Robotic manipulation of multiple objects as a POMDP

Joni Pajarinen and Ville Kyrki
Department of Electrical Engineering and Automation
Aalto University, Finland

I. Introduction

Service robots are being used more and more in complex unstructured environments. In these kind of environments, robotic manipulation has become important for performing a wide variety of tasks such as moving coffee cups into the dishwasher, clearing toys, or handing over medicine. However, manipulation of objects using imperfect sensors is hard, especially when objects occlude each other.

In this paper, we investigate multi-object manipulation in crowded scenes. Because of noise in the captured images, and object occlusions, the robot has incomplete information about the objects and the environment. Moreover, we assume that the dynamics of environment are partly unknown supporting the goal of long term robotic autonomy. A high-level research question is in which situations explicit planning under uncertainty is beneficial. Non-myopic information gathering actions could be beneficial in situations in which the robot does not directly observe object attributes, for example, when an object is occluded the robot could try to remove the occlusion in order to see the object better. Moreover, the robot could benefit from considering long term effects of actions. For example, when occlusions make grasping harder, the optimal combination of information gathering actions and actions that move objects away, may be non-trivial.

Because of the uncertainty about the state of the world, dynamics, and action effects we model multi-object manipulation of crowded occluded objects as a partially observable Markov decision process (POMDP). A POMDP takes into account both partial and noisy sensor readings, and the uncertainty in world transitions. In a POMDP, the objective is encoded using a reward function and therefore our multi-object manipulation approach can be used in various setups. Because occlusion is a dominating source of uncertainty in crowded scenes, the probabilities in our POMDP model depend on a model-free occlusion ratio parameter which indicates how much an object occludes another one. Furthermore, because objects differ and we have no prior object model, we adapt the grasp probability of each object according to grasp success observations, also during POMDP planning.

In order to plan manipulation actions, we present a new approximate POMDP method that produces compact policies for large complicated problems, allowing a human to inspect the policy and potentially modify problem preferences. In addition to offline planning, the method can be used for online planning which allows re-estimation of the current belief from sensor measurements at each time step, correcting for "mistakes" in the world model.

While POMDPs have been used in different robotic applications, to the best of our knowledge they have previously not been applied to multi-object manipulation in an (partly) unknown crowded high dimensional environment. In the experiments, we evaluate our approach in simulation and in physical experiments with a real robotic arm. The experimental results indicate that 1) multi-step planning under uncertainty is beneficial for multi-object manipulation in crowded environments, and that a greedy approach is not enough, 2) online planning allows adapting the model dynamics according to experience.

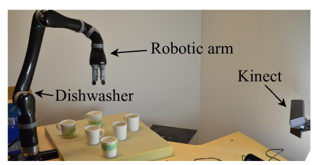

Fig. 1. Experimental setup. **Left:** A Kinova Jaco robotic arm manipulates objects on a table. A Microsoft Kinect captures RGB-D point clouds. In the experiments, robot tries to pick up dirty objects, here cups with green parts, from the table and place them into the "dishwasher" (blue box on the far left). **Right:** Cropped Kinect image. Object edges in blue.

II. Multi-object manipulation as a POMDP

In this paper, in each time step, a visual sensor captures an image, the robot decides which manipulation action to perform based on the captured image, and then a robotic arm executes the action. We consider an environment where multiple objects may occlude each other. In addition, the image captured by the visual sensor is noisy. We assume that a model of the objects is not available, meaning, that grasping an object may fail. We will now shortly describe what a POMDP is and then discuss how to model multi-object manipulation as a POMDP.

In a POMDP, at each time step: 1) the agent (the robot) executes an action, 2) the agent receives a reward depending on the current state and executed action, 3) the system moves to a new state according to the transition probability conditional on the current state and executed action, and 4) the agent makes an observation according to the observation probability conditional on the new state and executed action. When given the transition and observation probabilities, a reward function, and a probability distribution over current world states, a

POMDP defines the optimal action to execute. We will next describe how we model the state of the world in multi-object manipulation, and how we estimate transition and observation probabilities in crowded scenes.

State space and actions. In our multi-object manipulation model, the state space consists of semantic object locations (e.g. “on table”, “in a dishwasher”), object attributes (e.g. “clean”, “dirty”), and a history of observations and action successes for each object. The model assumes constant semantic object locations over time unless a manipulation action successfully changes those locations. Note that because we use online planning, the planning always restarts from a belief that takes into account the most recent sensory data.

Occlusion ratio. In crowded environments, occlusions can be the dominating source of uncertainty. Grasping an object when behind another object, and thus occluded, is usually harder than when the object is not occluded at all. Similarly, observing the object attributes correctly is harder when the object is occluded. Therefore, our transition and observation probabilities depend on a model-free occlusion ratio parameter: with more occlusion the probability of action success and the probability of making correct observations decrease. We compute the occlusion ratio using 2D object edge information.

Observation and grasp probabilities. We assume that the semantic locations and dependencies (which cup is in front of which cup) are fully observed and that grasp success is also fully observed. Both the probability to successfully grasp an object and the probability to correctly observe an object’s attributes decrease when the occlusion increases. The grasping success probability differs between objects. Therefore, in addition to the occlusion specific grasp probability, each object has an inherent grasp success probability which is not known a priori but, instead, learned from observations.

POMDP method. In order to construct compact plans for the high dimensional problem of multi-object manipulation, both offline and online, we present a new POMDP method [1]. The method is based on the offline monotonic policy graph improvement algorithm for “flat” state representations in [2]. It uses a particle based probability distribution representation to improve a fixed size policy graph iteratively. Because the policy graph has fixed size, computation time can be fixed in advance. Moreover, the resulting compact policy can be inspected by a domain expert to gain insight into the problem.

III. Dirty cups into the dishwasher

In the experiments, a 6-DOF Kinova Jaco robot arm tries to move dirty cups into a dishwasher; a Microsoft Kinect RGB-D sensor observes the scene. The robot can either lift an object, move an object into the dishwasher, or finish. Lifting an object allows the robot to see behind it accruing a reward of -1 for the time spent. Moving a dirty cup into the dishwasher yields a reward of $+5$, and moving a clean cup into the dishwasher -10. Finishing will yield a -5 for each dirty cup on the table. The left side of Fig. 1 shows the experimental setup. Note that our multi-object manipulation approach is not restricted to this specific setup or objects. The right side of Fig. 1 shows an example of an Kinect image. Object edges are marked with a blue color.

In the experiments, we initialized parameters of the grasp and observation probabily distributions by trying out grasps using the Kinova Jaco arm and observing objects for different occlusion ratios. We ran two sets of experiments. In the first set we verified world model assumptions using the Kinova Jaco arm (see Section 4.2 in [1] for details). In the second set we simulated world dynamics (which were estimated from real visual scenes and real robotic grasps) in order to get a large number of repetitions. For this set Fig. 2 shows experimental results for ten scenes captured by a Kinect sensor (Fig. 3 in [1] shows the scenes). The POMDP approach performed better than greedy heuristics. Perhaps surprisingly, planning three time steps into the future worked better than planning for only two time steps. Intuitively, one could imagine that simple action sequences such as lift cup up, and if cup behind lifted cup is dirty, then move it into dishwasher could perform as well as more complicated plans. However, the resulting conditional POMDP plans contain complex behavior (see Fig. 5 in [1] for an example).

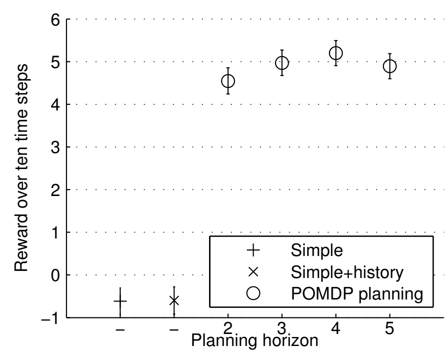

Fig. 2. The average reward sum and its 95% confidence interval (computed using bootstrapping) for the heuristic manipulation approach, heuristic manipulation approach utilizing grasp history information, and for the POMDP planning method.

IV. Conclusion

This paper shows how to model multi-object manipulation in crowded environments as a POMDP. In the proposed model observation and grasp success probabilities depend on object occlusions, a significant source of uncertainty in crowded scenes. The model adapts grasp probabilities according to observations. Experimental results show that greedy heuristic approaches are not sufficient, and that multi-step POMDP planning achieves higher performance.

References

[1] J. Pajarinen and V. Kyrki. Robotic manipulation of multiple objects as a POMDP. *ArXiv e-prints*, 2014.

[2] J. Pajarinen and J. Peltonen. Periodic Finite State Controllers for Efficient POMDP and DEC-POMDP Planning. In *Advances in Neural Information Processing Systems 24 (NIPS)*, pages 2636–2644. 2011.