# Robotic manipulation in object composition space

Joni Pajarinen and Ville Kyrki

***Abstract*— Manipulating unknown objects in a cluttered environment is difficult because object composition is uncertain. Because of this uncertainty, earlier work has concentrated on finding the "best" object composition and based on this composition decided on manipulation actions. Contrary to earlier work, we 1) utilize different possible object compositions in decision making, 2) take advantage of object composition information provided by robot actions, 3) take into account the effect of different competing object hypothesis on the actual task to be performed. We cast the manipulation planning problem as a partially observable Markov decision process (POMDP) which plans over possible hypotheses of object compositions. The POMDP model chooses the action that maximizes the long-term expected task specific utility, and while doing so, considers the value of informative actions and the effect of different object hypotheses on the completion of the task. In experiments with a physical robot arm and an RGB-D sensor, our approach outperforms an approach that only considers the most likely object composition.**

## I. INTRODUCTION

Service robots in domestic environments need the ability to manipulate objects without good prior models in order to cope with the variability of such environments. This need is usually approached by attempting to model the objects online using sensors based on stereopsis or structured light. When multiple measurements can be acquired around an isolated object, this approach works quite satisfactorily as the generated 3-D models can often be used for successful manipulation.

In cluttered scenes with multiple unknown objects, the segmentation of objects, also known as object discovery in perception research, becomes a major problem. Typically, the problem is to decide which of the segments in an oversegmented scene belong to the same object. This is challenging especially because objects can be partially occluded by others. A promising approach towards solving object discovery is interactive perception, where the object configuration is actively examined typically by manipulating the objects and observing the results. Another line of work is to use learned priors to find the most likely object composition. Despite the recent advances, manipulation of unknown objects in cluttered environments is still an open problem.



This work was supported by the Academy of Finland, decision 271394.

J. Pajarinen and V. Kyrki are with the Department of Electrical Engineering and Automation, Aalto University, Finland

E-mail: {Joni.Pajarinen, Ville.Kyrki}@aalto.fi

In this paper, a solution to manipulation planning is proposed, which plans over hypotheses of possible object compositions instead of trying to determine a single best hypothesis. The approach combines earlier ideas of interactive perception and learned composition priors in a planning under uncertainty framework. The manipulation planning problem is cast as a partially observable Markov decision process (POMDP), which integrates active exploration to planning. In contrast to earlier work, our approach 1) utilizes different possible object compositions in decision making, 2) takes into account the effect of competing hypotheses on the goal task, and 3) actively explores the hypothesis space if that benefits the task.

The paper is structured as follows: We begin by surveying works related to object discovery, interactive perception and planning under uncertainty in Section II. Our framework of manipulation planning over object compositions is then introduced in Section III. Section IV proposes our approach for estimating the distribution over object hypotheses. The state space of hypotheses is then used for manipulation planning as described in Section V. Experiments with a physical robot arm and an RGB-D sensor presented in Section VI demonstrate that the proposed approach is able to integrate perception to the manipulation task and that the use of multiple hypotheses improves system performance when compared to considering only the most likely hypotheses. Section VII concludes the paper.

## II. RELATED WORK

**Object discovery.** Scene segmentation is a classic problem in computer vision tightly coupled to object recognition so that it can be argued that the segmentation problem does not have unique solutions if the objects are not known. Nevertheless, there is a need to discover objects from scenes even when the objects are not known in advance. Recent works in the area are often based on learning general models used to recognize object classes from segments (e.g. segment labeling in [1]), to detect segments based on their "objectness" [2], or to choose which segments belong to a single object [3], [4]. Our work follows the line of work of [3], [4] but instead of trying to find a single optimal composition, considers the distribution of possible compositions.

**Active and interactive perception.** Instead of a passive approach (the scene is purely observed) such as the ones presented above, object discovery can be approached from the point of view of active perception [5]. Gaze control and foveation, which are purely perceptual processes, have been proposed for object discovery [6]. Furthermore, interactive perception has been proposed as a promising solution for object discovery with the goal of singulating [7] or clearing [8] a pile of objects. Both approaches use poking or pushing actions to estimate the object composition. This paper follows the interactive perception paradigm but in contrast to the works above, integrates the perception with goal-directed planning so that perceptual actions are only used when they are expected to support the task goal.

**Grasping unknown objects.** Grasping unknown objects has got significant attention in the research community, especially after Saxena's work [9] which proposed the use of machine learning for planning good grasps. Many of the current methods analyze point cloud data locally (e.g. [10], [11]) to determine parts of objects that afford good grasps. The point cloud can be first transformed to another domain for easier analysis, as for example in the surface height accumulated features of [12]. It is also possible to analyse a segmented object as a whole (e.g. [13], [14]), for example to align a robot hand according to the principal axes of the point cloud. In this work, grasping according to this latter line of work is used as a component for both informative and goal-directed actions.

**Manipulation planning under uncertainty.** In planning manipulation under uncertainty, such as planning where to grasp an object, classical deterministic planning can be used to reduce uncertainty. For example, Dogar et al. [15] plan pre-grasp pushing actions that collapse pose uncertainty of a target before executing a grasping action. This type of approach is usually only available for completely known objects.

With limited knowledge, POMDP-like approaches can be used to plan over a distribution of states. Hsiao et al. [16] proposed the partitioning of a one-dimensional configuration space to yield a discrete POMDP which can be solved for an optimal policy. In planning grasp locations, the state-of-the-art includes probabilistic approaches with a short time horizon. The goal can be formulated either as positioning the robot accurately as in [17] or maximizing the probability of a successful grasp as in [18]. The short-term planning can also be extended to include information gathering actions [19]. Besides pose uncertainty in grasping, Monso et al. have proposed to formulate clothes separation as a POMDP [20]. Of the above, only Monso et al. consider, similar to this paper, the case of a complex cluttered scene with more than one object. In contrast to them, this paper does not assume that each object is uniform in color, but instead, complex multi-colored and textured objects are considered for object discovery. Moreover, their state space model is clothes separation specific modeling the number of clothes in different areas. Our approach reasons about objects directly.

## III. Manipulating object compositions

We consider the scenario of a robot manipulating unknown objects based on RGB-D data. The manipulation goal is defined in terms of simple features that can be observed incompletely from the point clouds. For example, the goal could be to move all objects with a certain color to a particular location. Manipulating unknown objects is difficult because even if RGB-D data is available, the robot does not know in advance the shape or color of the objects. Thus, it has to guess which parts of the point cloud belong to the same object. Occlusion and noisy sensor readings make this task hard. Attempting to segment individual objects from the point cloud typically results in oversegmentation, which leads to the problem of deciding which segment belongs to which object, in other words, forming *object hypotheses*.

In previous works such as [12], the choice of an action is based on the most probable hypothesis of object composition. The shortcoming of this approach is that it does not take into account the long term effects of uncertainty or the value of information gathering actions. We propose to choose instead the action that maximizes reward over the distribution of possible compositions. By considering a temporally evolving system, the robot can infer from past grasp attempts the likelihood of object hypotheses.

### A. Overview

At each time instant, the robot performs the following steps:

1) Obtain RGB-D data of the current scene and segment the RGB-D data
2) For each pair of segments, estimate the probability of the segments being part of the same object
3) From the estimated probabilities create a probability distribution over possible object compositions that conforms with past grasp attempts (Section IV)
4) Use a POMDP to select the best long-term manipulation action for the current object distribution and execute the action (Section V)

For steps 1 and 2, for segmenting RGB-D data and estimating probabilities for segment pairs, we use an existing approach from [4]. The novel steps 3 and 4 are described in Sections IV and V, respectively.

### B. Robotic manipulation as a partially observable Markov decision process (POMDP)

Before going into belief estimation and manipulation planning, we describe how to model robotic manipulation as a POMDP. A POMDP defines optimal behavior for an agent in an uncertain world with noisy,partial measurements, when the stochastic world model is accurate and when the agent's goal has been defined precisely. Previously, POMDPs have yielded good results in different robotic applications [20], [21]. We utilize a POMDP because it takes uncertainty in action effects and observations into account. Moreover, a POMDP assigns the correct long-term value to informative actions which are needed when exploring object hypotheses.

The temporal model of a POMDP is defined by the transition probability $P(s'|s,a)$ from state $s$ to the next time step state $s'$, when action $a$ is executed, and the probability $P(o|s',a)$ of observing $o$, when action $a$ was executed and the world moved to the state $s'$. A real-valued reward $R(s,a)$ for executing action $a$ in state $s$ encodes the objective. An optimal policy $\pi$ maximizes the expected reward $E\left[\sum_{t=0}^{T-1} R(s(t),a(t))|\pi,b_0\right]$ over $T$ time steps, where

$b_0$ denotes the initial *belief*, that is, a probability distribution over world states. At each time step, the agent decides on an action $a$ based on its current belief. In principle, the belief can be kept up-to-date given an accurate temporal model. However, because an accurate model is in practice not available, we instead estimate the belief at each time step from current visual sensor data and past history, and use the *online* POMDP method introduced by us in [21] to compute a new policy. To cope with a huge state space, the POMDP method in [21] uses a state particle representation for the belief $b(s)$.

[21] defines robotic manipulation as a POMDP, in which the probabilities of successfully grasping an object, and observing its attributes (for example color), depend on how occluded the object is. The POMDP state $s = (s_1, s_2, \ldots, s_N)$ is a combination of object states $s_i = (s_i^{\text{loc}}, s_i^{\text{attr}}, s_i^{\text{hist}})$ where $s_i^{\text{loc}}$ is the semantic object location, $s_i^{\text{attr}}$ object attributes, and $s_i^{\text{hist}}$ historical information for object $i$. The POMDP model in this paper differs from the model in [21]: [21] assumes a single object composition, here we use a probability distribution over possible object compositions. For example, grasping actions occur in the space of object compositions which will be discussed in more detail in the following sections. Furthermore, in this paper, the grasp success probability is not a probability distribution conditional on the number of failed and succeeded grasps. Instead, we assume that a previously failed grasp can not succeed unless the occlusion on the grasped object changes.

## IV. Belief Estimation

The belief consists of state particles and their probabilities. In [21], a state consists of multiple objects and their relationships including information on which object is in front of which object, and object attributes. Here, instead of objects, each state consists of a set of *object hypotheses* called an *object composition*. An object hypothesis consists of a set of segments, where every segment is connected either directly or in-directly to each other. The segments are not connected to segments outside this set. Two segments can be directly connected if they occlude each other or if another object occludes both (the direct connection is then behind this occluding object).

The probability of a sampled belief state is proportional to the probability of the object composition, which is part of the belief state, to exist, when considering the grasp history. The key insight is that previously performed grasps must have failed for an object hypothesis, otherwise the object would have been moved. Furthermore, a grasp can only succeed for a wrong hypothesis, when the wrong hypothesis is part of a hypothesis, for which the grasp succeeds.

To estimate the belief, we segment the observed point cloud and compute the connectedness probability for each segment pair. Based on these probabilities, and based on whether segments can be directly connected, we define a Markov chain which converges to a distribution over object compositions. We sample object compositions from this Markov chain after a burn-in period. A belief state corresponds to a sampled object composition with sampled object attributes. The probability of the belief state is set proportional to the probability of the sampled object composition, which is computed (for uniform priors) as the probability of the observation/action history conditional on the object composition. This means that the belief over object compositions is shaped by past events: for example, if the robot fails to grasp an object hypothesis, which should be easy to grasp when the object hypothesis is correct, then the hypothesis is likely incorrect, and the belief will reflect this. Next, we will discuss how to sample object compositions and then show how to estimate the conditional probability of an object composition given past events.

### A. Markov chain sampling of object compositions

We sample object compositions from a Markov chain. The Markov chain moves from one object composition to another: the Markov chain randomly selects a direct connection candidate between two segments, and then samples whether the direct connection exists from a probability which depends on the probabilities of the composition with and without the direct connection. The computation is efficient because we only have to consider segments connected (indirecly) to either segment under consideration. In more detail, denote connectedness probabilities with $P(c_{i,j})$, where variable $c_{i,j}$ denotes whether segments $i$ and $j$ are part of the same object ($c_{i,j} = 1$) or not ($c_{i,j} = 0$). Moreover, denote with $h_k$ an object composition, the $k$th state of a Markov chain. The Markov chain starts from $h_0$, which contains only disconnected segments. To sample $h_{k+1}$ from $h_k$ we first sample uniformly randomly a segment pair $i, j$ from segment pairs which may be directly connected. Denote with $\delta_k^{i,j}$ the state of the direct connection between $i$ and $j$ in $h_k$. Second, assuming an uniform prior $P(\delta_{k+1}^{i,j})$ compute the conditional probability $P(\delta_{k+1}^{i,j} = 1|\hat{h}_k)$, where $\hat{h}_k$ denotes $h_k$ when $\delta_k^{i,j} = 0$:

$$P(\hat{h}_k|\delta_{k+1}^{i,j} = 1) = \frac{1}{Z(U,V,\hat{h}_k)} \prod_{u \in U} \prod_{v \in V} P(c_{u,v} = 1)$$

$$P(\hat{h}_k|\delta_{k+1}^{i,j} = 0) = \frac{1}{Z(U,V,\hat{h}_k)} \prod_{u \in U} \prod_{v \in V} P(c_{u,v} = 0)$$

$$P(\delta_{k+1}^{i,j} = 1|\hat{h}_k) = \frac{P(\hat{h}_k|\delta_{k+1}^{i,j} = 1)}{P(\hat{h}_k|\delta_{k+1}^{i,j} = 1) + P(\hat{h}_k|\delta_{k+1}^{i,j} = 0)}, \quad (1)$$

where $U = i$ and $V = j$, when $c_{i,j} = 1$ in $\hat{h}_k$, otherwise $U = i \cup (u|c_{i,u} = 1)$ and $V = j \cup (v|c_{j,v} = 1)$. $Z(U,V,\hat{h}_k)$ is a function containing a normalization term and the joint probability of all segment pairs $m,n$, where $m$ or $n$ is not in $U$ or $V$. Note that $Z(U,V,\hat{h}_k)$ cancels out in Equation (1). Finally, we sample $\delta_{k+1}^{i,j}$ using $P(\delta_{k+1}^{i,j} = 1|\hat{h}_k)$ to get $h_{k+1}$.

### B. Probability of an object composition given past events

In general, the probability of an object composition $h = (h_1, \ldots, h_N)$, where $h_i$ is a single object hypothesis, depends on the sequence of past actions and observations $\theta_t = (a(0), o(1), a(1), o(2), \ldots, a(t-1), o(t))$, where $t$ denotes the current time step. We assume uniform priors, independent object hypotheses, and independent history events:

$$\begin{aligned} P(h|\theta_t) &= \prod_{i=1}^{N} P(h_i|\theta_t) \\ &= \prod_{i=1}^{N}\prod_{k=1}^{t} P(a(k-1), o(k)|h_i). \end{aligned} \tag{2}$$

For object hypothesis manipulation, we maintain a history of unique executed grasps. In our model, there is no need to remember multiple identical grasps; a previously failed grasp cannot succeed again, unless the occlusion of the object, for which the grasp is optimized, changes because then the grasp also changes. The composition probability conditional on past independent grasps $(\text{grasp}_1, \ldots, \text{grasp}_M)$ is

$$P(h|\theta_t) = \prod_{i=1}^{N}\prod_{k=1}^{M} P(\text{grasp}_k|h_i). \tag{3}$$

## V. Manipulation planning

After the robot has estimated the current belief it decides on its next action based on the belief. As discussed earlier, we use a POMDP for decision making. Next, we discuss parts of our system that the POMDP requires for planning. We will discuss the actions available to the robot, how to sample a new state, how to sample an observation, and how to compute the observation probability.

### A. Actions.

In our problem setting, the robot may grasp an object and move it. We employ top-down grasping. For selecting the finger distance and rotation of the robot hand, we use a simple approach based on computing a vector at a narrow part of the unknown object with principal component analysis (PCA). The approach 1) projects the point cloud $PC_1$ of the target object hypothesis onto a plane, which is parallel to the wrist of the down-pointing robot hand, 2) makes the point density of the projected point cloud uniform to get the point cloud $PC_2$, and 3) projects the centroid of $PC_2$ towards $PC_1$ along the wrist-plane normal, by the distance between the centroids of $PC_1$ and $PC_2$, to get the grasp centroid. The approach computes the PCA decomposition of $PC_2$. In PCA, the first eigenvector aligns to the largest variance in the point cloud data and the second eigenvector aligns along the second largest variance in the data. For example, for a long object, the first eigenvector could be aligned along the length of the object, and the second eigenvector along the width of the object. The approach uses the second eigenvector in order to get a narrow grasp. In more detail, the approach projects two points in opposite directions from the grasp centroid, along the computed second eigenvector, far enough. Finally, the approach selects the two points from $PC_1$ which are closest to the projected points, as the two grasp contact points. In addition, the approach checks whether some part of another object hypothesis blocks the direct path up from the two grasp contact points, and if so, sets the probability of a successful grasp to zero.

*1) Restricted action set:* In order to restrict the computational load, we bound the number of possible grasps, and thus actions, by a predefined maximum number. Instead of restricting the number of possible object hypotheses, we select a subset of all hypotheses to use for grasping. Optimally, we would like to choose a set of grasps, which yields the best policy among all possible grasp sets. However, because we do not know the best policy, we settle for computing an action set that maximizes the expected grasp success probability. The grasp success probability $P_{\text{grasp prob}}(\text{SUCCESS}|h_i, A)$ defines the probability of successfully grasping and moving an object hypothesis $h_i$ when the robot chooses the best action from the action set $A$. The expected grasp success probability is

$$A = \arg\max_{A} \sum_{h_i} P_{\text{grasp prob}}(\text{SUCCESS}|h_i, A) P(h_i) \ , \tag{4}$$

where the number of actions is $|A|$. $A$ can be found by an integer linear program. Unfortunately, integer linear programming is in the worst case NP-hard. As an approximation, we use a greedy approach which incrementally selects object hypothesis which increase the expected grasp success probability the most. In the experiments, the expected grasp success probability using a restricted action set remained usually close to the probability with the complete set of possible actions.

*2) Grasp success probability:* A grasp is parameterized by the distance of the finger tips, rotation of the hand, and the location of the robot wrist. The grasp success probability is the product of the *grasp quality* and an occlusion specific grasp probability. When computing grasp probabilities we take previously executed grasps into account: a failed grasp cannot succeed again, unless the occlusion of the object changes for which the grasp is optimized (when the occlusion changes the grasp usually changes also). Grasp quality is intended to capture the quality of a grasp which is optimized for another object hypothesis. Grasp quality is equal to 1 when using a grasp which was computed for the same object hypothesis that the robot tries to grasp. The grasp quality decreases when the grasp centroid moves away from the optimized grasp centroid, and becomes zero when it is outside the object. We compute the grasp quality for grasping object hypothesis $X$ with a grasp optimized for object hypothesis $Y$ as follows

1) Inside $X$, find starting point $y_1$ and end point $y_2$ between the grasp points of $Y$
2) If there is no $y_1$ or $y_2$, then the grasp quality is zero because the line of grasp is outside object $X$
3) Compute centroid $c_Y$ of $y_1$ and $y_2$
4) Project $c_Y$ into the robot arm wrist plane along the plane normal
5) Project grasp centroid of $X$ into the wrist plane along the plane normal to get $c_X$
6) Denote with $\hat{X}$ the projection of $X$ onto the wrist plane. Project a point starting from $c_X$ through $c_Y$ so that it is outside $X$ and find the closest point $x_1$ in $\hat{X}$

7) Finally, compute the grasp quality as (distance from $c_Y$ to $x_1$) / (distance from $c_X$ to $x_1$), that is, the grasp quality decreases when the effective grasp centroid $c_Y$ moves closer to the surface, away from the grasp centroid $c_X$ optimized for $X$.

### B. Temporal model of the world

In order to use the POMDP method in [21] for planning, we need to model the evolution of the state of the world over time. We need state transition and observation probabilities. Because probability distributions use a state particle representation, we need, in particular, a way to sample states and observations, and a way to estimate the likelihood of a state particle given an observation. Next, we discuss how to accomplish these tasks.

*State sampling.* As discussed earlier, a world state consists of an object composition $h = (h_1, \ldots, h_N)$, and contains for each $h_i$ a semantic object location, attributes, and history. To sample a new state for a grasp action $a$, select the object hypothesis $h_i$ that has the highest grasp probability for $a$. Sample grasp success of $a$ on $h_i$ according to the grasp success probability. If the grasp fails, add the grasp to the grasp failure history of $h_i$. If moving an object succeeds, the semantic location of the object is changed to the destination location.

*Observation sampling.* After executing the grasp action, the robot observes which object was moved, and in the case of a successful move, the robot makes an observation about the attributes (color in the experiments) of a limited number of objects behind the moved object. Assuming independence between attribute observations, the observation probability is $\prod_i P(o_i|h_i)$, where $o_i$ is the observation of $h_i$. As in [21], $P(o_i|h_i)$ is computed from the occlusion of $h_i$ and the attribute instances of $h_i$.

*Observation probability.* The probability of making an observation $o$ in state $s$ is zero if the moved object hypothesis differs from the observed one, or if the move fails and the attribute observations do not match with previous attribute observations. Otherwise, the probability is defined by $\prod_i P(o_i|h_i)$ discussed above.

## VI. EXPERIMENTS

In the experiments, an RGB-D sensor (Microsoft Kinect) observes objects on a table and a 6-DOF Kinova Jaco arm with an integrated 3-fingered hand manipulates the objects. The setup is shown in Fig. 1. In the experimental task, the table contains different kinds of toys which occlude and can be very close to each other (see Fig. 2). The goal of the robot is to find and move an unknown number of fully red toys into a target area (Area 1). In 9 of 10 experimental scenes, only one of the toys is red, while one scene (Scene 9) has 3 red toys. To remove occlusions that hinder color detection and grasping, the robot can also move toys to a free area (Area 2). Moving a red object into Area 1 yields 1\$. Moving a non-red object into Area 1 or moving a red object into Area 2 costs 1\$. Note that both areas 1 and 2 are at the bottom in Fig. 1.

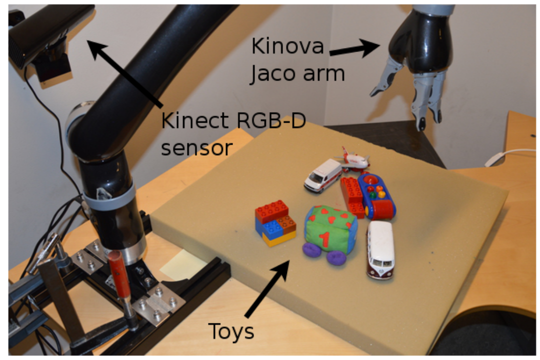


Fig. 1. Experimental setup. Based on RGB-D data from a Microsoft Kinect, a 6-DOF Kinova Jaco robotic arm tries to move red toys into a target area at the bottom of the picture.

In the experiments, segmentation and segment-pair probability computation was performed using the approach presented in [4]. In short, the approach assigns probabilities to segment pairs using support vector machines (SVMs) trained with RGB-D data of household items which are not in all ways similar to the toys we use in the experiments. For grasp probability we used parameters estimated for coffee cups in [21] and we set the first color observation parameter (see [21] for details) to $-0.5$ and the second to $-0.02$ for both red and non-red observations. The models were not optimized for the particular objects used in this paper because in the real-world the robot would need to be able to generalize to new objects.

We compare our POMDP based approach against a method called "Baseline" which uses the most probable object composition and tries to grasp the object which has the highest grasp success probability of all objects observed red. If such an object does not exist, then it finishes. The method does not condition the object composition on the grasp history. For the POMDP approach, we set rewards to the task specific goal and in addition assign a penalty of 0.01\$ for failed grasps. Note that our POMDP based approach is not restricted to the objective in this task but can optimize any goal. The POMDP policy size was set to 3x3 (see [21] for details) in the experiments.

For each scene shown in Fig. 2, we placed objects on a table and ran both methods on the placed objects. After each first run on a particular scene, we reconstructed the object positions manually. Table I shows the results. In four of the ten scenes, the POMDP outperformed the Baseline approach, while having otherwise identical performance.

There are two main reasons why the POMDP approach outperformed the Baseline approach. First, it planned its actions over the distribution of compositions. For example, in Scene 5 the POMDP succeeded while the Baseline approach finished execution prematurely because the most probable object composition did not contain red object hypotheses, although some other compositions did, as shown in Fig. 3. Second, the POMDP utilized information gathering actions. For example in Scene 1, it moved several non-red objects

Fig. 2. Cropped kinect images for each experimental scene. The order of images corresponds to the order of scenes in Table I.

away, thus reducing occlusion.

TABLE I

EXPERIMENTAL RESULTS FOR MOVING FULLY RED OBJECT(S) INTO AREA 1. $x$ IN THE ENTRY $V(x,y,z)$ DENOTES HOW MANY OBJECTS WERE MOVED TO THE CORRECT RED ($x$) OR NON-RED ($y$) AREAS, AND HOW MANY TO AN INCORRECT AREA ($z$). BOLD DENOTES HIGHER VALUES $V$ ($V = x - z$). POMDP OUTPERFORMED BASELINE IN FOUR SCENES AND HAD OTHERWISE IDENTICAL PERFORMANCE.

| Method \ Scene | 1 | 2 | 3 | 4 | 5 |
|---|---|---|---|---|---|
| Baseline | 0 (0,0,0) | 0 (0,0,0) | -1 (0,0,1) | 0 (0,0,0) | 0 (0,0,0) |
| POMDP | **1** (1,2,0) | 0 (0,2,0) | **0** (0,0,0) | 0 (0,0,0) | **1** (1,0,0) |

| Method \ Scene | 6 | 7 | 8 | 9 | 10 |
|---|---|---|---|---|---|
| Baseline | 0 (0,0,0) | 0 (0,0,0) | 0 (0,0,0) | 3 (3,0,0) | 0 (0,0,0) |
| POMDP | 0 (0,2,0) | **1** (1,1,0) | 0 (0,0,0) | 3 (3,0,0) | 0 (0,4,0) |

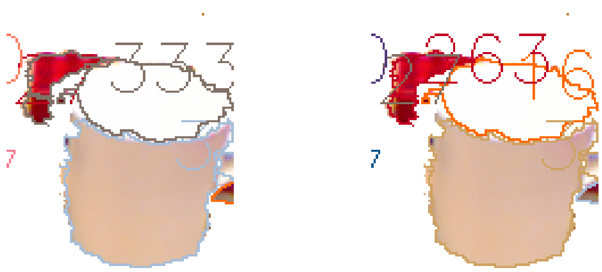

Fig. 3. Object compositions at time step 9 in scene 5 for the Baseline approach. The edges of each object hypothesis are marked with color. The baseline approach finishes prematurely because it considers only the most likely object composition and not the other compositions which may have red objects. **Left:** Most likely object composition: the (non-red) object hypothesis 333 contains both a red block and the top of a cup. **Right:** 2nd most probable object composition: the (red) object hypothesis 263 contains only the red block.

## VII. CONCLUSIONS

Manipulating unknown objects in a cluttered environment is a hard problem. A lack of object models and a noisy partial view make object discovery difficult. Instead of utilizing only the most likely object composition, we plan manipulation actions in the state space of object compositions. In experiments, our approach outperformed an approach based on the most likely object composition.

## REFERENCES


[1] H. Koppula, A. Anand, T. Joachims, and A. Saxena, "Semantic labeling of 3d point clouds for indoor scenes," in *Advances in Neural Information Processing Systems (NIPS)*, 2011, pp. 244–252.

[2] A. Karpathy, S. Miller, and L. Fei-Fei, "Object discovery in 3D scenes via shape analysis," in *Proc. of IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, 2013, pp. 2088–2095.

[3] A. Richtsfeld, T. Mörwald, J. Prankl, M. Zillich, and M. Vincze, "Segmentation of unknown objects in indoor environments," in *Proc. of IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE, 2012, pp. 4791–4796.

[4] ——, "Learning of perceptual grouping for object segmentation on RGB-D data," *Journal of visual communication and image representation*, vol. 25, no. 1, pp. 64–73, 2014.

[5] R. Bajcsy, "Active perception," *Proceedings of the IEEE*, vol. 76, no. 8, pp. 966–1005, 1988.

[6] M. Bjorkman and D. Kragic, "Active 3D scene segmentation and detection of unknown objects," in *Proc. of IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, 2010, pp. 3114–3120.

[7] L. Chang, J. Smith, and D. Fox, "Interactive singulation of objects from a pile," in *Proc. of IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, 2012, pp. 3875–3882.

[8] D. Katz, M. Kazemi, J. Bagnell, and A. Stentz, "Clearing a pile of unknown objects using interactive perception," in *Proc. of IEEE International Conference on Robotics and Automation (ICRA)*, 2013, pp. 154–161.

[9] A. Saxena, J. Driemeyer, J. Kearns, and A. Ng, "Robotic grasping of novel objects," in *Advances in Neural Information Processing Systems (NIPS)*, 2006, pp. 1209–1216.

[10] M. Popovic, G. Kootstra, J. A. Jørgensen, D. Kragic, and N. Krüger, "Grasping unknown objects using an early cognitive vision system for general scene understanding," in *Proc. of IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE, 2011, pp. 987–994.

[11] I. Lenz, H. Lee, and A. Saxena, "Deep learning for detecting robotic grasps," in *Proc. of Robotics: Science and Systems (RSS)*, 2013.

[12] D. Fischinger, M. Vincze, and Y. Jiang, "Learning grasps for unknown objects in cluttered scenes," in *Proc. of IEEE International Conference on of Robotics and Automation (ICRA)*. IEEE, 2013, pp. 609–616.

[13] D. Rao, Q. V. Le, T. Phoka, M. Quigley, A. Sudsang, and A. Y. Ng, "Grasping novel objects with depth segmentation," in *Proc. of IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE, 2010, pp. 2578–2585.

[14] J. Ilonen, J. Bohg, and V. Kyrki, "Three-dimensional object reconstruction of symmetric objects by fusing visual and tactile sensing," *The International Journal of Robotics Research*, vol. 33, no. 2, pp. 321–341, 2014.

[15] M. Dogar and S. Srinivasa, "A planning framework for non-prehensile manipulation under clutter and uncertainty," *Autonomous Robots*, vol. 33, no. 3, pp. 217–236, 2012.

[16] K. Hsiao, L. P. Kaelbling, and T. Lozano-Perez, "Grasping POMDPs," in *Proc. of IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, 2007, pp. 4685–4692.

[17] K. Hsiao, L. P. Kaelbling, and T. Lozano-Pérez, "Robust grasping under object pose uncertainty," *Autonomous Robots*, vol. 31, no. 2-3, pp. 253–268, 2011.

[18] J. Laaksonen, E. Nikandrova, and V. Kyrki, "Probabilistic sensor-based grasping," in *Proc. of IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 2012, pp. 2019–2026.

[19] E. Nikandrova, J. Laaksonen, and V. Kyrki, "Towards informative sensor-based grasp planning," *Robotics and Autonomous Systems*, vol. 62, no. 3, pp. 340–354, 2014.

[20] P. Monso, G. Alenya, and C. Torras, "POMDP approach to robotized clothes separation," in *Proc. of IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE, 2012, pp. 1324–1329.

[21] J. Pajarinen and V. Kyrki, "Robotic manipulation of multiple objects as a POMDP," *ArXiv e-prints*, 2014. [Online]. Available: http://arxiv.org/abs/1402.0649