Predictions of Human Task Performance and Handover Trajectories for Human-Robot Interaction

[Extended Abstract]

Justin W. Hart
Department of Mechanical Engineering
University of British Columbia
Vancouver, BC, Canada
justin.hart@mech.ubc.ca

Sara Sheikholeslami
Department of Mechanical Engineering
University of British Columbia
Vancouver, BC, Canada
sara.sheikholeslami@alumni.ubc.ca

Matthew K.X.J. Pan
Department of Mechanical Engineering
University of British Columbia
Vancouver, BC, Canada
mpan9@interchange.ubc.ca

Wesley P. Chan
Department of Creative Informatics
University of Tokyo
Tokyo, Japan
wesleyc@jsk.imi.i.u-tokyo.ac.jp

Elizabeth A. Croft
Department of Mechanical Engineering
University of British Columbia
Vancouver, BC, Canada
elizabeth.croft@ubc.ca

1. INTRODUCTION

When two or more agents collaborate to perform a task, situations arise which require for them to temporally coordinate their actions and motions. The Collaborative Human-Focused Assistive Robotics for Manufacturing (CHARM) project[1] seeks to develop robotic assistants for manufacturing which autonomously perform tasks alongside human workers; responding to human communicative cues and the state of the shared task in order to determine the appropriate action to take. An important component of this is the temporal coordination of behavior between the human worker and robotic assistant; contributing to interaction fluency and the efficiency with which the team is able to complete their shared task [9, 3]. Understanding how people coordinate their behaviors, modeling these mechanisms, and incorporating these models into human-assistive robotic devices has been a recent focus of the CHARM project [7].

One scenario performed as part of the CHARM project is a car door assembly task in which a human worker's efforts are supported by a Kuka Lightweight Robot (LWR) robotic arm mounted to an overhead gantry platform. This device is referred to as the "robot assistant." The robot assistant performs support tasks such as retrieving parts for the human worker during the collaborative assembly scenario. As such, the human worker's tasks have a temporal dependency on the actions of the robot, making the timing of the robot's behavior crucial to its utility in this task. Similarly, Hoffman and Breazeal noted that the selection of anticipatory actions based on predictions of the future state of the task at hand is an important skill for interaction fluency [9]. More recently, Hayes, Grigore, Litoiu, Ramachandran, and Scassellati [8] developed a system for estimating the duration of task performance by different agents. A related aspect of current work on CHARM is the construction of models of the human collaborator's task performance, enabling the system to temporally coordinate the robot's actions to those of the human worker, and an evaluation of how this affects the interaction.

We have also been working on the coordination of the specific behavior of human-robot handovers. As part of the car door assembly task, the robot assistant retrieves parts that are to be mounted on the door and hands them its human collaborator using a Robotiq gripper attached to its arm, as shown in Figure 1. The action of handing objects from robots to humans has been extensively studied by the Human-Robot Interaction community [15, 10, 11, 14, 13, 4, 5]. Shibata, Sahbi, Tanaka, and Shimizu [15] and Huber, Rickert, Knoll, Brandt, and Glasauer [10] both demonstrated that minimum-jerk arm trajectories for handovers appear safer and elicit shorter reaction times from human receivers [10, 15]. Moon, Troniak, Gleeson, Pan, Zheng, Blumer, MacLean and Croft [13] found that directing the robot's gaze to the intended handover location helps people to predict where the handover will occur. Admoni, Dragan, Srinivasa, and Scassellati found that adding delays to the robot's handover behavior, when the handover is preceded by a gaze cue, causes study participants to be more likely to heed the gaze cue [1]. Mason and Mackenzie[11] studied grip forces, finding that the giver's and receiver's grip forces and load forces exhibit a linear relationship during handovers. To perform handovers, the robotic assistant utilizes a han-

[1]CHARM is a collaboration between investigators at The University of British Columbia, Université Laval, and McGill University.

HRI 2015 Workshop on Human-Robot Teaming 2015 Portland, Oregon, USA

Figure 1: The robot assistant, which utilizes a Kuka Lightweight Robot (LRW) robotic arm mounted to an overhead gantry platform, performs a handover during a car door assembly scenario.

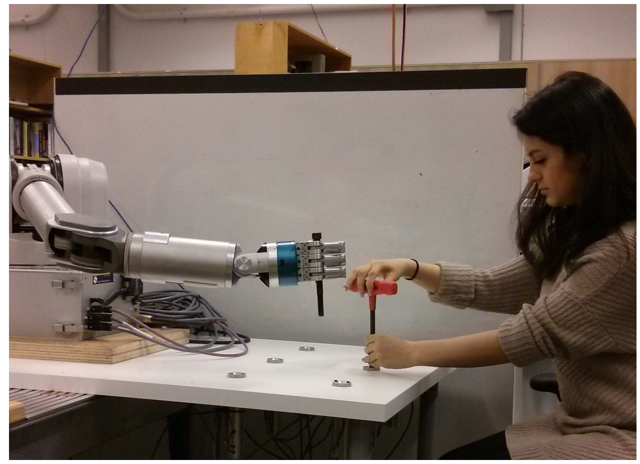


Figure 2: A Barrett Whole Arm Manipulator (WAM) prepares to insert a screw into a threaded hole while its collaborator tightens one of the already-inserted screws.

dover controller developed by Chan, Parker, Van der Loos, and Croft [4, 5]. We are now in the process of researching the reverse action, wherein a human hands an object to a robot. This process also requires coordination between the robot and human worker, coordinating the timing and location of the handover.

In this document, we will discuss two studies currently in progress at the Collaborative Advanced Robotics and Intelligent Systems (CARIS) Laboratory at The University of British Columbia, in which we attempt to coordinate a robot's behavior to that of a human collaborator working on a shared task. In the first of these studies, we have constructed a collaborative assembly task in which an automated system attempts to determine the duration of human task performance, allowing the robot to adapt its behavior accordingly. In the second, we are constructing an algorithm to predict the timing and location of the handover of an object from a human to a robot.

2. COORDINATION OF TASKS

To study the temporal coordination of human and robot behavior during collaborative assembly tasks, we have developed a simple task in which a Barrett Whole Arm Manipulator (WAM) robotic arm retrieves screws from a rack and places them into threaded holes in a table. A human collaborator follows up on this action by fastening the screw into the threaded hole. During repeated trials, the robot constructs a model of both its own performance and the human's performance on this task, developing a model of task duration. The task is purposely chosen as one involving a repetitive action, the action of fastening the screw, so that the system can construct models of the study participant's performance in only a few trials. Note that the system also explicitly measures the performance of the robotic agent, rather than computing its timing from pre-defined models. We consider learning and reasoning about the robot's "self" to be as important in human-robot interaction as reasoning about the human participant [6], and this design decision is reflective of this stance.

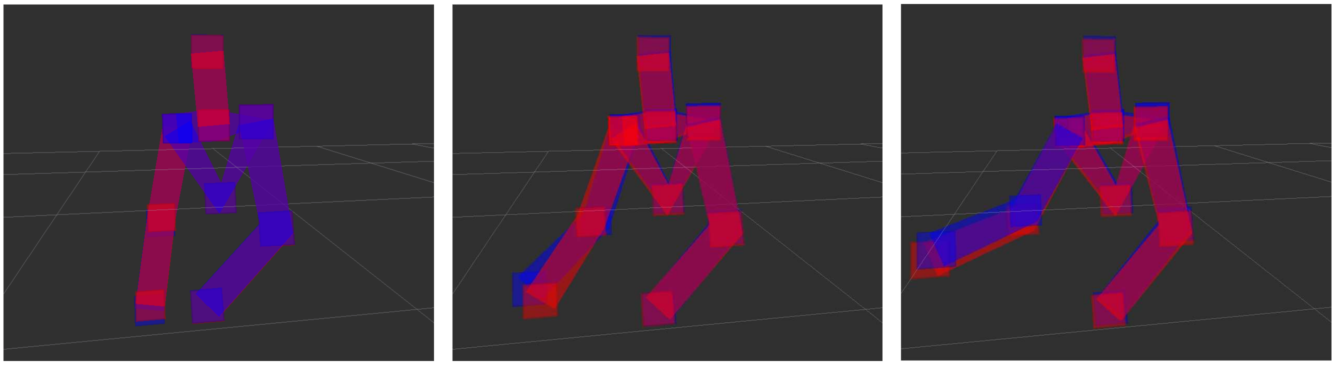
The experimental setup is designed to provide highly-accurate measurements of each study participant's performance on the screw fastening task. During the study, the participant sits at the table in Figure 2. The table has a series of six threaded holes arranged in two rows across its top, with the robotic arm facing directly across from the participant. The threaded holes are instrumented with sensors to indicate progress and completion of the process of fastening. Times are stored in a database specific to each participant, allowing the robot to compute an individualized performance model for each of its collaborators.

To plan its behaviors, the robot uses a custom-built partial-order planner. The planner is written to accept problem definitions in PDDL [12], allowing it to be used with a variety of hardware and in multiple scenarios. This is integral to the planner's design, as it is currently being used in both this task and the car door assembly task. It is intended to be used in future scenarios explored by the CHARM research program, allowing us to incorporate this functionality flexibly across platforms and tasks. The custom planner is designed to accept descriptions of the capabilities of each agent participating in the scenario, and to construct a plan and timeline for the performance of tasks by each agent. The planner can determine periods of overlapping behavior and optimize across various temporal and task-related metrics. For instance, the system can optimize either co-active time - the amount of time agents spend performing overlapping activities, or total time to completion of the task. In this format, it is possible to flexibly explore the use of models of human performance in collaborative tasks.

3. COORDINATION OF HANDOVERS

In another recent study, we have begun to investigate the task of a person handing an object to a robot. Part of this work has involve the construction of software to enable a human study participant to hand an object to a Willow Garage Personal Robot 2 (PR2), as in Figure 3.

The task of handing an object to a robot presents a unique set of challenges for fluent human-robot interaction. One



(a) First frame of reach trajectory. (b) Middle of reach trajectory. (c) End of reach trajectory.

Figure 4: Illustration of a least-squares fit linear trajectory model computed in the current version of the handover predictor, as compared to the sampled skeletal pose data from which it is constructed. The skeletal pose data was sampled using a Microsoft Kinect using the NITE skeleton tracker. The sampled pose data is presented in blue. Estimated poses from model trajectory are presented in red. Poses occurring in the same subfigure are synchronized by finding the pose in the trajectory model at the time corresponding to the timestamp of the presented sampled skeletal pose data. In these images, it can be seen that the model trajectory closely fits the sampled skeletal tracking data that it approximates.

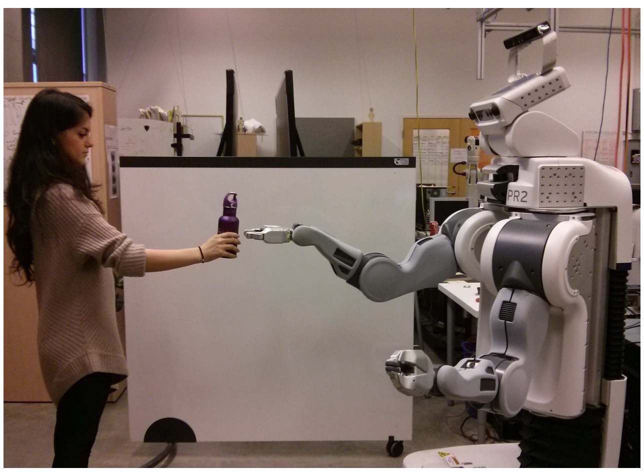


Figure 3: A Willow Garage Personal Robot 2 (PR2) reaches out to grasp a water bottle during a handover.

component of the planned study is the development of an algorithm for the prediction of the timing and location of handover events. In previous work on this task Strabala, Lee, Dragan, Forlizzi, Srinivasa, Cakmak, and Micelli developed a decision tree classifier which predicts the timing and location of a handover event from hand-annotated data in a human-human study [16, 17]. In another study, Basili, Huber, Brandt, Hirche and Glasauer [2] measured kinematic parameters of human-to-human handovers, finding that handovers are usually initiated by the giver approximately two meters from the receiver and tend to take place around the halfway point between the giver and receiver. Basili et al. also present evidence which suggests that these parameters are independent of the receiver's actions [2].

The intention of the system under construction at the CARIS Lab is to predict the time and location at the hand of a human giver attempting to hand an object to its intended recipient will stop its motion. The predictor works on a model based on the motion of the arm with which the giver attempts to hand the object over. Arm motion was chosen because of the wide availability skeleton trackers, such as the NITE skeleton tracker, which work with 3D sensors such as the Microsoft Kinect. Arm motion was also chosen because of the difficulty of obtaining a reliable signal from head pose, and the desire to avoid instrumenting study participants with head-mounted eye tracking equipment in order to utilize eye tracking data. The technique currently being pursued is the construction of a system that predicts the final position and timing of the handover from an extrapolation of the observed trajectory of the giver's arm.

Though the trajectory-based predictor is still under construction, preliminary results are promising. Figure 4 shows images of a recorded motion trajectory and an equivalent model trajectory based on a least-squares linear fit. Skeletal poses rendered in blue come directly from data sampled by a Microsoft Kinect 3D sensor using the NITE skeleton tracker. Poses rendered in red are generated from the fit trajectory model, and are chosen to match the timestamp corresponding to the sample data that they are rendered with.

To determine the performance of the handover predictor, a preliminary study was performed. For the study, 18 participants (14 male, 4 female; between the ages of 23 and 29) were recruited from the University of British Columbia student community. Prior to participation in the study, participants were provided with informed consent forms as per university Behavioral Research Ethics Board protocols. They were instructed to hand a small object, a foam, brain-shaped "stress ball," measuring 70mm×46mm to each other.

The results of the study indicated that the predictor is able to determine the final handover location to within $\sim$ 9cm and $\sim$ 0.15 seconds, but only after witnessing almost the entire sampled trajectory. In order to be of use, the predictor must be able to make predictions of the timing and location of the handover early in the process. Currently, the predictor is being refined to use a richer model of human motion and incorporate constraints based on the observations of other groups working in this area [16, 17, 2].

4. SUMMARY

This abstract has discussed two projects in progress at the Collaborative Advanced Robotics and Intelligent Systems (CARIS) Laboratory at The University of British Columbia as part of the Collaborative Human-Assitive Robotics for Manufacturing (CHARM) project. These projects are based on experiences developing a robot assistant, and a collaborative assembly scenario in which the robot assistant aids a human operator in the task of assembling a car door.

The first of these projects aims to construct models of human task performance in terms of time to completion in order to enable planning and scheduling software in a robotic system to utilize these data during the construction of plans for a collaborative assembly task. The second of these projects is intended to construct a predictor for the timing and location of the end of the arm motion of a human presenting an object to another agent for a handover. Both of these projects are intended to explore the use of predictive models of timing in order to improve fluency in human-robot interaction.

5. ACKNOWLEDGMENTS

This work has been possible thanks to support from General Motors and the Natural Sciences and Engineering Research Council of Canada under contract 140570247. The study was approved by the University Behavioural Research Ethics Board, approval number H10-00503-A020.

6. REFERENCES

[1] H. Admoni, A. Dragan, S. Srinivasa, and B. Scassellati. Deliberate delays during robot-to-human handovers improve compliance with gaze communication. In *Proceedings of the 9th ACM/IEEE International Conference on Human-Robot Interaction (HRI '14)*, Bielefeld, Germany, March 2014.

[2] P. Basili, M. Huber, T. Brandt, S. Hirche, and S. Glasauer. Investigating human-human approach and hand-over. In H. Ritter, G. Sagerer, R. Dillmann, and M. Buss, editors, *Human Centered Robot Systems*, volume 6 of *Cognitive Systems Monographs*, pages 151–160. Springer Berlin Heidelberg, Berlin, Heidelberg, 2009.

[3] M. Cakmak, S. S. Srinivasa, M. Lee, S. Kiesler, and J. Forlizzi. Using spatial and temporal contrast for fluent robot-human hand-overs. pages 489–496, New York, New York, USA, March 2011.

[4] W. P. Chan, C. A. C. Parker, M. H. F. Van Der Loos, and E. A. Croft. Grip forces and load forces in handovers: Implications for designing human-robot handover controllers. In *Proceedings of the 2012 ACM/IEEE International Conference on Human-Robot Interaction (HRI '12)*, pages 9–16, Boston, MA, USA, March 2012.

[5] W. P. Chan, C. A. C. Parker, M. H. F. Van Der Loos, and E. A. Croft. A human-inspired object handover controller. *International Journal of Robotics Research*, 32(8):971–983, 2013.

[6] J. W. Hart. *Robot Self-Modeling*. PhD thesis, Yale University, New Haven, CT, USA, August 2014.

[7] J. W. Hart, B. Gleeson, M. Pan, A. Moon, and E. A. Croft. Gesture, gaze, touch, and hesitation: Timing cues for collaborative work. In *HRI Workshop on Timing in Human-Robot Interaction*, Bielefeld, Germany, 2014.

[8] B. Hayes, E. C. Grigore, A. Litoiu, A. Ramachandran, and B. Scassellati. A developmentally inspired transfer learning approach for predicting skill durations. In *Proceedings of the 4th joint IEEE International Conference on Development and Learning and Epigenetic Robotics*, Genoa, Italy, October 2014.

[9] G. Hoffman and C. Breazeal. Cost-based anticipatory action selection for human-robot fluency. *IEEE Transactions on Robotics*, 23(5):952–961, 2007.

[10] M. Huber, M. Rickert, A. Knoll, T. Brandt, and S. Glasauer. Human-robot interaction in handing-over tasks. In *The 17th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN 2008)*, pages 107–112, Munich, Germany, August 2008.

[11] A. H. Mason and C. Mackenzie. Grip forces when passing an object to a partner. *Experimental brain research. Experimentelle Hirnforschung. Expérimentation cérébrale*, 163(2):173–187, May 2005.

[12] D. McDermott, M. Ghallab, A. Howe, C. Knoblock, A. Ram, M. Veloso, D. Weld, and D. Wilkins. Pddl — The Planning Domain Definition Language. Technical Report CVC TR98003/DCS TR1165, Yale Center for Computational Vision and Control, New Haven, CT, October 1998.

[13] A. Moon, D. M. Troniak, B. Gleeson, M. K. Pan, M. Zheng, B. A. Blumer, K. MacLean, and E. A. Croft. Meet me where I'm gazing: How shared attention gaze affects human-robot handover timing. In *Proceedings of the 2014 ACM/IEEE international conference on Human-Robot interaction (HRI '14)*, pages 334–341, New York, NY, USA, 2014.

[14] B. Mutlu, F. Yamaoka, T. Kanda, H. Ishiguro, and N. Hagita. Nonverbal leakage in robots: Communication of intentions through seemingly unintentional behavior. In *Proceedings of the 4th ACM/IEEE International Conference on Human-Robot Interaction (HRI '09)*, pages 69–76, La Jolla, CA, USA, 2009.

[15] S. Shibata, B. M. Sahbi, K. Tanaka, and A. Shimizu. An analysis of the process of handing over an object and its application to robot motions. In *Proceedings of the International Conference on Systems, Man, and Cybernetics*, pages 64–69, Orlando, Florida, USA, October 1997.

[16] K. Strabala, M. K. Lee, A. Dragan, J. Forlizzi, and S. S. Srinivasa. Learning the communication of intent prior to physical collaboration. In *The 21st IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN 2012)*, pages 968–973, Versailles, France, 2012.

[17] K. Strabala, M. K. Lee, A. Dragan, J. Forlizzi, S. S. Srinivasa, M. Cakmak, and V. Micelli. Towards seamless human-robot handovers. *Journal of Human-Robot Interaction*, 1(1):1–23, 2013.