Mobile Manipulation Object Search using Co-occurrences and Capacity

Lawson L.S. Wong, Leslie Pack Kaelbling, and Tomás Lozano-Pérez

***Abstract*— Object search is an integral part of daily life, and in the quest for competent mobile manipulation robots it is an unavoidable problem. Previous approaches have suggested that object co-occurrence information is useful in object search. We present a novel generative model for representing both co-occurrence structure and spatial constraints, and use this to perform belief-space planning in object search. We demonstrate the model on a detailed simulation involving a PR2 robot.**

I. Introduction

Consider the following example of searching for a large mixing bowl in the kitchen when preparing a meal. There are three cupboard shelves that have been partially viewed, two containing stacks of plates, the other some dish detergent. Other objects appear to be in the back, but they are occluded, and we need to remove the objects in front to continue searching the shelves. Which shelf should we look at?

Even though the bowl has not been observed yet, it seems intuitive to us to keep looking on a shelf with plates because they have closer function to bowls. Suppose we further observe that of the two shelves with plates, one is small, the other large. If we assume a bowl, if present, is equally likely to be anywhere in the cupboard, and that both regions incur the same exploration cost, then the large region is more desirable to look at since more objects can be expected. Moreover, given that mixing bowls are usually large, we may even determine that the bowl cannot fit in the small region and eliminate that from consideration.

The above example illustrates two aspects of object search we wish to capture in our model. First, certain categories of objects tend to co-occur with each other, such as plates and bowls in the example. We will be agnostic to the causes of similarities (e.g., function, shape, color, etc.), and simply rely on co-occurrences as an indicator of similarity. Second, in the latter half of the example, we also used geometric information to choose between unseen regions. Here we understood that object types have various volume distributions, and knew that unseen regions had some remaining capacity. Our model captures both intuitions illustrated above.

Works such as [**?**] and [**?**] have previously demonstrated that co-occurrence information is useful for guiding object search. In particular, both use object co-occurrences to determine the compatability of the target object with various household locations, given observations of other objects in the vicinity. However, neither consider geometric constraints, nor are they easily extendible to do so. Moreover, manipulation is not integrated in either framework.

Computer Science and Artificial Intelligence Laboratory, Massachusetts Institute of Technology, Cambridge, MA 02139 {lsw,lpk,tlp}@csail.mit.edu

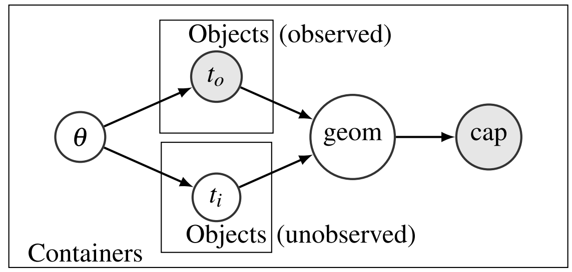

Fig. 1. Graphical representation of our model; see text below for details.

II. Model and Execution

We represent co-occurrence information and spatial constraints using the generative model shown in Figure **??**. We assume the environment is partitioned into a set of containers, each containing both seen and unseen object types. To model type similarity, we introduce the notion of a location's *composition* (θ), a latent distribution over object types, and supply a logistic-normal prior that captures type co-occurrences. The observed object types in the container then inform about the unknown ones via θ. By reasoning about the total volume that objects may occupy, we can check whether they fit within the known capacity of the container.

To find the target object, an agent needs to determine which containers likely contain the target object. This can be inferred from the above model, by determining in each container how likely the target object is one of the currently unobserved objects (if any exist). There are also costs associated with switching between containers and performing manipulation actions (e.g., moving occluding objects out of the way to reveal hidden objects). To tradeoff these costs, we use the online belief-space planner described in [**?**] to select next actions. The planner uses the probability that the query object type is in each container to determine whether to continue searching the current container, or to switch to another container if the current one seems unpromising.

The above object search framework was implemented in simulation with a mobile manipulator modeled on the Willow Garage PR2 robot. A sample execution can be found at: `http://youtu.be/cnWK8aBHmu0`

References

[1] T. Kollar and N. Roy, "Utilizing object-object and object-scene context when planning to find things," in *ICRA*, May 2009.

[2] M. Hanheide, C. Gretton, R. Dearden, N. Hawes, J. L. Wyatt, A. Pronobis, A. Aydemir, M. Göbelbecker, and H. Zender, "Exploiting probabilistic knowledge under uncertain sensing for efficient robot behaviour," in *IJCAI*, July 2011.

[3] L. P. Kaelbling and T. Lozano-Pérez, "Pre-image backchaining in the belief space for mobile manipulation," in *ISRR*, 2011.