POMDP based manipulation planning in object composition space

Joni Pajarinen and Ville Kyrki

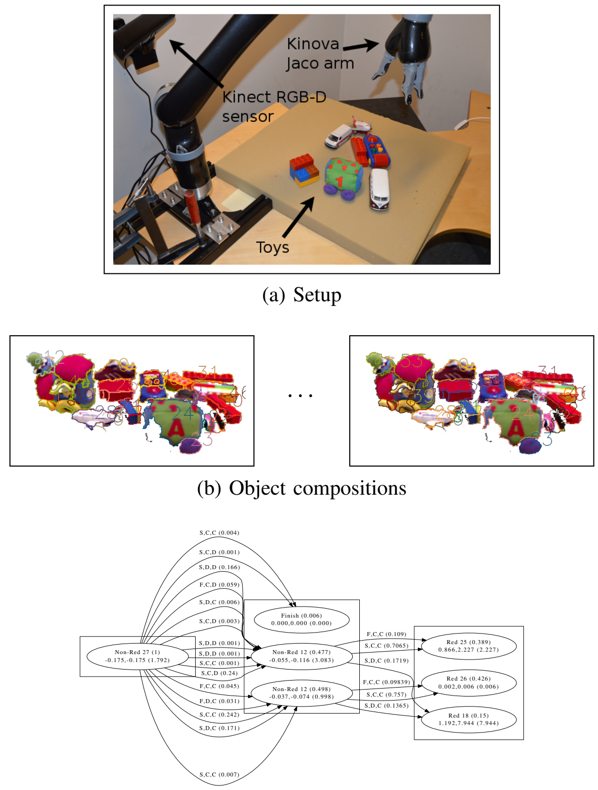
I. INTRODUCTION

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. When multiple measurements can be acquired around an isolated object, standard approaches work satisfactorily as the generated 3-D models can often be used for successful manipulation.

However, in cluttered scenes with multiple unknown objects, the segmentation of objects, also known as object discovery in perception research [1], [2], [3], [4], 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. Because of this uncertainty, earlier work has concentrated on finding the best manipulation action based on the most likely composition. Contrary to earlier work, we 1) utilize a probability distribution over 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 hypotheses on the actual task to be performed. We cast the manipulation planning problem as a partially observable Markov decision process (POMDP) [5] 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. By considering a temporally evolving system, the robot can infer from past grasp attempts the likelihood of object hypotheses. Our approach [6] combines earlier ideas of interactive perception [7], [8], [9], [10], [11] and learned composition priors [3], [4] in a planning under uncertainty framework.

II. 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. While manipulating unknown objects is difficult, occlusion and noisy sensor readings make the task even harder: the robot has to guess which parts of the captured RGB-D image belong to the same object. We propose to choose the manipulation 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) Setup

(b) Object compositions

(c) POMDP plan

Fig. 1. Overview. (a) At each time step, the robot uses an RGB-D sensor to observe unknown objects, which may occlude each other. (b) The robot over-segments the RGB-D image into patches and creates a probability distribution over all possible object compositions from patches. (c) Using this probability distribution the robot computes a POMDP plan. The plan takes into account uncertainty in object composition, observations, and grasp success. The POMDP plan is represented as a compact directed graph where each node corresponds to an action and each edge is a possible observation.

Fig. 1 shows an overview of the system. At each time instant, the robot performs the following steps:

1) Over-segment an RGB-D image of the current scene
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 II-A)
4) Use a POMDP to select the best long-term manipulation action for the current object distribution and execute the action (Section II-B)

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

For steps 1 and 2, for segmenting RGB-D data and estimating probabilities for segment pairs, we use an existing approach from [4]. Steps 3 and 4 are described below.

A. Robotic manipulation as a POMDP

We model robotic manipulation as a POMDP. A POMDP defines optimal behavior for an agent (robot) in an uncertain world with noisy, partial measurements. In particular, a POMDP assigns the correct long-term value to informative actions which are needed when exploring object hypotheses. Transition probabilities define how the world state may change when the robot performs an action, and observation probabilities define the sensor model. The robot does not observe the current world state directly but can make decisions based on a probability distribution over states, called the belief. A reward, given according to the current state and robot action, specifies the objective. The goal is to maximize the total reward over several time steps.

In our POMDP model [12], [6], the probabilities of successfully grasping an object hypothesis, and observing its attributes (for example color) depend on how occluded the object hypothesis is. A world state contains a composition of object hypotheses (called an object composition), object attributes, locations, and history information. Because each state may consist of different object hypotheses the transition and observation probabilities are state specific. For example, a certain world state may contain a set of three specific objects, while another state may contain a set of two different objects with different probabilities.

Because of the huge state space we use a particle representation for the belief, that is, a set of states and their probabilities. To create a belief from an RGB-D image, we sample particles using segment pair probabilities computed from the RGB-D image. For sampling we use a Markov chain Monte Carlo approach, which discards samples that do not conform with the observed history. The key insight is that previously performed grasps must have failed for a valid object hypothesis, otherwise the object would have been moved. Furthermore, a grasp can only succeed for a wrong object hypothesis, when the object is part of another object, for which the grasp succeeds.

B. Manipulation planning

In order to use the POMDP method in [12] for planning, we need to model the temporal evolution of the world state. Because our 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.

Actions. In our problem setting, the robot may grasp an object and move it. Our approach based on principal component analysis (PCA) grasps a narrow part of the target object top-down. In order to restrict the computational load, instead of allowing grasping of all possible object hypotheses, we greedily select a fixed amount of possible grasps according to the total grasp probability over all object hypotheses.

State sampling. As discussed earlier, a world state consists of an object composition, and contains for each composition a semantic object location, attributes, and history. To sample a new state for a grasp action, we select the object hypothesis that has the highest grasp probability for the action. Sample grasp success and if the grasp fails, increase the grasp failure of the object. If moving the object succeeds, change the semantic location of the object.

Observation sampling. After executing a grasp, 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. The observation probability for each object depends on how occluded the object is.

Observation probability. The probability of making an observation 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 depends on occlusion.

III. EXPERIMENTS & CONCLUSION

Fig. 1a shows the experimental setup. The robot tries to find and move an unknown number of fully red toys into a target zone (Box 1). To remove occlusions, the robot can move toys to a free zone (Box 2). Moving a red object into Box 1 yields 1$. Moving a non-red object into Box 1 or moving a red object into Box 2 costs 1$. Table I shows the results. In ten different scenes, the POMDP approach was better in five and a baseline approach, based on the most likely object composition, in one. 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 one scene 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. Second, the POMDP utilized information gathering actions. For example, in another scene, it moved several non-red objects away, thus reducing occlusion.

To summarize, a lack of object models and a noisy partial view make object discovery difficult. In experiments, our POMDP based approach which plans over different possible object compositions and takes uncertainty into account outperformed a baseline approach.

TABLE I

MOVING RED OBJECT(S) INTO A BOX. x IN $V(x,y,z)$ DENOTES HOW MANY OBJECTS WERE MOVED TO THE CORRECT RED (x) OR NON-RED (y) BOXES, AND HOW MANY TO AN INCORRECT BOX (z). BOLD DENOTES HIGHER VALUES V ($V = x - z$). POMDP PERFORMED BETTER IN FIVE AND BASELINE IN ONE SCENE.

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

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

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] L. P. Kaelbling, M. L. Littman, and A. R. Cassandra, "Planning and acting in partially observable stochastic domains," *Artificial Intelligence*, vol. 101, no. 1-2, pp. 99–134, 1998.

[6] J. Pajarinen and V. Kyrki, "Robotic manipulation in object composition space," in *Proc. of IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE, 2014, pp. 1–6.

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

[8] 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.

[9] 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.

[10] 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.

[11] H. van Hoof, O. Kroemer, and J. Peters, "Probabilistic segmentation and targeted exploration of objects in cluttered environments," *IEEE Transactions on Robotics*, vol. 30, no. 5, pp. 1198–1209, 2014.

[12] J. Pajarinen and V. Kyrki, "Robotic manipulation of multiple objects as a POMDP," *Artificial Intelligence*, 2015. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0004370215000570

[13] ——, "Decision making under uncertain segmentations," in *Proc. of IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, 2015.