Automated Plan Generation for Robotic Singulation from Mixed Bins

Krishnanand N. Kaipa$^1$ and Shaurya Shriyam$^1$ and Nithyananda B. Kumbla$^1$ and Satyandra K. Gupta$^1$

Abstract—In this paper, we address the problem of singulation in the context of robotic bin-picking applications, which involves grasping and tangle-free extraction of only one part at a time from a bin of randomized parts. In particular, rather than focussing only on interactions between the gripper and the part to be grasped in an isolated fashion, we are interested in problems where success rate is influenced by factors like perception uncertainty, grasp quality, approach quality, and extraction quality. Accordingly, we present a data-driven method that incorporates all these factors into the overall evaluation of singulation plans. In particular, we use a sampling based planner to generate a dense set of random plans. We evaluate each plan in a Monte Carlo simulator by computing its probability of failure. The plan which minimizes the probability of failure is chosen as the execution plan. We use illustrative experiments to show the working of our method.

I. INTRODUCTION

Robotic bin-picking is an important operation in many manufacturing and warehousing applications. Many research groups have addressed the problem of enabling robots, guided by machine-vision and other sensor modalities, to carry out bin-picking tasks [16], [7], [1], [29]. The perception problem in this context is very challenging and still not fully solved due to severe conditions commonly found in factory environments [18], [20]. In particular, unstructured bins present diverse scenarios affording varying degrees of part recognition accuracies: 1) Parts may assume widely different postures, 2) parts may overlap with other parts, and 3) parts may be either partially or completely occluded. The problem is compounded due to factors like sensor noise, background clutter, shadows, complex reflectance properties of parts made of various materials, and poorly lit conditions. All these factors result in perception errors in terms of part recognition and pose estimation uncertainties.

In this paper, we are interested in a class of bin-picking problems that manifest in the form of a part-order specifying multiple quantities of different parts to be singulated from a bin of randomly scattered pile of parts and transported to a destination location as rapidly as possible. Achieving this overall goal entails overcoming important challenges at various stages of task execution including part recognition, pose estimation, singulation, motion planning, and fine positioning. This paper is focused on the problem of singulation, which involves grasping and tangle-free extraction of only one part at a time from a bin of randomized parts. An illustration of the singulation task is shown in Fig. 1. Singulating a part from the bin, given a noisy estimate of part posture, presents several challenges in terms of how to plan the approach of the gripper toward the part such that it doesn’t collide with other nearby parts, how to determine grasp postures that result in force-closure of the grasped part, and how to perform tangle-free extraction.

In particular, rather than focussing only on interactions between the gripper and the part to be grasped in an isolated fashion, we are interested in problems where success rate is influenced by factors like perception uncertainty, grasp quality, approach quality, and extraction quality. Different failure scenarios during singulation of a desired part from the bin are shown in Fig. 2. Accordingly, we present a data-driven method that incorporates these factors into the evaluation of singulation plans. In particular, we use a sampling based planner to generate a dense set of random plans. We evaluate each plan in a Monte Carlo simulator by computing its probability of failure. The plan which minimizes the probability of failure is chosen as the execution plan. We use illustrative experiments to show the working of our method. We focus our attention on force-closure grasps [10] as opposed to form-closure [19] and caging grasps. [27].

II. RELATED WORK

Grasp planning literature is very vast. Approaches can be broadly divided into analytical and data-driven methods [28]. A review of analytical approaches to grasp synthesis can be found in [4]. A comprehensive survey on data driven grasp synthesis can be found in [6]. Data-driven approaches are increasingly becoming popular over the past decade with the advent of simulation based tools like Graspit! [21]. The grasp evaluation is usually based on a widely used force-closure quality metric for precision grips [10].

Given that pose estimation error impacts grasping performance in practice, many researchers have addressed the problem of grasp planning under perception uncertainty. Nguyen [25] incorporated contact location uncertainty into force-closure analysis. Roa and Suárez [26] proposed a grasp synthesis approach that accounted for the fact the real fingers can never contact the object at the computed points. Zheng and Qian [31] considered both friction uncertainty and

1Maryland Robotics Center, University of Maryland, MD USA
[kk|shriyam|nkumbla|skgupta]@umd.edu

Fig. 1. Illustration of the singulation task: (a) Robot gripper in the initial approach posture. (b) Part grasped. (c) Part successfully singulated.
III. PROBLEM FORMULATION

Let $\ell \in \mathbb{R}^6 = \{x, y, z, \alpha, \beta, \gamma\}$ represent a general posture where $(x, y, z)$ and $(\alpha, \beta, \gamma)$ represent the position and orientation, respectively in 3D.

**Definition 1.** A mixed-bin $B(\kappa, n, \{n_i\})$ is a bin of randomly scattered pile of $n$ parts, comprising different multiple instances $n_i$ of $\kappa$ different part types:

\[
B(\kappa, n, \{n_i\}) = \left\{ p_j^{(i)} : j = 1, 2, \ldots, n_i, i = 1, 2, \ldots, \kappa \right\}
\]

| $|$ $\sum_{i=1}^{\kappa} n_i = n$

where, part $p_j^{(i)}$ represents the $j^{th}$ instance of part type $i$.

**Definition 2.** Position-gripper refers as an action performed by the robot to position its gripper at an appropriate posture above the bin just before approaching a part to be grasped. We assume that this action is initiated from the robot’s initial position.

**Definition 3.** Approach-part refers as an action performed by the robot to move the gripper toward, and encompass, the part just before grasping takes place.

**Definition 4.** We say that a gripper encompasses a part when the intersection of the volume between the fingers with that of the part is non-zero and squeezing both the fingers in the pinch-direction results in a force-closure grasp. Note that in our context, encompassing is different from the notion of caging and only ensures satisfying the sufficient conditions for a force-closure.

**Definition 5.** Grasp refers to the act of grasping a part encompassed by the fingers. Note that by definition, force-closure is used as a constraint to evaluate candidate grasps.

**Definition 6.** Extract refers to the act of picking up a grasped part from the bin.

**Definition 7.** Singulation $\epsilon$ is defined as the concatenation of the four stages of positioning the gripper, approach, grasping, and extraction.

**Definition 8.** A Singulation plan consists of a sequence of grasp postures used to singulate a part.

**Definition 9.** We define tangle-free-singulation $\epsilon_{sf}$ as a singulation of a part from a bin such that it is not tangled with other neighboring parts in the bin during extraction, thereby ensuring singulation of only one part at a time.

We assume that there exists an automated perception system that scans and recognizes the set of all visible parts $\mathcal{P}_v \subseteq \mathcal{P}$, while reporting the following: (1) probability of successful recognition $\psi(p_j^{(i)}|\mathcal{D})$ of each part $p_j^{(i)} \in \mathcal{P}_v$, where $\mathcal{D}$ is the point-cloud data of the current scene and (2) estimate of part posture $\hat{\ell}_j^{(i)}$ and postural uncertainty $\sigma_\ell^{(i)}$ with a given confidence $p_j^{(i)}$, for each part $p_j^{(i)} \in \mathcal{P}_v$.

Now, we formulate our overall problem as follows: *Given a mixed bin $B(\kappa, n, \{n_i\})$ and an estimate of the posture $\hat{\ell}_j^{(i)}$ and postural uncertainty $\sigma_\ell^{(i)}$ of a part, generate a singulation plan that results in tangle-free singulation of the specified part from the bin.*
IV. SINGULATION

As defined in Section III (Definition 7), singulation is the concatenation of the four stages of positioning the gripper, approach, grasping, and extraction. The success of tangle-free singulation depends on postural uncertainty, grasp-approach quality $q_a$, grasp quality based on force-closure $q_g$, and whether the part is tangle-free during singulation or not. Accordingly, we present a method that incorporates all the above factors to generate and evaluate singulation plans. In particular, each singulation plan is evaluated by estimating the overall probability of successful tangle-free singulation $\mathcal{P}(s | p(j)_i, \hat{\ell}(j)_i, \sigma(j)_i, \rho(j)_i, \varepsilon)$ for each part instance $p(j)_i \in \mathcal{P}_v$. Figure 3 shows the overall system architecture used for plan generation and evaluation.

A. Grasp Quality

Each grasp candidate for a given part is determined as a function of the relative posture of the gripper with respect to the part. We use the widely used force-closure quality metric [10] to evaluate the grasp candidates. For this purpose, we first compute the set of all points on the part’s surface that have an orthogonal projection on one of the two fingers’ surface. Among these we find the subset of points that actually come into contact with the finger when the gripper fingers are closed. A point on the part’s surface is a contact point if the surface normal at that point makes $180^\circ$ with the inward normal making a semi-angle $\tan^{-1}(\mu)$, where $\mu$ is the coefficient of friction. Let $\hat{\mathbf{t}}$ be the unit normal at $a$, $\hat{\mathbf{L}}(a,b)$ be the unit vector lying along the line joining $a$ and $b$, and $\theta$ be the angle between these two vectors. Therefore, we have

$$\cos(\theta) = \hat{\mathbf{n}} \cdot \hat{\mathbf{L}}(a,b)$$

$$\sin(\theta) = \hat{\mathbf{t}} \cdot \hat{\mathbf{L}}(a,b)$$

where, $\hat{\mathbf{i}}$ is the unit tangent vector at $a$.

From Fig. 4(a), we can see that point $b$ lies in the friction cone of point $a$ if $|\tan(\theta)| < \mu$ with $\cos(\theta) > 0$ (since we consider only compressive grasps in this paper). Similar test can be done for point $a$ to check whether or not it lies in the friction cone of point $b$.

Now, we compute the grasp quality as the number of points that satisfy force-closure divided by the total number of points that project onto the finger surface.

For a particular grasp configuration, the pair of points where the center-axis of the gripper (along the pinching direction) intersects the part’s surface is uniquely determined (Fig. 4(b)). Therefore, we can represent each grasp candidate by such point pairs. Figure 4(c) shows the grasp quality of the best 20 grasp pairs evaluated using the above method for an industrial part. Sampled point clouds are generated for from the CAD models of the part and the gripper and used in the grasp quality computations. In the figure, the length of each line segment at a grasp point is proportional to the corresponding grasp quality of that grasp pair.

B. Plan Generation and Evaluation

Extracting parts from a bin involves grasping as well as positioning the gripper such that collision-free and robust
grasping can be achieved. Extraction operations are generally followed by transporting the grasped part to a desired destination posture. Hence extraction operations are generally concluded by taking the part safely away from the bin. From Definition 8, a singulation plan consists of a sequence of grasp postures used to singulate a part. In particular, each singulation plan is constructed by using four key postures: 1) Initial approach posture, 2) pre-grasp posture, 3) grasp posture, and 4) extraction posture. Intermediate waypoints are generated through linear interpolation between posture 1 and 2, posture 2 and 3, and finally posture 3 and 4. Note that only the position of the gripper changes during motion through the waypoints, but its orientation remains the same. Between pre-grasp and grasp, the location of the gripper remains constant, and the separation between the fingers decreases until the part is grasped.

We use a sampling based planner that generates several random plans. The initial approach posture is sampled at a safe height from the bin and in a small region around a nominal posture that corresponds to the grasp candidate that is ranked best by the above grasp quality metric. The pre-grasp posture is sampled in a small region around the estimated posture of the target part. The extraction posture is uniformly sampled at a safe height from the bin.

A Monte Carlo simulator evaluates each sampled plan by computing its probability of failure. We assume that point cloud obtained from the 3D sensor is split into two: one consisting only of the points in the bin excluding the part to be picked and the other point cloud consisting of those points of the part that were captured by the 3D camera as well as used by the perception system for matching it to CAD models in its repository. The simulation scene involves the gripper, part CAD model, point cloud of the bin excluding the part and the point cloud of the part. The part and gripper CAD models are also converted to sampled point clouds in Matlab before adding them to the scene. Thus, the simulation reduces to a problem involving point clouds only.

Each simulation run works in the following way. Given that an automated perception system provides an estimate of the pose of target part $\hat{\ell}^{(j)}_i$ with an error that follows a Gaussian distribution $\mathcal{N}(0, \sigma_i^{(j)})$, this is simulated by
placing a CAD model of the part at the estimated posture and shifting it by a value drawn from the above distribution of pose-error. Now, a candidate plan is evaluated by moving the gripper through the way-points, while checking for collision at each way-point. Point Cloud Library is used in C++ to check for collisions. A collision is said to occur between two point clouds when the minimum clearance between them falls below a certain threshold. If the way-point belongs to approach phase, then collision is checked between gripper and the entire bin. Patches missing in the point cloud that correspond to the desired part are accounted for by checking collisions also with the part CAD model in its estimated posture. If the way-point belongs to the grasping phase, then collision is checked between gripper and the bin excluding part. If the way-point belongs to extraction phase, then collision is checked between gripper and the bin excluding part, as well as between part and the rest of the bin. These collision check conditions ensure that we achieve tangle-free singulation of the part. If a collision is returned for at least one way-point during a trial, then that trial is classified as a failure. If there are $m$ such failure runs out of a total of $n$ runs, then the probability of failure for the specified plan is $\frac{m}{n}$. The plan which minimizes the probability of failure is chosen as the execution plan.

V. EXPERIMENTS

We report results from illustrative experiments to show the working of the sampling based plan generation and Monte Carlo based plan evaluation. We considered representative industrial parts (Fig. 5) that afford different recognition and grasping complexities in order to illustrate various challenges encountered during the singulation task. In this paper, we focussed our experiments with respect to the part shown in Fig. 4(c). This part presents both recognition as well as grasping complexities. In particular, the quality of the point cloud corresponding to this part is heavily influenced by its orientation relative to the 3D camera. Whereas the part is symmetric along its longitudinal axis, it is asymmetric along its lateral axis making the grasping problem nontrivial.

In particular, simulations were performed to compute failure probability $P_f$ of different plans generated by the Planner. Figure 6 shows snapshots from an animation of a sample Monte Carlo trial showing different stages of a successful singulation plan. Figure 7 shows snapshots from an animation resulting in a failed singulation plan where gripper collides with the part to be singulated during approach.

For each plan, a set of 100 trials were simulated with a postural uncertainty of $\text{Diag}(2\text{mm},2\text{mm},2\text{mm},4^\circ,4^\circ,4^\circ)$ added into the estimate of the target part in each trial. Figure 8 shows the graph of average clearance as a function of step number (1-5 Approach, 6-10 Pre-grasp to Grasp, 11-15 Extraction) for five sampled plans with varying probabilities of failure. When the extraction location was directly above the estimated location of the target part, $P_f$ was one ("□"-marked curve). This was due to collision with a neighboring part in the bin during extraction in every trial. But as the extraction point was moved away from this location, $P_f$ reduced to zero ("○"-marked curve). Whenever the average minimum clearance dips below a threshold of $\approx 3 \text{mm}$, we flag the state as collision and the plan is aborted. The clearance values after this point for each plan is only averaged over the trials which have been reported as success by the simulator. For another plan with $P_f = 1$ ("×"-marked curve), some of the trials failed during approach and the remaining during extraction at step 12. For the plan with $P_f = 0.86$ ("○"-marked curve), most of the trials failed indicating that it was a bad plan for the current uncertainty model. The plan with $P_f = 0$ is representative of an ideal plan for this uncertainty model. At every step in the plan, the average minimum clearance is safely above the threshold value. For the plan with $P_f = 0.166$ ("△"-marked curve), some of the trials failed due to collision during approach as a result of uncertainty in pose estimation.

In Fig. 9, the average minimum clearance was plotted for 25 different plans sampled by the planner and simulated over 100 trials. For good plans the average minimum clearance along with the standard deviation never goes below the collision threshold. All these plans have $P_f = 0$. For plans with $0 < P_f < 1$ average minus standard deviation values dips below the threshold signifying that some of the trials resulted in a collision during the singulation process. For plans with $P_f = 1$, the average $\pm$ standard deviation values lie entirely below the threshold line indicating that all the trials in these plans failed due to collision.

VI. CONCLUSIONS

We presented a data driven approach for automated generation and evaluation of singulation plans for mixed bins. We designed experiments to illustrate different aspects of how our system works. In this paper, we used a single example of a mixed bin. More empirical evaluations on bins with increasing levels of complexity are in order for systematically testing the ideas presented in the paper. The current sampling-based method can be extended to an asymptotically optimal planner that monotonically improves the plan in the sense of reducing the probability of failure over time. Further evaluation needs to be carried out by testing the generated
plans using real robot experiments. We assumed that the pose estimates and the uncertainty information is given. Integration between the singulation method presented here, an automated perception system, a human-aided perception system [16], and a fine-positioning method [14] is currently under progress. In our previous work, we have developed other modules including ontology for task partitioning in human-robot collaboration for kitting operations [2], sequence planning for complex assemblies [22], instruction generation for human operations [15], ensuring human safety [24], and a framework for replanning to recover from errors [23]. Future work consists of investigating how to integrate them in order to realize hybrid work cells where humans and robots collaborate to carry out industrial tasks.

VII. ACKNOWLEDGEMENTS

This work is supported in part by Office of Naval Research award N000141010886 and the Maryland Robotics Center.

REFERENCES


