Large-Scale Object Model Generation for Learning Contact-Rich Robotic Manipulation Tasks

Ryan Diaz, Chahyon Ku, Karthik Desingh

April 2023

1 Project Overview

1.1 Background and Motivation

In daily household contexts, there may be simple tasks that some people may begin to struggle with, such as putting a cap on a bottle. Our long-term goal is to develop robotic agents that can be placed in these contexts and are able to accomplish these simple tasks that might give people some trouble. In order to train these robotic agents to interact with the objects around them in a useful way, we will make use of data-driven methods.

Data-driven methods involving the use of large amounts of data to train deep neural networks to solve problems have been developed and successfully applied to various fields in computer science such as computer vision and natural language processing. Due to this success, large-scale databases of object models such as ShapeNet [1] have been developed to train and benchmark models for various applications in computer vision and computer graphics. We apply the data-driven method to the field of robotic perception and manipulation by developing a method of programmatically generating a large-scale data set of 3D object models for the "cap-the-bottle" task for use in simulation in order to train a robotic agent to perform the task.

1.2 Methods and Results

We model the "cap-the-bottle" task as a simplified task of putting two objects with a geometric relationship together without any auxiliary actions such as twisting or excessive pushing. The "cap" in this context will be a cuboid with an extrusion of a particular shape, and the "bottle" will be a cuboid with an intrusion of the same shape.

The generation of the objects in the data set was done via the Blender Python API [2]. 9 shapes were used for the intrusions and extrusions of the objects. These shapes were grouped into 3 distinct categories (Figure 1): shapes with one symmetry and thus one way to fit them together (arrow, key, U), shapes with two to four symmetries with few ways to fit them together (line, diamond, cross), and shapes with many symmetries and ways to fit together (pentagon, hexagon, circle). The task is considered successfully completed if the intrusion and extrusion shapes are lined up; the cuboid bodies do not necessarily need to be lined up as well.

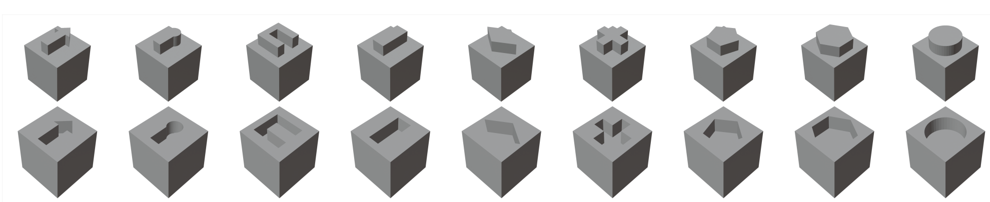

Figure 1: Base cap and bottle models for 9 shapes

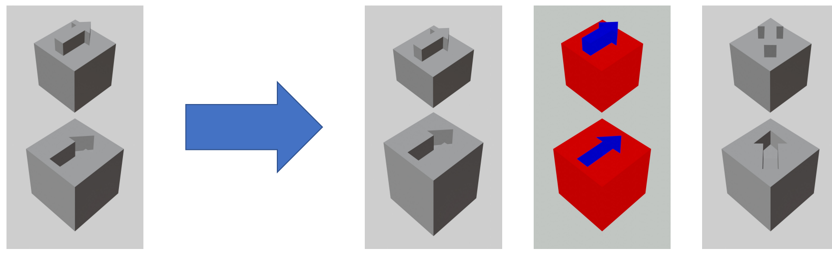

Figure 2: Various transformations of the base object models

The cubes have dimensions of 8 x 8 x 8 cm, with the extrusions having a height of about 2.5 cm. The intrusions are slightly increased in size to give a uniform tolerance (that is, along each side of the shape) of 4 mm compared to the extrusions. This tolerance is introduced in order to afford some error when these objects are 3D-printed and used in a real-life version of the simulated task. In order to introduce a higher degree of diversity in the eventual large-scale data set, parameters are implemented that can be applied to the base object models and transform them in some manner (Figure 2). These transformations include rotating the extrusions and intrusions on the cube face, adding textures to the cubes such that the intrusions and extrusions have a distinct color from the cuboid bodies, and adjusting the heights of the cube bodies to be either taller or shorter than the standard 8 x 8 x 8 cm model.

In order to eventually evaluate the generalization ability of a model trained on this generated data set, cubes with more complex intrusions and extrusions were created. These new intrusion and extrusion shapes are selected from the ShapeNet database, with convex models being chosen from the data set such that the extrusion shape is able to physically fit inside of its corresponding intrusion. These ShapeNet models will give us a test data set that will allow us to evaluate the quality of our generated objects used for training.

With the base object models and the parameters that can be used to transform them, we have a method of generating a large and varied data set of objects specialized for the "cap-the-bottle" task. Our method provides easy extensibility by allowing for new intrusion and extrusion shapes to be added, with a single new shape allowing for a large variety of new models to be generated with the transformation parameters we have developed. Ultimately, this large data set represents our application of the data-driven method in training a robotic agent to perform this specific task

1.3 Objective Accomplishments

This project was developed in three distinct parts. The first phase, lasting about two weeks, involved learning the Blender software as well as how to develop with the provided Python API. During this time, the idea of modeling the "cap-the-bottle" task as fitting together two cuboid shapes was also being developed. The next month was spent creating a program to automatically generate the base intrusion and extrusion models in Blender given a manually created 3D model of the intrusion and extrusion shape. The following month was then allocated towards developing the parameters that would be used for transforming the base object models, as well as implementing a standardized and uniform tolerance for the intrusions. Throughout the process, several of the generated models were 3D-printed in order to check their feasibility in the real-world task, and to match their dimensions with the robotic gripper that would be interacting with them. Progress checks consisted of weekly individual and group meetings, where we discussed the work that was done over the previous week and formed a plan for the work that would be done over the following week.

2 Evaluation of Experience

I greatly appreciate the Undergraduate Research Scholarship program for providing me with the opportunity to connect with a faculty member and graduate students and learn how the research process in computer science really works. Throughout this project I have learned many valuable skills such as conducting literature reviews, discussing and iterating upon the work that I am doing, and summarizing and presenting my research to a general audience. Joining Professor Desingh's Robotics: Perception and Manipulation (RPM) lab has exposed me to the new and exciting research projects going on in the field, and has allowed me to collaborate and make connections with other lab members. As a result of this experience, I am more confident in my decision to pursue graduate school in computer science, as now I have a better idea of what research in computer science entails, and this project has motivated me to pursue similar lines of work in the future. I would again like to thank the Office of Undergraduate Research for their generosity in allowing me to pursue this opportunity. I would also like to specifically thank Professor Desingh and graduate student Chahyon Ku for their mentorship, guidance, and their efforts to make me feel included in the lab. I would not have gotten as much out of this project as I did without their help!

References

[1] Angel X. Chang, Thomas Funkhouser, Leonidas Guibas, Pat Hanrahan, Qixing Huang, Zimo Li, Silvio Savarese, Manolis Savva, Shuran Song, Hao Su, Jianxiong Xiao, Li Yi, and Fisher Yu. Shapenet: An information-rich 3d model repository, 2015. cite arxiv:1512.03012.

[2] Blender Online Community. Blender - a 3d modelling and rendering package. 2018.