Abstract

Robot manipulation in clutter with objects in physical contact remains a challenging problem till date. The challenge is posed by interaction involved among the objects at various levels of complexity. Understanding positional semantics of the environment plays an important role in such tasks. The interaction with surrounding objects in the environment must be considered in order to perform the task without causing the objects fall or get damaged. In our work, we learn the semantics in terms of support relationship among different objects in a cluttered environment by utilizing various photometric and geometric properties of the scene. To manipulate an object of interest, we use the inferred support relationship to derive a sequence in which its surrounding objects should be removed while causing minimal damage to the environment. We believe, this work can push the boundary of robotic applications in grasping, object manipulation and picking-from-bin, towards objects of generic shape and size and scenarios with physical contact and overlap.

In the first part of the thesis, we aim at learning semantic interaction among objects of generic shapes and sizes lying in clutter involving both direct and indirect physical contact. Three types of support relationships are inferred: "Support from below", "Support from side", and "Containment". Subsequently, the learned semantic interaction or support relationship is used to derive a sequence or order in which the objects surrounding the object of interest should be removed without causing damage to the environment. The generated sequence is called Support Order. We have proposed and analysed two alternative approaches for support inference. In the first approach "Multiple Object Support Inference", support relations between all possible pairs are inferred. In the second approach "Hierarchical Support Inference", given an object of interest, its support relationship with other graspable objects is inferred hierarchically. The support relationship is used to predict the "support order" or the order in which the surrounding objects need to be removed in order to manipulate the target object.

In the second part of the thesis, we attempt to learn the semantic interaction among different objects in clutter using multiple views. At first, support relationship among objects in each view is inferred. Then the inferred support relationships are combined to define support relationships across multiple views. The combined global support relationship is used to recover missing support relations and predict the support order. Support order is the order in which objects surrounding an object of interest should be removed. The support order predicted using global support relationship incorporates hidden objects and missing spatial support relations.

We have created two RGBD datasets consisting of various objects used in day-to-day life present in clutter. In "Indoor dataset for clutter", 50 cluttered scenes are captured from frontal view using 35 objects of different shapes and sizes. In "Indoor multiview dataset", 7 cluttered scene are captured.

Each scene each captured from multiple views. In this dataset, total 67 images are captured using 9 objects of different shapes and sizes. The dataset is made publicly available for the research community around the world. We explore many different settings involving different kind of object-object interaction. We successfully learn support relationships and predict support order in these settings. It can play significant role in extending the scope of manipulation to cluttered environment involving both direct and indirect physical contact, and generic objects.

Keywords: Robotic Vision, Support Relations, Support Order, RGBD, Semantic Interaction, Clutter, Multiple Views