Mobile Robots and Marching Humans: Measuring Synchronous Joint Action While in Motion

Tariq Iqbal, Michael J. Gonzales, and Laurel D. Riek
Department of Computer Science and Engineering,
University of Notre Dame, Notre Dame, IN, USA.

Abstract

It is challenging to build socially-aware robots due to the inherent uncertainty in the dynamics of human behavior. To become socially-aware, robots need to be capable of recognizing activities in their environment to make informed actions in concert with co-present humans. In this paper, we present and validate an event-based method for robots to detect synchronous and asynchronous actions of humans when working as a team in a human-social environment. Our results suggest that our method is capable of detecting synchronous and asynchronous actions, which a step towards building socially aware robots.

Introduction

Robots are becoming a ubiquitous technology, working alongside humans as team members in many fields, from manufacturing and assembly processes, to assistive technologies that help people with disabilities (Wilcox, Nikolaidis, and Shah 2012; Fasola and Mataric 2013). However, for robots to become capable team members in human-social environments (HSEs), they must have a clear understanding of the people and events happening around them (Riek 2013; Hayes, O'Connor, and Riek 2014).

In particular, robots require the ability to interpret and predict team member activities in order to inform their own actions. If a robot can make better sense of its environment, its interactions with humans in HSEs can reach higher levels of coordination, leading to a *fluent* meshing of their actions (Hoffman and Breazeal 2007; Iqbal, Gonzales, and Riek 2014; Cakmak et al. 2011). If a robot does not understand the activities occurring around it and the context of those activities, then it is more likely to be error-prone, and less likely to be accepted by co-present humans (Riek 2013).

Strides have been made in the fields of artificial intelligence, robotics, and computer vision to improve robots' understanding of their environment (Ryoo and Matthies 2013; Sung et al. 2011). However, recognizing high-level human actions occurring in the environment is still a difficult problem for robots due to the inherent uncertainty of the dynamics within an HSE. This problem is even more difficult when the HSE involves robots and humans in motion.

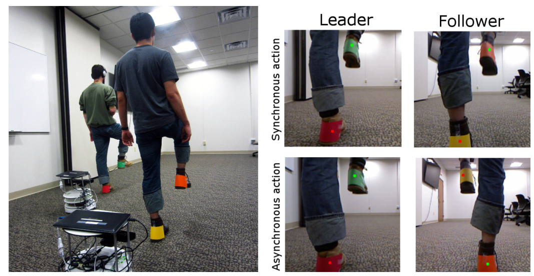

Figure 1: A) Two humans marching synchronously, while two mobile robots following them. B) The synchronous and an asynchronous joint action conditions.

In human-human interaction, synchronous motion, or joint action, is a common phenomena. Joint action is a form of social interaction where two or more participants coordinate their actions both in space and time while making changes to their environment (Sebanz, Bekkering, and Knoblich 2006). In the context of a human-robot interaction (HRI) scenario, understanding joint action is particularly important as it may improve the overall engagement of a robot as a team member. For example, in collaborative manipulation tasks, a robot can be a more effective teammate by being able to anticipate a co-human's motion (Unhelkar, Siu, and Shah 2014; Strabala et al. 2013). Similarly, to move "in-step" with a group of people (e.g., while collaboratively climbing, running, or dancing), a robot needs to understand the motion of its co-humans and predict their future moves.

In this paper, we describe a novel method to automatically detect human-human synchronous and asynchronous actions, and discuss its application and validation on a human-robot teamwork scenario. Our method takes multiple types of task-level events into consideration, and can detect both synchronous and asynchronous human motion. Our method is also robust to both static and mobile sensors, i.e., when both the robots and humans are in motion.

Methodology

We designed an experimental scenario where two humans and two robots performed a team activity while in motion. The humans performed a dynamic and rhythmic activity (marching), while the robots acted as observers. We measured the synchronous and asynchronous actions of the human performers based on this observation by the robots.

In the experiment, the humans performed a "high march" action, either synchronously or asynchronously, depending upon the experimental condition. One person acted as the leader, and marched in a consistant pace. The second person acted as a follower, and stood to the right and approximately two feet behind the leader (See Fig. 1-A). The follower was instructed to march either synchronously or asynchronously with the leader based on the scenario (See Fig. 1-B).

We used two Turtlebot robots in this experiment. Each Turtlebot followed behind one of the human performers by approximately two feet (See Fig. 1-A). The robots recorded RGB and depth data from the environment using their attached Kinect sensors. The Turtlebots ran the Robot Operating System (ROS) on Ubuntu Linux.

We recorded data from a total of four scenarios during this experiment. Each marching scenario lasted approximately 35 seconds, and was timed using a stopwatch. In the first experimental scenario, we instructed the follower to perform the same marching pattern in synchrony with the leader. In the second scenario, the performers marched asynchronously for entire duration of the scenario (marching with opposite steps from the leader). For the third scenario, the follower started marching synchronously with the leader. After 12 seconds of marching, we instructed the follower to become asynchronous with the leader, then to again march synchronously after 24 seconds. For the fourth scenario, we instructed the follower to perform the same actions as the third scenario, but in reverse order.

We first detected the feet positions of the human marchers from the recorded RGB data of the mobile robots. We attached a unique small square of colored paper to the performers' left and right feet (See Fig. 1-B). We used a total of four different colors: orange, yellow, green, and red. We used a standard blob tracking technique from the ROS *cmvision* package on the recorded RGB data to track the feet of the performers.

We then defined two types of task-level events based on these positions. The first type of event was the point in time when a person began to raise their leg from the ground. The second type of event was the point in time when a leg reached its maximum height. As a result, a total of four types of events occurred (two for each leg).

To measure the synchronous and asynchronous actions from these events, we developed a method based on the work by Quian Quiroga et al. (Quiroga, Kreuz, and Grassberger 2002), which considers multiple types of task-level events together. Our method is more accurate than other methods from the literature, which take a single type of event for the measurement of synchrony. A detailed description of our method can be found in Iqbal and Riek (Iqbal and Riek 2014a; 2014b).

For our method, first we express the events associated with each person with a time series, measure the pairwise synchronization index involving a single type of event, and finally measure the overall pairwise synchronization index while taking multiple types of events together. To explain this mathematically, suppose, x_n and y_n are the two time

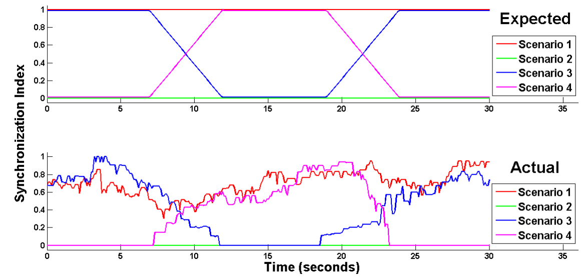

Figure 2: A) The expected synchronization indices over time of our experimental scenarios in an ideal setting. B) Measured synchronization indices using our method.

series, where $n = 1, \ldots, N$ (N samples). For each event type $e_i \in E$, $m_x(e_i)$ and $m_y(e_i)$ are the number of events occurring in x and y respectively, where E is the set of all events. Now, for each event type $e_i \in E$, we calculate the pairwise synchronization index ($Q(e_i)$). Then, the overall pairwise synchronization index (Q) considering all events is calculated as:

$$\forall e_i \in E : Q = \frac{\sum [Q(e_i) \times [m_x(e_i) + m_y(e_i)]]}{\sum [m_x(e_i) + m_y(e_i)]} \quad (1)$$

Results

The synchronization indices for these four experimental scenarios are presented in Fig. 2-A. We used a sliding window of five seconds for this experiment.

We expected to see a high value for a synchronization index for the entire duration of a session for Scenario 1, and a value of zero for Scenario 2. For Scenario 3, we expected to see our measured synchronization index decrease beginning around seven seconds to a value of zero at 12 seconds, and increase again at about 20 seconds. For Scenario 4, we expected similar results, however in reverse order. From Fig. 2-B, one can see that the measured synchronization indices over time reasonably match what we expected.

Discussion

Our work addresses the problems of detecting multiple task-level events and measuring the synchrony of a human-robot team while both the humans and robots are in motion. Results of this study suggest that despite the difficulties a robot experiences in recognizing high-level group tasks, our method is capable of detecting these task-level events and measuring synchrony successfully in scenarios involving movement.

This work is directly applicable to a number of fields, including HRI, social signal processing, and artificial intelligence. In addition, this work may directly support researchers exploring human-robot fluency, a first step in enabling the automatic interpretation of synchronized human-robot interaction, including gesture mimicry or non-verbal expression. Moving forward, we aim to apply this research toward the development of real-time robotic systems with the capability of understanding high-level group behavior to inform more appropriate actions in HSEs.

References

Cakmak, M.; Srinivasa, S. S.; Lee, M. K.; Kiesler, S.; and Forlizzi, J. 2011. Using spatial and temporal contrast for fluent robot-human hand-overs. In *Proceedings of the 6th International Conference on Human-robot interaction*, 489–496. ACM.

Fasola, J., and Mataric, M. 2013. A Socially Assistive Robot Exercise Coach for the Elderly. *J. Human-Robot Interact.* 2:3–32.

Hayes, C. J.; O'Connor, M. F.; and Riek, L. D. 2014. Avoiding robot faux pas: Using social context to teach robots behavioral propriety. In *Proceedings of the 2014 ACM/IEEE International Conference on Human-robot Interaction*, HRI '14, 168–169.

Hoffman, G., and Breazeal, C. 2007. Cost-based anticipatory action selection for human–robot fluency. *IEEE T Robotics*.

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

Iqbal, T., and Riek, L. D. 2014b. Role distribution in synchronous human-robot joint action. *In Proc. of the 23rd IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN), Towards a Framework for Joint Action Workshop*.

Iqbal, T.; Gonzales, M. J.; and Riek, L. D. 2014. A Model for Time-Synchronized Sensing and Motion to Support Human-Robot Fluency. In *ACM/IEEE International Conference on Human-Robot Interaction (HRI), Workshop on Timing in HRI*.

Quiroga, R. Q.; Kreuz, T.; and Grassberger, P. 2002. Event synchronization: a simple and fast method to measure synchronicity and time delay patterns. *Phys Rev E* 66:041904.

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

Ryoo, M. S., and Matthies, L. 2013. First-Person Activity Recognition: What Are They Doing to Me? *2013 IEEE Conf. Comput. Vis. Pattern Recognit.*

Sebanz, N.; Bekkering, H.; and Knoblich, G. 2006. Joint action: bodies and minds moving together. *Trends Cogn. Sci.*

Strabala, K. W.; Lee, M. K.; Dragan, A. D.; Forlizzi, J. L.; Srinivasa, S.; Cakmak, M.; and Micelli, V. 2013. Towards seamless human-robot handovers. *Journal of Human-Robot Interaction* 2(1):112–132.

Sung, J.; Ponce, C.; Selman, B.; and Saxena, A. 2011. Human Activity Detection from RGBD Images. *Plan, Activity, and Intent Recognition*.

Unhelkar, V. V.; Siu, H. C.; and Shah, J. A. 2014. Comparative performance of human and mobile robotic assistants in collaborative fetch-and-deliver tasks. In *Proceedings of the 2014 ACM/IEEE international conference on Human-robot interaction*, 82–89. ACM.

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