Applying a Learning Framework for Improving Success Rates in Industrial Bin Picking

L.-P. Ellekilde*, J. A. Jørgensen†, D. Kraft†, N. Krüger†, J. Piater‡ and H. G. Petersen†

Abstract—In this paper, we present what appears to be the first studies of how to apply learning methods for improving the grasp success probability in industrial bin picking. Our study comprises experiments with both a pneumatic parallel gripper and a suction cup. The baseline is a prioritized list of grasps that have been chosen manually by an experienced engineer. We discuss generally the probability space for success probability in bin picking and we provide suggestions for robust success probability estimates for different sizes of experimental sets. By performing grasps equivalent to one or two days in production, we show that the success probabilities can be significantly improved by the proposed learning procedure.

I. INTRODUCTION

A significant part of the tasks done in production facilities involves transportation and putting of objects into feeding stations [3]. An easy way of moving objects is to use standard containers (pallets and bins), but unfortunately it is often difficult or impossible to keep the objects structured. A key challenge is therefore to empty the containers in a safe, efficient and economically feasible way. In mass production a specialized solution might be possible, but for small and medium sized batches, one often need to rely on human labour. An alternative is to use a bin picker comprised of: a sensor system for detecting the objects and a robot for picking them individually from within the container. An example of such a system is shown in Figure 1.

A particular property of the bin-picking scenario (in contrast to most other industrial robot applications) is that grasp errors are allowed to occur: Since bin-pickers usually fill feeding stations with a buffer of a number of objects in front of it, occasional errors do not disturb the main production process. In case such a grasping error occurs, another grasping will simply be tried. The large number of individual grasping trials (several thousands a day) are however not yet used to improve the performance of the grasping process. In this paper, we will show how this rich amount of grasping experience can be utilized to further improve the bin picking process through learning.

The problem of bin picking is rather complex: 1) The sensor system needs to do full 6D pose estimation and be robust towards occlusion. 2) The grasping process needs to handle arbitrary orientations of objects and be robust towards objects lying on top of each other, and 3) the motion of the robot needs to be collision free and efficient. Systems are furthermore faced with challenges such as cost, cycle time and the ability to detect and pick up all objects in the container. With these characteristics and requirements it is not possible to guarantee that all grasp attempts are successful, hence the system needs to handle when a grasp fails.

Even though the system can handle failures, it will negatively influence the average cycle time, which often has a large impact on the value of the bin picker. Or more precisely, a user will typically decide whether or not to deploy a robotic bin picker by primarily looking at the cycle time and the price. The cycle time is influenced by all aspects of the system. One factor is the sensor systems that determines which object to pick and what pose the object has. In the past this has been well researched [1], [7], [12], [19] and today there exists commercially available systems. Another important factor is the efficiency of the robot motions, which is the challenge addressed in [6].

Last but not least, cycle time is influenced by the average grasp success probability. Grasp failures may occur due to several reasons including conceptually incorrect or imprecise pose estimates, hindering placements of neighboring objects.

* L.-P. Ellekilde is with Scape Technologies A/S, Denmark, Email: lpc@scapetechnologies.com
† J. A. Jørgensen, D. Kraft, N. Krüger and H. G. Petersen are with The Maersk Mc-Kinney Møller Institute, Faculty of Engineering, University of Southern Denmark, Denmark, Email: (jimali, kraft, norbert, hgp)@mmmi.sdu.dk
‡ J. Piater is with the Institute of Computer Science, University of Innsbruck, Austria, Email: justus.piater@uibk.ac.at

1 See e.g. Scape Technologies (www.scapetechnologies.com)
and the chosen grasp strategy. These effects can not be modelled directly, hence appear as unmeasurable confounders in the system. The contribution of this paper is to show that the overall system performance can be improved by applying a learning strategy, able to update grasps based on real world experiments and thereby at least partially compensate for these unmeasurable confounders.

The paper is organized in the following way: In Section II we specify the grasp selection problem in bin picking of known objects. In Section III we give an overview of the related work concerning this problem. In Section IV we give a formal definition of the space that determines the outcome of a grasp in bin picking, discuss how to adequately sample this space and outline appropriate learning schemes. In Section V we present experiments with two different industrial bin picking setups. We show that with rather few experiments we can already achieve a significant improvement in grasp success probability in both setups.

II. THE GRASP SELECTION PROBLEM FOR BIN PICKING

Consider a set of known (e.g. through 3D CAD models) objects randomly piled in a bin. We shall for simplicity assume that all objects are of the same type. The basic task in bin picking consists of the following steps:

S1 Use a sensor system (typically a camera or range scanner or a combination of these) to detect one object in the bin and its pose
S2 Select an appropriate way to grasp the object
S3 Execute the grasp
S4 If the grasp was successful move the object to a desired location

In this paper, we address S2. We assume therefore that the object to be picked has been selected and its pose \( p \in SE(3) \) (position and orientation relative to some fixed coordinate frame) has been estimated. The grasp selection problem is quite generally to choose a strategy for grasping the object. We only consider two types of grasping devices, namely parallel grippers and suction cups. For both devices, we may define the grasping strategy as a point \( g \in SE(3) \) defining the gripper pose relative to a frame fixed in the object. For the parallel gripper, this is defined as the gripper pose immediately before the jaws are closed whereas for a suction cup, it is the pose where suction is applied.

In a bin-picking situation the object can only be grasped from above. Thus, for a given object pose \( p \), we should restrict the set of feasible grasp poses \( g \) to a set \( \Omega(p) \) defined by the “from above” constraint. We must therefore have a set of good grasps \( \{g_\alpha\} \) with \( \alpha = 1, \ldots, n \) so that for each object pose \( p \) there is preferably at least one \( g_\alpha \in \Omega(p) \).

More specifically, we may divide the grasp selection problem into two issues:

11 Choose a set of “good grasps” \( G = \{g_1, \ldots, g_n\} \) that covers the object in \( SE(3) \) as well as possible.

12 For each grasp \( g_\alpha \), a priority \( \pi_\alpha \) should be defined based on an estimate of the success probability of that grasp. The priorities are needed to have some way of selecting a grasp, when several options are available. With a solution for these two issues, the grasp selection problem for an object with pose \( p \) can be reduced to selecting the grasp \( g_\alpha \in \Omega(p) \) with the highest priority \( \pi_\alpha \).

In this paper, we will present a framework for choosing and refining priorities based on learning and present results which show a significant improvement of the overall success probability of executed grasps on two industrial bin picking platforms.

As the two issues 11 and 12 depend somewhat on each other and as both issues have been subject to several studies, we shall briefly review these studies in the next section.

III. RELATED WORK

The problem 11 of choosing a set of good grasps for an object is well known within data-driven grasp planning [2], [8]–[10], [14] i.e. planners that select a feasible grasp from a previously generated set of grasps (a grasp database). These methods use grasp quality measures [17], [18] to label “good” grasps and are typically based on some approximation of the form- or force-closure of a grasp. Another approach complements the force closure quality of a grasp with a measure based on the expected tactile feedback and the local density of successful grasps [11].

None of the above data-driven grasp planning approaches modify the database once it has been created. This is potentially a flaw since the virtual or heuristic environment within which the database is created, does not necessarily model the complexities and uncertainties of a real system. In this paper we refine the grasp database by re-adjusting quality labels and success probability estimates on the individual grasps and shows that it increases the overall robustness of grasps in the database.

The focus of data-driven approaches is often on the online grasp selection or synthesis part. Naturally this is where the context dependent search and refinement of grasps take place. [2] introduces a grasping score that prioritizes grasps that have a higher clearance to the environment, to minimize collisions. The authors also sort their data with a wrench based quality measure and prioritize selecting grasps that are “closer” to the gripper.

Another data-driven planning approach presented in [4] introduced the concept of a Grasp Knowledge Base (GKB). The GKB is essentially a grasp database which can be modified when new objects are sensed. However, the grasps in the GKB are generated in offline simulation and failed grasp attempts are not used to refine the GKB. Hence, their work essentially targets only 11.

In [16] a different approach was taken. To target 12 the control strategy to reach the final grasp was modified for each failed grasp using Dynamic Movement Primitives (DMPs). Learning a better grasp control strategy seems feasible however, if grasps are simply not reachable, picking another grasp would present a better solution.
Work on so called grasp densities has been introduced in [5]. These express grasp affordances associated to objects probabilistically and each experimental batch on a specific object updates the grasp affordances. Grasp densities can express a complete set of grasping options associated to an object with associated success probabilities. They are learned from successful grasping attempts, and therefore target both 11 and 12. Based on these grasp densities, efficient and flexible grasp strategies can be developed.

In contrast this paper presents a simple learning method which successfully incorporates the knowledge of failed grasp attempts into the offline generated grasp-database. The method is specifically designed to be applicable in industrial contexts, which does not allow for online learning strategies. We document its performance in two industrial setups and demonstrate its applicability in industrial bin-picking.

The probability space on which grasp densities are defined is the relation between object and gripper pose and thus essentially $SE(3)$ or a subregion hereof. As will be discussed in the next section, this probability space is only a very small subset of the space of all parameters determining the outcome of a given bin picking operation.

IV. SAMPLING GRAPS IN THE BIN PICKING PROBABILITY SPACE

In section [V-A] we give a mathematical formalization of the problem of learning the probabilities $P_a$ of a grasp $g_a$ being successful and sampling approximations of these. These probabilities are then used in the experiments in section [V] to prioritise grasps with estimated high success likelihood. This learning is done by drawing samples from a space which is much too complex to be modeled explicitly since effects like pose uncertainty, distribution of other objects in the bin and even illumination conditions play an important role.

It is important to define a sampling strategy that is applicable in industrial bin-picking. Such a strategy should not lower the overall success of the system too much and at the same time should allow for a reasonable fast convergence of the estimates of the success probabilities $P_a$. This is dealt with in section [V-B] where an appropriate sampling strategy is defined.

A. The Bin Picking Probability Space

As mentioned above, the parameter space in bin picking determining the outcome of a given pick comprises much more than object pose errors and the gripper pose $g$ relative to the object. To illustrate this, consider a hypothetical example where the same object residing at exactly the same pose in the bin is picked twice by the robot with exactly the same picking method. Clearly, the outcome of the two picks may be different due to variations in the neighborhood of the object. Differences in the neighborhood may or may not make the access to the object feasible with the given pick strategy and may or may not lead to a drag of the object so that the gripper loses it.

The probability space, which we shall formally refer to as $X$, determining the outcome of a given grasp, is defined as the exact location of the bin and all the objects in the bin including the object that we wish to pick. Notice, that this definition also indirectly includes uncertainties in the pose estimate and in the gripper pose relative to the bin and other factors such as, e.g. illumination conditions. Although this space is huge and intractable, we will provide a couple of formal definitions of probabilities using this space.

As described in Section II, we may assume that we have chosen a finite set of possible grasps $G = \{g_1, \ldots, g_n\}$ to be applied for picking a given type of object. The exact success probability for a given grasp $g_a$ with a given object pose estimate $p$ is then given by the formula

$$P_a(p) = \int_X o(g_a, x) p_p(x) dx$$

where $o(g_a, x)$ is the given outcome (success=1, failure=0). The function $p_p(x)$ is the probability density of $x$ given that a pose $p$ was estimated. It should however be noticed that the outcome $o(g_a, x)$ may be significantly more influenced by the placements of other objects, than small pose errors. Moreover, the grasps $g_a$ may have been selected to be robust towards these small pose errors.

For an arbitrary object pose $p$, only a subset of the elements in $G$ will be applicable because of the “above from above” criteria mentioned in Section II. Vice versa, any grasp $g$ will only be applicable from a continuous set of poses $p$. We define the region of object poses in which $g_a$ is applicable as $\Omega_a$ and the set of grasps $g$ which are applicable in a certain object pose $p$ is still referred to as $\Omega(p)$.

The formula for the overall success probability of the grasp is then given as

$$P_a(\Omega_a) = \int_{\Omega_a} P_a(p) z(p) dp$$

where $z(p)$ is the object pose probability density which in bin picking is often far from uniform.

1) Sampling in the Bin Picking Probability Space: Clearly, it is very difficult to define any good approximation for $p_p(x)$ because this would require in-depth knowledge of the probability distribution of how the objects align, but we can approximate the $P_a$s in a very simple way by performing grasp trials. Assume that we obtain $N_a$ pose estimates $p_1, \ldots, p_{N_a} \in \Omega_a$ and for each of these perform the grasp $g_a$. We may then estimate $P_a(\Omega_a)$ as

$$\hat{P}_a(\Omega_a) = \frac{1}{N_a} \sum_{k=1}^{N_a} o(g_a, x_k)$$

(1)

where $x_k \in X$ is the (unknown) element in the probability space when associated to the pose estimate $p_k$. By definition, the $x_k$s are sampled from the density $p_p(x)$. Thus we have asymptotic convergence, i.e.

$$\hat{P}_a(\Omega_a) \to P_a(\Omega_a) \text{ for } N_a \to \infty$$

(2)

For large $N_a$, we should therefore expect a rather precise approximation using sampling.

[3] Notice that this is a formal generalization of the situation in the grasp density concept [5], in which $x$ is just the true pose of the single object to be picked (deviating from $p$ with a small error).
B. Learning methods

After defining the learning space in section IV-A, we now deal with the problem of drawing samples efficiently in the bin picking context. The starting point of the learning methods is again a set of $n$ grasps $g_\alpha$, but now also a set of initial priorities $\pi_\alpha$. We shall now study two methods performing experiments for learning. The two methods only differ in the way a grasp is selected for a given estimated object pose $p$.

1) Method based on weighted random selection: An obvious method is to consider all the grasps $g_\alpha$ in $\Omega(p)$ and select one randomly based on weights chosen from the current priorities so that even the grasp with the lowest priority has a nonzero probability of being chosen. Since all grasps thus have a nonzero probability of being chosen in each experiment, Eq. 2 then yields the result that

$$\bar{P}_\alpha(\Omega_\alpha) \rightarrow P_a(\Omega_\alpha)$$

holds for all $\alpha$.

Thus, as long as each grasp $g_\alpha$ has a nonzero probability of being chosen during the learning phase, we will for large $N$ obtain good approximations of the correct $P_a(\Omega_\alpha)$. However, the rate of convergence for a given grasp depends on how often the grasp is chosen. On the other hand, if we too often choose grasps predicted to be inferior, we may obtain unnecessary slow convergence of the grasps that a priori seemed most promising.

2) Method based on selecting the highest priority: If, for some reason, we can only use a rather short experimental sequence ($N/n$ ratio not being sufficiently large), the property described in Theorem 1 becomes rather useless as we will not have sampled dense enough in the huge space $X$. If we furthermore have a set of initial priorities that have been chosen carefully (e.g. by an experienced engineer) and thus may not be far from optimal, we will have a significant risk of decreasing the overall success probability of the system due to large deviations between the estimates $\bar{P}_\alpha(\Omega_\alpha)$ and the true values $P_a(\Omega_\alpha)$.

Therefore, for short experimental runs, we propose to only use the grasps with the highest priorities during the experiments. It should be mentioned that even in this case, low priority grasps are occasionally executed in cases where they turn out to be the only executable ones. We leave the priorities unchanged during the experiments. After execution of all $N$ experiments, we adjust the priorities according to the new success probabilities. We may then possibly repeat the procedure with the new priorities.

In this method, many of the grasps $g_\alpha$ will never or rarely be chosen. In order to also assign success probabilities for such grasps, we use the following more robust estimate of the success probability.

\[
\hat{P}_\alpha(\Omega_\alpha) = \min\left(\frac{N_\alpha}{N_{\text{crit}}}, 1\right) \hat{P}_\alpha(\Omega_\alpha) + \left[1 - \min\left(\frac{N_\alpha}{N_{\text{crit}}}, 1\right)\right] \hat{P}
\]

where $\hat{P}$ is the unweighted average over all the $N$ grasps. We thus rely on the estimate $\hat{P}_\alpha$ when $N_\alpha \geq N_{\text{crit}}$. Alternatively, if the initial priorities would be stated as success probabilities, we might replace $\hat{P}$ with $\pi_\alpha$ in Eq. 3.

Clearly, we no longer have asymptotic convergence as many of the $g_\alpha$s will never be tried. Despite its drawbacks, the method will ensure that unforeseen unsuccessful $g_\alpha$s will have their priority lowered and thus it is very likely that the overall success probability will increase.

For medium experimental batch sizes, such as using bin pickers, running production may be treated by occasionally trying grasps with lower priorities. Then all grasps will occasionally be tried without negatively affecting the average cycle time too much by introducing grasps with relatively low expected success rate.

V. EXPERIMENTAL VALIDATION

The feasibility of the proposed learning approach has been verified using two test platforms. Test platform 1 (see Figure 1) is equipped with two Gimatic GS-20 pneumatic parallel grippers. These grippers have a stroke of 10.4 mm, which significantly reduces the number of ways the object can be grasped. The two grippers are thus equipped with different jaws, enabling the robot to grasp objects both in the narrow and wide directions. Test platform 2 (see Figure 2) uses a standard 2.5 bellow 32 mm suction cup. Both setups use tool mounted sensors based on structured light and pose estimation software developed by Scape Technologies. The test objects are shown in Figure 3 and consist of a T shaped forged steel object (test platform 1) and a bent aluminum sheet object (test platform 2).

Each of the two tests contained 4 distinct steps which were

1) Selecting a set $G = \{g_1, \ldots, g_n\}$ of grasps and an associated set of baseline priorities. The selection is done by an experienced engineer using a graphical interface, in which the user can select poses and set priorities of the grasps. Figures 3(a) and 3(c) illustrate the grasp poses selected for the two objects. The sets for object 1 and 2 contained $n = 163$ and $n = 263$ grasps, respectively.

2) Performing bin picking with grasps selected from $G$ using the baseline priorities. In the online process, we use the approach in Section IV-B.2 to select the grasp.

3) Updating the priorities of the individual grasps based on the method proposed in Eq. 3 in Section IV-B.2.

4) Performing bin picking while randomly switching back and forth between using the baseline and the updated priorities.

We shift back and forth in 4) to suppress slowly changing systematic artifacts such as how objects have packed in the
bin, the number of objects, slow changes in illumination (may affect the sensor system) and the overall calibration.

A. Benchmarking methodology

The effect of the learning can be measured by the comparing the overall success probabilities \( \hat{P}_{\text{base}} \) and \( \hat{P}_{\text{updated}} \) for the baseline and updated priorities respectively. We compute all success probabilities using Eq. (1).

The experiments which are to be benchmarked have the following characteristics

- They contain sets of experiments corresponding to Bernoulli trials, meaning that each has two outcomes, success and failure, which are mutually exclusive.
- Using an autocorrelation test, we have shown that there are no correlations between the outcome of each trial.

With these characteristics, the sampled mean value \( \hat{P} \) of the overall success probability (total number of successes divided by total number of trials) is approximately normally distributed around the true mean with a variance \( \hat{P}(1-\hat{P})/N \) where \( N \) is the total number of trials [15]. This property can be used to calculate confidence intervals of the sampled means, thereby providing statistical evidence that the results are significant.

B. Results for platform 1

A total of \( n = 163 \) grasps was selected for the first tests. These grasps are illustrated in Figure 3(a), where the markers show the poses of the gripper relative to the object and the colors shows the selected priority. Green, yellow and red refer to priority 1 (high), 2 and 3 (low) grasps, respectively.

Initially \( N = 1029 \) grasp attempts were performed based on which, new priorities were learned. The change in priorities can be seen by comparing the original and updated grasps illustrated in Figures 3(a) and 3(b) respectively.

(c) Grasps with baseline priorities. (d) Grasps with updated priorities.

Fig. 3. Grasps selected for test platform 1 (top row) and test platform 2 (bottom row). The markers indicate the poses of the gripper and the color the priorities, where green, yellow and red are for priority 1, 2 and 3 grasps, respectively (where 1 is the best priority).

The results of the benchmarking are summarized in Table 4. The differences in the numbers of successful and failed grasps result in success probabilities of \( \hat{P}_{\text{base}} = 0.521 \) and \( \hat{P}_{\text{updated}} = 0.635 \), which indicates an improved success probability of 11.4%. Computing the 95% confidence intervals shows that the two sets are disjoint as this level of significance, which provides evidence that the improvements are not due to chance. This can also be illustrated graphically by plotting the distributions, as shown in Figure 4. A more in depth statistical evaluation of the results can be obtained using a hypothesis test for testing two proportions [13]. Initially we pose the null-hypothesis that \( \hat{P}_{\text{base}} = \hat{P}_{\text{updated}} \). Our test statistics then becomes

\[
Z = \frac{\hat{P}_{\text{base}} - \hat{P}_{\text{updated}}}{\sqrt{\frac{\hat{P}_c(1-\hat{P}_c)}{N_{\text{base}}} + \frac{\hat{P}_c(1-\hat{P}_c)}{N_{\text{updated}}}}}\]

with \( P_c \) being the combined success probability over all experiments, and \( N_{\text{base}} \) and \( N_{\text{updated}} \) being the number of samples in the baseline and the updated sample sets, respectively. Inserting the associated values gives \( z = -3.50498 \). Trying to reject the null-hypothesis we can compute the area under the standard normal distribution and to the left of \( z = -3.50498 \) which gives 0.00025. This gives the probability of having rejected a true null-hypothesis, hence the reject seems fair, thus we can conclude that \( \hat{P}_{\text{updated}} > \hat{P}_{\text{base}} \).

C. Results for platform 2

The \( n = 263 \) selected grasps for platform 2 are illustrated in Figure 3(c). A batch of \( N = 3063 \) grasp attempts were
<table>
<thead>
<tr>
<th></th>
<th>Total</th>
<th>Success</th>
<th>Failure</th>
<th>$\hat{p}$</th>
<th>Confidence Interval (95%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Baseline priorities</td>
<td>449</td>
<td>234</td>
<td>215</td>
<td>0.521</td>
<td>[0.478; 0.567]</td>
</tr>
<tr>
<td>Updated priorities</td>
<td>469</td>
<td>298</td>
<td>171</td>
<td>0.635</td>
<td>[0.592; 0.679]</td>
</tr>
</tbody>
</table>

**TABLE I**

RESULTS OF EXPERIMENTS WITH TEST PLATFORM 1.

![Probability Density Function](image)

Fig. 4. Normal distributions representing the expected values for $P_{\text{base}}$ (blue) and $P_{\text{updated}}$ (red) for test platform 1.

Fig. 5. Normal distributions representing the expected values for $P_{\text{base}}$ (blue) and $P_{\text{updated}}$ (red) for test platform 2.

Initially performed from which new priorities were learned (see Figure 3(d)).

The benchmarking results are summarized in Table II. Even though the success probability of the baseline set is significantly higher (starting at 78.8%), the learning was still able to improve it, but in this case only with 4.9%. The associated 95% confidence intervals are again separated sets providing evidence that the difference is significant. This is graphically illustrated in Figure 5. Repeating the hypothesis test from Section V-B, we get $z = -3.00056$. We will again reject the null-hypothesis, this time with a 0.00135 probability of it being true.

### VI. CONCLUSION

In this paper, we studied the probability space determining the outcome of grasp attempts in bin picking. We presented a theoretical formalism describing the success probability for each grasp type and associated approximations based on sampling. We then discussed different learning techniques for improving the success probability. Finally, we have—on two different industrial bin picking platforms—shown that by using one of the proposed learning procedures, and performing a number of grasps equivalent to a very short production time, we were able to increase success ratios and the cycle time significantly. By that we gave an example of utilizing the large amount of grasp experience in industrial bin picking for learning.

There are several promising options for further increasing the success probabilities. First, the information obtained about the individual $P_{\text{gs}}$ could be used in selecting which object to pick next. This is currently chosen solely based on the sensorial information. The experiments for setting the priorities were carried out with two industrial grade setups. Logging grasps in real production would result in much larger data sets. However, the options of choosing also lower priority grasps occasionally should then be incorporated into the system to fully exploit the learning capabilities.

We are currently finalizing the preparations to be able to simulate bin picking within our dynamic simulator Rob-WorkSim [11]. This will open for huge experimental sets with millions of grasp attempts, which could lead to some understanding in how the grasp outcome $\phi(g,x)$ could be modeled. Such an understanding could lead to new and even better options for both choosing the grasp list $G$ and for setting the priorities.

**ACKNOWLEDGMENT**

This work was supported by the LearnBiP project (an experiment funded in the context of the ECHORD project (FP7-ICT-231143)) and The Danish National Advantage Technology Foundation through the project Bin Picker.

**REFERENCES**


