Understanding the Role of Haptics in Affordances

Vivian Chu*, Andrea L. Thomaz*,
*School of Interactive Computing
Georgia Institute of Technology, Atlanta, Georgia 30332–0250
Email: vchu@gatech.edu, athomaz@cc.gatech.edu

***Abstract*—Humans manipulate and learn about objects by not only using vision, but also physical sensory input. To better understand how physical input improves the learning of objects affordances, the relationship between objects and agents, this work aims to characterize forces and torques felt by a force/torque sensor (FTS) mounted on a robot while performing five unique tasks. We simplified the search space by demonstrating to the robot trajectories to perform and recorded 20 interactions where a human operator moved the object in different positions to achieve 10 successful and 10 failed interactions for each task. The visualized information from the FTS shows that it is possible to distinguish successful vs. failed interactions, which indicates there is potential to use this information to build an unified model using both haptic and visual feedback.**

I. Introduction

Robots are moving from factories and highly structured environments to homes where the environment is often cluttered and unpredictable. This trend opens up a new realm of applications that range from assisting the elderly in homes to working alongside workers in manufacturing plants. One specific area that begins to address this growing need is the study of affordances in relation to robots. The term affordance was first coined by psychologist researcher J.J. Gibson in 1977 and is defined as the "action possibilities" that appear between the environment and the agent [4]. More specifically, we can look at the relationship between objects and agents and the discovery of these affordances in an unstructured environment.

This paper will look at actions and their effects that a robot can perform on an object and focus on visualizing and understanding initial signals captured by the robot's force torque sensors. This information can later be used to solve several major goals of affordance learning, which include learning affordances with a given object and testing a new object for specific affordances (knowledge transfer).

Some of the early and influential work in affordance learning for robotics came from Fitzpatrick et al. [3], where they used four parametrized primitive actions to interact with different objects that could be pushed or rolled and Stoytchev [10], where he looked at the problem of using a tool as an extension of the robot arm to bring objects within reach. These works established a framework for affordance learning using exploration and interaction of the environment. However, these works were very much proof-of-concept systems and focused on visual cues for learning affordances.

Following these influential works, many groups demonstrated complex systems that could learn affordances. Dogar et al. [2] tackled the traversability affordance using three simple primitives (turn-left, turn-right, and move-forward). However, most of the work included visual cues and the only physical feedback included in learning were wheel encoders. Montesano et al. [9] looked at learning affordances to be used for imitation. The work showed that knowledge transfer for imitation was possible, but the only features used aside from visual involved just the duration of contact and position of the arm. In an effort to use and plan with affordances, Kruger et al. [8] developed a rich framework that allowed for affordances to be defined as low-level primitives as well as chained to performed high-level tasks. While the framework supports using sensory input other than visual, the demonstrated system focused on characterizing visual features. Hermans et al. [5] investigated primitives for pushing objects on a flat surface. While this work performed extensive experimentation, once again the only perceptual cues used were visual.

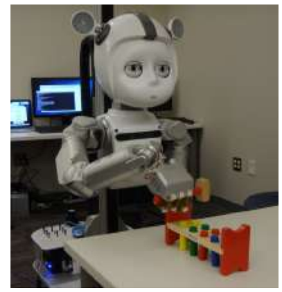


Fig. 1: "Curi" the robot after hammering a peg.

More recent work have looked at interesting problems related to affordances. Koppula and Saxena [7] used affordances to predict and anticipate human activities using visual heatmaps. Katz et al. [6] used grasping affordances to learn the best way to clear rubble in a pile. While Katz et al. used force compliant primitives, the primitives were hand tuned, specific to pile removal, and were not learned by the system.

II. Approach

In order to determine how haptics can be used to improve affordance discovery and transfer, we focused on looking at interactions with objects and how successful interactions differ from failed interactions. This knowledge can later be fused with visual information to build a model of object affordances.

A. Platform

For this work, the robot platform used was "Curi", a robot found in the Socially Intelligent Machines Lab. Curi has two 7 degree-of-freedom (DOF) arms and a movable torso mounted

(a) (b) (c) (d) (e)

Fig. 2: Five objects that Curi performs skills on: (a) bottle and cap (b) box (c) scoop and beans (d) insertion object (e) hammer toy set

on a mobile base. Curi also has two ATI Mini40 force/torque (F/T) sensors attached to each wrist. As seen in Figure 1, Curi is placed in front of table that has an object on the surface. Above the table is a mounted ASUS Xtion Pro RGB-D sensor. The point cloud information is used to segment the object on the table and information about the object pose (position and orientation) and the bounding box is recorded.

B. Controller

To simplify the learning of primitives to explore the object, we used kinesthetic learning from demonstration using keyframes to show Curi the trajectory to perform. For more details on the specific algorithm on trajectory learning from keyframe demonstration, please seee work by Akgun et al. [1].

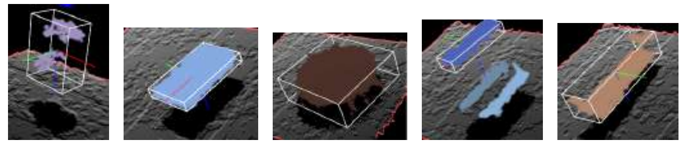
Curi has two main states: **demonstration** and **exploration**. During the **demonstration** state, a human operator kinesthetically demonstrates a skill for Curi to perform. Curi takes the keyframes, generates a trajectory, and then repeats the trajectory with no other input about the world. This repetition is the **exploration** phase. Rather than have Curi vary the trajectory, the human operator moves the object around. This decision is based on prior work from Thomaz and Cakmak [11] where they showed that humans are particularly good at placing objects in unique configurations to find object affordances. This decision also allows us to focus on characterizing the F/T data between successful and unsuccessful explorations.

C. Data Collection

To classify the different forces that occur during interaction, we selected five different task for Curi to perform. They are listed below and all of the objects are shown in Figure 2. All of the tasks were done with the right arm and can also be seen in the supplemental video.

1) **Closing the box** - Curi closes a simple plastic box that clicks shut when properly closed.
2) **Hammering** - The toy set seen in Figure 2 is designed for children to hammer down different colored cylinders that can be pulled up. Curi's task is to successfully hammer a single cylinder that has been pulled up already.
3) **Scooping** - Curi uses a ladle to scoop beans from a bowl
4) **Insertion** - Curi inserts the handle of the hammer from the hammering task into a structure with a hole
5) **Capping** - Curi puts the cap on a laundry bottle

Each interaction was performed 20 times, where 10 interactions were successfully performing the task and 10 were unsuccessful. During the interaction, force information from wrist F/T sensors, object information, as well as all joint positions were recorded and stored in ROS bags.



(a) (b) (c) (d) (e)

Fig. 5: Point cloud and bounding box: (a) bottle and cap (b) box (c) scoop and beans (d) insertion object (e) hammer toy set

III. Initial Results and Discussion

After collecting the data from the different interactions, we can begin to visualize the F/T information and explore methods to integrate this data with visual cues. Figure 5 shows the five objects in Figure 2 segmented from the table. Figure 3 and Figure 4 each show force and torque values for one successful trial and one failed trial for each task, respectively. Furthermore, by using the keyframes given to Curi during the demonstration, the portion of the interaction where Curi actually touches the object can be segmented out and is highlighted in both figures. All of the data is normalized by subtracting the mean of the first 10 values of the interaction.

Looking at both Figure 3 and Figure 4, we can see differences between successful and failed runs within the highlighted sections of the data for the tasks of hammering, closing the box, and capping the bottle. Interestingly, the scooping task looks relatively similar during the actual scoop against the beans, but after the highlighted portion, we can see clear differences in the force felt by the sensor. This force represents the extra weight felt by Curi when the scoop has beans vs. when empty. The insertion task has relatively similar signals between success and failure. This can be attributed to the fact that during a failure case where the robot misses the hole completely or knocks the object over, relatively little force is felt either because the robot missed or the object is light. This looks identical to a successful case where the robot goes through the hole and does not knock anything over.

The visualized data demonstrates that there is potential in using F/T sensors to help distinguish and augment exploration. For example, learning that the weight of an object can change when scooping, adds a dimension to the affordance of scoop that would be difficult to see using just visual cues. Another example includes understanding that certain boxes latch closed vs. just resting closed, which can enrich a simple box closing affordance to one that includes the box being tightly shut or loosely shut. Future work in this area will look at modeling this information in a single representation with visual cues. One possibility is to use the model of F/T information to help Curi increase or decrease the possible affordances of an object. Exploration of a novel object using existing models could be a two-tiered system where an initial set of visual affordances are generated and once Curi touches the object, the set of affordances is refined. Furthermore, once this model is developed we can begin testing ways to have Curi explore this space with less help from the human operator.

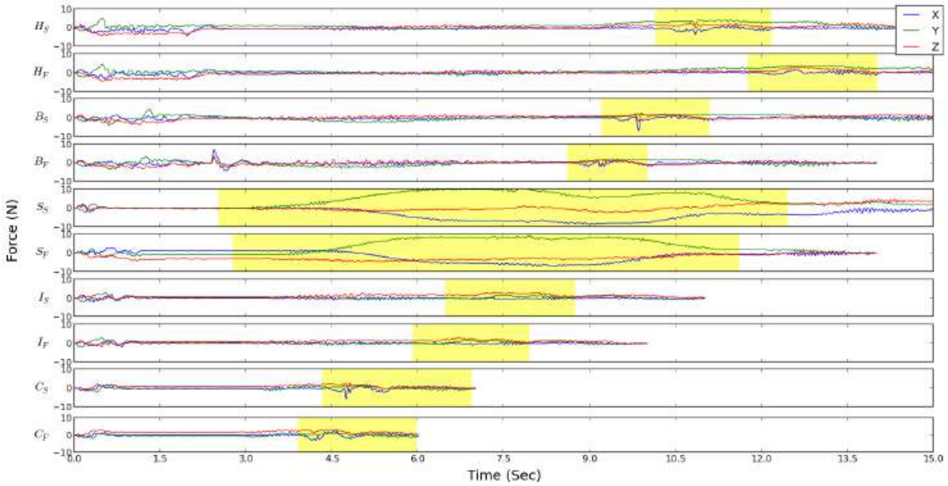


Fig. 3: Force data captured during exploration. Each task is displayed twice, one for success and one for failure. The labels is as follows: H_S and H_F - hammer success and failure, B_S and B_F - box closing success and failure, S_S and S_F - scooping success and failure, I_S and I_F - insertion success and failure, and C_S and C_F - capping the bottle success and failure. Highlighted regions indicate when Curi was in contact with the object

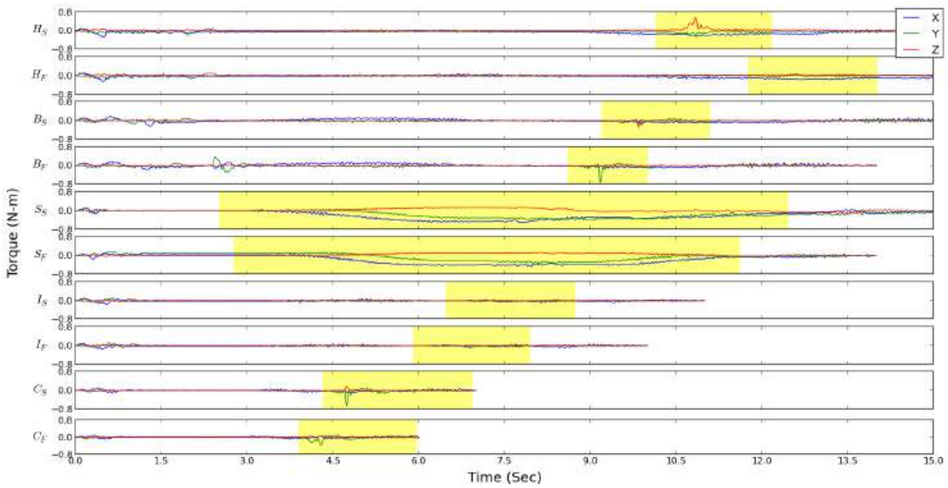


Fig. 4: Torque data captured during exploration. See Figure 3 for label descriptions.

REFERENCES

[1] Baris Akgun, Maya Cakmak, Karl Jiang, and Andrea L Thomaz. Keyframe-based learning from demonstration. *Int'l Journal of Social Robotics*, 4:343–355, 2012.

[2] Mehmet Remzi Dogar, Emre Ugur, Erol Sahin, and Maya Cakmak. Using learned affordances for robotic behavior development. In *International Conference on Robotics and Automation (ICRA)*, pages 3802–3807, 2008.

[3] Paul Fitzpatrick, Giorgio Metta, Lorenzo Natale, Sajit Rao, and Giulio Sandini. Learning about objects through action - initial steps towards artificial cognition. In *International Conference on Robotics and Automation (ICRA)*, pages 3140–3145, Sept 2003.

[4] JJ Gibson. The concept of affordances. *Perceiving, acting, and knowing*, pages 67–82, 1977.

[5] Tucker Hermans, Fuxin Li, James M Rehg, and Aaron F Bobick. Learning stable pushing locations. In *Development and Learning and Epigenetic Robotics (ICDL) 2013*, pages 1–7, 2013.

[6] Dov Katz, Arun Venkatraman, Moslem Kazemi, J. Andrew (Drew) Bagnell, and Anthony (Tony) Stentz. Perceiving, learning, and exploiting object affordances for autonomous pile manipulation. In *RSS Berlin*, June 2013.

[7] Hema S Koppula and Ashutosh Saxena. Anticipating human activities using object affordances for reactive robotic response. *RSS, Berlin*, June 2013.

[8] Norbert Krüger, Christopher Geib, Justus Piater, Ronald Petrick, Mark Steedman, Florentin Wörgötter, Aleš Ude, Tamim Asfour, Dirk Kraft, Damir Omrčen, Alejandro Agostini, and Rüdiger Dillmann. Object–action complexes: Grounded abstractions of sensory–motor processes. *Robotics and Autonomous Systems*, 59(10):740–757, 2011.

[9] Luis Montesano, Manuel Lopes, Alexandre Bernardino, and José Santos-Victor. Learning object affordances: From sensory–motor coordination to imitation. *Transactions on Robotics*, 24(1):15–26, 2008.

[10] Alexander Stoytchev. Behavior-grounded representation of tool affordances. In *International Conference on Robotics and Automation (ICRA)*, pages 3060–3065, April 2005.

[11] Andrea L Thomaz and Maya Cakmak. Learning about objects with human teachers. In *International Conference on Human Robot Interaction (HRI)*, pages 15–22, 2009.