 200444, China.

This article has supplementary downloadable material available at <https://doi.org/10.1109/LRA.2025.3580327>, provided by the authors.

Digital Object Identifier 10.1109/LRA.2025.3580327

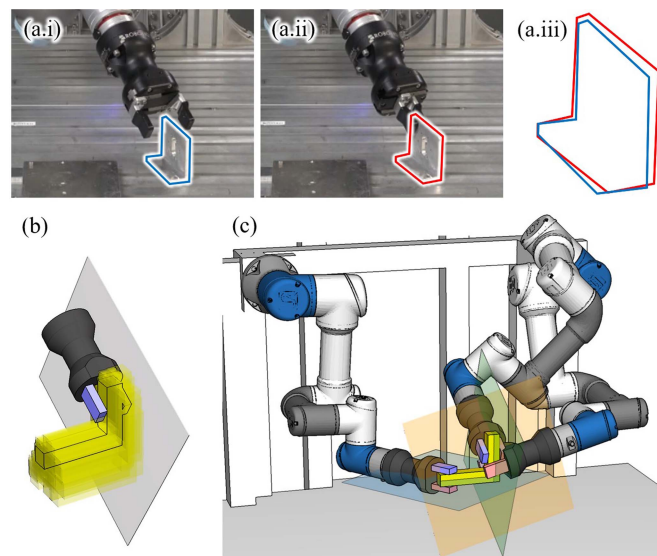


Fig. 1. (a) Precisely grasping an object is difficult. (a.i) Before grasping. (a.ii) After grasping. (a.iii) Comparison. (b) Flat finger pads of a parallel gripper help constrain uncertainty to a plane. (c) Three approximately orthogonal parallel grasps constrain three planes, leading to an exact pose.

sensors or tactile sensors, to detect and mitigate uncertainties. Fixture-based methods, on the other hand, rely on structuring the environment or using special-purpose structures to constrain object pose. However, both approaches have notable limitations, as they often require additional hardware adjustments or environment modifications.

This letter proposes a bimanual regrasp planning method that leverages flat finger pad constraints and admittance control to reduce uncertainties. It does not require external fixtures or additional sensors and is more flexible for diverse robotic manipulation scenarios. The fundamental concept of the proposed method is illustrated in Fig. 1(b) and (c). It is based on the observation that the flat finger pads of parallel grippers can reduce uncertainty along the hand's opening/closing direction through flat surface contact. Three such grasps collectively constrain the object's position and orientation to a unique state.¹ Guided by the concepts, we develop a regrasp planning approach [5] that

¹This unique pose constraint holds under the following assumptions: i) The object is sufficiently rigid such that the grasping force is transmitted and sensed through the force sensors, rather than causing deformation. ii) The contact between the gripper fingers and the object is assumed to be flat surface contact with sufficient friction to prevent slippage. iii) The maximum uncertainty in the object pose is within the gripper's opening range, ensuring valid contact during grasping.

sequentially executes a series of regrasping motions using two robotic arms while selecting approximately orthogonal grasp triplets to simultaneously reduce uncertainties in the object pose. However, grasp planning alone does not fully achieve our goal, as simply executing the planned motion sequence using position control may exert excessive grasping forces on the uncertain object and lead to deformation or even damage. To address this issue, we develop an admittance control [6] method that utilizes force/torque sensors at the robots' wrists to adjust the grasping pose and conform to the uncertain object pose. By integrating regrasp planning with admittance control, our method iteratively reduces uncertainties along three approximately orthogonal axes, thus achieving an accurate object pose estimation.

In the experiments, we implemented the proposed method and evaluated its performance on different initial object uncertainties. The results demonstrate that the method achieves satisfactory repeatability. The deviation levels of remaining uncertainties were the same order of magnitude as those of an optical tracking system, demonstrating that the method provides strong relative inference performance. By leveraging the robot's inherent geometric constraints and force sensors, the proposed approach can help actively reduce uncertainties in the grasped object.

II. RELATED WORK

A. Categories of Regrasp Planning Methods

Regrasp planning aims to automatically generate grasp sequences to accomplish object manipulation. It involves two main categories: discrete and continuous.

Discrete regrasp planning focuses on identifying discrete pick-and-place sequences. For example, Liu et al. [7] developed regrasp planning using multiple mobile robots. The method achieves flipping manipulation with the help of regrasping. Kim et al. [8] developed a randomized planning method for regrasp planning. Baek et al. [9] considered non-prehensile grasping in regrasp planning and secured power grasps for stable holding. Qin et al. [10] presented a regrasp planning method for installing a deformable linear object. Rigidity constraints were considered when planning regrasp motion. Mitrano et al. [11] studied a similar problem and proposed considering avoiding environment-grasp-body loops to accelerate grasp selection and planning. Sundaralingam et al. [12] extended regrasp planning to a multi-fingered hand. Murooka et al. [13] proposed the foot-step and regrasping planning method for a humanoid robot to simultaneously determine foot steps and hand switches.

Continuous regrasp methods, on the other hand, aim to adjust object poses through controlled rolling or sliding motions while maintaining contact. For instance, Maximo et al. [14] presented the grasp space method for continuous regrasp (rolling or sliding) that maintains force closures. Hu et al. [15] formulated regrasp planning as a multi-modal problem, where three manifolds, including the regrasp manifold, the transfer manifold, and the slide manifold, were planned across to determine action sequences and thus plan motion. Cheng et al. [16] further proposed the contact mode method to guide manipulation planning across multiple manifolds.

The work presented in this letter is based on iterative pick-and-place actions using two robot hands, and thus it falls under the first type of regrasp planning. It can be seen as a concrete realization of a flipped funnel [17], where the object, through interaction, progressively guides the three structured grasping poses, allowing its uncertain state to be estimated toward a determined pose.

B. Key Challenges and the Focus of This Work

One of the most critical challenges in regrasp planning is the planning speed. Researchers have explored various methods to enhance efficiency. For example, Cho et al. [18] developed a fast regrasp planning algorithm using look-up tables and two-finger parallel grippers. Chiu et al. [19] proposed a reinforcement learning method to accelerate the planning of regrasp poses. Their actions were defined in the local frame of a robot hand, thereby resolving the dependencies on robotic kinematics. Wada et al. [20] significantly improved planning speed by using a neural network to filter waypoint poses and quickly select grasp poses. Xu et al. [21] proposed learning reorientation object poses and then planning regrasp motion while considering the learned middle reorientation poses. Learning-based methods have demonstrated effectiveness in improving regrasp planning efficiency. While there are numerous other publications, this letter does not focus heavily on this aspect and thus will not delve into further details.

Another critical challenge in regrasp planning is uncertainty. Repeated regrasping can result in error accumulation, causing the grasp pose to deviate significantly from the planned outcome. The deviation can prevent the robot from successfully executing subsequent tasks. To address the uncertainty problems, Bauza et al. [22] integrated tactile information into regrasp planning and achieved precise and robust pick and placement for regrasping. Bronars et al. [23] proposed a simultaneous tactile estimation and control method to plan sliding regrasp motion while maintaining grasping precision continuously. Gualtieri et al. [24] estimated the quality of grasp candidates against uncertainty based on point cloud completion and optimized regrasp sequences to be robust. Zhang et al. [25] used reinforcement learning to learn a regrasp policy that optimizes grasp stability. Both vision and visual-tactile information were used for learning. This study can be partially classified as research to reduce uncertainties in grasping poses during the regrasp planning process. However, our primary goal is not to achieve precise pose estimation for regrasping. Instead, we focus on actively leveraging the consecutive geometric constraints inherent in the regrasping process [26], [27], [28], [29] to reduce uncertainty.

From a broader perspective, our research also falls within the category of methods that utilize interactive perception [30] to estimate the pose of grasped objects. In addition to the grasping-based interaction approaches, recent literature on non-prehensile interaction has been receiving growing attention. For example, Jankowski et al. [31] discussed the stochastic contact dynamics of pushing and proposed a pushing planning method that minimizes uncertainty considering informed priors. Dong et al. [32], on the other hand, modeled the contact dynamics of

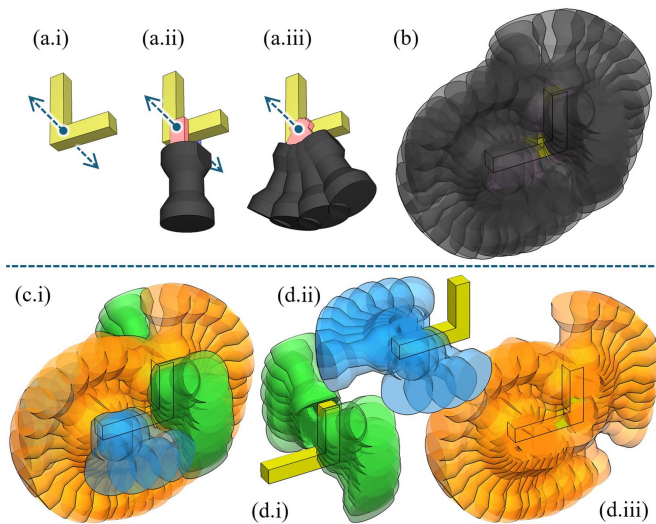


Fig. 2. (a) Antipodal grasp planning. (b) Candidate grasp poses planned using the antipodal method. (c) Separating the candidate grasp poses into three orthogonal groups considering their opening and closing directions. Each color in the subfigure represents one group.

pushing as an energy model and used constraints on the energy margin to plan robust pushing.

III. REGRASP PLANNING USING ORTHOGONAL TRIPLETS

To achieve the goal of this study, we need to identify grasp poses along three approximately orthogonal directions. Given a 3D model of an object, our method begins with antipodal grasp planning [33]. During the grasp planning, we sample contact points on the object’s surface. For each sampled contact point, a ray is cast in the direction opposite to the surface normal of the point. The intersection between this ray and the object’s surface is determined using ray tracing. If the surface normal at the intersection point is opposite in direction to the normal at the ray’s starting point (satisfying the parallel grasp constraint), this intersection point and the starting point are stored as a candidate pair. For all candidate pairs, we align the gripper fingers’ contact centers with these points and rotate the gripper about the axis passing through the pair. Gripper poses that do not result in collisions are extracted as candidate grasps. Fig. 2(a.i)~(a.iii) illustrates the grasp planning process on a single contact point. The greater the number of sampled contact points, the larger the number of generated candidate grasps. Fig. 2(b) demonstrates the generated candidate grasps when the number of contact points is set to 5.

Subsequently, based on the results of the grasp planning, we separate the grasps into groups. Since our goal is to identify three approximately orthogonal grasp poses, the grouping criterion is defined by the opening and closing directions of the gripper fingers. Grasp poses with the same opening and closing direction are grouped together, and each group can be represented by a unique opening and closing vector, denoted as \mathbf{v}_i^2 . Through the grouping process, we can obtain several groups represented by $\mathbf{v}_1, \mathbf{v}_2, \dots$. The candidate grasp poses in Fig. 2(b) are classified

²In this context, the direction of \mathbf{v}_i is not considered.

into three strictly orthogonal groups. They are visualized in Fig. 2(c) using red, green, and blue.

After obtaining the grasp groups, we perform triplet combinations to identify three groups that satisfy the approximately orthogonal constraint. To evaluate whether the constraint is satisfied, we use the following formula to compute a score for each triplet combination.

$$\text{score} = |\mathbf{v}_i \cdot \mathbf{v}_j| + |\mathbf{v}_i \cdot \mathbf{v}_k| + |\mathbf{v}_j \cdot \mathbf{v}_k| \quad (1)$$

The $\mathbf{v}_i, \mathbf{v}_j$, and \mathbf{v}_k in the above equation represent the opening and closing directions of three different groups in a triplet. We sort the scores of all triplets and prioritize them with a score closest to 0 for regrasp planning and uncertainty reduction. The closer the score value is to 0, the more orthogonal the three grasps from the groups are to each other. By performing successive regrasp using the prioritized triplets, object uncertainties can be effectively minimized.

It is important to note that strict 90-degree orthogonality is not required to determine the object’s pose. The example shown in Fig. 2 represents a special case where the planned grasp poses are right divided into three orthogonal groups with a score of zero. In more general cases, there may be multiple candidate groups and triplet combinations. The triplets may have non-zero scores. However, as long as the grasp direction vectors $\mathbf{v}_i, \mathbf{v}_j$, and \mathbf{v}_k form a non-singular matrix when arranged as columns, the pose can be estimated. This point will be clearer in Section V when we introduce the pose estimation algorithm. Here, the reason we sort triplet grasp combinations by their score is not for enforcing strict orthogonality, but rather to account for potential changes of grasp directions caused by subsequent admittance control. Prioritizing grasp triplets with higher orthogonality helps minimize the impact of grasp direction deviations. We will discuss this aspect in more detail in the admittance control section.

After identifying the approximately orthogonal groups, we leverage the two arms of a bimanual robot to perform regrasp planning interactively for the three grasp groups. We use an incremental method to explore combinations of grasp triplets from the orthogonal groups and output a successful motion once found. Each triplet comprises three groups of grasp poses. During planning, one pose from each group needs to be selected to form a combination of three grasps for computation. In this approach, we do not explicitly extract the three-grasp combinations as a separate step. Instead, we leverage the two arms of a bimanual robot to perform regrasp planning interactively for the three grasp groups. The grasps in each group are sequentially assigned to the two robotic arms for planning. In this way, we incrementally explore the combinations of the grasp triplets and output successful motion once found. The detailed algorithmic flow is shown in Fig. 3.

It should be noted that our grasping and planning methods are probabilistic, and their success rates inherently depend on the granularity, number, and diversity of the grasp samples. At the same time, increasing the number of grasps inevitably impacts planning time. Identifying the optimal trade-off between these factors is an important challenge that could be addressed using predictive approaches. If a large amount of data

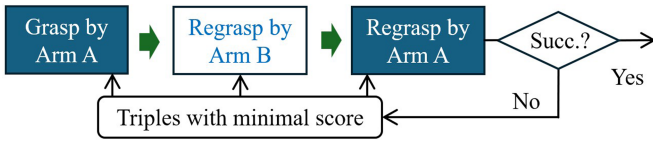


Fig. 3. Regrasp planning used in this study is a simple incremental search method, which sequentially traverses grouped grasp triplets until a valid solution is found.

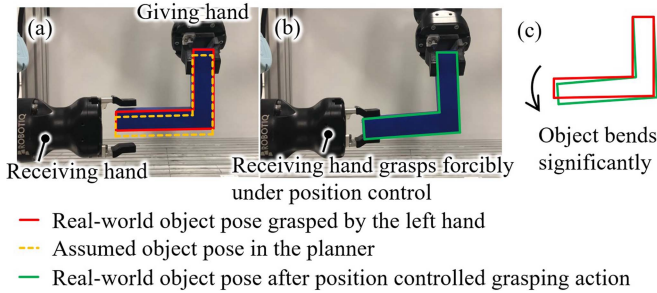


Fig. 4. (a) The receiving hand approaches to grasp the object. Due to uncertainties, the object is not positioned exactly at the center of the two fingers, as assumed, but is instead slightly shifted upward. (b) At this point, if position control is used, the upper-side finger will make contact with the object first, exerting excessive force on it and causing deformation. (c) Comparison of object deformation before and after grasping.

can be accumulated using the proposed method in the future, we are interested in exploring probabilistic modeling methods to learn grasp triplets and thus improve planning efficiency. However, this is beyond the current scope and will not be elaborated.

IV. CONFORMING TO OBJECT POSE UNCERTAINTY

After planning a feasible sequence of bimanual robot motions, the next step is execution. However, due to uncertainties in the object's pose, a significant discrepancy may exist between the real-world pose and the planning simulation, making pure position control infeasible. This subsection discusses this limitation and highlights the necessity of conforming to object pose uncertainty. It then introduces the admittance control method and explains how it enables the system to adapt to uncertain object poses during handover regrasping.

Consider the scenario shown in Fig. 4(a), where the giving hand has already moved the object from its initial pose to the handover pose, and the receiving hand is approaching to perform the regrasp. The target pose of the giving hand is planned based on an assumed pose in the simulation. However, due to uncertainties in the real world, the object's assumed pose in the simulation deviates from its actual pose, as illustrated by the orange dashed frames and red solid frames in Fig. 4(a). Consequently, the receiving hand's grasping pose becomes misaligned with the actual object pose. Applying position control for grasping in this scenario may lead to undesirable outcomes. In mild cases, as shown in Fig. 4(b) and (c), the object may undergo deformation. In more severe cases, it may sustain damage. Therefore, rigidly executing the planned grasping motion solely with position control is not advisable.

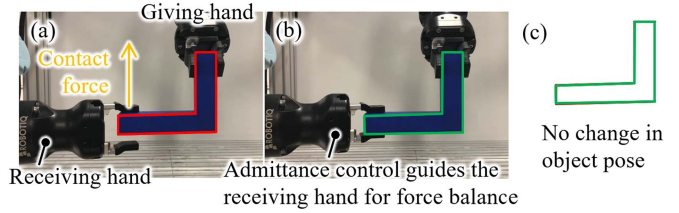


Fig. 5. Admittance control helps avoid deformation or damage. (a) A reaction force f is detected when the upper-side finger contacts the object. (b) Admittance control adjusts the receiving hand by moving it to balance f . (c) After admittance control, the object is firmly grasped without change.

To address this issue, this study employs admittance control during robotic grasping and enables the receiving hand's grasping pose to conform to the uncertain object pose. Admittance control establishes a mass-damper-spring system's motion equation based on the reaction force exerted on the receiving hand and the desired target force [6]. The receiving hand follows the solution of the equation to adjust its motion accordingly, thereby accommodating reaction force. For a desired receiving hand position x_d , the admittance control equation is

$$M(\ddot{x}_p - \ddot{x}_d) + B(\dot{x}_p - \dot{x}_d) + K(x_p - x_d) = f + f_d. \quad (2)$$

Here, x_p , \dot{x}_p , and \ddot{x}_p are respectively the measured current end-effector position, velocity, and acceleration of the receiving hand. M , B , and K denote the virtual inertia, damping, and stiffness parameters. f is the reaction force applied to the receiving hand when grasping. It can be measured using the force/torque sensor at the wrist of the receiving manipulator. f_d represents the target exerting force of the receiving hand after applying admittance control. It is essentially zero, as our goal is to ensure that the receiving hand does not exert any external force on the object beyond the necessary grasping force, thereby preventing unintended movement, deformation, or damage. Given the robot's current x_p , \dot{x}_p , \ddot{x}_p , f , the desired trajectory x_d , \dot{x}_d , and \ddot{x}_d can be computed from the admittance control equation. These computed values are then fed to the admittance control law to drive the receiving hand to conform to the object's uncertain pose during grasping.

Fig. 5 illustrates the effect of applying admittance control to the grasping process shown in Fig. 4. In Fig. 5(a), the upper-side finger first makes contact with the object and experiences a reaction force f . The admittance control then adjusts the robotic hand by moving it in a direction that balances this force, as shown in Fig. 5(b). The adjustment continues until both fingers are in contact with the object and their grasping forces reach equilibrium. At this point, f disappears, and the grasping forces become internal forces within the whole gripper-object system. The final grasping result is depicted in Fig. 4(c). Unlike position control, admittance control ensures that the object's pose remains unchanged after grasping.

Note that the above equation describes only the position-based admittance control. A similar formulation also applies to rotational admittance control. However, the admittance controller may redirect the grasp toward an adjacent surface due to rotational errors. In extreme cases, this can result in parallel regrasps, making the object's pose indeterminate. As a safeguard

against such situations, we suggest prioritizing near-orthogonal grasps following (1). The design minimizes the likelihood that the three grasps lie on the same plane, even when one of the grasp directions is redirected due to unexpected admittance behavior.

V. POSE ESTIMATION USING CONFORMED GRASP POSES

A. Error Parameterization

We denote the pose of the object and first grasp in the world frame of the simulation as $({}^w\mathbf{p}_o^{\text{sim}}, {}^w\mathbf{R}_o^{\text{sim}})$, and $({}^w\mathbf{p}_{g_1}^{\text{sim}}, {}^w\mathbf{R}_{g_1}^{\text{sim}})$, respectively. The relative pose of the object with respect to the first grasp can be computed as

$$\begin{cases} g_1 \mathbf{R}_o^{\text{sim}} = ({}^w\mathbf{R}_{g_1}^{\text{sim}})^\top {}^w\mathbf{R}_o^{\text{sim}} \\ g_1 \mathbf{p}_o^{\text{sim}} = ({}^w\mathbf{R}_{g_1}^{\text{sim}})^\top ({}^w\mathbf{p}_o^{\text{sim}} - {}^w\mathbf{p}_{g_1}^{\text{sim}}) \end{cases} \quad (3)$$

Meanwhile, we denote the first grasp's actual pose after impedance control as $({}^w\mathbf{p}_{g_1}^{\text{real}}, {}^w\mathbf{R}_{g_1}^{\text{real}})$. Without considering uncertainty, the pose of the object in the world frame of the real world can be computed as

$$\begin{cases} {}^w\mathbf{R}_o^{\text{ideal}}(g_1) = {}^w\mathbf{R}_{g_1}^{\text{real}} \cdot g_1 \mathbf{R}_o^{\text{sim}} \\ {}^w\mathbf{p}_o^{\text{ideal}}(g_1) = {}^w\mathbf{p}_{g_1}^{\text{real}} + {}^w\mathbf{R}_{g_1}^{\text{real}} \cdot g_1 \mathbf{p}_o^{\text{sim}} \end{cases} \quad (4)$$

The subscript (g_1) indicates the computations are based on the first grasp. The superscript "ideal" indicates the results are an estimation without considering uncertainty.

To account for uncertainty introduced during the grasping process, we model the object pose error as comprising a rotational component $\mathbf{R}_o^{\text{err}}(g_1)$ and a translational component $\mathbf{p}_o^{\text{err}}(g_1)$. Similarly, (g_1) indicates that the errors are associated with the first grasp. Due to the geometric constraints between the gripper and the object surface, $\mathbf{R}_o^{\text{err}}(g_1)$ is constrained to a rotation about the second column of the real-world grasp frame, ${}^w\mathbf{r}_{g_1,2}^{\text{real}} = {}^w\mathbf{R}_{g_1}^{\text{real}}[:, 1]$, and $\mathbf{p}_o^{\text{err}}(g_1)$ is restricted to the plane spanned by the first and third columns ${}^w\mathbf{r}_{g_1,1}^{\text{real}} = {}^w\mathbf{R}_{g_1}^{\text{real}}[:, 0]$ and ${}^w\mathbf{r}_{g_1,3}^{\text{real}} = {}^w\mathbf{R}_{g_1}^{\text{real}}[:, 2]$. The three arrows in Fig. 6(a.ii) show these columns. The errors can thus be parameterized as

$$\begin{cases} \mathbf{R}_o^{\text{err}}(g_1) = \exp\left(\theta [{}^w\mathbf{r}_{g_1,2}^{\text{real}}]_\times\right) \\ \mathbf{p}_o^{\text{err}}(g_1) = \delta_{g_1,1} {}^w\mathbf{r}_{g_1,1}^{\text{real}} + \delta_{g_1,3} {}^w\mathbf{r}_{g_1,3}^{\text{real}} \end{cases} \quad (5)$$

Accordingly, the true pose of the object in the real world is

$$\begin{cases} {}^w\mathbf{R}_o^{\text{real}} = {}^w\mathbf{R}_o^{\text{ideal}}(g_1) \cdot \mathbf{R}_o^{\text{err}}(g_1) \\ {}^w\mathbf{p}_o^{\text{real}} = {}^w\mathbf{p}_o^{\text{ideal}}(g_1) + \mathbf{p}_o^{\text{err}}(g_1) \end{cases} \quad (6)$$

B. Rotation Error Identification

We can use the second grasp to identify $\mathbf{R}_o^{\text{err}}(g_1)$. Let $({}^w\mathbf{p}_{g_2}^{\text{sim}}, {}^w\mathbf{R}_{g_2}^{\text{sim}})$ denote the simulated pose of the second grasp (equally the grasp pose before admittance control), and let $({}^w\mathbf{p}_{g_2}^{\text{real}}, {}^w\mathbf{R}_{g_2}^{\text{real}})$ denote its real-world pose after admittance control. The rotation difference from the simulation to the real execution is $\mathbf{R}_{g_2}^{\text{diff}} = ({}^w\mathbf{R}_{g_2}^{\text{sim}})^\top \cdot {}^w\mathbf{R}_{g_2}^{\text{real}}$, which captures the rotational deviation introduced by the admittance control. Since $\mathbf{R}_o^{\text{err}}(g_1)$ is constrained to only occur around $\mathbf{r}_{g_1,2}^{\text{real}}$, we extract the component of $\mathbf{R}_{g_2}^{\text{diff}}$ along $\mathbf{r}_{g_1,2}^{\text{real}}$ to get θ :

$$\theta = \text{axis_angle}(\mathbf{R}_{g_2}^{\text{diff}}, {}^w\mathbf{r}_{g_1,2}^{\text{real}}), \quad (7)$$

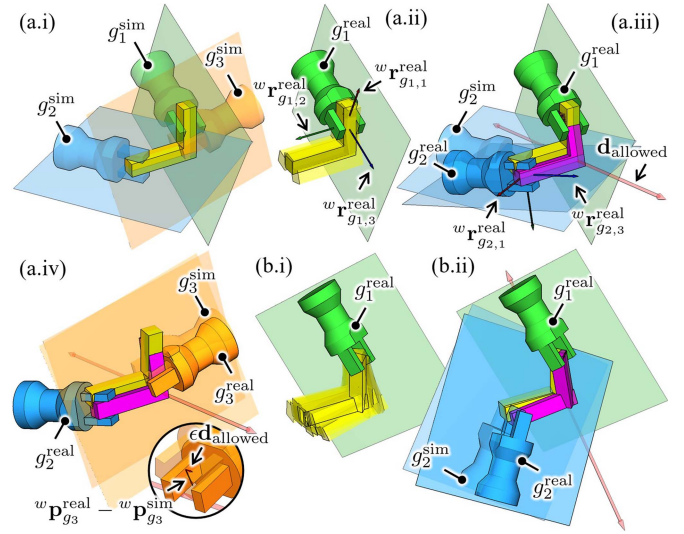


Fig. 6. (a) Three grasps of an L-shaped object with a square cross-section (i) and the corresponding constrained uncertainties (ii ~ iv). (b) The case of an L-shaped object with a diamond cross-section. The three grasps are not strictly orthogonal. Uncertainty can be estimated as long as they are non-parallel.

and consequently $\mathbf{R}_o^{\text{err}}(g_1)$. The object's true rotation in the real world can be uniquely identified as

$$\begin{aligned} {}^w\mathbf{R}_o^{\text{real}} &= {}^w\mathbf{R}_o^{\text{ideal}}(g_1) \cdot \mathbf{R}_o^{\text{err}}(g_1) \\ &= {}^w\mathbf{R}_o^{\text{ideal}}(g_1) \cdot \exp\left(\theta [{}^w\mathbf{r}_{g_1,2}^{\text{real}}]_\times\right). \end{aligned} \quad (8)$$

The rotation of the purple object in Fig. 6(a.iii) illustrates an identified ${}^w\mathbf{R}_o^{\text{real}}$.

C. Translation Error Identification

Next, we address $\mathbf{p}_o^{\text{err}}(g_1)$. In simulation, the pose of the object relative to the second grasp is

$$\begin{cases} g_2 \mathbf{R}_o^{\text{sim}} = ({}^w\mathbf{R}_{g_2}^{\text{sim}})^\top {}^w\mathbf{R}_o^{\text{sim}} \\ g_2 \mathbf{p}_o^{\text{sim}} = ({}^w\mathbf{R}_{g_2}^{\text{sim}})^\top ({}^w\mathbf{p}_o^{\text{sim}} - {}^w\mathbf{p}_{g_2}^{\text{sim}}) \end{cases} \quad (9)$$

Let $({}^w\mathbf{p}_{g_2}^{\text{real}}, {}^w\mathbf{R}_{g_2}^{\text{real}})$ be the second grasp's real-world pose. If there is no translational uncertainty, the position of the object in the real world can be computed using the second grasp as

$${}^w\mathbf{p}_o^{\text{ideal}}(g_2) = {}^w\mathbf{p}_{g_2}^{\text{real}} + {}^w\mathbf{R}_{g_2}^{\text{real}} g_2 \mathbf{p}_o^{\text{sim}}. \quad (10)$$

Since ${}^w\mathbf{R}_o^{\text{real}}$ has already been identified, the corrected relative position $g_2 \mathbf{p}_o^{\text{real}}$ instead of $g_2 \mathbf{p}_o^{\text{sim}}$ is used in the equation. $g_2 \mathbf{p}_o^{\text{real}}$ is computed as

$$g_2 \mathbf{p}_o^{\text{real}} = g_2 \mathbf{R}_o^{\text{real}} (g_2 \mathbf{R}_o^{\text{sim}})^\top g_2 \mathbf{p}_o^{\text{sim}}, \quad (11)$$

where

$$g_2 \mathbf{R}_o^{\text{real}} = ({}^w\mathbf{R}_{g_2}^{\text{real}})^\top {}^w\mathbf{R}_o^{\text{real}}. \quad (12)$$

The translational uncertainty caused by the second grasp lies within the plane spanned by the first and third columns of ${}^w\mathbf{R}_{g_2}^{\text{real}}$ and can be parameterized as

$$\mathbf{p}_o^{\text{err}}(g_2) = \delta_{2,1} {}^w\mathbf{r}_{g_2,1}^{\text{real}} + \delta_{2,3} {}^w\mathbf{r}_{g_2,3}^{\text{real}}, \quad (13)$$

where $w_{g_{2,1}}^{\mathbf{r}^{\text{real}}} = w_{g_2}^{\mathbf{r}^{\text{real}}}[:, 0]$ and $w_{g_{2,3}}^{\mathbf{r}^{\text{real}}} = w_{g_2}^{\mathbf{r}^{\text{real}}}[:, 2]$. Fig. 6(a.iii) illustrates them. The true position of the object can be expressed using $\mathbf{p}^{\text{err}}(g_2)$ as

$$w_{\mathbf{p}_o}^{\text{real}} = w_{\mathbf{p}_o}^{\text{ideal}}(g_2) + \mathbf{p}_o^{\text{err}}(g_2). \quad (14)$$

It is important to note that the translational errors associated with the first and second grasps are respectively constrained to planes determined by $(w_{g_{1,1}}^{\mathbf{r}^{\text{real}}}, w_{g_{1,3}}^{\mathbf{r}^{\text{real}}})$ and $(w_{g_{2,1}}^{\mathbf{r}^{\text{real}}}, w_{g_{2,3}}^{\mathbf{r}^{\text{real}}})$. Their normal vectors are $\mathbf{n}_{g_1} = w_{g_{1,1}}^{\mathbf{r}^{\text{real}}} \times w_{g_{1,3}}^{\mathbf{r}^{\text{real}}}$ and $\mathbf{n}_{g_2} = w_{g_{2,1}}^{\mathbf{r}^{\text{real}}} \times w_{g_{2,3}}^{\mathbf{r}^{\text{real}}}$. Since the translational error must simultaneously satisfy the constraints imposed by both planes, it is restricted to lie along their intersection line. The direction of this line can be computed using

$$\mathbf{d}_{\text{allowed}} = \frac{\mathbf{n}_{g_1} \times \mathbf{n}_{g_2}}{\|\mathbf{n}_{g_1} \times \mathbf{n}_{g_2}\|}. \quad (15)$$

Thus, the translational error can essentially be parameterized using a single scalar value as $\epsilon \mathbf{d}_{\text{allowed}}$, and the true real-world object position simplifies to

$$w_{\mathbf{p}_o}^{\text{real}} = w_{\mathbf{p}_o}^{\text{ideal}}(g_2) + \epsilon \mathbf{d}_{\text{allowed}}. \quad (16)$$

The scalar value ϵ can be easily determined from the displacement of the third grasp after admittance control. Let $w_{\mathbf{p}_{g_3}}^{\text{sim}}$ and $w_{\mathbf{p}_{g_3}}^{\text{real}}$ denote the simulated position of the third grasp (position before admittance control) and the real-world position after admittance control, respectively. The displacement is computed as $w_{\mathbf{p}_{g_3}}^{\text{real}} - w_{\mathbf{p}_{g_3}}^{\text{sim}}$. The scalar ϵ can be obtained using the following projection

$$\epsilon = \mathbf{d}_{\text{allowed}}^\top (w_{\mathbf{p}_{g_3}}^{\text{real}} - w_{\mathbf{p}_{g_3}}^{\text{sim}}), \quad (17)$$

as indicated within the circle in Fig. 6(a.iv). Substituting the obtained ϵ into (16) yields the object's true position.

In addition to the L-shaped object with a square cross-section discussed thus far, Fig. 6(b) presents an L-shaped object with a diamond cross-section. Although this object does not contain any pair of strictly orthogonal grasping directions, it is still possible to select three non-parallel grasps and estimate the final object pose using the method described in this section. Fig. 6(b.i) shows the uncertainty after the first grasp, while Fig. 6(b.ii) illustrates the orientation determined after the second grasp, along with the direction of the remaining translational error. A detailed 3D visualization of both the L-shaped objects can be found in the supplementary video.

VI. EXPERIMENTS AND ANALYSIS

Our experimental platform consists of two UR3e robots. Their setup is shown in Fig. 7. Each robot is installed with a Robotiq Hand-E gripper. The opening range of the gripper is 0~50 mm. The regrasp planning is performed using the WRS simulation environment. The coordinate system of the platform is defined at the back of the robot, with its origin at the intersection of the central plane between the two robot arms and the ground. The x -axis points toward the monitor in the figure. The y -axis points towards the first arm. The z -axis points upward. For comparison, we attached optical markers to both the objects and the robots, and tracked the poses of the resulting rigid

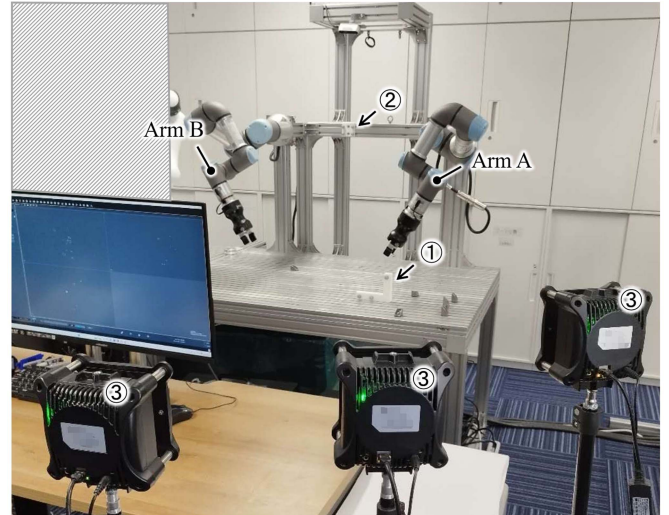


Fig. 7. Bimanual platform used in the experiments. ① Target object with optical markers. ② Markers for global reference. ③ Cameras used for detecting optical markers.

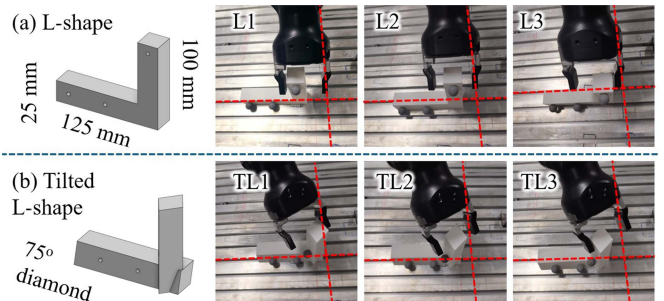


Fig. 8. (a) L-shape object and its random initial placements. (b) TL-shape object and its random initial placements.

bodies using three tracking cameras. The poses estimated by the proposed method were compared with those captured by the optical tracking system to evaluate the method's performance. The tracking system provides a detection accuracy of ± 0.5 mm.

Fig. 8 shows the objects used in the experiments, their dimensions, and the randomly assigned placements in experimental trials. The objects are L-shaped items with square and diamond cross-sections, as described earlier. Their long arm measures 125 mm, the short arm 100 mm, and the distance between the parallel edges of the cross-section is 25 mm for both. The diamond-shaped cross-section has a 75° acute interior angle. Optical tracking markers were attached at pre-drilled holes on the objects. During grasp planning, we can automatically generate grasps and motions that avoid collisions with the markers. For each object, we randomly assigned three different initial placements, labeled L1~3 and TL1~3 in Fig. 8, and tasked the robots with moving the objects to predefined target positions. The supplementary video accompanying this letter contains details of the grasp selection, motion planning, and robot actions. We recommend readers to refer to this video for an intuitive understanding of the experimental procedures.

Table I shows the results. The data section has two main column groups reporting the positional differences and rotational

TABLE I
EXPERIMENTAL RESULTS WITH STATISTICS

	Position Differences ($\Delta\mathbf{p}$, mm)									Rotation Vector Differences ($\Delta\mathbf{w}$, degrees)								
	second grasp			third grasp			object			second grasp			third grasp			object		
L1	0.40	-1.60	5.20	-8.90	2.00	-0.80	-33.70	-61.80	32.20	0.55	0.91	0.02	0.34	0.75	-0.91	0.01	2.82	-1.39
L2	0.30	-1.70	5.20	0.20	2.00	-0.80	-33.60	-61.70	32.30	0.51	0.14	0.03	0.33	-0.25	0.06	0.12	1.83	-1.95
L3	0.40	-1.60	5.20	9.00	1.90	-0.40	-33.20	-61.50	31.70	0.30	0.13	0.00	0.20	-0.51	1.11	-0.00	0.98	-1.07
Mean	-	-	-	-	-	-	-33.50	-61.67	32.07	-	-	-	-	-	-	0.04	1.87	-1.47
Std	-	-	-	-	-	-	0.25	0.15	0.25	-	-	-	-	-	-	0.06	0.95	0.44
TL1	-1.81	0.99	0.38	8.02	8.05	-7.20	-59.35	32.75	52.41	0.64	1.07	-1.00	0.45	0.36	0.78	-0.01	2.05	0.42
TL2	-0.08	-0.18	1.31	0.40	-1.60	5.20	-59.66	34.42	51.76	-0.00	-0.00	0.00	0.30	0.13	0.00	0.13	0.73	1.09
TL3	0.07	-0.21	1.17	-1.71	-0.36	-1.47	-56.37	36.34	52.26	-0.04	-0.01	-0.00	1.25	-0.29	0.56	0.86	0.27	2.11
Mean	-	-	-	-	-	-	-58.46	34.50	52.14	-	-	-	-	-	-	0.33	1.02	1.21
Std	-	-	-	-	-	-	1.86	1.50	0.36	-	-	-	-	-	-	0.47	0.38	0.85

differences. Within each group, there are three subcolumns. The “second grasp” columns show the positional and rotational differences of the g_2 grasp before and after admittance control. The “third grasp” columns present the corresponding differences for the g_3 grasp. The “object” columns report the differences between the object pose estimated by the proposed method and the pose measured by the optical tracking system. The “first grasp” uses position control and is thus excluded.

From the table, we observe that both the g_2 and g_3 grasps exhibit substantial changes before and after admittance control, indicating that the robot actively adjusts the regrasp pose to estimate uncertainties for each task. However, the final estimation results are not satisfactory, showing positional differences of up to 61.67mm and rotational differences of up to 1.86 degrees compared to the optical tracking system.

Upon closer examination, we found that these errors are primarily systematic. For the L-shaped object, the standard deviations in $\Delta\mathbf{p}$ across tasks were [0.25, 0.15, 0.25] mm, all below 0.3mm, and the deviations in $\Delta\mathbf{w}$ were [0.06, 0.95, 0.44] degrees, all below 1.0 degrees. Similarly, for the TL-shaped object, the deviations in $\Delta\mathbf{p}$ and $\Delta\mathbf{w}$ were [1.86, 1.50, 0.36] mm and [0.47, 0.38, 0.85] degrees, respectively, remaining below 2 mm and 0.9 degrees. These results suggest that while the overall estimation errors appear large, their internal consistency is high, indicating that the system’s inaccuracies stem largely from systematic biases. Likely sources include calibration errors in the optical tracking system, installation errors in the two robot arms and the mounting aluminum frame, and absolute control errors of the robots. In addition, when comparing the deviations between the L-shaped and TL-shaped objects, we found that the inference results for the L-shaped object were better than those for the TL-shaped object. This outcome differs from our theoretical expectations. We can conclude from the equations in Section V that even if the three grasps are not perfectly orthogonal, the final inference accuracy should not degrade, and the L- and TL-shaped objects should exhibit similar deviations. We believe that the observed difference arises because the TL-shaped object involves rotations around axes unaligned with body frames. The large difference indicates that the robot setup has greater rotational misalignment. This raises the question of compensating for system bias to improve absolute accuracy. While theoretically feasible, such correction would

require complex estimation of both positional and rotational errors, which we leave for future work.

In conclusion, although the results exhibit large absolute errors, the deviation levels are on the same order of magnitude as the optical tracking system. The results indicate that the proposed method has strong relative pose estimation performance.

VII. CONCLUSIONS AND FUTURE WORK

This letter proposes a method for reducing object uncertainties by leveraging the grasping constraints of parallel grippers with flat finger pads and a bimanual regrasp planning method. We applied this method to estimate grasping errors for a roughly placed L-shaped object. The results demonstrate that the proposed approach can effectively reduce uncertainties and achieves high repeatability accuracy.

In our experiments, we used a dual-arm setup built from two single-arm robots mounted on an aluminum frame, which inevitably introduced systematic errors and led to relatively large absolute errors. As a next step, we plan to validate the method in a dynamic simulation environment, where no absolute error is present, and use a dual-arm robot with commercially-calibrated body integration for practical deployment.

REFERENCES

- [1] G. Du, K. Wang, S. Lian, and K. Zhao, “Vision-based robotic grasping from object localization, object pose estimation to grasp estimation for parallel grippers: A review,” *Artif. Intell. Rev.*, vol. 54, no. 3, pp. 1677–1734, 2021.
- [2] Y. Yan et al., “Soft magnetic skin for super-resolution tactile sensing with force self-decoupling,” *Sci. Robot.*, vol. 6, no. 51, 2021, Art. no. eabc8801.
- [3] N. C. -Dafle and A. Rodriguez, “Prehensile pushing: In-hand manipulation with push-primitives,” in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst.*, 2015, pp. 6215–6222.
- [4] F. Fiedler et al., “Jigs and fixtures in production: A systematic literature review,” *J. Manuf. Syst.*, vol. 72, pp. 373–405, 2024.
- [5] W. Wan, M. T. Mason, R. Fukui, and Y. Kuniyoshi, “Improving regrasp algorithms to analyze the utility of work surfaces in a workcell,” in *Proc. IEEE Int. Conf. Robot. Automat.*, 2015, pp. 4326–4333.
- [6] R. Mae and R. Kikuuwe, “An admittance controller with a jerk limiter for position-controlled robots,” *J. Robot. Mechatron.*, vol. 36, no. 2, pp. 483–493, 2024.
- [7] W. Liu, M. Ren, K. Song, M. Y. Wang, and Z. Xiong, “A planning framework for complex flipping manipulation of multiple mobile manipulators,” *IEEE Robot. Automat. Lett.*, vol. 10, no. 5, pp. 5162–5169, May 2025.

- [8] J. Kim, I. Ko, and F. C. Park, "Randomized path planning for tasks requiring the release and regrasp of objects," *Adv. Robot.*, vol. 30, no. 4, pp. 270–283, 2016.
- [9] I. Baek, K. Shin, H. Kim, S. Hwang, E. Demeester, and M. -S. Kang, "Pre-grasp manipulation planning to secure space for power grasping," *IEEE Access*, vol. 9, pp. 157715–157726, 2021.
- [10] Y. Qin, A. Escande, F. Kanehiro, and E. Yoshida, "Dual-arm mobile manipulation planning of a long deformable object in industrial installation," *IEEE Robot. Automat. Lett.*, vol. 8, no. 5, pp. 3039–3046, May 2023.
- [11] P. Mitrano and D. Berenson, "The grasp loop signature: A topological representation for manipulation planning with ropes and cables," in *Proc. IEEE Int. Conf. Robot. Automat.*, 2024, pp. 10888–10894.
- [12] B. Sundaralingam and T. Hermans, "Geometric in-hand regrasp planning: Alternating optimization of finger gaits and in-grasp manipulation," in *Proc. IEEE Int. Conf. Robot. Automat.*, 2018, pp. 231–238.
- [13] M. Murooka, I. Kumagai, M. Morisawa, F. Kanehiro, and A. Kheddar, "Humanoid loco-manipulation planning based on graph search and reachability maps," *IEEE Robot. Automat. Lett.*, vol. 6, no. 2, pp. 1840–1847, Apr. 2021.
- [14] M. A. Roa and R. Suarez, "Regrasp planning in the grasp space using independent regions," in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst.*, 2009, pp. 1823–1829.
- [15] J. Hu, Z. Tang, and H. I. Christensen, "Multi-modal planning on regrasping for stable manipulation," in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst.*, 2023, pp. 10620–10627.
- [16] X. Cheng, E. Huang, Y. Hou, and M. T. Mason, "Contact mode guided motion planning for quasidynamic dexterous manipulation in 3D," in *Proc. 2022 Int. Conf. Robot. Automat.*, 2022, pp. 2730–2736.
- [17] M. Mason, "The mechanics of manipulation," in *Proc. IEEE Int. Conf. Robot. Automat.*, 1985, pp. 544–548.
- [18] K. Cho, M. Kim, and J. -B. Song, "Complete and rapid regrasp planning with look-up table," *J. Intell. Robotic Syst.*, vol. 36, pp. 371–387, 2003.
- [19] Z. -Y. Chiu, F. Richter, E. K. Funk, R. K. Orosco, and M. C. Yip, "Bimanual regrasping for suture needles using reinforcement learning for rapid motion planning," in *Proc. IEEE Int. Conf. Robot. Automat.*, 2021, pp. 7737–7743.
- [20] K. Wada, S. James, and A. J. Davison, "Reorientbot: Learning object reorientation for specific-posed placement," in *Proc. IEEE Int. Conf. Robot. Automat.*, 2022, pp. 8252–8258.
- [21] K. Xu, H. Yu, R. Huang, D. Guo, Y. Wang, and R. Xiong, "Efficient object manipulation to an arbitrary goal pose: Learning-based anytime prioritized planning," in *Proc. IEEE Int. Conf. Robot. Automat.*, 2022, pp. 7277–7283.
- [22] M. Bauza, A. Bronars, Y. Hou, I. Taylor, N. Chavan-Dafle, and A. Rodriguez, "Simple, a visuotactile method learned in simulation to precisely pick, localize, regrasp, and place objects," *Sci. Robot.*, vol. 9, no. 91, 2024, Art. no. eadi8808.
- [23] A. Bronars, S. Kim, P. Patre, and A. Rodriguez, "Texterity: Tactile extrinsic dexterity," in *Proc. IEEE Int. Conf. Robot. Automat.*, 2024, pp. 7976–7983.
- [24] M. Gualtieri and R. Platt, "Robotic pick-and-place with uncertain object instance segmentation and shape completion," *IEEE Robot. Automat. Lett.*, vol. 6, no. 2, pp. 1753–1760, Apr. 2021.
- [25] Z. Zhang et al., "Learning to regrasp using visual-tactile representation-based reinforcement learning," *IEEE Trans. Instrum. Meas.*, vol. 73, 2024, Art. no. 3534311.
- [26] N. C. Dafle et al., "Extrinsic dexterity: In-hand manipulation with external forces," in *Proc. IEEE Int. Conf. Robot. Automat.*, 2014, pp. 1578–1583.
- [27] J. Zhou, M. T. Mason, R. Paolini, and D. Bagnell, "A convex polynomial model for planar sliding mechanics: Theory, application, and experimental validation," *Int. J. Robot. Res.*, vol. 37, no. 2/3, pp. 249–265, 2018.
- [28] P. Mannam, A. Volkov Jr., R. Paolini, G. Chirikjian, and M. T. Mason, "Sensorless pose determination using randomized action sequences," *Entropy*, vol. 21, no. 2, 2019, Art. no. 154.
- [29] F. von Drigalski, K. Kasaura, C. C. Beltran-Hernandez, M. Hamaya, K. Tanaka, and T. Matsubara, "Uncertainty-aware manipulation planning using gravity and environment geometry," *IEEE Robot. Automat. Lett.*, vol. 7, no. 4, pp. 11942–11949, Oct. 2022.
- [30] J. Bohg et al., "Interactive perception: Leveraging action in perception and perception in action," *IEEE Trans. Robot.*, vol. 33, no. 6, pp. 1273–1291, Dec. 2017.
- [31] J. Jankowski, L. Bruder Müller, N. Hawes, and S. Calinon, "Robust pushing: Exploiting quasi-static belief dynamics and contact-informed optimization," *Int. J. Robot. Res.*, early access, 2025, doi: [10.1177/02783649251318046](https://doi.org/10.1177/02783649251318046).
- [32] Y. Dong, X. Cheng, and F. T. Pokorny, "Characterizing manipulation robustness through energy margin and caging analysis," *IEEE Robot. Automat. Lett.*, vol. 9, no. 9, pp. 7525–7532, Sep. 2024.
- [33] I. -M. Chen and J. W. Burdick, "Finding antipodal point grasps on irregularly shaped objects," *IEEE Trans. Robot. Automat.*, vol. 9, no. 4, pp. 507–512, Aug. 1993.