CIS400/401 Project Proposal Specification - BartenderBot

Dept. of CIS - Senior Design 2010-2011

Seth Shannin
sshannin@seas
Univ. of Pennsylvania
Philadelphia, PA

Kevin Xu
kevinxu@seas
Univ. of Pennsylvania
Philadelphia, PA

Camillo Jose Taylor
cjtaylor@cis
Univ. of Pennsylvania
Philadelphia, PA

ABSTRACT

Programming a robot to perform human tasks has been the focus of many research papers. The task has been traditionally very challenging, as it involves heavy computation and complicated coordination between many different joints. Here we present a method to teach the Willow Garage PR2 robot how to mix and serve a drink through imitation. By using immersive teleoperation, it is possible to issue complex commands to the PR2 and have it shadow human motion. The teleoperation will be provided by the Microsoft XBox Kinect, and communication with the PR2 will be handled with ROS. The goal of this project is to have the PR2 not only shadow human motion captured by the Kinect, but also learn from this motion and adapt to new situations via reinforcement learning techniques. We will outline our plan for achieving both shadowing and learning with the Kinect, PR2, and ROS.

1. INTRODUCTION

Constructing a fully autonomous and adaptive robot has been a long-time goal of robotics research. There have been many different attempts at overcoming the challenges involved in developing such a robot. The ability to learn is a powerful intermediate step towards full autonomy. A common problem involves choosing how exactly to demonstrate a desired behavior in such a way that a robot can learn that behavior. In this paper we propose to teach a Willow-Garage PR2 Robot how to perform a reasonably complex task (mixing a drink) through shadowing of human motion capured by the Microsoft XBOX Kinect. The XBOX Kinect sensor from Microsoft provides real-time depth information from a scene at 30 FPS. Combined with various open source libraries[1], the Kinect has been used in many projects involving real-time tracking of human motion, many of which can be found online[2]. The PR2 is a humanoid robot developed by Willow Garage[3] for the purpose of robotics research. It is has been taught by different teams to do many things, including baking cookies[4], scanning and bagging groceries[5], and fetching a sandwich from Subway[6].

We propose to use the Kinect to map human movement to teach the PR2 how to mix and pour a drink. The Kinect sensor provides a convenient way to demonstrate the desired behavior by tracking human motion. The captured data can be relayed to the PR2 via ROS, an open-source Robot Operating System[4]. ROS provides a convenient framework for inter-process communication and coordination between different sensors and components of the PR2. ROS enables relatively short programs to issue surprisingly sophisticated commands to the PR2, such as continually tracking a moving point over time[2]. By using ROS, Kinect, and the PR2, we will demonstrate the effectiveness of teleimmersive demonstration learning in teaching a robot new behavior. This method has several advantages over existing approaches. First of all, the Kinect sensor provides accurate real-time human motion tracking that can be translated to joint movement in the PR2 thanks to ROS. Secondly, teleimmersion better enables a human teacher to show a robot learner exactly how to move in a given situation compared to kinesthetic learning, which involves manipulating the robot learner directly by physical contact. Teleimmersion will also allow demonstrations for robots that cannot be subject to kinesthetic learning easily, such as very large or very small robots. Our method, if successful, would allow for rapid introduction of all kinds of different behavior to the PR2 all from human motion. This technique could be generalized to other humanoid robots besides the PR2 to teach them different behavior.

2. RELATED WORK

There have been many other projects involving autonomous robots and handling drinks. Hillenbrand *et al.* [3] designed a semi-autonomous hand-arm robot for serving drinks. The robot was capable of responding to user input by choosing a drink from multiple choices, opening it, and pouring it into a glass, and then offering the drink to the user. The hand was capable of not only picking up bottles and cups, but also unscrewing bottle caps. The robot combined stereo processing and object recognition to identify the drinks, and then used grasp planning to pick up the drink itself. Bohren *et al.* [1] used the PR2 and ROS to build a robotic system for retrieving a beer from a refrigerator. In their work, they developed a task-level execution system known as SMACH for rapidly prototyping robotic applications. The PR2 had to navigate an obstacle map to reach the refrigerator, use object recognization and grasp planning to identify the door handle and

[1]http://openkinect.org/wiki/Main_page
[2]http://www.freenect.com/
[3]http://www.willowgarage.com/pages/pr2/overview
[4]http://spectrum.ieee.org/automaton/robotics/home-robots/pr2-learning-to-bake-cookies-humanity-surrenders-to-yumminess
[5]http://spectrum.ieee.org/automaton/robotics/robotics-software/pr2-can-scan-and-bag-your-groceries
[6]http://spectrum.ieee.org/automaton/robotics/artificial-intelligence/pr2-can-now-fetch-you-a-sandwich-from-subway

the drinks, and ultimately use facial recognition to deliver the beer to a human recipient. Each step of the process contained detail planning and image processing in order to carry out the expected behavior. Srinivasa *et al.* [5] designed an autonomous robot capable of navigating a household-like environment and manipulating a wide variety of household objects. Consisting of an arm mounted on a segway, HERB used a powerful array of six multi-core processors to successfully traverse its environment and interact with objects around it. All of these robots relied on vision processing and path planning to carry out their tasks. However, there have been other approaches involving demonstration and learning to allow a robot to perform a specific job. (Insert stuff here describing work with kinesthetic teaching and demonstrative teaching, maybe with some reinforcement learning too?)

3. PROJECT PROPOSAL

Now is the time to introduce your proposed project in all of its glory. Admittedly, this is not the easiest since you probably have not done much actual research yet. Even so, setting and realizing realistic research goals is an important skill. Begin by summarizing what you are going to do and the expected benefit it will bring.

3.1 Anticipated Approach

Having summarized *what* you are going to do, its time to describe *how* you plan to do it. Our factorization example does not work so well here (its likely impossible to realize) – so let us suppose you are going to create a service that takes a cell-phone picture of a building and returns via text-message, the name of that building[7].

In this case you might want to talk about establishing a server to receive pictures via MMS. Once the picture is received, you will run an edge extraction algorithm over it. Then, similarity between the submitted picture and those stored (and tagged) in a MySQL database will be computing using algorithm XYZ. Finally, the tag of the most similar image will be returned to the user. Do not bore the reader with trivial details, but give them an overview; a block-flow diagram may prove extremely helpful.

Also note where you anticipate having <u>novel</u> difficulty. Maybe you have never setup a MySQL database or even used SQL before at all – yes, that is a challenge – but not one reader's care about. More novel would be the fact that many buildings on Penn's campus look similar and your classifier may be inaccurate in such instances.

3.2 Evaluation Criteria

Suppose you have implemented your approach, and it is functioning. Now how are you going to convince readers your approach is better than what exists? In the factorization example, you could just compare run-times between algorithms run on the same input. The image recognition example might use a percentage of accurate classifications. Other fields may have established testing benchmarks.

No matter the case, you need to prove you have contributed to the field. This will be easier for some than others. In particular, those with 'sensory' projects involving visual or sonic elements need to think this point through – objectivity measures are always better than subjective ones.

[7]Do not use this idea – someone did it last year.

4. RESEARCH TIMELINE

Finally, we would like you to speculate about the pace of your research progress. This section need not be lengthy, we would just like you to specify some milestones so we can gauge your progress during our intermediate interviews. Let us follow through with our image recognition example:

- ALREADY COMPLETED: Preliminary reading. Began implementation of image-recognition algorithm.
- PRIOR-TO THANKSGIVING : Photograph buildings for DB. Make algorithm more efficient, tune parameters.
- PRIOR-TO CHRISTMAS : Create server-MMS interface. Expand tagged DB collection.
- COMPLETION TASKS : Verify implementation is bug-free. Conduct accuracy testing. Complete write-up.
- IF THERE'S TIME : Investigate image pre-processing techniques to improve accuracy.

5. REFERENCES

[1] Jonathan Bohren, Radu Bogdan Rusu, E. Gil Jones, Eitan Marder-Eppstein, Caroline Pantofaru, Melonee Wise, Lorenz MÃĺosenlechner, Wim Meeussen, and Stefan Holzer. Towards autonomous robotic butlers: Lessons learned with the pr2. In *ICRA 2011*. Willow Garage, 2011.

[2] Steve Cousins. ROS on the PR2. *IEEE Robotics & Automation Magazine*, pages 23–25, September 2010.

[3] Ulrich Hillenbrand, Bernhard Brunner ad Christoph Borst, and Gerd Hirzinger. The robutler: a vision-controlled hand-arm system for manipulating bottles and glasses. Technical report, Institute of Robotics and Mechatronics, 2004.

[4] Morgan Quigley, Brian Gerkey, Ken Conley, Josh Faust, Tully Foote, Jeremy Leibs, Eric Berger, Rob Wheeler, and Andrew Ng. ROS: An open-source robot operating system. Technical report, Stanford University, 2009.

[5] Siddhartha S. Srinivasa, Dave Ferguson, Casey J. Helfrich, Dmitry Berenson, Alvaro Collet, Rosen Diankov, Garratt Gallagher, Geoffrey Hollinger, James Kuffner, and Michael VandeWeghe. HERB: a home exploring robotic butler. Technical report, Carnegie Mellon University, 2009.

APPENDIX

A. OTHER SPECIFICS

Your proposal need not have appendices like this section and the next, but we still have critical info to share:

1. PROPOSAL LENGTH: We require that your proposal be 4–5 pages in length, bibliography included. Be careful, LaTeX and our style-file in particular are *extremely* space efficient. An 9-page MS-Word document could easily become a 5-page LaTeX one.
2. PLAGARISM: **DO NOT** plagarize. If you are caught, you will fail the class (*i.e.*, not graduate), or worse.

B. LaTeX EXAMPLES

At this point, the proposal specification is complete. From here on out, we are just going to show off some commonly used LaTeX technique. Be sure to look at the 'code behind' and see Tab. 1, Eqn. 1 and Fig. 1 for the output!

$$M(p) = \int_0^\infty (1 + \alpha x)^{-\gamma} x^{p-1} dx \tag{1}$$

User Type	Cleanup%	Honesty%
Good	90-100%	100%
Purely Malicous	0-10%	0%
Malicious Provider	0-10%	100%
Feedback Malicous	90-100%	0%
Disguised Malicous	50-100%	50-100%
Sybil Attacker	0-10%	Irrelevant

Table 1: Example Table

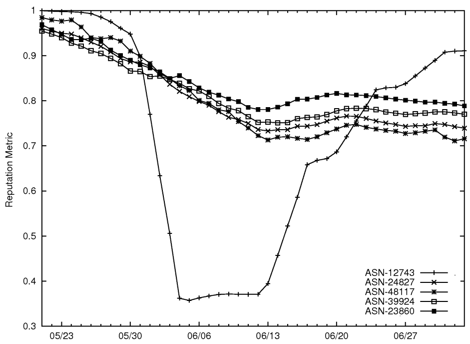

Figure 1: Example Figure/Graph