Sketching Affordances for Human-in-the-loop Robotic Manipulation Tasks

Sina Masnadi[1], Joseph J. LaViola Jr.[1],
Jana Pavlasek[2], Xiaofan Zhu[2], Karthik Desingh[2] and Odest Chadwicke Jenkins[2]

I. Introduction

In order for robots to perform complex autonomous manipulation tasks in natural human environments, they must be able to carry out dexterous actions, beyond pick-and-place, on a wide variety of objects. The actions which are available to a given agent over a set of objects are called *object affordances*. The concept of affordance was first introduced by psychologist J. J. Gibson [1] and has garnered much interest in the robotics community in recent years [2], [3]. How to best acquire these affordances is an open problem in robotics. Leveraging human knowledge of object affordances can enable robots to interact with a variety of common household objects to perform complex tasks, such as making a cup of coffee.

We propose to enable a human user, without expert knowledge about robotics and programming, to transfer knowledge about affordances in a given scene to a robot. To this end, we propose an easy-to-use system to acquire object geometries and their associated affordances through sketching on a graphical interface. This allows users to interact with robotic systems by utilizing sketch-based techniques to provide a straightforward user interface, as shown in Figure 2. The user sketches the geometry of the object and its affordances. During task execution, when the robot encounters the objects for which it has affordance information, it can execute the affordances by registering the object geometries to its RGB-D data and then performing actions sequentially to achieve the goal.

Sketching is a low overhead means of rapid communication with the robot which can quickly provide information about the environment and the actions available to the robot. It provides the user with the ability to quickly demonstrate unseen tasks to the robot in the form of affordance templates [4]. Our system facilitates a human-in-the-loop approach which enables the robot to perform a large variety of manipulation tasks on different objects.

Access to object 3D mesh models can assist in object pose estimation, but many methods for this task [5]–[7] rely on synthetically designed mesh models. In cluttered scenes, sketching over RGB-D data from the robot can provide precise geometries with very little overhead. Maghoumi et al. introduced a sketch-based system to extract 3D geometries from point clouds in cluttered scenes which recovers geometries even when objects are occluded [8]. We are extending this technique to handle more complex objects. The system allows us to segment an object into its parts and can provide the geometries and affordances over object parts. Affordance templates can later be registered onto objects from the depth image data in arbitrary environments using their parts-based geometries using previously developed pose estimation techniques [9].

[1]Department of Electrical Engineering & Computer Science, University of Central Florida, {masnadi, jjl}@cs.ucf.edu
[2]Department of Electrical Engineering & Computer Science, Robotics Institute, University of Michigan, Ann Arbor {pavlasek, zhuxf,kdesingh,ocj}@umich.edu

II. Sketch-Based Affordance Templates

Our current work provides tools to sketch on the 3D point cloud or the 2D image projected on the 3D point cloud from the robot's RGB-D data. The extracted meshes are automatically fitted to their location on the scene and overlayed on the point cloud. The user can sketch affordances on a desired object and generate its corresponding affordance template. The sketched affordances are represented as *affordance waypoints*. An affordance waypoint has the following attributes: gripper state (position, orientation, and open/closed state), type, and axis.

The waypoint's *type* represents whether it is a *Direct Affordance Waypoint (DAW)* or an *Assisted Affordance Waypoint (AAW)*. In the sketching system, *DAWs* are created by directly circling a point of interest on the scene's 3D point cloud or 2D image. The sketched circle is used to cast a ray on the point cloud and add the affordance waypoint in the location that the ray hits the point cloud. *AAWs* are created using *helpers*. A *helper* is a set of points, lines, and axes which helps the user to precisely move the affordance waypoints or create them based on constraints. These helpers are: *perpendicular line helper*, *parallel line helper*, *rotation axis helper*, and *surface helper*. Using a collection of helpers, affordance waypoints can be moved precisely to create complex object part articulations. Figure 1 shows how an axis helper is used to extract a coffee maker's lid opening affordance.

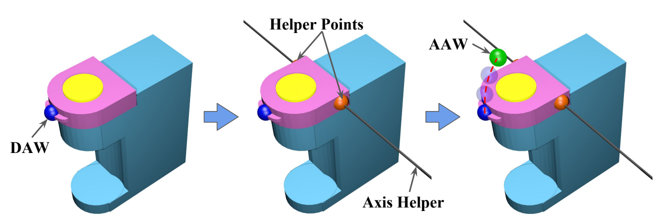

Fig. 1: Extracting the lid opening affordance for a coffee maker. The initial grasp point is defined by a DAW. Using sketched helper points, the user creates an axis helper. The axis helper will be a constraint for moving an AAW. A DAW will be converted to an AAW when it is constrained by a helper. In this case the existing initial grasp point is duplicated and rotated around the axis helper to create an AAW which is the final state waypoint.

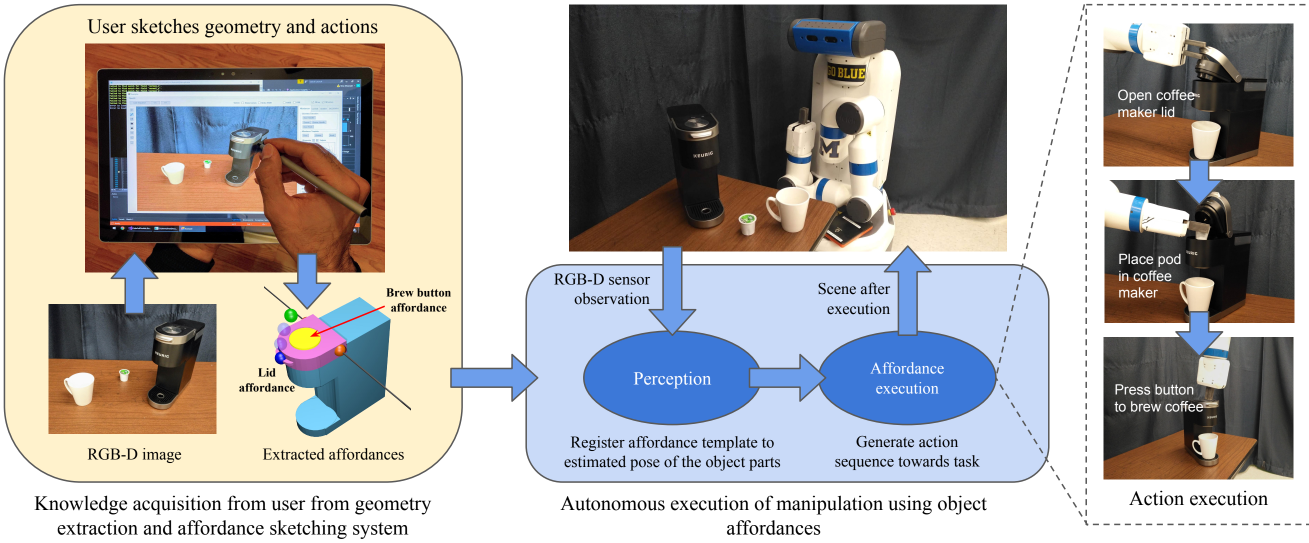

Fig. 2: An overview of the proposed sketch-based affordance extraction and object manipulation for the task of making coffee. The user sketches the affordances of the coffee maker which describe how to open the lid, insert the pod, and press the brew button. The robot executes a sequence of actions to perform the task of brewing coffee using these affordances. For this illustration, we ignore the placement of the coffee cup and assume that action is given.

The waypoints and helpers can be combined to generate more complex interactions such as rotating a door handle to open a door, opening and closing drawers, and pushing a button. Creating waypoints for such interactions is not feasible using DAWs, since DAWs are limited to the point cloud's point locations. Using helpers, waypoints can be precisely moved in the 3D space based on the helper's constraints.

This sketched information fits naturally into the affordance template framework. Object geometries, action descriptors and pre- and post-conditions for the actions can be saved to a database and referenced for action execution at a later time.

III. Preliminary Results

We conducted preliminary experiments to open a cabinet drawer and manipulate a door handle. We extracted geometries of the door handle and drawer handle using the sketching interface. Additionally, we extracted waypoints describing how these parts move to perform an action. For example, the door handle requires a rotating action and drawer handle requires a translation action. Waypoints describing these actions along with object geometries gives us affordance templates that can be used by a robot to perform action. We tested this 10 times with 10 different users providing us with 10 geometries and associated affordance templates. Robot completed the task successfully for all 10 different affordance templates for both the door handle and the drawer handle. In these trials, the affordance templates were manually registered onto the point clouds before action execution in RViz. In future work, autonomous registration of the templates onto the point cloud can be done by performing pose estimation using geometrical models [7], [9].

IV. Conclusion and Future Work

We propose a system for human-in-the-loop task execution which allows a user to sketch affordance templates on an easy-to-use GUI. The robot can then use these affordance templates to interact with a variety of objects. We have demonstrated the ability to execute manipulation of a door handle and drawer handle on a mobile manipulator using this system. We intend to build upon these successes to provide more complex information about objects to the robot which can be used towards full autonomous task completion, such as making coffee. We will also autonomously register the affordance templates to the point cloud observations. We plan to explore how we can leverage the information from the sketching system to generalize affordances over multiple instances of an object. Sketch-based affordance extraction can be done quickly and is therefore highly scalable. It can potentially be used to acquire large amounts of labeled data of object affordances which can be applied to data-driven techniques.

References

[1] J. J. Gibson, *The Theory of Affordances*. Hillsdale, NJ: Lawrence Erlbaum Associates Publishers, 1977, ch. 8, pp. 67–82.

[2] H. Min, C. Yi, R. Luo, J. Zhu, and S. Bi, "Affordance research in developmental robotics: A survey," *IEEE Transactions on Cognitive and Developmental Systems*, vol. 8, pp. 237–255, 2016.

[3] L. Jamone, E. Ugur, A. Cangelosi, L. Fadiga, A. Bernardino, J. Piater, and J. Santos-Victor, "Affordances in psychology, neuroscience, and robotics: A survey," *IEEE Transactions on Cognitive and Developmental Systems*, vol. 10, pp. 4–25, 2018.

[4] S. Hart, P. Dinh, and K. Hambuchen, "The affordance template ROS package for robot task programming," in *Proceedings of the IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, May 2015, pp. 6227–6234.

[5] Z. Sui, O. C. Jenkins, and K. Desingh, "Axiomatic particle filtering for goal-directed robotic manipulation," in *International Conference on Intelligent Robots and Systems (IROS)*. IEEE, Sept 2015, pp. 4429–4436.

[6] V. Narayanan and M. Likhachev, "Discriminatively-guided deliberative perception for pose estimation of multiple 3d object instances." in *Robotics: Science and Systems*, 2016.

[7] Z. Sui, Z. Zhou, Z. Zeng, and O. C. Jenkins, "SUM: Sequential scene understanding and manipulation," in *International Conference on Intelligent Robots and Systems (IROS)*. IEEE, 2017.

[8] M. Maghoumi, J. J. LaViola, K. Desingh, and O. C. Jenkins, "Gemsketch: Interactive image-guided geometry extraction from point clouds," in *2018 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, 2018, pp. 2184–2191.

[9] K. Desingh, S. Lu, A. Opipari, and O. C. Jenkins, "Factored pose estimation of articulated objects using efficient nonparametric belief propagation," in *International Conference on Robotics and Automation (ICRA, to appear)*. IEEE, 2019.