Enabling Synchronous Joint Action in Human-Robot Teams

Samantha Rack, Tariq Iqbal, and Laurel D. Riek
Dept. of Computer Science, University of Notre Dame
{srack,tiqbal,lriek}@nd.edu

ABSTRACT

Joint action is an increasing area of interest for HRI researchers. To be effective team members, robots need to be able to understand, anticipate, and react appropriately to high-level human social behavior. We have designed a new approach to enable an autonomous robot to act fluently within a synchronous human-robot team. We present an initial description and validation study of this approach. Using a synchronous dance scenario as an experimental testbed, we found that our robot was able to execute appropriate actions using our method. Moving forward, we aim to extend this method by developing predictions for the robot's actions using an understanding of the group's dynamics. Our method will be helpful to other researchers working to achieve fluency of action within human-robot groups.

Categories and Subject Descriptors

I.2.9 [**Artificial Intelligence**]: Robotics

1. INTRODUCTION

In the robotics community, we have gained great proficiency at building robots to solve problems autonomously in fully controlled environments, such as manufacturing. However, as robots leave controlled spaces and begin to work alongside people, many assumptions roboticists make about perception and action do not apply. Unlike factories, people act unpredictably - they occlude sensors, reconfigure their living spaces, and otherwise "break the rules" when it comes to what a robot can expect *a priori* [10].

In order for robots to effectively integrate in Human Social Environments (HSEs), they must be able to comprehend high-level social signals and respond appropriately [10]. Joint action is one naturally present high-level social signal that occurs when multiple people coordinate their actions in time and space to make a change to their environment [11]. In the HRI community, researchers have been working in the area of *joint action*, where a robot can interpret and predict the intentions of groups to better inform actions, and lead to more *fluent* interactions [12, 3, 2, 9, 1, 7, 8, 5, 6].

HRI'15 Extended Abstracts, Mar 02-05 2015, Portland, OR, USA
ACM 978-1-4503-3318-4/15/03.
http://dx.doi.org/10.1145/2701973.2702064

Recent work has focused on developing predictive methods for improving the fluency of interaction between a robot and one or more humans. For example, a method developed by Hawkins et. al. [2] uses a probabilistic model that accounts for human inconsistencies to determine appropriate actions for a robot providing assembly parts. Further, Hoffman et. al. [4] developed a robot that improved its performance on a human-robot team by considering both its current perceptions and its anticipations based on what it had already experienced. While this work improves the ability of robots to have fluent interactions with one or more individuals in HSEs, the method we present incorporates an understanding of the behavior of the group as a whole in order to inform robot action.

We have previously proposed an event-based method for measuring the degree of synchrony of a group when members are performing a joint action task [7, 6]. Using this method, a robot can recognize the presence of joint action in HSEs, but in order to be effective and capable collaborators, robots must also be capable of response and participation. We plan to use this method to enable robots to make appropriate decisions that consider high-level group dynamics during a joint action collaboration with a multi-human team.

Our method is part of an ongoing research effort to develop approaches for enabling robots to plan and execute appropriate behaviors in HSEs, through real-time observations of the environment. As a first step in validating the method, we describe an experimental testbed involving a robot dancing synchronously with multiple human performers. In later sections, we describe the system setup, and discuss initial validation findings from a pilot study.

2. METHODOLOGY

When creating our experimental testbed, we had two major goals. The first was to design a scenario that allowed us to explore research questions regarding group motion and synchronization, while taking into account the capabilities of our non-holonomic drive mobile robot (Turtlebot). The second was to create a realistic data-capture situation, where people could freely perform tasks within the context of the scenario while still being observable by four external Kinect sensors. To that end, we selected an iterative, synchronous dance scenario, where a robot observes several iterations of a human performing the dance alone (Phase 1) and then itself joins in (Phase 2).

To design the dance, we consulted with several amateur dancers, and created a choreographed dance routine set to Michael Jackson's "Smooth Criminal". Smooth Criminal is a song in 4/4 timing, and the choreographed routine includes simple forward, backward, and rotational movements. The dance is iterative, and performed cyclically in a counter-clockwise manner. Each phase includes four repetitions of the following steps: two forward and backward movements, two claps, and a 90-degree turn.

To enable the robot's motion, four extrinsic sensors were placed counterclockwise around the dance floor. These were tightly-synchronized Microsoft Kinect v2 sensors, each of which captured depth, skeletal, and audio data [5]. The sensors were connected to four computers that served as clients in our system architecture. From the captured sensor data, the clients detected the high level movements performed by the human (e.g., when a person started moving forward).

After detecting an event, each client sent an event message to a server, indicating the type, timing, and attributes of the event (e.g., distance traversed). The server maintained synchronized time across all clients and the robot. The server aggregated event data for prediction, accepting event messages only from the client that the human and robot were facing at the given time. During Phase 2, the server sent movement commands to the robot at appropriate times.

For this dance routine, we only considered forward and backward motion and claps as potential movement patterns for the human performer. These motions were measured based on the change in joint positions over time in skeletal data. Forward and backward motion was detected when there was a sufficient change in the z component of the performer's spine base joint. A clap was detected when the hand joint positions of the performer reached a sufficiently small local minima in a time window. To dictate robot motion, the server sent commands with movement parameters.

In this experimental setup, the server sent movement instructions to the robot during the second phase of the dance, which were derived from the timing of events during the first phase of the dance. The robot executed those patterns during Phase 2 in synchronization with the human dancer. In Fig 1, we present Phase 2 of an experimental session where a human performer is dancing with a robot.

3. RESULTS AND DISCUSSION

To test our system architecture, as well as the movement patterns of the robots, we performed a pilot study with 5 participants (mean age = 27 yrs), 2 were female. Participants were opportunistically recruited.

Upon arrival, participants provided demographic information, viewed a short video explaining the dance routine, and had time to practice dancing along with the music. Following the practice session, we conducted the main study.

In the study, we sought to measure how accurately the robot performed movement patterns along with the human performer, as well as how timely action messages were received by the robot from the server. To acquire these measurements, we recorded the odometry data of the robot. We compared this data to velocity, distance, and time data that was used to dictate the actions of the robot.

From our results, we found that the robot received command messages from the server with a very short delay, likely due to network latency. From the data, we observed similar movement pattens performed by the human and the robot. However, in some cases, we observed some deviations in the robot's movement per movement step. These deviations in motion may have resulted from friction between carpet surface and the robot's wheels, or the physical limitations of the Turtlebot to perform at the speed required by the song.

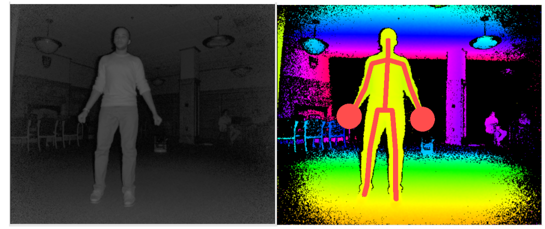

Figure 1: An experimental session where a person is dancing.

These preliminary findings will enable us to develop methods for generating more appropriate robot motion that considers the limitations of existing hardwares. Moving forward, we are working on developing methods to enable robots to predict future activities of team members and plan their actions using an understanding of the group's dynamics. Our method will be helpful to other researchers interested in achieving fluency of actions within human-robot teams.

4. ACKNOWLEDGEMENTS

This material is based upon work supported by the National Science Foundation under Grant No. IIS-1253935.

5. REFERENCES

[1] M. Cakmak, S. S. Srinivasa, M. K. Lee, S. Kiesler, and J. Forlizzi. Using spatial and temporal contrast for fluent robot-human hand-overs. In *ACM/IEEE Intl. Conf. on Human-Robot Interaction*, 2011.

[2] K. P. Hawkins, S. Bansal, N. N. Vo, and A. F. Bobick. Anticipating human actions for collaboration in the presence of task and sensor uncertainty. In *IEEE Int. Conf. Robot. Autom*, 2014.

[3] G. Hoffman and C. Breazeal. Cost-based anticipatory action selection for human–robot fluency. *T-RO*, 2007.

[4] G. Hoffman and C. Breazeal. Effects of anticipatory perceptual simulation on practiced human-robot tasks. *Autonomous Robots*, 2010.

[5] T. Iqbal, M. J. Gonzales, and L. D. Riek. A Model for Time-Synchronized Sensing and Motion to Support Human-Robot Fluency. In *ACM/IEEE Intl. Conf. on HRI, Workshop on Timing in HRI*, 2014.

[6] T. Iqbal, M. J. Gonzales, and L. D. Riek. Mobile robots and marching humans: Measuring synchronous joint action while in motion. *In AAAI Fall Symp. on AI-HRI*, 2014.

[7] T. Iqbal and L. Riek. Assessing group synchrony during a rhythmic social activity: A systemic approach. In *6th Conference of the International Society for Gesture Studies (ISGS)*, 2014.

[8] T. Iqbal and L. D. Riek. Role distribution in synchronous human-robot joint action. *IEEE RO-MAN, Towards a Framework for Joint Action Workshop*, 2014.

[9] T. Lorenz, A. Mortl, B. Vlaskamp, A. Schubo, and S. Hirche. Synchronization in a goal-directed task: human movement coordination with each other and robotic partners. In *IEEE RO-MAN*, 2011.

[10] L. D. Riek. The social co-robotics problem space: Six key challenges. In *Robotics: Science, and Systems (RSS), Robotics Challenges and Visions*, 2013.

[11] N. Sebanz, H. Bekkering, and G. Knoblich. Joint action: bodies and minds moving together. *T. Cogn. Sci.*, 2006.

[12] R. Wilcox, S. Nikolaidis, and J. A. Shah. Optimization of temporal dynamics for adaptive human-robot interaction in assembly manufacturing. In *Robotics: Science, and Systems (RSS)*, 2012.