Prediction and Production of Human Reaching Trajectories for Human-Robot Interaction

Sara Sheikholeslami*
sara.sheikholeslami@ubc.ca

Justin W. Hart†
hart@cs.utexas.edu

Wesley P. Chan*
wesley.chan@ubc.ca

Camilo P. Quintero*
camilo.quintero@ubc.ca

Elizabeth A. Croft*
elizabeth.croft@ubc.ca

ABSTRACT

In human-human interactions, individuals naturally achieve fluency by anticipating the partner's actions. This predictive ability is largely lacking in robots, leading to stilted human-robot interactions. We aim to improve fluency in human-robot reaching motions using a unified predictive model of human reaching motions. Using this model, we allow the robot to infer human intent, while also applying the same model to generate the robot's motion to make its intent more transparent to the human. We conducted a study on human reaching motion and constructed an elliptical motion model that is shown to yield a good fit to empirical data. In future studies, we plan to confirm the effectiveness of this model in predicting human intent and conveying robot intent for achieving fluency in human-robot handovers.

ACM Reference Format:
Sara Sheikholeslami, Justin W. Hart, Wesley P. Chan, Camilo P. Quintero, and Elizabeth A. Croft. 2018. Prediction and Production of Human Reaching Trajectories for Human-Robot Interaction. In *HRI '18 Companion: 2018 ACM/IEEE International Conference on Human-Robot Interaction Companion, March 5–8, 2018, Chicago, IL, USA.* ACM, New York, NY, USA, 2 pages. https://doi.org/10.1145/3173386.3176924

1 INTRODUCTION

When two or more humans collaborate to perform a task, they naturally reach a high level of temporal coordination, contributing to a fluent meshing of their actions [9]. In contrast, Human-Robot Interaction (HRI) is often characterized by stop-and-go pauses. To attain fluent meshing in human-robot team activities, robotic teammates must 1) be able to infer the intentions of their human counterparts in order to determine the appropriate action to take [6], and 2) move in a way that is legible (intent-expressive) to the human [3].

A great deal of work in neuroscience [1, 4, 5, 8] has sought to model the principle underlying human arm reaching movements. Motor commands for these movements are represented in terms of spatial hand trajectories rather than joints angles [5]. The hand follows a predictable path that is elliptical with smooth and symmetric bell-shaped velocity profiles [2, 7].

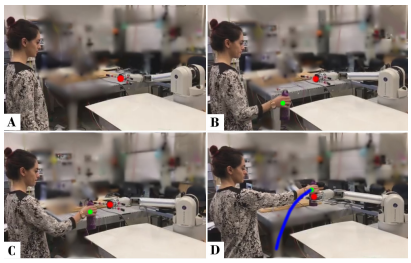

Figure 1: Human initial hand motion (A, B) is used to predict a suitable trajectory for the robot (C, D).

We focus on single-arm reaching motions as a natural communication channel in physical human-robot collaboration. Our goal is to enable robots to understand the intention behind the motion of their human counterparts, and to generate intent-expressive motions for achieving fluent human-robot collaboration, see Figure 1.

Previous research has focused on goal-directed human reaching motions with only two possible goals at most [3, 6]. The ways in which legible motion can be generated is limited in the presence of too many possible goals. Also, Acquiring a cost function which models the observer's expectation on how the robot moves legibly is still an active research challenge that could benefit from investigating some regularities of biological motion control [5].

2 RESEARCH GOALS AND HYPOTHESES

Enabling fluent action meshing in human-robot team activity requires understanding and predicting how people coordinate their reaching motions, modeling these mechanisms, and incorporating the models into robots to produce trajectories that synchronize naturally with human motion.

- **G1:** Prediction of human reaching movements
- **H1:** For human reaching movements, the hand follows a predictable path that is roughly elliptical.

Therefore, an elliptical fit to the motion early in the path is predictive of the remainder of the path, and models of velocity and acceleration can then be used to determine the final timing and location trajectory early in the process. This principle can be used to predict in real-time the intended target of a human performing a reaching motion with a physical target (eg. pick-and-place tasks)

- **G2:** Production of legible motion with predictable outcomes.

*The authors are with Faculty of Mechanical Engineering, University of British Columbia, Vancouver, BC, Canada
†The author is with Faculty of Computer Science, University of Texas at Austin, Austin, Texas, USA

HRI '18 Companion, March 5–8, 2018, Chicago, IL, USA

ACM ISBN 978-1-4503-5615-2/18/03.
https://doi.org/10.1145/3173386.3176924

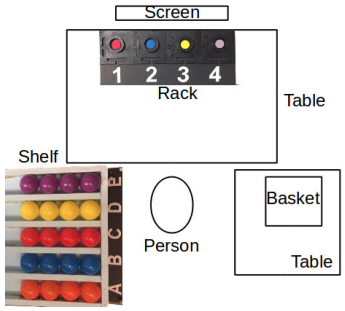

Figure 2: The top view of experimental setup for Study 1

- **H2:** Robot motion following an elliptical model is more legible.

We aim to create a single, unified model for describing reaching motions that can be used for both motion prediction/understanding, and motion production, in order to facilitate fluent human-robot collaborations.

3 METHODOLOGY

Study 1: Building Models of Human Reaching Motions
We use Vicon motion capture data to first construct high-fidelity human reaching motion models. The study simulated a "box packing" task in which participants were asked to fulfill the role of a packer in a shipping operation while we recorded their reaching motions. Each participant sat at a desk with a monitor in front. Surrounding them was a shelf with color-coded ball supplies, a rack with four spaces for intermediate ball placement, and a square basket with four spaces for final packing, Figure 2.

In each trial, the monitor provided pictographic instructions of ball placements onto the rack and the basket, and the participant followed the instruction to pack the balls into the rack, then into the basket, then back to the shelf as a reset for the next trial.

The task was repeated 25 times, yielding a total of 200 reaching motion samples per participant. 30 participants were recruited.

Study 2: Generating an Intent-Expressive Model of a Human-Like Reach for the Robot
A motion planner was developed based on the elliptical motion model constructed in Study 1, and was implemented onto a 7-DOF Barrett Whole Arm Manipulator (WAM)[1] equipped with a three-fingered Barrett Hand[2]. This involved the construction of an algorithm based on "spatial control" hypothesis by Morasso [7] that transforms the cartesian trajectory of the end effector (hand) in task space into coordinated trajectories in joint space in real-time.

Study 2 will be an entertainment study determining whether the robot incorporating our elliptical path planner will improve the quality and speed of interaction. This study will simulate a collaborative human-robot "box packing" task that will require the agents to share resources, and resolve unwanted resource conflicts. The task will specifically invoke situations that would require the agents to work based on intentions behind each other's motions such as:

[1] WAM^{TM}, Barrett Technologies, Cambridge, MA, USA
[2] Barrett Hand^{TM}, Barrett Technologies, Cambridge, MA, USA.

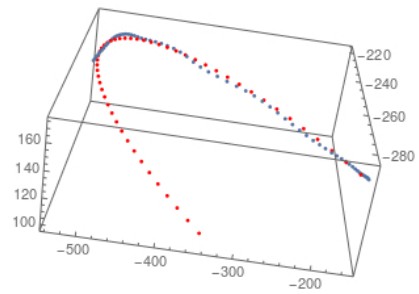

Figure 3: The cartesian positions of the hand in cm as it moves through space for a random participant (blue), and our model of the best fit ellipse to the hand trajectory (red)

- Handing objects from the human to robot, or the reverse action, wherein the human hands objects to the robot.
- Both agents reaching for the same object at the same time

4 PRELIMINARY RESULTS OF STUDY 1

Preliminary data analysis of Study 1 indicates that human reaching motion follows a predictable elliptical path. Figure 3 illustrates a sample of spatial trajectories of the hand (in blue) overlaid by the elliptical motion model. The predicted ellipse closely fits the reach.

5 SUMMARY

This abstract has discussed two studies. The first of these studies aims to understand and construct models of human reaching motions to facilitate prediction of the timing and location of the end of human arm reaching motions in tasks such as grasping or object handover. Based on the identified model, we have constructed a robot trajectory generator. In the second study, we will test our trajectory generator to confirm that it produces predictable and legible motion, and improves fluency and speed in human-robot interaction.

REFERENCES

[1] E. Burdet, R. Osu, D. W. Franklin, T. E. Milner, and M. Kawato. 2001. The central nervous system stabilizes unstable dynamics by learning optimal impedance. *Nature* 414 6862 (2001), 446–9.
[2] H. Collewijn, C. J. Erkelens, and R. M. Steinman. 1988. Binocular co-ordination of human horizontal saccadic eye movements. *The Journal of Physiology* 404, 1 (1988), 157–182.
[3] A. D. Dragan, K. C. T. Lee, and S. S. Srinivasa. 2013. Legibility and predictability of robot motion. In *2013 8th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*. 301–308.
[4] T. Flash and N. Hogans. 1985. The Coordination of Arm Movements: An Experimentally Confirmed Mathematical Model. *Journal of neuroscience* 5 (1985), 1688–1703.
[5] C. Harris and D. M. Wolpert. 1998. Signal-dependent noise determines motor planning. 394 (09 1998), 780–4.
[6] G. Hoffman and C. Breazeal. 2007. Cost-Based Anticipatory Action Selection for Human-Robot Fluency. *IEEE Transactions on Robotics* 23, 5 (Oct 2007), 952–961.
[7] P. Morasso. 1981. Spatial control of arm movements. *Experimental Brain Research* 42, 2 (01 Apr 1981), 223–227.
[8] E. Todorov and M. Jordan. 2002. Optimal feedback control as a theory of motor coordination. Nat. Neurosci. *Nature Neuroscience* 5 (12 2002), 1226–35.
[9] M. Tomasello, M. Carpenter, J. Call, T. Behne, and H. Moll. 2005. Understanding and sharing intentions: The origins of cultural cognition. *Behavioral and Brain Sciences* 28, 5 (2005), 675–691.