Robot Learning from Demonstration: Kinesthetic Teaching vs. Teleoperation

Baris Akgun and Kaushik Subramanian

May 5, 2011

Abstract

We are interested in developing learning from demonstration systems that are suitable to be used by everyday people. We compare two interaction methods, *kinesthetic teaching* and *teleoperation*, for the users to show successful demonstrations of a skill. In the former, the user physically guides the robot and in the latter the user controls the robot with a haptic device. We evaluate our results using skill dependent quantitative measures, timing information and survey questions. We find that kinesthetic teaching is faster in terms of giving a single demonstration and the demonstrations are more successful. However, the learned skill does not perform better as expected. The survey results show that users think kinesthetic teaching is easier and more accurate and an open-ended question suggests that people would prefer kinesthetic teaching over teleoperation for everyday skills.

1 Introduction

The use of robots in everyday human environments has long been a goal for researchers and scientists all over the world. This dream has been elusive and it remains an open and hard problem. In reality, it is hard to imagine robots being pre-programmed with all the necessary skills for real-world problems. The framework of Learning from Demonstration (LfD) [1] has been widely used to alleviate this problem. LfD takes advantage of humans and uses their guidance to make the problem tractable for the robot learner. However, most of the literature have concentrated on algorithms and representation, neglecting the user side. Their developers usually provided demonstrations to evaluate them. However, potential users of these systems will not be experts in the field of robotics or machine learning. This serves as a motivation for our research agenda where, in many of the practical LfD applications, the teacher will be an everyday end-user. Thus, our research explores the ways in which Machine Learning can exploit human social learning interactions—*Socially Guided Machine Learning* (SG-ML) [2].

Within the scope of this problem, the goal for our project is to learn low-level skills on a robotic platform using inputs from humans. We focus on realizing a system where humans are comfortable in teaching robots different skills. We attempt to approach this problem by comparing different forms of human-robot interaction for LfD. We focus on the two following broad questions:

- How would humans like to teach robots?
- How can robots efficiently use the inputs given by humans?

The robot we will use is the PR2 from Willow Garage. We choose two modes of interactions, kinesthetic teaching and teleoperation, and compare their utilities in a user study. We begin by describing the individual methods used in our approach, the choices we made, the experiment with humans and finally the results.

2 Description of Methods

2.1 Interaction Modes

In our system, the human interacts with the PR2 using two modes of interaction:

- Kinesthetic Teaching (KT): In this mode, the teacher physically maneuvers the robot. The desired joints of the robot are set in gravity compenstation mode to allow for easy control. An advantage of this method is that it helps avoid the correspondence problem, induced by the human to robot mapping function. The method is shown in Figure 1(a).
- Teleoperation (TO): This mode is performed with the help of Sensable's Phantom Omni®. It is a haptic device which has 6-degree of freedom positional sensing and allows for limited force-feedback (which we do not use). The human manipulates the

(a) Kinesthetic Teaching (b) Teleoperation

Figure 1: Modes of Interaction

Phantom Omni and using a predetermined mapping ((Jacobian transpose), the joints of the robot arm are appropriately translated and rotated. The Omni also has two buttons which allow the human to open and close the gripper of the robot. The method is shown in Figure 1(b).

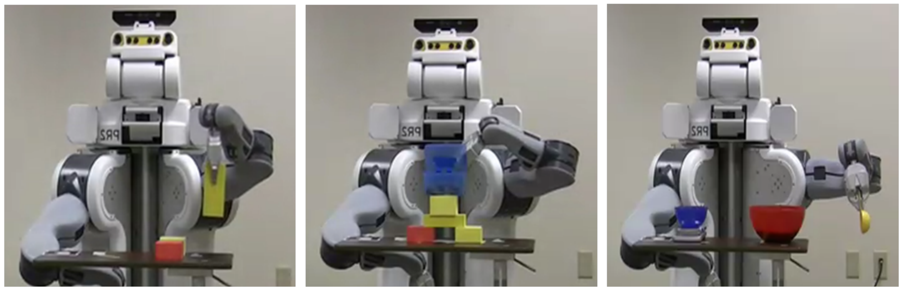
These are chosen since they are among the most probable interaction methods for LfD systems. Kinesthetic teaching is suitable when the human and the robot are co-located. It is also intuitive since this form of interaction is commonplace in human-human interaction (e.g. teaching how to finger paint to an infant). Teleoperation is also being widely used from the state-of-the-art medical robots to excavator arm control. The advantage of this approach is that it is safer (user does not have to stand near the robot). Moreover, it is a viable, if not the only, choice when the robot is not situated at the same place as the human (distance) or its size does not allow for kinesthetic teaching (scale). We compare the utilities of these mothods in a everyday LfD setting.

2.2 Skills

In our study, we have two skills that we are interested in learning with the robot along with a practice skill. The practice skill was designed to help the human get familiar with the modes of interaction as well as the general flow of the experiment. The practice task shown in Figure 2(a) is called Orient and Place (OP). The goal of this task is to manipulate the arm of the robot in a way that makes the yellow cuboid fit within the gap of the two blocks placed on the table. It requires the user to both manipulate the position and orientation of the robot's end-effector. The first skill we wanted the robot to learn is called Close the Lid (CL), shown in Figure 2(b). The goal is to move the robot arm such that it closes the lid of the open box. The second task is called Scoop and Place (SP), shown in Figure 2(c). In this task, a spoon is placed in the robot's gripper. The goal is to transfer as many M&M's as possible from the big bowl to an adjacent smaller bowl. We chose these tasks as they have direct correlations to real-world scenarios. For example, in a kitchen scenario, one might want a bowl of cereal. This involves a combination of scooping and placing as well as closing open boxes.

2.3 Skill Learning

In order to facilitate the learning of the above skills, we make use of a supervised learning approach called Gaussian Mixture Models (GMM). It has been used in the LfD setting [3] and was found to achieve reasonable results. The learner takes in a set of sample demonstrations of a skill and computes a representative generalized model. To summarize, the demonstrations are given to the learner in the form of time-stamped robot joint angles and velocities. These are first time-warped to ensure that each of them has a similar time scale. After this pre-processing step, k-means algorithm is run to cluster the data. The cluster means (m_0) and covariances (Σ_0) are used as the initial values for the Expectation-Maximization (EM) algorithm, which learns a GMM from X_p. Note that there are k components in the model. In our study we used a k value that was derived empirically. The outputs of GMM are sub-population means (m_f) and covariances (Σ_f). These are considered as the model of the skill. Gaussian Mixture Regression (GMR) is then used to generate the skill. GMR accepts the model of the skill and a time vector and calculates the corresponding joint positions and velocities required to generate the new demonstration.



(a) Orient and Place (b) Close the Lid (c) Scoop and Place

Figure 2: Skill used in our experiment

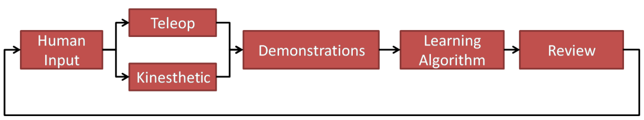


Figure 3: High-level sample scenario

2.4 Scenario

Figure 3 provides an example of the scenario we will be using in our experiments. The human will interact with the PR2 using one of the modes of interaction. He/She will show the robot an example of a specific task following which the learner generates a generalized model of the task. This generalized model is then reproduced by the robot and this shows the human what the robot has learnt. The human reviews the robot's performance and then proceeds to either correct the task or teach another task. The process of correction is facilitated by giving multiple demonstrations from start to end of the task to be learned until the robot performs the task accurately.

3 Experiments

We have done user studies to evaluate our approach LbD and compare *kinesthetic teaching* (KT) and *teleoperation* (TO) mode of interactions for LbD.

3.1 Research Questions and Hypotheses

In our experiments, we are interested in addressing the following questions:

- Which, among kinesthetic teaching and teleoperation, is more preferable (e.g. ease of use, enjoyability) from an every-day user perspective?
- Which of the teaching methods achieve quantitatively better results on the robot?

Our hypotheses are:

- Users will find KT to be easier in terms of interaction.
- Users will find TO to be more enjoyable.
- KT leads to better skill performance.

3.2 Variables and Measurement

We have the interaction method as the only independent variable, being either KT or TO. We have several quantitative measures and survey questions as our dependent variables.

We measure the amount of M&M's transferred, in ounces, from one bowl to the other for the SP skill. Note that we treat this as a continuous variable. We use paired t-test to analyze this variable.

We have the end state of the box, either closed or open, as our measure for the CL skill. This is a binary variable and we use the McNemar's test to analyze it.

We also measure the time taken, in seconds, per single demonstration for each skill. We use unpaired t-test to analyze this variable. The reason for choosing the unpaired test is that the number of samples provided for different skills are not the same for each user.

We asked 7-point Likert-scale survey questions to the participants at the end of their interactions with the robot. These questions are about *ease of use*, *enjoyability*, *accuracy of demonstrations* and *improvement given time*. We use Wilcoxon signed rank test to evaluate the survey results. We also asked an open-ended question to get the overall impression of the user. We furthermore discuss the user responses and our personal anecdotes.

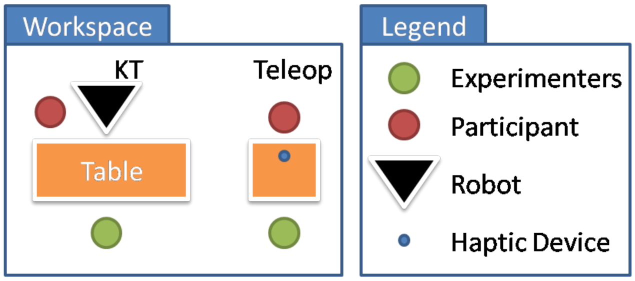


Figure 4: Schematic view of the experiment environment.

3.3 Participation and Controls

As mentioned above, we have two interaction methods and two main skills. We performed a within-user study i.e., all the participants used both methods to interact with the robot. Moreover, all the participants demonstrated both of the skills for each method. We counterbalanced the order of the methods and the skills as a control. Note that the minimum number of participants for the counterbalanced ordering is $2 \times 2 \times 2 = 8$, ((Method 1—2) -¿ (Skill 1—2), (Other Method) -¿ (Skill 1—2))).

We had 9 participants[1], 5 females and 4 males. Their ages were between 23 and 32 with a median of 25. None of the participants were experts at robotics or machine learning and none of them had used a haptic device before.

3.4 Experimental Setup

Our experimental setup can schematically be seen in figure 4. We have two experimenters that give instructions and explanations (how to interact, what the skills are etc.) to the participant and guide the experimental flow (start-stop interaction etc.). Note that only the right arm of the robot is used.

4 Results and Discussion

We present our results with box and whisker plots[2] because the number of samples are low (¡ 10) and calculating distribution parameters would not be appropriate. We do not report any statistical test scores due to the same reason.

[1]Having a slight imbalance in counterbalancing order did not affect the results

[2]In a box and whisker plot, red line corresponds to the median, lower and upper box bounds correspond to 25% and 75% percentile ranges of the data and the whiskers correpond to 5% and 95% percentile ranges of the data. Red plus signs correpond to outliers.

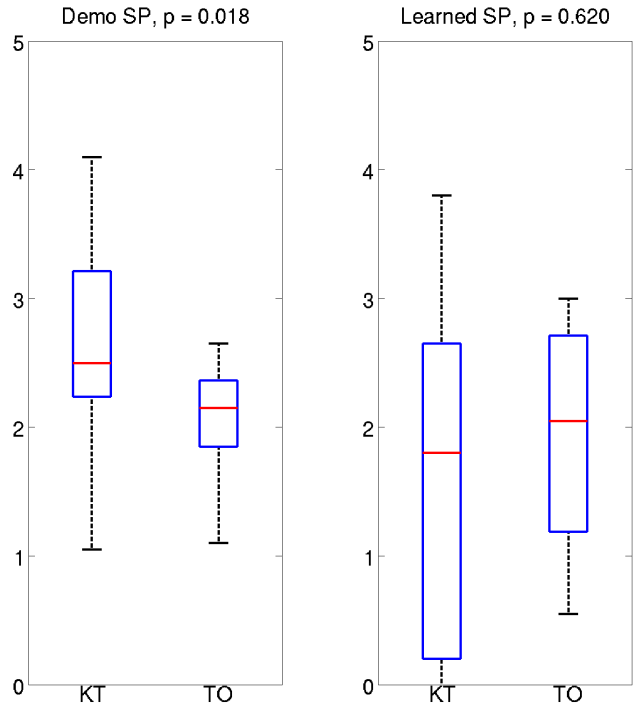


Figure 5: Box and whisker plots of amount of M&Ms transferred. Left: demonstrations, right: learned model. p values are obtained with

4.1 Quantitative Results

Users are able to give more successful demonstrations with kinesthetic teaching for the SP skill. We evaluate the amount of M&M's transferred per condition both for demonstrations and for the learned model. The results are shown in figure 5. From the left part of the figure, it can be seen that the participants managed to transfer more during kinesthetic demonstrations ($p = 0.018$), but with more variability. However, this is not reflected in the learned skills and the performances are very similaras seen from the right part of the figure. We provide two probable causes for this. First, users can make subtle but useful (e.g. rocking the spoon) during kinesthetic teaching since they are more accoustomed to this form of interaction. However, our learning algorithm treats these as noise and smooths them, thereby losing the information. Second, we did not control the distribution of the M&M's after a demonstration. Note that after a user demonstration, a dent is left in the distribution. The learned skill will try to scoop from around the demonstrated region but will not get as many M&M's due to the dent.

For the CL skill, all the participants gave successful demonstrations and only one learned skill (with KT method) failed. Our McNemar's test revealed $p = 1$.

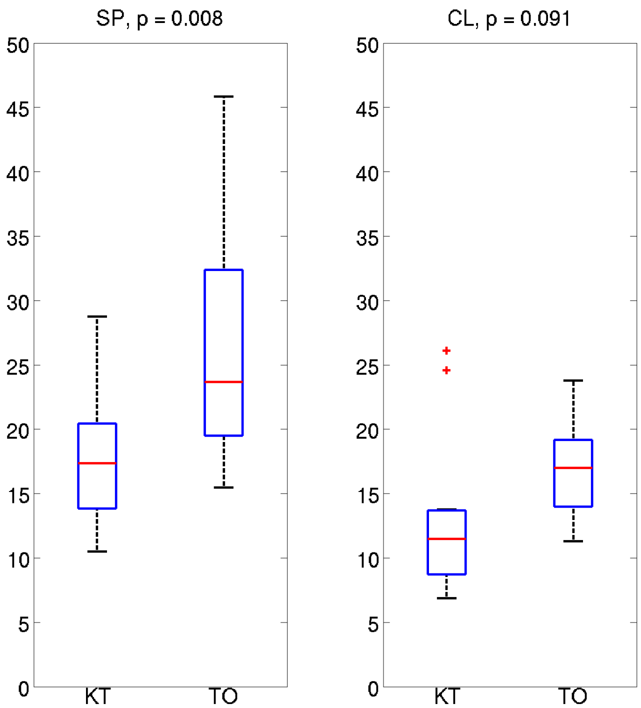


Figure 6: Box and whisker plots of time taken per single demonstration. Left: Scoop and place skill, right: close the box skill.

We therefore conclude that the either the skill was not challenging or our measure was not adequate.

Users are faster at giving demonstrations with kinesthetic teaching. The figure 6 shows the box and whisker distribution of the time data. Left part of the figure shows that the difference is significant in favor of the kinesthetic method ($p = 0.008$) and teleoperation method had more variance in the SP skill. For the CL skill, the results are not as significant but there is a trend ($p = 0.091$), even in the presence of outliers. The two outliers are there because it took some time for these users to realize they needed to move some of the robot joints (shoulder joints) that are away from the end effector.

4.2 Survey Results

Figure 7 shows the survey responses of the users. We will base our discussion on this figure without referring to it explicitly in the text.

Users find kinesthetic teaching to be easier. The median answer to the ease-of-use question was 6 for the KT case, whereas it was 5 for the TO case. Note that the answers are significantly different than each other ($p = 0.05$). We expected this result due to the fact that every people are more accustomed to a kinesthetic type of teaching, i.e., it occurs naturally in human-human interactions. Moreover, the users can adjust their perspective better, see more of the workspace and be more "situated" with this interaction method.

Users enjoyed both methods. There was no difference in the enjoyability from the users' perspective and they mostly enjoyed both methods. This could be attributed to the *novelty effect* since none of the users has played with a robot before.

Users tend to think that they can give more accurate demonstrations with the kinesthetic teaching method. There is a strong tendency from the users' perspective that KT method is more accurate ($p = 0.077$). The distribution for the KT method has a median 6 and is tighter compared to the TO method with median 5 and more variability.

Users think that they could improve their demonstrations if they were given more time. The lowest score for this question corresponded to "not much" which means that any score higher than 1 would be considered as room for improvement. Although the results are not statistically significat, we can say that TO might have more room improvement from the users' perspective due to the its distribution being tighter and more close to the upper scores.

4.3 Open-Ended Question

We asked the question *Which mode would you prefer if you bought a robot and why?* as an open-ended question at the end of our survey. 7 of the participants replied KT whereas 2 of the participants replied TO. We present the user choice and the comments in 1. Some of the users noticed the degree of freedom difference between the methods. They cited "different parts" and the motion of the elbow. Note that with TO, user can control 6 dofs (end effector in the 3D Cartesian space) whereas with KT the user can control the entire arm which has 7 dofs. Other reponses include time, ease of use, accuracy and preference for day to day skills.

4.4 Anecdotes

In this section, we briefly mention some of the things that users mentioned during their interactions. Some of the users referred to the robot as being cute when he performed the learned skill. This suggests that the way the robot moves and its apperance have an influence on the user. Some of the users asked more space around the robot to be able to move its arm more freely, especially for the TO method. Some of the users were curious about trying out different ways to teach the same skill. they made use of the rejection option if they did not like what they had thought. Many users mentioned

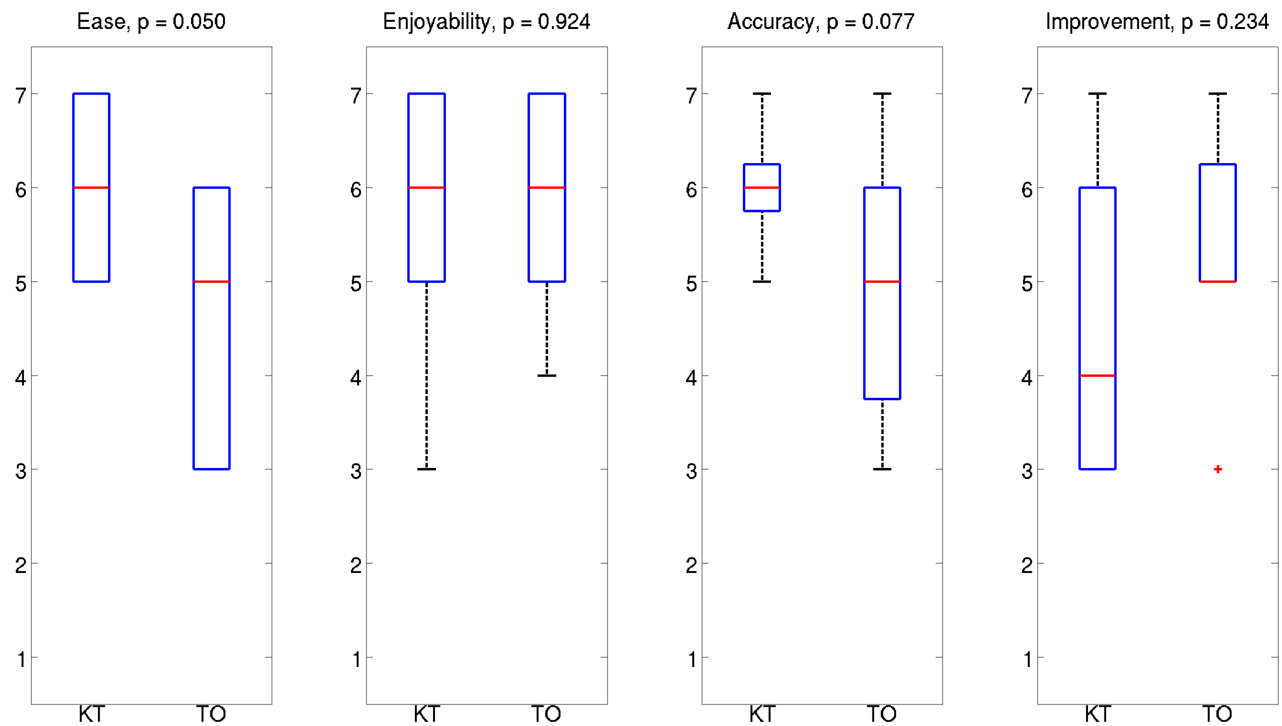


Figure 7: Box and whisker plots of survey answers.

Table 1: User responses to the open-ended question

Choice	Comment
KT	"It would require less time to teach the robot because it was an easier method to use initially."
KT	"Better for day to day operations."
KT	"It seemed more reliable because of the ability to move different parts appropriately for the task which was difficult when the teleoperation teaching mode was used."
KT	"It allowed me to use the arm and place the arm in ways which seemed more human like, felt the Teleoperation mode to be a bit more difficult because the feedback from the operation was not as natural in terms of use, rotating the wrist can be difficult too in the Teleoperation mode. In the teleoperation mode the Robot arm's elbow didn't move as expected and felt it became difficult to move it when the object was to placed close or too far away. The Teleoperation mode also made it difficult to maneuver the object when it was to be placed with accuracy."
KT	"It is much easier. Haptic device needs a lot more time to get familiar with."
KT	"Easy and fun to interact."
KT	(No answer)
TO	"I would go with teleoperation. It seems like, in terms of usability, it is more convenient."
TO	"It is more accurate."

teleoperation as being hard during their use of the haptic device. We noticed that the robot's joint space can cause problems for some of the users who concentrate mostly on the end-effector of the robot during KT.

5 Conclusions

In this study, we developed an iterative LfD system based on our previous experience with two different interaction modes, namely kinesthetic teaching and teleopration. We have done a user study to compare the two in a probable every day setting. As expected, kinesthetic teaching was preferred by most of the participants, based on timing, ease of use and accuracy of demonstrations.

However the results for teleoperation are optimistic in that we believe there is scope for further improvements using this mode of interaction. Moreover, this type interaction might be the only option for some scenarios (mostly involving domain experts such as surgeons). We treat our current experiments as a pilot study and aim to device further experiments that will try to bridge the ease of use gap between kinesthetic teaching and teleoperation.

Overall we are interested in further developing LfD systems that are tailored to be used by everyday people and potentially augmented for domain expert. We will be using information gathered from this study for our participation in the AAAI 2011 LfD Challenge. We plan to improve the work and allow the robot to learn low-level skills and high-level tasks through an ongoing social dialog with a human partner.

6 Video

A video of our implementation is available at `http://www.youtube.com/watch?v=muuRFmM_oyA` and a video of interesting insights from our experiment is available at `http://www.youtube.com/watch?v=LbEsYsaegWM`. The video provides an overview about the different modes of interactions with two sample scenarios, their associated demonstrations and learned behaviors.

References

[1] B. Argall, S. Chernova, M. M. Veloso, and B. Browning. A survey of robot learning from demonstration. *Robotics and Autonomous Systems*, 57(5):469–483, 2009.

[2] C. Breazeal and A. L. Thomaz. Learning from human teachers with socially guided exploration. In *ICRA*, pages 3539–3544, 2008.

[3] S. Calinon, F. Guenter, and A. Billard. On learning, representing and generalizing a task in a humanoid robot. *IEEE Transactions on Systems, Man and Cybernetics, Part B. Special issue on robot learning by observation, demonstration and imitation*, 37(2):286–298, 2007.