Title	Dig-Grasping via Direct Quasistatic Interaction Using Asymmetric Fingers: An Approach to Effective Bin Picking
Authors	Tong, Zhekai; Ng, Yu Hin; Kim, Chung Hee; He, Tierui; Seo, Jungwon
Source	IEEE Robotics and Automation Letters, v. 6, (2), April 2021, article number 9363618, p. 3033-3040
Version	Accepted Version
DOI	10.1109/lra.2021.3062250
Publisher	IEEE
Copyright	© 2021 IEEE. Personal use of this material is permitted. Permission from IEEE must be obtained for all other uses, in any current or future media, including reprinting/republishing this material for advertising or promotional purposes, creating new collective works, for resale or redistribution to servers or lists, or reuse of any copyrighted component of this work in other works.

This is the author's version of an article that has been published in this journal. Changes were made to this version by the publisher prior to publication.
The final version of record is available at http://dx.doi.org/10.1109/LRA.2021.3062250

Dig-Grasping via Direct Quasistatic Interaction Using Asymmetric Fingers: An Approach to Effective Bin Picking

Zhekai Tong, Yu Hin Ng, Chung Hee Kim, Tierui He, and Jungwon Seo

***Abstract*—This letter introduces a new method for simultaneously singulating and picking objects from clutter. The method can lead to effective robotic bin picking, which still remains elusive despite its importance in many industrial and domestic applications, especially for objects with a thin profile. We leverage planar quasistatic pushing manipulation as a way of standardized physical interaction between a robot and the object to pick. A gripper designed with digit asymmetry, realized as a two-fingered gripper with different finger lengths, is suggested as the key to successful singulating and picking through the controlled pushing maneuver. A detailed account of the manipulation process and design principles will be presented. An extensive set of experiments validate the effectiveness of our approach in three-dimensional bin picking tasks. Beyond picking, more complex manipulation capabilities such as autonomous pick-and-place/pack will also be presented.**

***Index Terms*—Grasping, Grippers and Other End-Effectors**

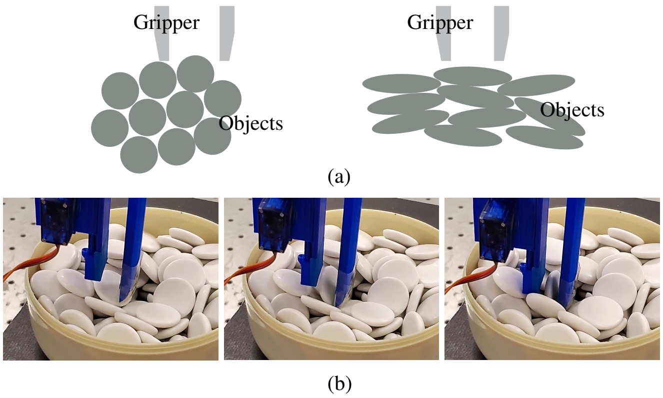

Fig. 1: (a) Planar bin picking: circular disks vs. elliptical objects. (b) Our dig-grasping manipulation applied to the bin picking of Go stones.

I. Introduction

CONSIDER common planar bin picking scenarios where the goal is to pick items one after the other out of clutter. See Fig. 1(a). Given exceptional, well-rounded object shapes such as a circular disk, this task of simultaneous singulating and picking might not come across as a difficult problem. It can be rather straightforward to devise a winning manipulation strategy that can address the clutter; for example, one can simply consider setting the gripper aperture to the diameter of the disk (or slightly larger than that) and moving the gripper directly towards an object instance in an effort to capture it between the fingers. For other shapes with a possibly large width-to-thickness ratio, such as elliptical objects, there may even be a multitude of options to set the gripper aperture, let alone methods of manipulation. Large (small) gripper aperture will contribute to picking (singulating) but will complicate singulating (picking): a trade-off.

This paper presents a robotic manipulation framework that we call *dig-grasping*. Taking advantage of quasistatic pushing and novel gripper design, dig-grasping addresses the challenges of simultaneous singulating and picking in the bin picking of relatively thin objects, with a large width-to-thickness ratio. We show that direct physical interaction with the object to pick that happens while the gripper "digs" the clutter, realized as controlled pushing, is important to successfully capturing it from the clutter, unlike many traditional approaches in which direct contact interaction is supposed to be avoided. A gripper designed with digit asymmetry plays a critical role in making the idea of grasping through digging possible. Fig. 1(b) illustrates the progress of dig-grasping using a two-fingered gripper with digit asymmetry, that is, different finger lengths. The robot pushes the Go stone to pick, with the longer finger. The stone is then expected to be funnelled into the gap between the fingers according to the mechanics of pushing. By having the other finger shorter, unwanted collision between the stone and the gripper is avoided in the meantime, while other stones are refused to enter the gripper's workspace. Overall, this looks similar to the way a human Go player would singulate and pick a stone from a cluttered bowl using the index and the middle finger.

As previewed in Fig. 1(b), our approach is validated through challenging and practical three-dimensional picking tasks, with a minimalist hardware setting featuring a two-fingered gripper. Moreover, we will also demonstrate more

Manuscript received: October, 15, 2020; Revised January, 12, 2021; Accepted February, 7, 2021.

This paper was recommended for publication by Editor Hong Liu upon evaluation of the Associate Editor and Reviewers' comments. This work was supported by Hong Kong Innovation and Technology Fund ITS/240/17FX, ITS/018/17FP, and ITS/104/19FP.

Z. Tong, Y. H. Ng, C. H. Kim, T. He, and J. Seo are with The Hong Kong University of Science and Technology (HKUST), Clear Water Bay, Hong Kong {ztong, yhngad, chkimaa, theae}@connect.ust.hk and junseo@ust.hk.

This letter has supplemental downloadable multimedia material available at http://ieeexplore.ieee.org, provided by the authors. This includes one multimedia MP4 format movie clip, showing detailed dig-grasping operations on various bin picking scenarios. This material is 19.6MB in size.

Digital Object Identifier (DOI): see top of this page.

complex manipulation tasks contingent on successful picking: autonomous pick-and-place/pack.

II. Related Work

Pushing is a fundamental form of manipulation and has been a topic of great interest in the recent developments in robotic manipulation. See [1] for an overview. The mechanics and planning of quasistatic planar pushing was first addressed in [2]. In [3], building on [4], an edge-to-edge contact between a pusher and a slider was leveraged to execute stable pushing motions. More recently, an approach based on a convex polynomial representation was applied to refine the mechanics of sliding and pushing [5]. In [6], planar pushing under the influence of gravity was examined and an algorithm for planning in-hand manipulation by pushing was presented. Data-driven approaches have also proved useful in pushing manipulation [7]. A wide variety of object handling tasks can be facilitated contingent on successful pushing. For example, pre-grasp manipulation realized as planar pushing is applied to robustifying object grasping [8], [9]. Our presented work here confirms this point too. Another example related to our work can be seen in the literature on object singulation [10] and picking through quasistatic regrasping [11].

The practical applications of the presented work include bin picking—picking items one by one from a cluttered bin. It is an important material handling capability that has received considerable attention (including our recent effort [12]), but still remains a great challenge. Recently, bin picking solutions presented in high-profile competitions, such as Amazon Picking Challenge, have received considerable attention [13]. Gripper hardware configuration is critical to bin picking. These competitions reconfirmed the usefulness of the common suction gripper, the two-fingered gripper, or the combination of both [14]. Another critical issue in bin picking is how to detect items and their pose. A range of data-driven, learning-based image processing techniques, instance/semantic segmentation [15] for example, have recently been applied to the problem of item detection. The pose of an item is usually resolved by matching its RGB-D image and solid model, through point cloud registration [16]. Grasp planning for bin picking is concerned with the determination of the relative pose between a gripper and an item that will result in successful picking. During the last decade, learning-based image processing approaches have successfully been applied to the planning problem [17], [18]. Based on a vast amount of three-dimensional object models, a series of works presented in [19] and its earlier versions feature a grasp planning approach applicable to multiple gripper options.

III. Preliminaries: Planar Pushing

This section reviews the planar motion of a rigid object subjected to a quasistatic push by a motion-controlled point pusher, on a rigid flat surface. Fig. 2(a) shows the setting.

The concept of the *limit surface* [1] models the relationship between the quasistatic motion of the slider (that is, the object being pushed) and the resulting frictional wrench imparted by the support surface. The limit surface is the boundary of the set of all the possible frictional wrenches, and is represented as a convex, closed surface in the three-dimensional wrench space related to planar sliding. In case the slider is stationary, any wrench contained in the set is realizable. When slip occurs, the frictional wrench lies on the limit surface and the twist of the slider is perpendicular to the surface at the point of the wrench. As discussed in [20], the limit surface is often approximated as an ellipsoid in the three-dimensional wrench space, $f_x f_y m_z$-space:

$$\left(\frac{f_x}{f_{\max}}\right)^2 + \left(\frac{f_y}{f_{\max}}\right)^2 + \left(\frac{m_z}{m_{\max}}\right)^2 = 1$$

where $f_{\max}$ ($m_{\max}$) is the maximum frictional force (moment) that occurs when the slider is in pure translation (in pure rotation about its center of mass).

The resulting quasistatic motion of the slider induced by the motion-controlled point pusher can be resolved using the limit surface. Here we summarize the procedure presented in [2], [21]. First, the *motion cone* at the contact between the pusher and the slider is established based on the limit surface. The simple ellipsoid approximation facilitates this. The motion of the slider is then determined by considering whether the contact is fixed or sliding. If the velocity of the pusher is within the motion cone, the contact is fixed and thus the slider moves such that the velocity of the contact point is equal to that of the pusher. Otherwise, the contact is sliding and the friction force lies on an edge of the friction cone. The velocity of the slider is then normal to the limit surface at the point representing the cone edge. The magnitude of the velocity is determined such that the pusher slides in the direction tangent to the slider's boundary at the contact.

We implemented a pushing motion simulator (Fig. 2) based on the ellipsoid limit surface model, with the assumption of uniform pressure distribution between the support surface and the object. It will be used throughout the paper.

IV. Dig-Grasping for Simultaneous Singulating and Picking in Clutter

This section presents our manipulation method for the simultaneous singulating and picking of target objects in a cluttered workspace.

A. Manipulation and Design for Dig-Grasping

In Fig. 2(a), our planar quasistatic pushing simulator is tasked with pushing an elliptical object using a linear finger. We will use the elliptical object model to conceptualize our manipulation technique targeted at relatively thin objects, with a large width-to-thickness ratio. The finger lies along a line denoted ℓ fixed in space and passing through the point of initial contact $\mathbf{p}_0 = (x_0, y_0)$, and is also controlled to move forward along ℓ, the line of pushing. ψ_0 denotes the initial orientation of ℓ in the object frame $\{b\}$. Our objective is to move the object proximally towards the base of the finger (that is, away from the fingertip) as quickly as possible. Pure translation would then be the most undesirable outcome. It

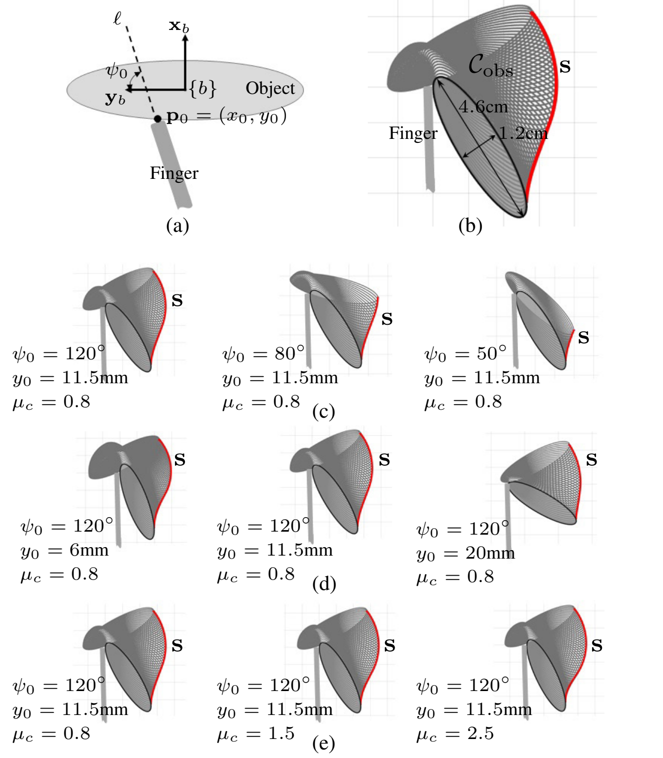

Fig. 2: (a) Finger about to push the elliptical object. $\{b\}$ is a frame attached to the centroid of the object; its axes are aligned with the axes of the ellipse. $\mathbf{p}_0 = (x_0, y_0)$, expressed in $\{b\}$, is the point of initial contact. (b) Evolution of the configuration of the elliptical object as being pushed by the finger, seen from the finger. (c)-(e) Different shapes of $\mathcal{C}_{\text{obs}}$ obtained with the elliptical object in (b) while varying (c) ψ_0, (d) y_0 (x_0 is fixed given y_0 on the ellipse), and (e) μ_c, friction coefficient at the finger-object contact.

is thus necessary to be able to rotate the object, let's say, clockwise without loss of generality. We assume that $\mathbf{p}_0$ and ψ_0 have been chosen to induce the desirable clockwise rotation; the search process here can be executed by directly applying the *voting theorem* [2]. Fig. 2(b) then shows a result, the trace of the object rotating clockwise seen from an observer moving together with the finger. The trace, or the area swept by the object, will be called the obstacle space $\mathcal{C}_{\text{obs}}$, for a reason to be explained soon. The shape of $\mathcal{C}_{\text{obs}}$ and thus part of its boundary $\mathbf{s}$, the collection of the point on the object that is the most distant from ℓ at each instant, change with the direction of pushing ψ_0, the location of the initial contact $\mathbf{p}_0$, and the friction between the object and the finger, as shown in Fig. 2(c-e). It can be seen that the s-shaped curve $\mathbf{s}$ eventually flattens out as the slider (that is, the object) moves out of the pusher's way. $\mathbf{p}_0$ closer to the object center or larger friction at the slider-pusher contact results in increasing the curvature of $\mathbf{s}$.

Now we form a gripper by adding another linear finger that is parallel to and moves together with the original one, as can typically be seen in two-fingered grippers. See Fig. 3(a), where the original (newly added) finger is denoted Finger #1 (#2). The shape of $\mathbf{s}$ suggests the relationship between the gripper's aperture a and finger length difference δ for the successful capturing of the object through quasistatic

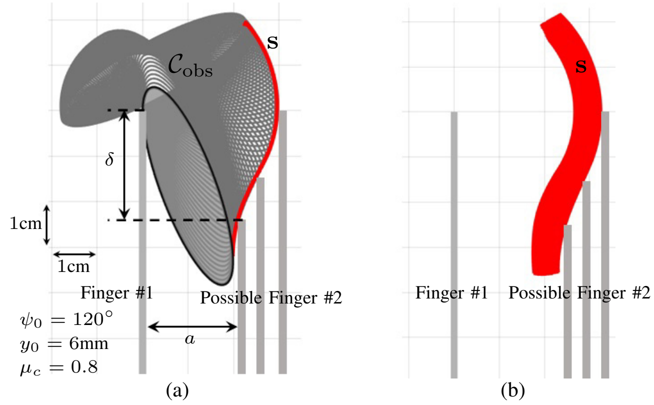

Fig. 3: (a) Possible form factors of a gripper as the configurations of Finger #2 relative to Finger #1. The tip of Finger #2 is situated on $\mathbf{s}$. The size of the elliptical object is the same as the one in Fig. 2(b). (b) Collection of $\mathbf{s}$ where 3mm$< y_0 <$ 9mm and $115° < \psi_0 < 125°$ (extended around the initial condition, y_0 and ψ_0, in (a)), resulting in enlarged $\mathcal{C}_{\text{obs}}$.

pushing. Specifically, $\mathbf{s}$ delimits the space that Finger #2 is not allowed to penetrate, in case a is set to be less than the largest distance from ℓ to a point on $\mathbf{s}$. Otherwise, collision occurs between the object and Finger #2 during the course of pushing by Finger #1. This justifies why the shaded areas in Figs. 2 and 3 were denoted $\mathcal{C}_{\text{obs}}$. Therefore, δ is bounded from below by $\mathbf{s}$: δ needs to be set large enough for Finger #2 not to penetrate into $\mathbf{s}$, given a. Fig. 3(a) shows some of the possible configurations of Finger #2. The gripper designs in Fig. 3(a) are obtained for a single initial condition—a pair of ψ_0 and $\mathbf{p}_0$. By varying the choice of the initial condition and putting the simulation results together, $\mathcal{C}_{\text{obs}}$ is enlarged and we obtain more conservative gripper designs that are also able to address errors in positioning the gripper (Fig. 3(b)).

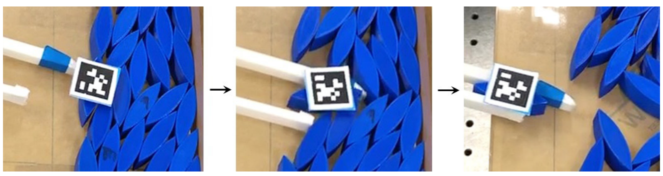

Fig. 4: Dig-grasping for planar bin picking: align (left) - dig (middle) - pinch (right). The marked object instance is being dig-grasped from the clutter.

The gripper design process above naturally suggests a method for capturing an object. We apply a course of manipulation composed of the following three primitive operations *align*, *dig*, and *pinch* to be performed sequentially (Fig. 4):

- *Align*: The gripper is positioned on a target object and oriented along a desired line of pushing. The gripper design process already produces sufficient information on how to perform this operation: the location of the initial contact $\mathbf{p}_0$ and the line of pushing ℓ (for example, the motion parameters ψ_0 and y_0 in Fig. 3).
- *Dig*: The gripper is controlled to translate forward to push the object with Finger #1. The amount of translation, to be referred to as "digging depth," is determined by the pushing simulation as explained in Fig. 5.
- *Pinch*: The gripper closes to pinch-grasp the object.

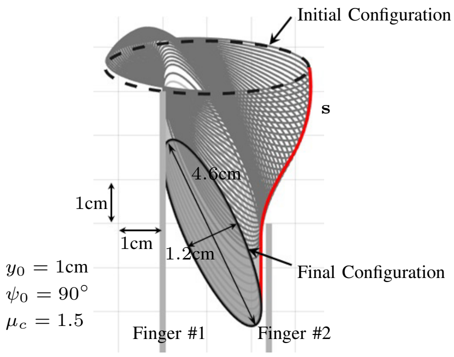

Fig. 5: Trace of the object being pushed by Finger #1 from the initial configuration (see y_0 and ψ_0 above) to the final configuration where the tip of Finger #2 is located around the same position as the object's centroid in the direction of the line of pushing (that is, the line of Finger #1). The digging depth here is defined as the distance that Finger #1 travels.

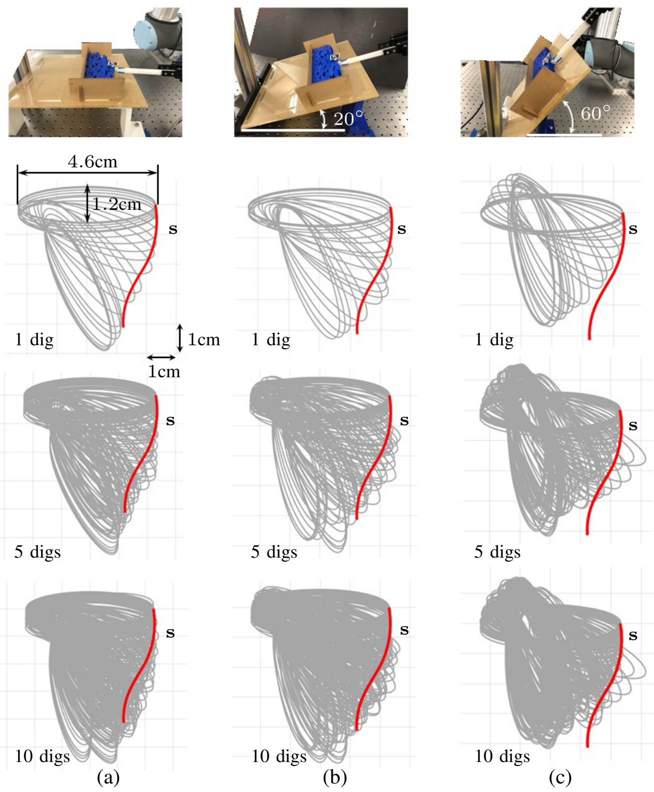

Fig. 6: Reconstructed $\mathcal{C}_{\text{obs}}$ as the sampled traces of an elliptical object collected from digging experiments where the bin is on (a) a level or (b, c) an inclined surface, as can be seen in the panels in the top row (shared motion parameters: $y_0 = 13$mm, $\psi_0 = 90°$, and digging depth 4.0cm). Rows 2-4 show the samples overlaid during one, five, and ten digs, respectively. In simulation, we additionally set $\mu_c = 1.5$ to obtain $\mathbf{s}$.

We call this process *dig-grasping*, considering the resulting "digging" behavior of the gripper in bin picking scenarios, which is our main application, as depicted in Fig. 4. During the align and dig operations, there is no need to control the relative configuration of the fingers. Note, when the finger length difference δ is increased, the digging depth also needs to be increased. This may be undesirable under environmental constraints. In other words, δ may need to be set closer to its lower bound determined by $\mathbf{s}$. In our experiments to follow, Finger #2 is set to be exactly on $\mathbf{s}$.

B. Dig-Grasping in 3D Cluttered Workspaces

We now discuss how dig-grasping, based on the planar pushing of an isolated object, can effectively be applied to simultaneous singulating and picking in a three-dimensional cluttered workspace, which is typical of bin picking tasks, our target application.

First, we take account of a cluttered workspace. Recall the trade-off between singulating and picking in a cluttered space, back in Sec. I. In principle, the inherent effectiveness of dig-grasping in resolving the trade-off can be seen by noting that it is possible to set the gripper's aperture relatively small, which is favorable for singulating, without losing the capability of capturing an object owing to the appropriate finger asymmetry δ that can address $\mathcal{C}_{\text{obs}}$. In addition, $\mathcal{C}_{\text{obs}}$ helps prevent other object instances from entering the space between the fingers. Whether this conceptual effectiveness matters in practice or not depends on the predictive power of the pushing simulation for the actual $\mathcal{C}_{\text{obs}}$ in clutter. Our experiments show that the predictive power can be sufficiently strong. See Fig. 6. It features pushing experiments performed in a cluttered bin; the overall setting is the same as Fig. 4, except that only one finger was used as the pusher. The target object with the fiducial marker is first placed in front of the clutter, and the trace of the object being pushed is recorded. Fig. 6 shows that $\mathbf{s}$ obtained in the pushing simulation is in qualitatively good agreement with empirically obtained $\mathcal{C}_{\text{obs}}$. This can be explained by that the random interactions between objects in dense clutter counteract each other statistically, and thus the actual configuration of the slider evolves similarly to the simulation. Moreover in Fig. 6(c), it can be seen that the presence of gravity contributes to the conservativeness of our analysis (that is, overestimated $\mathcal{C}_{\text{obs}}$ in simulation). Apparently, increased gravity functions similarly to larger friction (Fig. 2(e)). This can be accounted for by increased pressure in the object cluster caused by gravity. We also performed dig-grasping with six different elliptical objects using two-fingered gripper configurations (Fig. 4 was excerpted from this experiment). The results are summarized in Table I. In the treatment group, dig-grasping was performed with the gripper configured according to our approach. In the control group, we tested other gripper designs with larger apertures or no finger length difference. Table I confirms the advantage of our approach with the highest success rate (successes/attempts), where a success is defined as successful singulation and subsequent picking.

Given a three-dimensional object shape, the gripper design for dig-grasping can be determined using a cross-section of the shape as explained in Sec. IV-A. Dig-grasping is then performed by aligning the gripper along the cross-section of an object instance. The effectiveness of dig-grasping in this three-dimensional setting depends on how well the object's motions out of the plane of the cross-section (that is, the differential twists that have a nonzero linear velocity component in the direction normal to the plane or a nonzero

TABLE I: Experiment results showing the effectiveness of our gripper design in simultaneous singulating and picking in a cluttered workspace.

Object Size (cm)	y_0 (cm)	ψ_0 (°)	Digging Depth (cm)	Treatment Group			Control Group								
				a (cm)	δ (cm)	success rate	a (cm)	δ (cm)	success rate	a (cm)	δ (cm)	success rate	a (cm)	δ (cm)	success rate
4.0×1.0	0.9	90	4.0	2.0	2.6	20/25	4.0	2.6	7/20	2.0	0	0/20	4.0	0	8/20
5.0×1.0	0.7	70	4.8	2.0	3.3	19/25	5.0	3.3	10/20	2.0	0	0/20	5.0	0	8/20
4.6×1.2	1.0	90	4.6	2.4	3.0	20/25	4.6	3.0	7/20	2.4	0	0/20	4.6	0	10/20
6.0×1.2	1.0	70	5.7	2.4	4.2	19/25	6.0	4.1	11/20	2.4	0	0/20	6.0	0	10/20
4.5×1.5	1.2	90	4.6	3.0	2.5	22/25	4.5	2.5	6/20	3.0	0	0/20	4.5	0	5/20
6.0×1.5	1.3	70	5.2	3.0	4.0	19/25	6.0	4.0	10/20	3.0	0	0/20	6.0	0	12/20
						119/150			**51/120**			**0/120**			**53/120**

angular velocity component about an axis on the plane) can be suppressed. Feedback control can be one way to attenuate the effects of the unwanted motions, for example, through visual servoing to keep track of the object to pick. However, the cluttered bin setting can render visual situational awareness difficult. Alternatively, we design fingers that benefit from multiple contact wrenches and caging. Specifically, the gripper used for our experiments (see Fig. 9) has fingers protruding from the plane, to be able to make contacts on both sides of the plane, and featuring a fingertip concavity, to be able to cage the object partially (see [22] for a relevant study). In the next section, the applicability and limitation of this design approach will be shown through experiments.

V. Applications of Dig-Grasping: Implementation, Experiments, and Discussion

This section presents our implementation of dig-grasping and its application to bin picking and more complex tasks. We also discuss the range of application of dig-grasping.

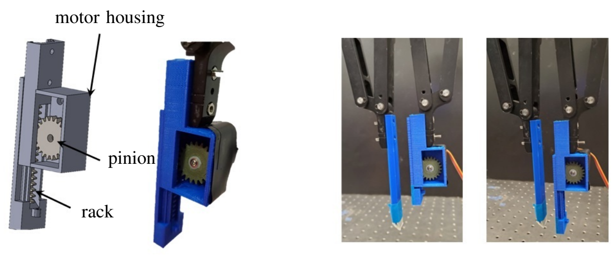

Fig. 7: (a) CAD model (left) and prototype (right) of our variable-length finger module. (b) The finger module in action on a two-fingered gripper, changing relative finger lengths.

A. Our Gripper

Dig-grasping necessitates the capability of changing relative finger lengths. Fig. 7 shows our variable-length finger module fabricated using a rack and pinion actuated by a servo motor, which can be retrofitted into existing grippers. As will be shown with the experiments to follow, the morphology of the fingertip can be customized to suit the design principles presented in Sec. IV-B. The wrist of the gripper accommodates a camera for vision.

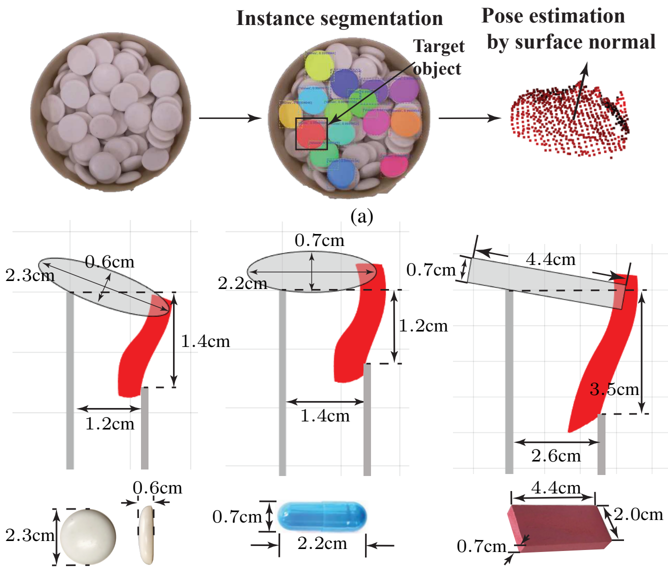

Fig. 8: (a) Instance segmentation and pose detection for Go stones in a cluttered bowl. Determining gripper design for (b) the Go stone with the parameters $\mu_c = 1.1$ (friction between the pusher and the slider), 0.2cm$< y_0 <$ 0.8cm, and $65° < \psi_0 < 75°$, (c) the capsule with the parameters $\mu_c = 0.9$, 0.2cm$< y_0 <$ 0.8cm, and $85° < \psi_0 < 95°$, and (d) the domino with the parameters $\mu_c = 1.2$, 0.7cm$< y_0 <$ 1.3cm, and $75° < \psi_0 < 85°$. In (b-d), the ranges of y_0 and ψ_0 were set around the values used in the experiments (like Fig. 3).

B. Autonomous Bin Picking

We first tested dig-grasping in autonomous bin picking. Our test objects are Go stones, capsules, and domino blocks.

Given an image of a cluttered bowl, object instances are detected and delineated through instance segmentation (Fig. 8(a)). Among the detected instances whose footprint is larger than a threshold, we choose the one closest to the center of the bowl as the one to pick. Finally, the pose of the selected instance is identified by estimating the normal at its centroid, using Open3D. It took around 4 seconds to process an RGB-D image input and to finally obtain the pose of a target instance, on our computer with NVIDIA GTX1050 Ti and 2.8 GHz Intel Core i7-7700HQ.

Next, the gripper design and the motion parameters for dig-grasping are determined as explained in Sec. IV. Fig. 8(b) illustrates the process for the Go stone, whose cross section is modeled as an ellipse. The gripper aperture is set to be 1.2cm (twice the thickness of the stone such that no more than one stone can be accommodated), and then the finger

This is the author's version of an article that has been published in this journal. Changes were made to this version by the publisher prior to publication.
The final version of record is available at http://dx.doi.org/10.1109/LRA.2021.3062250

TABLE II: Results of bin picking experiments.

	Object	Description	Fingertip Concavity	a (cm)	δ (cm)	Success Rate	PPH
Treatment group	Go stone	Gripper designed according to our approach	⊔	1.2	1.4	82/100	140
	Capsule	Gripper designed according to our approach	∨	1.4	1.2	75/100	108
	Domino	Gripper designed according to our approach	⊔	2.6	3.5	67/100	96
Control group	Go stone	Gripper with no fingertip concavity	–	1.2	1.4	82/100	
		Large gripper aperture	–	2.3	1.4	42/100	
		Symmetric fingers	–	1.2	0	17/100	
		Symmetric fingers	–	2.3	0	29/100	
		Ad hoc (normal pinch)	–	2.9	0	31/100	
		Ad hoc (parallel pinch)	–	0.8	0	65/100	
	Capsule	Gripper with no fingertip concavity	–	1.4	1.2	55/100	
		Large gripper aperture	∨	2.2	1.2	43/100	
		Symmetric fingers	∨	1.4	0	38/100	
		Symmetric fingers	∨	2.2	0	21/100	
		Ad hoc (parallel pinch)	∨	1.4	0	56/100	
	Domino	Gripper with no fingertip concavity	–	2.6	3.5	43/100	
		Chopstick-like slender fingers	–	2.6	3.5	36/100	

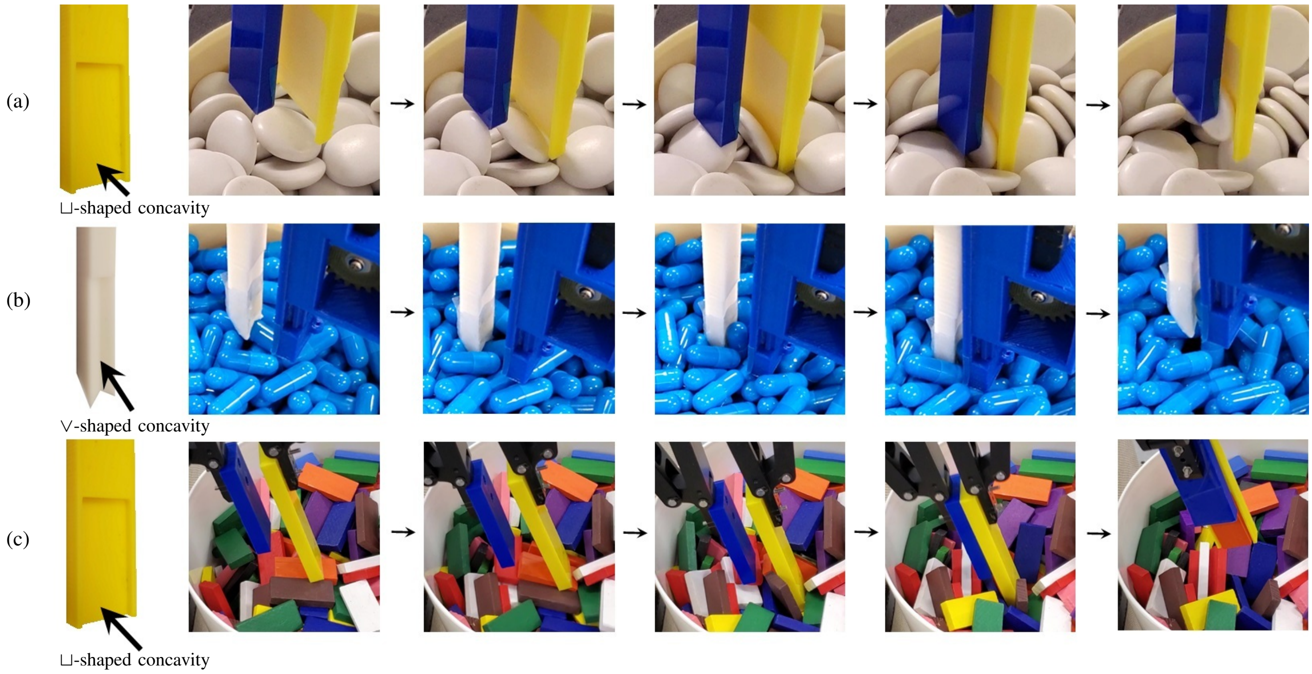

Fig. 9: Dig-grasping applied to bin picking from a cluttered bin: (a) Go stone, (b) capsule, and (c) domino. The fingertips used are shown on the left.

length difference and the digging depth are 1.4cm and 2.7cm (given $y_0 = 0.5$cm and $\psi_0 = 70°$), respectively. Likewise, the gripper designs for the capsule, modeled as an ellipse, and the domino block, modeled as a thin rectangle, are shown in Fig. 8(c-d). The digging depths for the capsule and the domino block were determined to be 2.8cm ($y_0 = 0.5$cm, $\psi_0 = 90°$) and 5.4cm ($y_0 = 1.0$cm, $\psi_0 = 80°$), respectively.

The steps of dig-grasping, align-dig-pinch, are then executed according to the determined motion parameters in an open-loop manner based on RGB-D geometric information (Fig. 9 and the video attachment). See Table II for the results. It presents two measures, success rates (successes/attempts) and picks per hour (PPH), for a range of scenarios. A success is defined as successful singulation and subsequent picking as in Table I. PPH should be understood as "successful" picks per hour (it is thus the product of the rate of picking and the success rate). In addition, PPH is inclusive of both perception and manipulation time, and is measured over a period of bin picking, during which the bin was being replenished to keep the size of the object cluster sufficiently large (for example, 100 or more Go stones). Our best performances were 140 PPH, 108 PPH, and 96 PPH for Go stones, capsules, and domino blocks respectively (note that our implementation features neither a high-end GPU nor an optimized software system). As shown in Fig. 9 and specified in Table II, these results were obtained with the fingertips that are at least as wide as the object instance, that is, protruding from the plane of the cross section, and have a ⊔-shaped (Go stone and domino) or a ∨-shaped concavity (capsule), customized according to the approach discussed in Sec. IV-B. The PPH values show that the domino blocks can be the most challenging, possibly due to their non-smooth shape. The control group experiments feature (1) other gripper configurations not following our design approach in full (i.e., with no fingertip concavity as marked "–" in Table II, a and δ not based on our approach, or slender fingers not protruding significantly from the plane) and (2) *ad hoc* picking in which

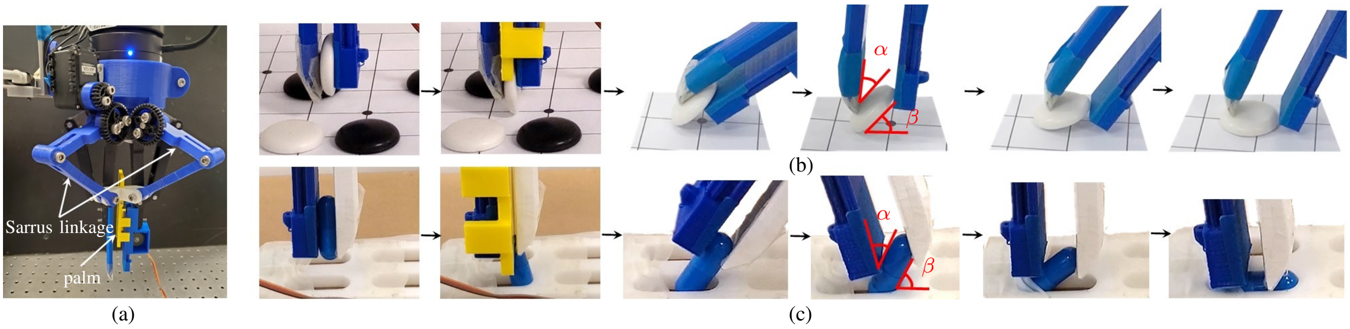

Fig. 10: (a) Two-fingered gripper retrofitted with the variable-length finger and our extendable palm device, the blue parallel linkage connecting the palm and the wrist of the gripper. (b) Placing a Go stone. (c) Packing a capsule into a partially covered cavity. α (β) refers to the angle that the left finger's face (the object) forms on the object (the ground). α and β are planned by shallow-depth insertion. See also the video attachment.

a gripper with conventional symmetric fingers was controlled to obtain a parallel (the major axis of the ellipse modeling the Go stone or the capsule is parallel to the finger's face) or a normal (the axis normal to the fingers) pinch grasp. In the ad hoc picking, we first set the gripper aperture to be slightly larger than the object's profile, align the gripper such that the profile can directly be projected into its workspace with no collision, and finally move the gripper straight towards the object. The best results of these experiments, reported in Table II, were outperformed by our approach. In the failed attempts, we witnessed that the gripper aligned poorly, due to errors in perception, or the object exhibited an unexpected translational motion.

C. *More Complex Tasks: Pick-and-Place/Pack*

Recall the process of our robotic dig-grasping applied to the Go stone, presented in Fig. 9(a). It bears a resemblance to the way human Go players singulate a stone from a cluttered bowl and pinch-grasp it usually between the index and the middle finger, through some in-hand manipulation. The player would then place the stone on a Go board again through in-hand manipulation using the two fingers. Our work on *shallow-depth insertion* [23] enables robots to emulate this placement capability. We integrated dig-grasping and shallow-depth insertion to deliver an end-to-end robotic system for pick-and-place/pack tasks.

Our invention, the extendable palm device in Fig. 10(a), is essential for bridging the two techniques. The yellow part is the palm end-effector located between the fingers. It is driven proximally and distally to the wrist of the gripper, by the Sarrus linkage that connects it with the wrist. It can thus be used to push out the object held between the fingers.

After successful picking by dig-grasping, the relative configuration between the fingers and the picked object is adjusted using the palm such that shallow-depth insertion can be initiated. The first two panels in Fig. 10(b) illustrate this with a Go stone, which is to be placed on a flat surface. The following four panels in Fig. 10(b) depict a motion path from an initial pinch grasp to the target configuration in which the Go stone is placed on the board, generated by the shallow-depth insertion technique, which resolves how to change the relative configuration of the object-gripper-environment system (parametrized by the angles α and β in Fig. 10(b)) without losing force-closure. In 20 attempts, 15 successful Go stone pick-and-places were achieved, where a success is defined as successful dig-grasping and no loss of grasp during the placement operation. Unsuccessful picking was the sole cause of all the failed attempts; in other words, after a successful pick, the subsequent placement was always successful. Fig. 10(c) shows the course of pick-and-pack, where the picked capsule is supposed to be packed into a cavity on a blister pack. We particularly address a mean scenario in which the cavity is partially covered with a rubber sheet as can be seen in Fig. 10(c), rendering traditional peg-in-hole inapplicable. This necessitates pushing the capsule horizontally into the covered space before terminating packing. Shallow-depth insertion enables the push maneuver with force-closure grasps. The last four panels in Fig. 10(c) depict the packing process through shallow-depth insertion, which was preceded by the initialization process shown in the first two panels. Out of 20 attempts, there were 13 successful and 7 unsuccessful pick-and-packs: unsuccessful picking in five attempts and unsuccessful packing in two attempts.

D. *Discussion: Range of Application*

The experiments show the applicability of dig-grasping, which is based on a two-dimensional analysis, to three-dimensional bin picking. We finally discuss the range of application in terms of object shape and the properties of a given object cluster.

a) Object Shape: Our findings that the difference in finger lengths facilitates simultaneous singulating and picking are based on the fact that $\mathcal{C}_{\text{obs}}$, the area between s and the pushing finger, becomes noticeably narrower. This is the case with relatively thin objects with a large width-to-thickness ratio, such as Go stones or domino blocks presented in Fig. 8. The specific shape of the side cross section profile, whether it is elliptical or thin rectangular as also shown in Fig. 8, is not critical, considering what matters is the outer boundary of the area it sweeps. For other group of objects whose width-to-thickness ratio is close to 1, for example objects with a square or a circular cross section (Fig. 11), our method using digit asymmetry is not expected to prove superior to symmetric

This is the author's version of an article that has been published in this journal. Changes were made to this version by the publisher prior to publication.
The final version of record is available at http://dx.doi.org/10.1109/LRA.2021.3062250

fingers because $\mathcal{C}_{\text{obs}}$ does not become noticeably narrower. However, these objects may be addressed in a straightforward manner as discussed in Sec. I. Quantitatively, the "overshoot" of $\mathbf{s}$, defined to be the maximum rise of $\mathbf{s}$ above its final flat part in the direction normal to the line of pushing ℓ (Fig. 11), can be used as a criterion for dig-grasping: do a dig-grasp if the overshoot of $\mathbf{s}$ is larger than a threshold.

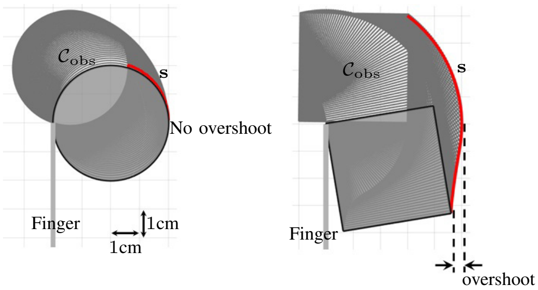

Fig. 11: $\mathcal{C}_{\text{obs}}$ obtained with a circle (left) and a square (right).

b) Properties of Cluster: The presented dig-grasping technique is applicable to

- Object clusters in an unjammed, malleable state: This is because the finger needs to dig the clutter. Although a detailed treatment of jamming is outside our scope, nonconvex object geometry or high friction between objects can result in jamming. Therefore, the current form of dig-grasping based on straight pushing would be ineffective in dealing with a cluster of metal screws, for example.
- Object clusters of sufficiently large size: The amount of objects in a bin needs to be sufficiently large such that the digging maneuver is not interrupted by the bottom of the bin. In Go stone picking for example, we were able to reproduce the best performance (around 80% success rate, Table II) consistently when there were 95 or more stones in the bin that we used. However, the success rate dropped to around 60% when there were around 75 stones, which were insufficient for complete digging as planned. Dig-grasping thus requires a constantly replenished supply of items to pick.

If these conditions are satisfied, it is also possible to apply dig-grasping to a cluster of heterogeneous objects. Our finger module makes it possible to accommodate different object shapes and sizes in a bin by adjusting the amount of digit asymmetry as needed. However, not all cases of possibly complex interactions between the objects in clutter can be treated by our approach featuring finger design for partial caging, as witnessed in the experiments. But the cost of picking failure in this bin picking setting is mitigated by that there are plenty of other objects. Therefore, in case a dig-grasp fails, the robot can simply attempt it again. Our adoption of the PPH measure addresses this scenario.

References

[1] M. T. Mason, *Mechanics of robotic manipulation*. MIT press, 2001.
[2] M. T. Mason, "Mechanics and planning of manipulator pushing operations," *The International Journal of Robotics Research*, vol. 5, no. 3, pp. 53–71, 1986.
[3] K. M. Lynch and M. T. Mason, "Stable pushing: Mechanics, controllability, and planning," *The International Journal of Robotics Research*, vol. 15, no. 6, pp. 533–556, 1996.
[4] M. A. Peshkin and A. C. Sanderson, "The motion of a pushed, sliding workpiece," *IEEE Journal on Robotics and Automation*, vol. 4, no. 6, pp. 569–598, 1988.
[5] J. Zhou, M. T. Mason, R. Paolini, and D. Bagnell, "A convex polynomial model for planar sliding mechanics: theory, application, and experimental validation," *The International Journal of Robotics Research*, vol. 37, no. 2-3, pp. 249–265, 2018.
[6] N. Chavan-Dafle, R. Holladay, and A. Rodriguez, "Planar in-hand manipulation via motion cones," *The International Journal of Robotics Research*, vol. 39, no. 2-3, pp. 163–182, 2020.
[7] M. Bauza and A. Rodriguez, "A probabilistic data-driven model for planar pushing," in *2017 IEEE International Conference on Robotics and Automation (ICRA)*, pp. 3008–3015, IEEE, 2017.
[8] R. C. Brost, "Automatic grasp planning in the presence of uncertainty," *The International Journal of Robotics Research*, vol. 7, no. 1, pp. 3–17, 1988.
[9] M. R. Dogar and S. S. Srinivasa, "Push-grasping with dexterous hands: Mechanics and a method," in *2010 IEEE/RSJ International Conference on Intelligent Robots and Systems*, pp. 2123–2130, IEEE, 2010.
[10] Z. Dong, S. Krishnan, S. Dolasia, A. Balakrishna, M. Danielczuk, and K. Goldberg, "Automating planar object singulation by linear pushing with single-point and multi-point contacts," in *2019 IEEE 15th International Conference on Automation Science and Engineering (CASE)*, pp. 1429–1436, IEEE, 2019.
[11] C. Mucchiani, M. Kennedy, M. Yim, and J. Seo, "Object picking through in-hand manipulation using passive end-effectors with zero mobility," *IEEE Robotics and Automation Letters*, vol. 3, no. 2, pp. 1096–1103, 2018.
[12] Z. Tong, T. He, C. H. Kim, Y. Hin Ng, Q. Xu, and J. Seo, "Picking thin objects by tilt-and-pivot manipulation and its application to bin picking," in *2020 IEEE International Conference on Robotics and Automation (ICRA)*, pp. 9932–9938, 2020.
[13] N. Correll, K. E. Bekris, D. Berenson, O. Brock, A. Causo, K. Hauser, K. Okada, A. Rodriguez, J. M. Romano, and P. R. Wurman, "Analysis and observations from the first amazon picking challenge," *IEEE Transactions on Automation Science and Engineering*, vol. 15, no. 1, pp. 172–188, 2018.
[14] D. Morrison, A. W. Tow, M. McTaggart, R. Smith, N. Kelly-Boxall, S. Wade-McCue, J. Erskine, R. Grinover, A. Gurman, T. Hunn, D. Lee, A. Milan, T. Pham, G. Rallos, A. Razjigaev, T. Rowntree, K. Vijay, Z. Zhuang, C. Lehnert, I. Reid, P. Corke, and J. Leitner, "Cartman: The low-cost cartesian manipulator that won the amazon robotics challenge," in *2018 IEEE International Conference on Robotics and Automation (ICRA)*, pp. 7757–7764, 2018.
[15] K. He, G. Gkioxari, P. Dollár, and R. B. Girshick, "Mask R-CNN," *CoRR*, vol. abs/1703.06870, 2017.
[16] P. J. Besl and N. D. McKay, "A method for registration of 3-d shapes," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 14, no. 2, pp. 239–256, 1992.
[17] I. Lenz, H. Lee, and A. Saxena, "Deep learning for detecting robotic grasps," *The International Journal of Robotics Research*, vol. 34, no. 4-5, pp. 705–724, 2015.
[18] S. Levine, P. Pastor, A. Krizhevsky, J. Ibarz, and D. Quillen, "Learning hand-eye coordination for robotic grasping with deep learning and large-scale data collection," *The International Journal of Robotics Research*, vol. 37, no. 4-5, pp. 421–436, 2018.
[19] J. Mahler, M. Matl, V. Satish, M. Danielczuk, B. DeRose, S. McKinley, and K. Goldberg, "Learning ambidextrous robot grasping policies," *Science Robotics*, vol. 4, no. 26, 2019.
[20] R. D. Howe and M. R. Cutkosky, "Practical force-motion models for sliding manipulation," *The International Journal of Robotics Research*, vol. 15, no. 6, pp. 557–572, 1996.
[21] K. M. Lynch, H. Maekawa, and K. Tanie, "Manipulation and active sensing by pushing using tactile feedback.," in *IROS*, vol. 1, 1992.
[22] J. Seo, M. Yim, and V. Kumar, "A theory on grasping objects using effectors with curved contact surfaces and its application to whole-arm grasping," *The International Journal of Robotics Research*, vol. 35, no. 9, pp. 1080–1102, 2016.
[23] C. H. Kim and J. Seo, "Shallow-depth insertion: Peg in shallow hole through robotic in-hand manipulation," *IEEE Robotics and Automation Letters*, vol. 4, pp. 383–390, April 2019.