HUMANOIDS 2017 - Workshop “Human-Humanoid collaboration: the next industrial revolution?”

Human-Robot Physical Co-Manipulation of an Extended Object

Erich A. Mielke Eric C. Townsend David Wingate Marc D. Killpack

I. Introduction

Human teams are able to perform co-manipulation tasks with minimal communication. In the future, robots will work alongside humans in many applications including logistics, healthcare, agriculture, disaster response, and others. The advantage of human-robot collaboration in these areas is that humans provide intelligence and dexterity while robots may provide strength and stability [1]. In these situations, robots will need to work safely and intuitively, in order to be an asset when interacting with people. Specifically, they will need to be able to predict and respond to human intent in an effective manner. This paper proposes methods for predicting human intent in a co-manipulation carrying task that are based on human-human data [2]. For clarity, we define human motion intent as the intent to move in a particular direction with a particular velocity.

In order to understand human-human co-manipulation of rigid, extended objects, we ran an exploratory study with 21 human dyads. Each dyad moved a long board representing a table as we measured their motion and forces on the board. These trials are detailed in [2], and a video of the most complex task can be seen at `https://youtu.be/DAbLRDN20yE`. During the experiment, each partner was assigned a role as leader or follower. The leader was then given instructions on where to move the table. The performance in these trials could be considered the long-term objective for performance and something against which we can benchmark human-robot controllers for the application shown in Figure 1. Other past research studies have been completed analyzing human movement, but most past studies have only examined a limited number of degrees of freedom or were haptic simulations instead of including large, real, extended objects that needed to be rotated.

In this workshop paper, we outline our method for predicting human intent using neural nets. We then describe a preliminary human-robot user study that we are currently in the process of performing where we are comparing a controller that uses the neural net prediction versus a controller that uses thresholds and trends from our human-human data to extend past work on variable impedance control for human-robot collaboration.

II. Neural Networks For Estimating Human Motion Intent

As a first approach to developing a nonlinear estimator of human intention, we formulated a neural network using the Google Tensorflow API. The objective of the estimator is to determine how the robot should respond to human intent, i.e. how it should move to achieve the human partner’s goal. The intent estimator should allow us to control the object with a control loop similar to that seen in Fig. 2.

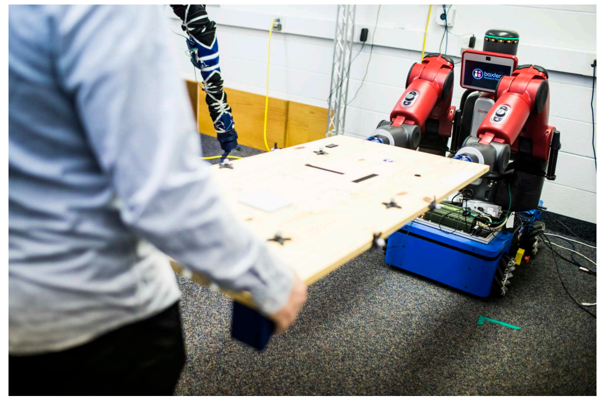

Fig. 1: Rethink Robotics Baxter robot mounted on a holonomic base and co-manipulating a table.

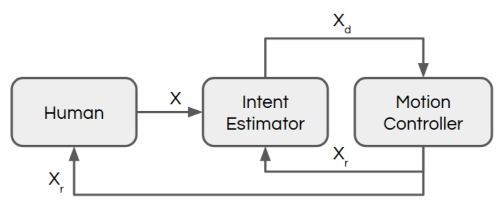

Fig. 2: Control loop structure of intent estimation in co-manipulation. The human moves the co-manipulated object, and object motion, x, is fed to an intent estimator, which determines a desired motion of the robot, x_d.

Because our human-human data that we collected in [2] considered the interaction between a human leader and a human follower, the input x, could be considered what the leader did–applied forces or moved the object–to indicate their intent to the follower. The follower then deciphered the intent, x_d, and moved as they believed appropriate, x_r. The leader then reacted to this motion.

Our approach to developing the Neural Net follows most closley work by Martens et al. which showed how Recurrent Neural Networks (RNNs) could be used in predicting text. In considering what form our neural network should take, this model proved to be the most applicable to our work. From our exploratory study, we had sequences of forces applied to and motion of a table that could be used as inputs to an RNN. We determined that we could input the force and motion data and get a motion prediction as an output, similar to how RNNS are used for predicting text. This prediction encapsulates the human intent, and provides a goal for the robot to achieve.

After deciding on this specific RNN structure, we determined that we would only previous motion data as inputs. The reason for using only motion data is that neural networks we trained that included force data had poor convergence and prediction accuracy. It is possible that using a more specific structure for how forces influence dynamic models could have improved the prediction accuracy when including force data, but we have left this for future work. The basic structure of the neural network is shown in Fig. 3.

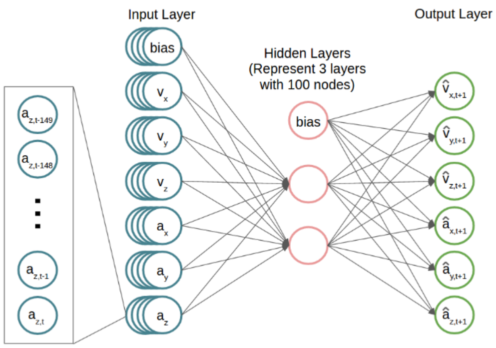

Fig. 3: Basic neural network structure. Time-series motion data inputs (left) enter the network and predicted velocities are given as outputs (right).

Better structures, including methods other than neural networks, may exist but our purpose in this paper is to show that human intent estimation is possible based on the data collected from our user study. It was shown by Chipalkatty et al. that more complex predictions of future movement can actually decrease performance if they do not agree with what the human is trying to do. They found that it was more important that the human understand what the robot will do next, meaning that our controller should be intuitive for a human partner in a human-robot dyad [**?**]. The inputs to the neural network, as seen in Fig. 3, are 150 past steps of velocity and acceleration in the x, y, and z direction, $\{x_{t-149}, x_{t-148}..., x_{t-1}, x_t\}$. The outputs are the predicted velocity and acceleration of the table in the x, y, and z direction for 1 time step into the future, $\hat{x}_{t+1}$, where $\hat{x}$ indicates a predicted value. Details of the RNN inputs, outputs, and training implementation, although important, are outside the scope of this workshop paper. However, Figure 4 shows the neural network predictions of velocity in the x direction for a single representative human-human co-manipulation task. The actual velocity is shown for the whole task in blue. The predicted velocity is shown in red starting every second and each one continues for 50 time steps or .25 seconds. As seen, the predictions are accurate for that time scale. Here we only show velocity, but the acceleration data must also be predicted because each velocity prediction depends on the prediction of acceleration for the time step before it. Without acceleration data, the neural net performance degrades.

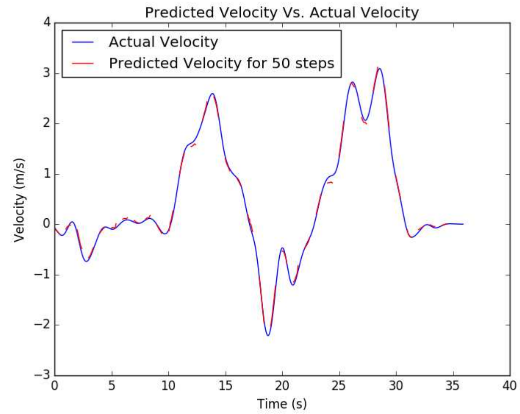

Fig. 4: Comparison of velocity prediction to actual future data in the x direction while a human dyad moves a table. Each red line is a separate 50 step prediction.

III. Pilot Study

The purpose of predicting human intent was to have the robot use the prediction. Our robot platform for this research is a Rethink Robotics Baxter robot mounted on an AMP-I holonomic base from HStar Technologies as seen in Figure 1. We chose to use a holonomic base with mecanum wheels instead of something like a bipedal robot in order to validate the human intent prediction at speeds similar to two humans moving an object in every day life. This is important to ensure that our method works in real world applications as limiting speed may affect the dynamics of the interaction. For our experiments the Baxter arms were rigidly attached to the table. Each arm runs a low impedance controller with a commanded joint angle specified by the position of the arm relative to the table before the task begins.

We are currently testing two different types of control. One that uses the RNN prediction of intent as a velocity command to the mobile base (as seen at `https://youtu.be/L0uy7-DMM3s`). The other controller is based on force and torque thresholds from the human-human data we collected and that the robot can measure and react to using force-torque sensors at its end effectors. The details of the control development is unfortunately also outside the scope of this workshop paper.

References

[1] H. Kazerooni, "Human-Robot Interraction via the Transfer of Power and Information Signals," *IEEE Transactions on Systems, Man, and Cybernetics*, vol. 20, no. 2, pp. 450–463, 1990. [Online]. Available: http://bleex.me.berkeley.edu/wp-content/uploads/hel-media/Publication/SMC.HumanRobotInteractPowerInfoSigs.V20N2.03.1990.pdf

[2] E. A. Mielke, E. C. Townsend, and M. D. Killpack, "Analysis of Rigid Extended Object Co-Manipulation by Human Dyads: Lateral Movement Characterization," 2017. [Online]. Available: http://arxiv.org/abs/1702.00733